CENTRE FOR
FINANCIAL RISK
*Faculty of Business
and Economics*

MACQUARIE
University

A Transparent Parametrization of the Lee-Carter Model Based on "Needed Exposure"

WORKING PAPER 16-08

Piet de Jong, Leonie Tickle and Jianhui Xu

A transparent parameterization of the Lee–Carter model based on "Needed Exposure"

Abstract

An alternate parameterization of the Lee–Carter model is introduced, called the Transparent Lee–Carter (TLC) model. The TLC model facilitates direct and transparent interpretation of fitted parameters in terms of "needed–exposure" (NE) quantities. The NE is the number required in order to get one expected death and is closely related to the "needed–to–treat" measure used to communicate risks and benefits of medical treatments. The TLC model structures time series behaviour in terms of an overall across–age NE. Age parameters are interpretable as age–specific elasticities: percentage changes in the NE at a particular age in response to a percent change in the overall NE. The TLC model is informative, intuitive and simple to apply for both mortality analysis and forecasting.

Keywords: Mortality, Lee–Carter, needed–exposure, age–response, age–specific elasticities, needed–to–treat

1. Introduction

The Lee–Carter (LC) model (Lee and Carter 1992) has become the "*leading statistical model of mortality in the demographic literature*" (Deaton and Paxson 2004, p.264). It has served as a foundation for numerous related models and forecasting methods (e.g. Booth et al. 2002; Brouhns et al. 2002; Cairns et al. 2011; de Jong and Tickle 2006; Hyndman and Shahid Ullah 2007; Hyndman et al. 2013; Lee and Miller 2001; Li and Lee 2005; Li 2013; Plat 2009; Renshaw and Haberman 2006). Lee–Carter and related models are used by demographers, actuaries, statisticians, economists and others to produce forecasts that inform government decision–making, that enable the effective operation of insurers, pension funds and financial markets, and that enhance understanding of mortality risk (e.g. Deng et al. 2012; Hollmann et al. 1999; Janssen et al. 2013; Li and Hardy 2011; Niu and Melenberg 2014; Stoeldraijer et al. 2013; Tuljapurkar et al. 2000; Yang and Wang 2013; Zhou et al. 2014). Reviews, including those by Booth and Tickle (2008), Wong-Fupuy and Haberman (2004), and Cairns et al. (2008), outline these methodological developments and applications.

Although widely used in forecasting, the LC model is formulated in a way that hampers interpretation and comparison between populations. The LC model expresses the log–central mortality rate at age x in year t, $\ln m_{xt}$, as a function of age parameters a_x and b_x, an underlying time series process k_t, and error terms ϵ_{xt}:

$$\ln m_{xt} = a_x + b_x k_t + \epsilon_{xt} \ , \qquad x = 0, \ldots, p \ , \qquad t = 1, \ldots, N \ . \tag{1}$$

Here there are $p + 1$ ages x and N time periods t of observation. In applications the age range may start from for example 1 or some other age.

The equations in (1) must be augmented with constraints on the a_x, b_x and k_t to ensure b_x and k_t are identified. The usual constraints (Lee and Carter 1992) are

$$\sum_x b_x = 1 \ , \qquad \sum_t k_t = 0 \ . \tag{2}$$

These two constraints together with $\sum_t \epsilon_{xt} = 0$ imply the estimated a_x is the average log mortality at age x across time.

Lee and Carter (1992) used the constraints (2) in conjunction with the singular value decomposition (SVD) for estimation. They also advocated a second stage adjustment whereby k_t is re–estimated to equate observed and estimated total deaths in each year so as to avoid large discrepancies that may result from modeling the log–rates. Alternative adjustments have been proposed by Booth et al. (2002) and Lee and Miller (2001). Brouhns et al. (2002) apply the same constraints and avoid the need for k_t adjustments and SVD through direct use of a Poisson model for deaths fitted by maximum likelihood estimation.

With the constraints (2), k_t is interpreted as an index of the level of mortality at time t, and b_x represents the response at age x to changes in the overall level of mortality over time. These constraints imply the k_t and b_x parameters are interpretable only in combination with or in relation to other values. For example, while the change in k_t over time reflects the overall change in mortality, the meaning of the value of k_t at a single point in time and the way that different ages have been combined in its calculation are not made obvious.[1] Likewise, a large absolute value of b_x indicates that the relevant age group is highly–responsive to changes in overall mortality levels, but the actual mortality change over time can only be determined by combining information about b_x and the change in k_t. The parameter values have meaning only relative to or in conjunction with other values and a different but equally valid normalization would result in completely different parameter values.

The Lee–Carter model is usually applied for the purpose of forecasting, and in this context, the lack of interpretability of the individual parameter values is not generally of concern. Provided that the time parameter k_t is forecast in a way that is invariant to the chosen normalization, the forecast rates will be appropriate. However, selection of a normalization that does allow interpretability would bring the advantage of clarifying the meaning of model parameters, assisting in the appropriate application of the model for forecasting. More importantly, interpretable Lee–Carter parame-

[1] The constraints (2) and $\sum_x \epsilon_{xt} = 0$ imply k_t is the sum over all ages of the excess of $\ln m_{xt}$ over the average (over time) $\ln m_{xt}$.

ters would facilitate the potential use of Lee–Carter models to encompass mortality modeling, analysis and comparison in addition to forecasting.

