Research Article

Investigating Mortality Uncertainty Using the Block Bootstrap

Xiaoming Liu and W. John Braun

Department of Statistical and Actuarial Sciences, University of Western Ontario, London, ON, Canada N6A 5B7

Correspondence should be addressed to Xiaoming Liu, xliu@stats.uwo.ca

This paper proposes a block bootstrap method for measuring mortality risk under the Lee-Carter model framework. In order to take account of all sources of risk (the process risk, the parameter risk, and the model risk) properly, a block bootstrap is needed to cope with the spatial dependence found in the residuals. As a result, the prediction intervals we obtain for life expectancy are more accurate than the ones obtained from other similar methods.

1. Introduction

For actuarial pricing and reserving purposes, the mortality table needs to be projected to allow for improvement in mortality to be taken into account. It is now a well-accepted fact that mortality development is difficult to predict; therefore, a stochastic mortality modelling approach has been advocated [1, 2]. For a review of recent developments in stochastic mortality modelling, interested readers are referred to Pitacco [3] and Cairns et al. [4].

Under the framework of a stochastic mortality model, future mortality rates are random as are the other quantities derived from the mortality table. In order to manage the mortality risk properly, we need to assess the uncertainty coming from the mortality dynamics carefully.

There are three types of risk embedded in adopting a stochastic mortality model (see [5]):

(a) the "process risk", that is, uncertainty due to the stochastic nature of a given model;

(b) the "parameter risk", that is, uncertainty in estimating the values of the parameters;

(c) the "model risk", that is, uncertainty in the underlying model (actual trend is not represented by the proposed model).

It is important to take account of all sources of variability to gain an accurate understanding of the risk. Since the quantities of interest, such as life expectancies and annuity premiums, are normally related to the mortality rates at different ages in a nonlinear format, a theoretical analysis is intractable. In the literature, simulation techniques have been proposed to measure the mortality risk under such circumstances.

In this paper, we focus on investigating the mortality risk under the framework of the Lee-Carter model using a block bootstrap method. Lee and Carter [6] proposed to describe the secular change in mortality via a stochastic time-varying index that is not directly observable. In this context, future mortality rates (as well as the quantities derived from the mortality rates) are all influenced by this stochastic index. In the original paper of Lee and Carter, it was suggested that the variability from the index dominates all the other errors in forecasting mortality rates and the associated life expectancies, especially when it comes to long-term forecasting. We are interested in verifying this statement, and in providing accurate risk assessment by considering all sources of variability.

In particular, we aim at obtaining prediction intervals for forecasted life expectancies using a residual-based block bootstrap. Recently, various bootstrap methods have been proposed to measure mortality risk, as seen in Brouhns et al. [7] for the parametric bootstrap, in Brouhns et al. [8] for the semiparametric bootstrap, and in Koissi et al. [9] for the ordinary residual bootstrap. In those papers, the implicit assumption is that the residuals after fitting the model to the data are independent and identically distributed. However, in our experience, correlations across age and year can be observed in the residuals. When calculating prediction intervals by bootstrap methods, there may be an underestimation of the mortality risk if correlations in residuals are not properly handled.

Though it may not be easy to see how different sources of uncertainty may be correlated with (and affect) one another, the block bootstrap partially retains the underlying dependence structure in the residuals and generates more realistic resamples [10]. As shown in Section 5, the prediction intervals we obtained from the block bootstrap are wider and more accurate than the ones obtained from the aforementioned methods.

This paper is organized as follows. In Section 2, we introduce the data and notation used in this paper. In Section 3, we describe the Lee-Carter model and the estimation method. We also discuss the challenges related to measuring mortality risk in the associated projections. The (block) bootstrap method and its application in mortality risk are provided in Section 4. Our proposed method is applied to Swedish male mortality data, and the results are presented in Section 5, followed by a discussion in Section 6.

2. Mortality Data, Definitions, and Notation

2.1. Data Description

As a test case for our proposed technique, we use the mortality data on Swedish males from 1921 to 2007, freely provided by the "Human Mortality Database" [11]. The data from 1921 up to 1960 are used to estimate the model parameters, and the data since 1961 are used to check if our proposed technique has improved the calibration of mortality risk in the long-term prediction.