A further barrier to interpretability of Lee–Carter, and indeed many mortality indices and models, is that these are often based on measures of the mortality rate or probability of death. It is well–known that people find probabilities, particularly when low, difficult to interpret (Camerer and Kunreuther 1989; Kunreuther et al. 2001). Although demographers are intimately familiar with the use of rates and generally regard them as the fundamental unit of analysis in their study of mortality (Wilmoth 2000), "*even an expert in the mortality field has difficulty interpreting the meaning of an improvement in the mortality rate*" (Pollard 1988, p. 265). It is therefore worth exploring whether Lee–Carter can be reformulated in terms of alternative mortality measures that are more intuitive, more readily interpretable and transparent.

In this paper, we propose a reformulation and normalization of the Lee–Carter model to make its parameters more transparent and interpretable. The Transparent Lee–Carter (TLC) model based on needed–exposures is introduced in section 2 and interpreted in section 3. Applications are presented in sections 4 and 5, and are discussed in section 6.

2. Transparent Lee–Carter (TLC) model parameterization

The reparameterization of the LC model proposed in this paper consists of three changes. Each change is, in and of itself, trivial. Together they permit more transparent and insightful analysis of the LC output.

The first change is to replace m_{xt} by $n_{xt} = 1/m_{xt}$. The quantity n_{xt} is the "needed exposure" (NE) or population size at age x in year t to produce one expected death aged x. In terms of NE, (1) is

$$\ln n_{xt} = -a_x - b_x k_t - \epsilon_{xt} \ . \tag{3}$$

The second change is to replace the usual constraints (2) with flexible constraints which can be tailored to suit. For any scalars μ and $\sigma \neq 0$

$$-a_x - b_x k_t = -a_x - \mu b_x - \sigma \left(\frac{k_t - \mu}{\sigma} \right) b_x = \alpha_x + \beta_x \ln n_t \ , \tag{4}$$

where

$$\alpha_x = -a_x - \mu b_x \ , \qquad \beta_x = \sigma b_x \ , \qquad n_t = \mathrm{e}^{-(k_t - \mu)/\sigma} \ . \tag{5}$$

Combining the first and second changes results in

$$\ln n_{xt} = \alpha_x + \beta_x \ln n_t + \epsilon_{xt} \ . \tag{6}$$

The third change is to chose μ and σ in $\ln n_t = (\mu - k_t)/\sigma$ so that n_t is interpretable as an average (across-age) n_{xt} measure. There are a number of possibilities with the actual choice guided by the aims of the mortality analysis. The choice of μ and σ is conveniently cast in a weighted least squares framework:

$$\min_{\mu,\sigma} \sum_{x,t} w_{xt} (\ln n_{xt} - \ln n_t)^2 \ , \qquad \ln n_t = \frac{\mu - k_t}{\sigma} \ , \tag{7}$$

where the $w_{xt} \geq 0$ are weights. Since k_t is independent of age x the least squares criterion in (7) reduces to

$$\min_{\mu,\sigma} \sum_t w_t \left\{ \mathcal{E}(\ln n_{xt}) - \frac{\mu - k_t}{\sigma} \right\}^2 \ , \tag{8}$$

where $\mathcal{E}(\ln n_{xt})$ is the weighted average of the $\ln n_{xt}$ at time t using the weights w_{xt} and $w_t = \sum_x w_{xt}$ is the total weight at time t.

The weights w_{xt} determine the emphasis given to the experience at different ages x and times t in the normalization of the overall n_t. The choice of weights is guided by the purpose of the modeling. Possible choices include:

1. If all $w_{xt} = 1$ then all time periods and ages are equally important in the determination of the $\ln n_t$ and

$$\mathcal{E}(\ln n_{xt}) = \frac{1}{1+p} \sum_{x=0}^{p} \ln n_{xt} \ , \qquad t = 1, \ldots, N \ ,$$

   and $\ln n_t$ is the predicted average $\ln n_{xt}$ from a least squares fit of the latter on k_t. That is k_t is linearly transformed so it best predicts, in a least squares sense, the average of $\ln n_{xt}$ at each t.

2. The weights w_{xt} can be exposures at different ages. In this case $\mathcal{E}(\ln n_{xt})$ is an exposure weighted average of the $\ln n_{xt}$ and $\ln n_t$ is the linear prediction of the exposure weighted average using k_t.
3. With exposure weighting and an ageing population the weighted average $\mathcal{E}(\ln n_{xt})$ is increasingly influenced by older ages. A decline in the rate of mortality improvement may thus be a consequence of an ageing population rather than changes in actual mortality. To correct for such confounding the $w_{xt} = w_x$ may be held constant for each t with the w_x reflecting a fixed and standard population with respect to which mortality improvement is measured.
4. In the case where mortality is compared across populations and over time, using weights w_{xt} from a common "benchmark population" in a single year ensures differences can be attributed to mortality differentials rather than differences in weights.
5. The time series $\ln n_t$ can be tuned to particular ages or periods of time. For example setting all $w_{xt} = 0$ except for age $x \geq 65$ aligns $\ln n_t$ to the average $\ln n_{xt}$ at ages 65 and over. If $w_{xt} = 0$ except for two time points then $\ln n_t$ is forced through the weighted average $\ln n_{xt}$ at two time points.

The fitted n_t, when based on the linear alignment of k_t to $\mathcal{E}(\ln n_{xt})$, is interpreted as the NE to get one expected death where the one year mortality is the weighted geometric mean of the m_{xt}:

$$n_t = \prod_x \frac{1}{\hat{m}_{xt}^{w_{xt}}} \,, \tag{9}$$

where the $\hat{m}_{xt}$ are estimates based on the fitted LC model and it is assumed $\sum_x w_{xt} = 1$ for each t.

The NE (9) is formalised by considering a population with integer $w_{xt} \geq 0$ in age group x. Then the reciprocal of (9) is the probability the whole population dies within one year and hence (9) is the expected number of populations required to get one such event. Taking the $\sum_x w_x$ root of this number standardises on the population size and hence (9) is the number of populations required to get one expected complete population dying out, normalising on the population size.

Different normalizations of k_t, that is choices of μ and σ, will produce different parameter values but will not affect fitted log-mortalities or – if forecasts of the time series parameter are based on location–scale preserving models such as the random walk with drift (Nielsen and Nielsen 2014) – forecast log-mortalities. Further the k_t can be from any estimation method and normalization basis – for example they may be SVD estimates, Lee–Carter estimates, or Poisson Lee–Carter estimates.

Lee–Carter modeling and forecasting also typically involves fitting time series models to k_t. A usual time series model choice is the random walk with drift:

$$k_{t+1} = d + k_t + \eta_t \ , \tag{10}$$

where d is the drift and η_t is zero mean noise. It follows from (10) and (5)

$$\ln n_{t+1} = -\frac{k_{t+1} - \mu}{\sigma} = -\frac{d + k_t + \eta_t - \mu}{\sigma} = -\frac{d}{\sigma} + \ln n_t - \frac{\eta_t}{\sigma} \ ,$$

and hence n_t is a geometric random walk with drift $-d/\sigma$. From (8), $-1/\sigma$ is the weighted least squares coefficient when regressing $\mathcal{E}(\ln n_{xt})$ on k_t. The drift $-d/\sigma$ is thus a more cogent index of mortality improvement than d since n_t is aligned to actual mortality experience.

3. Interpreting the parameters of the TLC

A large value of n_t indicates low mortality and vice versa while increasing n_t indicates improving mortality, that is, fewer deaths. A linear increase in $\ln n_t$ indicates a constant percentage improvement in mortality. In particular if $\ln n_{t+1} - \ln n_t \approx \delta$, independent of t, then n_t is increasing by percentage δ per time period t. In other words weighted geometric mortality is declining by percentage δ. Note that percentage increases in n_t are equivalent to percentage decreases in weighted geometric average mortality and vice versa.

From (4) and (5)

$$\hat{n}_{xt} = \mathrm{e}^{\alpha_x} n_t^{\beta_x} \ , \qquad \beta_x = \frac{\mathrm{d} \ln(\hat{n}_{xt})}{\mathrm{d} \ln(n_t)} \ . \tag{11}$$

Here $\hat{n}_{xt}$ is the predicted NE at age x in year t, modeled as an age–specific transform of a common n_t where β_x is the "elasticity" at age x: percentage change in the NE at age x and given a one percent change in n_t. Any x such that $\beta_x > 1$ indicates superior percentage improvement at the given age x while $\beta_x < 1$ indicates inferior percentage improvement. Further

$$\frac{\mathrm{d}\ln(n_t\hat{n}_{xt})}{\mathrm{d}\ln(n_t)} = 1 + \beta_x$$

where $n_t\hat{n}_{xt}$ is the required population size in year t to produce one death aged x. Hence $1 + \beta_x$ is the percentage change in NE at age x given a one percent change in NE in year t.[2]

The actual NE at age x is higher or lower depending on α_x, interpreted as an adjustment factor with $\alpha_x > 0$ indicating NE at age x is higher: that is the mortality is lower. Using total derivatives

$$\mathrm{d}\ln\hat{n}_{xt} = \mathrm{d}\alpha_x + \ln n_t \mathrm{d}\beta_x + \beta_x \mathrm{d}\ln n_t \ . \tag{12}$$

This is the percentage change in $\hat{n}_{xt}$ given a unit change in both x and t, that is the cohort effect. If (12) is positive then increases in mortality due to increasing age are offset by a secular improvement in overall mortality. The separate percentage changes with respect to age x and year t are $\mathrm{d}\alpha_x + \ln n_t\mathrm{d}\beta_x$ and $\beta_x\mathrm{d}\ln n_t$, respectively.