2.2. Definitions and Notation

The following terminology will be used at various points in this paper. For an individual at integer age x and calendar time t, we have the following:

(i) D_{xt} is the number of deaths at integer age x in calendar time t.

(ii) ETR_{xt} is the exposure-to-risk at integer age x in calendar time t, that is, the person-years lived by people aged x in year t.

(iii) $\mu(x + s, t)$, $0 \leq s \leq 1$ is the force of mortality. This is assumed to be constant within each year of age, that is, $\mu(x + s, t) = \mu_{xt}$, $0 \leq s \leq 1$.

(iv) m_{xt} is the central mortality rate and is defined as D_{xt}/ETR_{xt}. Under the constant force of mortality assumption, $m_{xt} = \mu_{xt}$.

(v) $q_x(t)$ is the probability that an individual aged x in year t dies before reaching age $x+1$. $p_x(t) = 1 - q_x(t)$. Under the constant force of mortality assumption, $p_x(t) = e^{-\mu_{xt}}$.

(vi) e_0 is the life expectancy at birth and is calculated as

$$e_0 = \frac{1}{2} + \sum_{k=1}^{\infty} \prod_{x=0}^{k-1} p_x(t). \tag{2.1}$$

In this paper, we investigate how e_0 is influenced by the uncertainty coming from projections under the Lee-Carter model.

3. The Lee-Carter Modelling Framework

3.1. Model Description

Lee and Carter [6] proposed a two-stage dynamic procedure to describe the secular change in the logarithm of age-specific central rates:

$$\log m_{xt} = a_x + b_x k_t + \epsilon_{xt}, \tag{3.1}$$

$$k_t = k_{t-1} + c + \xi_t \quad \text{with i.i.d. } \xi_t \sim N(0, \sigma^2). \tag{3.2}$$

In the first stage, (3.1) is fit to the historical mortality data (from year 1 to t_0, i.e., $t = 1, \ldots, t_0$) to estimate the parameters a_x, b_x, and k_t. The interpretation of the parameters is as follows. a_x describes the general age shape of the log central rates, while the actual log rates in a specific year change according to a time-varying mortality index k_t modulated by an age response variable b_x. In other words, the profile of b_x indicates which rates decline rapidly and which decline slowly over time in response to changes in k_t. ϵ_{xt} contains the error that is not captured by the model at age x in year t.

In the second stage, the estimated values of k_t from the first stage are fit to a random walk (3.2). According to (3.2), the dynamics of k_t follow a random walk with drift parameter c and normal noise ξ_t with mean 0 and variance σ^2.

3.2. Estimation of the Parameters

Instead of using least-squares estimation via a singular value decomposition (SVD) to estimate a_x, b_x, and k_t as in the original Lee and Carter paper, we adopt the methodology proposed by Brouhns et al. [12]. The main drawback of the SVD method is that the errors ϵ_{xt} in (3.1) are assumed to be homoscedastic, which is not realistic. In order to circumvent the problems associated with the SVD method, Brouhns et al. [12] proposed the Poisson assumption to be applied to the death counts D_{xt}. That is,

$$D_{xt} \sim \text{Poisson}(\lambda_{xt}), \tag{3.3}$$

with

$$\lambda_{xt} = \text{ETR}_{xt} m_{xt}, $$
$$m_{xt} = \exp(a_x + b_x k_t). \tag{3.4}$$

We now can write the log-likelihood function for the parameters a_x, b_x, and k_t:

$$l(a_x, b_x, k_t) = \sum_{x,t} \left[D_{xt}(a_x + b_x k_t) - \text{ETR}_{xt} \exp(a_x + b_x k_t) \right] + \text{constant}. \tag{3.5}$$

Maximizing (3.5) iteratively provides the MLE estimates $\widehat{a}_x$, $\widehat{b}_x$, and $\widehat{k}_t$ [12].