4. Application to United States mortality data

Figure 1 shows estimated parameters of the LC model (1) with the usual constraints (2), and of the reformulated and normalized TLC model (4) and (5), applied to male and female United States mortality data. In both cases, estimation of parameters is by SVD. The data used is from the Human Mortality Database (2016) and comprises central mortality rates, m_{xt}, and mid-year populations by sex and single years of age $x = 0, 1, \ldots, 90$ for years 1933 to 2014 inclusive. For the normalization of the TLC

[2] An actuarial interpretation of n_t is the total amount payable per death over the year t to $t + 1$ (ignoring interest) if everyone in the population contributes at the rate of \$1 p.a. while alive over the year t to $t + 1$ to a "term insurance" fund.

model, weights in every year are set equal to the United States combined (male and female) mid-year population in the final year 2014, and a least squares fit is used.

The Lee–Carter k_t displays the familiar downward–sloping trend indicating the general decline in male and female mortality over the period. Although shown here in the same panel, the male and female parameters are not comparable due to the fact that they have been independently and arbitrarily scaled. For the same reason, the actual values of k_t do not have a direct and intuitive interpretation.

The $\ln n_t$ parameter of the reformulated and normalized TLC model shows an upward–sloping trend, indicating an increase in the exposure required to generate one expected death, that is, declining overall mortality. In this case the model specification and normalization ensures that n_t does have an straightforward and intuitive interpretation. For example, the male $n_{2014} = \mathrm{e}^{6.066} = 431$ indicates that a population of 431, spread across all ages, is required to generate one expected death in that year. Furthermore, the model specification and normalization in conjunction with the use of a consistent across–age weighting in the normalization of n_t ensures that results are comparable across populations and over time. For example, the increase in male n_t from $n_{1933} = \mathrm{e}^{4.750} = 116$ to $n_{2014} = \mathrm{e}^{6.066} = 431$ is an interpretable and meaningful change. The difference in male $n_{2014} = 431$ and female $n_{2014} = \mathrm{e}^{6.628} = 756$ is a meaningful measure of the overall male / female mortality differential.

The β_x parameter of the TLC model is also meaningful and interpretable: as the percentage change in the NE at age x in response to a one percent change in the overall NE. We see here that males aged up to 15 have experienced faster rates of mortality improvement than the average over the period, as have females aged up to 36. As indicated in (5), the TLC β_x and Lee–Carter b_x parameters are identical apart from scaling. However whereas the Lee–Carter b_x parameter value is arbitrary and can be interpreted only relative to other values, the TLC normalization ensures that the β_x parameter has an absolute interpretation.

The α_x parameter of the TLC model can be regarded as an adjustment factor to pitch the NE at the right level, analagous to the intercept in regression models.

We see here that α_x is positive, that is, NE is adjusted up and mortality is lower, at ages 15 to 51 for males and 12 to 65 for females. The TLC α_x parameter has an interpretation only in combination with other parameter values in contrast to the LC a_x parameter which is interpretable as the average log–mortality over the whole period at each age x. In the TLC normalization the contraints have been used to ensure that the $\ln n_t$ and β_x parameters have an absolute meaning whereas in the Lee–Carter model they have instead been directed to the interpretability of the a_x intercept parameter.

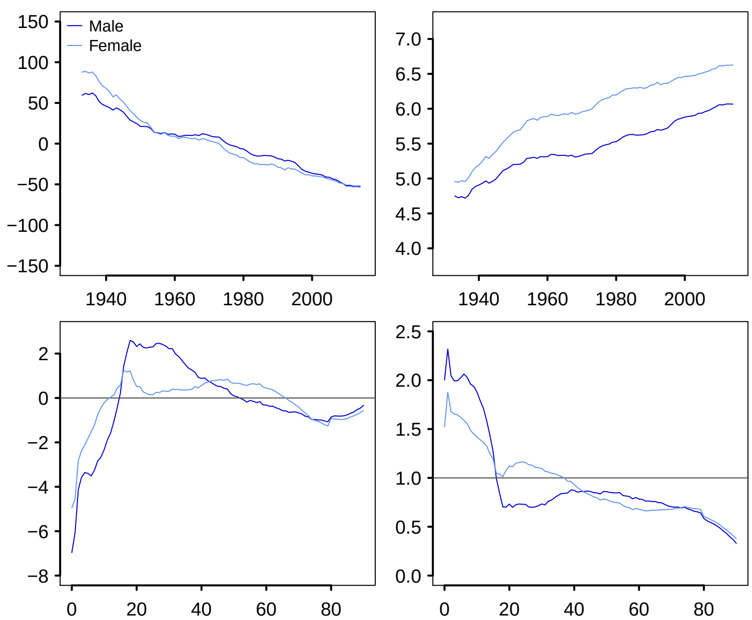

Fig. 1: Top left panel displays k_t of the standard LC model fitted to United States 1933–2014 male and female mortality. Remaining panels are $\ln n_t$ (top right), α_x (bottom left) and β_x (bottom right) of the TLC model fitted to the same data using a least squares normalization and 2014 combined mid-year population weights.

5. Cross–country and male / female comparisons using the TLC model

The TLC model is applied to produce cross–country and male / female comparisons for six countries. Data is from the Human Mortality Database (2016) and comprises central mortality rates m_{xt} and the United States mid-year population, by sex and years of age $x = 0, 1, \ldots, 90$. Different time series of data are available for each country and the full dataset has been used in each case: Australia (AUS) 1921 to 2011, Canada (CAN) 1921 to 2011, England and Wales (ENW) 1841 to 2013, Japan (JPN) 1947 to 2012, the Netherlands (NLD) 1850 to 2012 and the United States (USA) 1933 to 2014. To ensure comparability, the normalization weights for every country, sex and year combination are set equal to the United States combined 2014 mid-year population, and a least squares fit is used.

Fig. 2 shows the fitted $\ln n_t$ and β_x for males and females for the six countries. For comparison with $\ln n_t$, also shown is the logarithm of the weighted geometric mean of the actual observed age–specific NEs in each year.

Parameter $\ln n_t$ shows the expected increase over time for all countries, reflecting declining mortality over the period. Downward spikes are evident due to the first and second World Wars, and for England and Wales and the Netherlands in 1918 due to the Spanish influenza epidemic. In the TLC model the absolute value and slope of $\ln n_t$ are meaningful indicators of the level and rate of change in mortality. A regression of $\ln n_t$ on t over the period since 1947 gives an average growth of 0.019, 0.018, 0.016, 0.030, 0.017 and 0.014 for males and 0.021, 0.021, 0.018, 0.037, 0.018 and 0.015 for females for the six countries. The rapid improvement in Japanese male and particularly female mortality over the period is apparent. Female mortality has shown slightly faster improvement than male mortality in all countries for the entire period, but it is well known that this has reversed in many countries in recent decades (Glei and Horiuchi 2007; Thorslund et al. 2013).

A comparison of the left and middle panels reveals that n_t and the calculated weighted geometric mean of the observed age–specific n_{xt} are very similar. That is, normalizing the new model n_t to the weighted geometric mean produces a series that is very similar to the actual calculated weighted geometric mean over all years.

Parameter β_x shows a broadly similar pattern across all countries, with the youngest ages responding the most to the overall decline in mortality, and the oldest ages the least. This parameter is interpretable as the percentage change in NE at age x in response to a one percent change in overall NE. In broad terms, ages below around 40 have been more responsive than average to mortality change (with some exceptions for males) and ages above 40 have been less responsive than average. For the purposes of this illustration, different periods have been used for each country, limiting the extent to which conclusions should be drawn from the results. For example, the countries with the longest time series of data, England and Wales and the Netherlands, have the lowest β_x at the oldest ages, which is unsurprising given that declines in mortality at the oldest ages accelerated since 1950 in many industrialized countries (Kannisto et al. 1994).

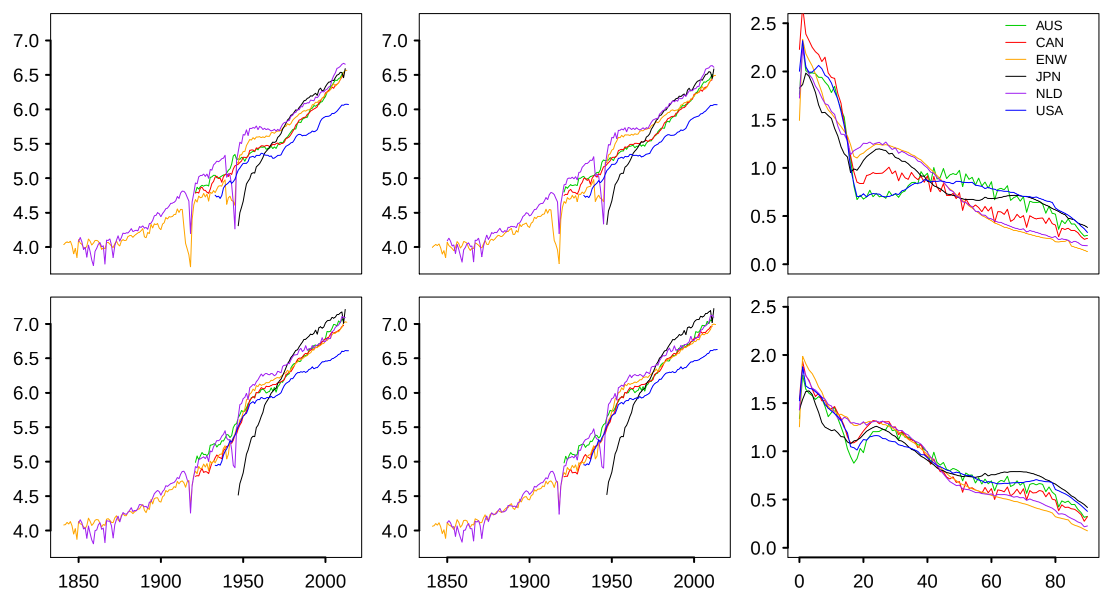

Fig. 2: Calculated logarithm of the weighted geometric mean of observed age–specific NEs (left), $\ln n_t$ (middle) and β_x (right) for males (top) and females (bottom), for six countries

Comparisons of overall mortality in different populations can be made by taking the ratio of overall indices such as the weighted geometric mean of age–specific mortality rates or the direct standardized mortality rate: this is discussed in section 6. The same approach is used here with the normalized n_t, which is valid given its interpretation as an overall mortality measure and the use of equal weights across populations. Figure 3 shows the ratio of n_t in Japan to n_t in each country, separately for males and females. The results therefore indicate the ratio of overall mortality in each country to that in Japan, with a figure greater than one indicating heavier mortality. The rapid improvement in Japanese mortality and the low current levels, especially for females, are again evident. The deteriorating position of the United States relative to other developed countries is also clear.

Also shown in Fig. 3 is the mortality sex ratio in each country. This ratio is calculated as the female n_t divided by the male n_t, which represents the overall ratio of male to female mortality. The ratio shows a general increasing trend to a maximum during 1970 to 2000 followed by a decline. This pattern has been observed and examined by several authors (Glei and Horiuchi 2007; Thorslund et al. 2013).

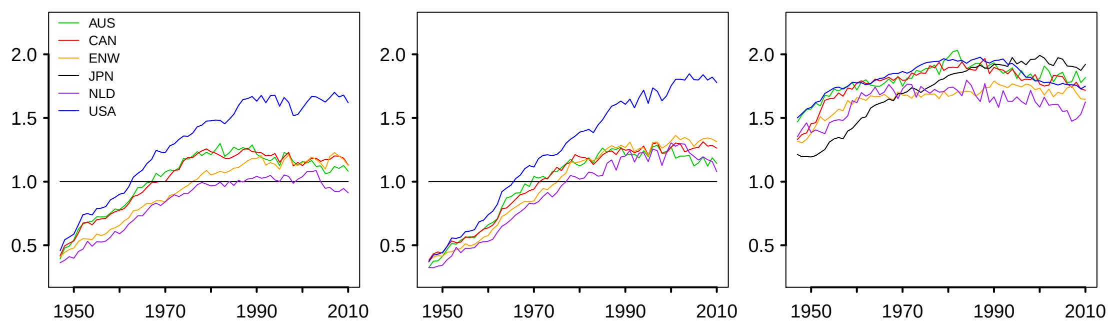

Fig. 3: The ratio of overall mortality in each country to overall mortality in Japan for males (left) and females (middle), and the ratio of male to female mortality in each country (right), 1947–2010

6. Discussion

The field of mortality forecasting was changed by the publication of the pioneering Lee–Carter method in 1992 (Lee and Carter 1992). The method was quickly adopted and extended, and Lee–Carter–based methods have now been dominant for some time. Among the forecasting advantages are the minimal subjective judgement required (and the relative accuracy of forecasts compared to those based on methods incorporating greater judgement), and the production of probabilistic prediction intervals. The fact that the Lee–Carter parameterization is not unique is not an issue for forecasting. As expressed by Girosi and King (2007): "*This is not a conceptual obstacle; it merely means that the likelihood associated with the model … has an infinite number of equivalent maxima, each of which would produce identical forecasts.*

In practice, we merely pick an arbitrary but consistent parameterization sufficient for identification" (p. 2). It does, however, mean that parameter values do not on their own have an inherent meaning and interpretation, and cannot be used to compare mortality across populations.

Another issue with the interpretability of Lee–Carter, along with other mortality models, is that it models the logarithm of the mortality rate. Studies have established that people have difficulties interpreting probabilities (Kunreuther et al. 2001; Reyna and Brainerd 2008), and have suggested that extremely low probability events are interpreted as "*essentially nil risk*" (Stone et al. 1994). The logarithm of small probabilities is an even more abstract construct. While rates offer many advantages as demographic measures and are widely used, it would be beneficial to be able to express them in a more readily and naturally understood form.

In this paper we address both of these aspects of interpretability of the Lee–Carter model. The proposed model is based on the needed–exposure (NE) measure as a more interpretable alternative to the mortality rate. The reformulated model is normalized so that the age and time parameters are meaningful, interpretable and comparable across populations.

The NE is the reciprocal of the mortality rate and therefore simply reflects the number of individuals that would need to be alive in a year to generate one expected death over that year. Interpretation of NEs is intuitive and avoids difficulties associated with small probabilities. The improvement in overall (between age 0 and age 90) female United States mortality between 1933 and 2014 can be expressed as either a reduction in the overall mortality rate from 0.007042 to 0.001322 or, more intuitively, as an increase in the NE from 142 to 756. Advantages are even more evident at young ages where mortality rates are very small.

There are other convenient properties associated with the NE. Similar to the mortality rate, it is meaningful to refer to overall, age–specific and cause–specific NE measures whereas such variants are not as conveniently defined for other mortality measures such as the expectation of life. Existing mortality models for the logarithm of the mortality rate require minimal adjustment to be based on NEs, since the logarithm of the mortality rate is the negative logarithm of the NE. Relative mortality risk measured by the ratio of two rates can be readily reproduced by taking the ratio of the two NEs with numerator and denominator reversed.

The potential usefulness of the NE measure is consistent with the finding that risk perceptions are affected by re–expressing the probability of death in terms of the time interval during which a single death is expected (Weinstein et al. 1996). The NE is closely tied to the "number needed to treat" (NNT) measure used to communicate effectiveness of medical treatments. In a trial comparing a new with an existing treatment, the NNT is the estimated number of patients who would need to receive the new rather than the existing treatment for one additional patient to benefit, that is, the reciprocal of the absolute risk reduction (Altman 1998). An NNT of 1 indicates the ideal situation where all patients improve with the new treatment and none with the existing treatment, and higher NNTs indicate declining effectiveness of the new treatment. The NNT, introduced in the late 1980s (Laupacis et al. 1988), has been found to be an effective tool for communication (Cook and Sackett 1995; Tramèr and Walder 2005) and is increasingly used (Altman 1998).

After re–expressing the Lee–Carter model in terms of the NE, it is normalized to ensure interpretability of parameters. This normalization is made possible by first replacing the Lee–Carter parameter k_t by the parameter $\ln n_t$. A consequence of this reformulation is that the parameter n_t can now – with appropriate normalization – be interpreted as a measure of the same type as the age–specific measure being modeled, a *needed–exposure*. In the original model the k_t parameter is instead related to the log of the modeled measure. With normalization, n_t can be interpreted as an overall (across–age) combination of the age–specific n_{xt}.

In this paper, we have implemented the normalization by aligning n_t with the weighted geometric mean of the n_{xt}, using least squares regression. This approach replaces the usual two arbitrary Lee–Carter constraints $\sum_x b_x = 1$ and $\sum_t k_t = 0$. In both models, the time parameter is an indication of overall mortality in each year, but the Lee–Carter k_t is related to the age–specific log–mortality rates in a non-intuitive and indirect way whereas the normalized n_t approximates the weighted geometric mean of the age–specific n_{xt}. Because the NE is the reciprocal of the mortality rate, this is equivalent to the overall mortality rate approximating the weighted geometric mean of the age–specific mortality rates.

The (weighted or unweighted) geometric mean is one of a number of measures used for condensing age–specific mortality rates into a single overall measure (see Hinde 1998, Schoen 1970, Smith 2013 and Wunsch 2012 for a discussion of various indices and their advantages and disadvantages). It offers a number of advantages in this context which have been discussed by Schoen (1970). The geometric mean is "*the kind of average . . . which best reflects the nature of the underlying mortality function*" (Schoen 1970, p. 318), namely, exponential. It addresses the shortcoming of the widely–used direct standardized rate of attaching too much weight to the older ages (Yerushalmy 1951). Geometric means – or weighted geometric means with common weights – can be compared across populations. Further, the ratio of two (weighted) geometric means is a valid and meaningful way to compare overall mortality in two populations, and the ratios are themselves comparable across populations.

The proposed time parameter n_t therefore has a clear interpretation as an overall mortality measure normalized to the (weighted) geometric mean. Comparisons of n_t parameters across populations are meaningful, and give similar results to direct comparisons of calculated (weighted) geometric means, as illustrated in Fig. 2. Relative mortality in two populations can be calculated as the ratio of the n_t parameters, as illustrated for cross–country and male / female comparisons in Fig. 3.

The β_x parameter of the normalized model is also interpretable. It represents the elasticity at age x: that is, the percentage change in the NE at age x given a one percent change in the overall NE. As in the original Lee–Carter, it is a measure of the responsiveness of age–group x to overall time trends in mortality in the population, but in this case it has a clearly defined value rather than being an index on an arbitrary scale. A β_x less than (greater than) one indicates that age x is less (more) responsive than average to overall time trends in mortality in the population. The α_x parameter of the proposed model has a relative interpretation whereas the comparable Lee–Carter parameter has an absolute interpretation as the across–time average of the log–mortality rate at age x, as illustrated in Fig. 1. In the proposed model, constraints have been used to ensure the interpretability of n_t and β_x, which is considerably more useful than ensuring interpretability of α_x.

7. Conclusion

We have replaced the arbitrary constraints used in the estimation of Lee–Carter parameters with constraints that ensure that the time and age-response parameters are meaningful, interpretable, and comparable across populations. Further, we have recast the model to be based on the needed–exposure as a more interpretable alternative to the mortality rate. Forecasts are unchanged but the time and age-response parameters now have a clear and intuitive definition and can be compared across populations. The same approach can be used to confer interpretability on parameters of many other mortality models. The advances pave the way for widely–used mortality forecasting models to be used for mortality modeling, analysis and comparison.

References

Altman, D. G. (1998). Confidence intervals for the number needed to treat. *BMJ: British Medical Journal*, 317(7168):1309–1312.

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

Booth, H. and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 3:3–43.

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

Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., and Khalaf-Allah, M. (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48(3):355–367.

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

Camerer, C. F. and Kunreuther, H. (1989). Decision processes for low probability events: Policy implications. *Journal of Policy Analysis and Management*, 8(4):565–592.

Cook, R. J. and Sackett, D. L. (1995). The number needed to treat: A clinically useful measure of treatment effect. *BMJ: British Medical Journal*, 310(6977):452–454.

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

Deaton, A. S. and Paxson, C. (2004). Mortality, income, and income inequality over time in Britain and the United States. In *Perspectives on the Economics of Aging*, pages 247–286. University of Chicago Press.

Deng, Y., Brockett, P. L., and MacMinn, R. D. (2012). Longevity/mortality risk modeling and securities pricing. *Journal of Risk and Insurance*, 79(3):697–721.

Girosi, F. and King, G. (2007). Understanding the Lee–Carter mortality forecasting method. *http://gking.harvard edu/files/lc.pdf.*

Glei, D. A. and Horiuchi, S. (2007). The narrowing sex differential in life expectancy in high–income populations: Effects of differences in the age pattern of mortality. *Population Studies*, 61(2):141–159.

Hinde, A. (1998). *Demographic Methods*. Hodder Arnold Publication, London, UK.

Hollmann, F. W., Mulder, T. J., and Kallan, J. E. (1999). *Methodology & Assumptions for the Population Projections of the United States: 1999 to 2010*. US Department of Commerce, Bureau of the Census, Population Division, Population Projections Branch.

Human Mortality Database (2016). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Retrieved from http://www.mortality.org, accessed on 21 July 2016.

Hyndman, R. J., Booth, H., and Yasmeen, F. (2013). Coherent mortality forecasting: The product–ratio method with functional time series models. *Demography*, 50(1):261–283.

Hyndman, R. J. and Shahid Ullah, M. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, 51(10):4942–4956.

Janssen, F., van Wissen, L., and Kunst, A. (2013). Including the smoking epidemic in internationally coherent mortality projections. *Demography*, 50(4):1341–1362.

Kannisto, V., Lauritsen, J., Thatcher, A. R., and Vaupel, J. W. (1994). Reductions in mortality at advanced ages: Several decades of evidence from 27 countries. *Population and Development Review*, 20(4):793–810.

Kunreuther, H., Novemsky, N., and Kahneman, D. (2001). Making low probabilities useful. *Journal of Risk and Uncertainty*, 23(2):103–120.

Laupacis, A., Sackett, D. L., and Roberts, R. S. (1988). An assessment of clinically useful measures of the consequences of treatment. *New England Journal of Medicine*, 318(26):1728–1733.

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

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

Li, J. (2013). A Poisson common factor model for projecting mortality and life expectancy jointly for females and males. *Population Studies*, 67(1):111–126.

Li, J. S.-H. and Hardy, M. R. (2011). Measuring basis risk in longevity hedges. *North American Actuarial Journal*, 15(2):177–200.

Li, N. and Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee–Carter method. *Demography*, 42(3):575–594.

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

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

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

Pollard, J. H. (1988). On the decomposition of changes in expectation of life and differentials in life expectancy. *Demography*, 25(2):265–276.

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

Reyna, V. F. and Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, 18(1):89–107.

Schoen, R. (1970). The geometric mean of the age-specific death rates as a summary index of mortality. *Demography*, 7(3):317–324.

Smith, D. (2013). *Formal Demography*. The Springer Series on Demographic Methods and Population Analysis. Springer US.

Stoeldraijer, L., van Duin, C., van Wissen, L., and Janssen, F. (2013). Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: The case of the Netherlands. *Demographic Research*, 29(13):323–354.

Stone, E. R., Yates, J., and Parker, A. M. (1994). Risk communication: Absolute versus relative expressions of low-probability risks. *Organizational Behavior and Human Decision Processes*, 60(3):387–408.

Thorslund, M., Wastesson, J., Agahi, N., Lagergren, M., and Parker, M. (2013). The rise and fall of women's advantage: A comparison of national trends in life expectancy at age 65 years. *European Journal of Ageing*, 10(4):271–277.

Tramèr, M. R. and Walder, B. (2005). Number needed to treat (or harm). *World Journal of Surgery*, 29(5):576–581.

Tuljapurkar, S., Li, N., and Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. *Nature*, 405(6788):789–792.

Weinstein, N. D., Kolb, K., and Goldstein, B. D. (1996). Using time intervals between expected events to communicate risk magnitudes. *Risk Analysis: An International Journal*, 16(3):305–308.

Wilmoth, J. (2000). Demography of longevity: Past, present, and future trends. *Experimental Gerontology*, 35(9–10):1111–1129.

Wong-Fupuy, C. and Haberman, S. (2004). Projecting mortality trends: Recent developments in the United Kingdom and the United States. *North American Actuarial Journal*, 8(2):56–83.

Wunsch, G. (2012). *Introduction to Demographic Analysis: Principles and Methods*. Springer US.

Yang, S. S. and Wang, C.-W. (2013). Pricing and securitization of multi–country longevity risk with mortality dependence. *Insurance: Mathematics and Economics*, 52(2):157–169.

Yerushalmy, J. (1951). A mortality index for use in place of the age–adjusted death rate*. *American Journal of Public Health and the Nation's Health*, 41(8 part 1):907–922.

Zhou, R., Wang, Y., Kaufhold, K., Li, J. S.-H., and Tan, K. S. (2014). Modeling period effects in multi–population mortality models: Applications to Solvency II. *North American Actuarial Journal*, 18(1):150–167.