3.3. Uncertainties in the Lee-Carter Mortality Projections

Let $\widehat{a}_x$, $\widehat{b}_x$, $\widehat{k}_t$, $\widehat{c}$, and $\widehat{\sigma}$ represent all the estimated parameters obtained from fitting the Lee-Carter model to the data (from year 1 to t_0). In order to forecast the time-varying index k_{t_n} years ahead given all the data up to time t_0, formula (3.2) is extrapolated as

$$k_{t_0+n} = k_{t_0} + \widehat{c} \cdot n + \sum_{j=1}^{n} \xi_j, \quad \text{with i.i.d. } \xi_j \sim N\left(0, \widehat{\sigma}^2\right). \tag{3.6}$$

Here, ξ_j represents the uncertainty coming from forecasting k_t.

Accordingly, the future central mortality rates m_{xt} in the calendar year $t_n (= t_0 + n)$ are given by

$$\log m_{xt_n} = \widehat{a}_x + \widehat{b}_x k_{t_n}, \tag{3.7}$$

where $\widehat{a}_x$ and $\widehat{b}_x$ remain constant. From these rates, we can construct projected life tables and calculate the associated life expectancies using the relations given in Section 2.2.

Under the Lee-Carter model, k_t follows a stochastic process. Further, according to model (3.7), all future mortality rates in the same year are affected by the same time-varying random variable k_t, where ξ_t is the sole error source of the model. Under this circumstance, the projected life expectancy given in formula (2.1) can be viewed as a sum

of a sequence of comonotonic variables, since all survival probabilities can be written as monotone transformations of the same underlying random variable (which is the noise ξ_t in this context) so that they are perfectly positively dependent. Comonotonicity allows us to express the quantiles for projected life expectancies in terms of the quantiles of ξ_t analytically. According to Denuit [13], the p-percentile for e_0 in the forecast year t_n taking account of only the uncertainty from k_t can be computed as

$$F_{e_0(t_n)}^{-1}(p) = \frac{1}{2} + \sum_{h \geq 1} \exp\left(-\sum_{x=0}^{h-1} \exp\left(\hat{a}_x + \hat{b}_x\left(k_{t_0} + n\hat{c} + \hat{\sigma}\sqrt{n}\Phi^{-1}(1-p)\right)\right)\right). \qquad (3.8)$$

The Denuit quantile formula (3.8) will be used in this paper as a benchmark to compare with our simulated prediction intervals.

Of course, uncertainty may also come from errors in parameter estimation and/or from model discrepancy. If the model is properly specified, it is reasonable to expect that these other sources of error are minor. Under the assumption that different sources of error at a given age are uncorrelated and that the error sources other than k_t are uncorrelated across age, Lee and Carter [6, in Appendix B] estimated that these other types of error contribute less than 2% of the forecast error for life expectancy e_0 for forecast horizons greater than 10 years. However, this assertion is not supported by the data; inspection of the residuals reveals substantial correlations across age and year [9, 14]. In other words, spatial dependence has been found in the residuals. It has also been found that the actual coverage of the prediction intervals based solely on the process risk of k_t is lower than their nominal level (see [14]). This indicates the need for appropriate methods to take into account different types of uncertainty.

3.4. The Deviance Residuals

As mentioned above, we need to check the residuals to assess model adequacy. Under the (Poisson) Lee-Carter model, deviance residuals can be used for this purpose in a similar manner to how ordinary residuals are used in regression analysis. For reference, see Maindonald [15]. Deviance residuals are calculated as

$$r_D = \text{sign}\left(D_{x,t} - \hat{D}_{x,t}\right)\left[D_{x,t}\ln\left(\frac{D_{x,t}}{\hat{D}_{x,t}}\right) - (D_{x,t} - \hat{D}_{x,t})\right]^{1/2}. \qquad (3.9)$$

4. Bootstrap Method

4.1. Ordinary Bootstrap Method

It is not possible to analytically obtain prediction intervals for e_0 taking account of all sources of variability. One numerical method that has been employed for this purpose is the bootstrap. The basic idea of the bootstrap is to artificially generate resamples of the data which can be used as a basis for approximating the sampling distribution of model parameter estimates. Koissi et al. [9] give a brief overview of the bootstrap and also some references in the context of applying the bootstrap method to the Lee-Carter model. They use an ordinary deviance residual-based bootstrap for this problem.

Briefly, their approach is to fit the Lee-Carter model to the data and to obtain deviance residuals. They then sample with replacement from these residuals to obtain resampled residuals. These resampled residuals are transformed back to get resampled death counts. The Lee-Carter model parameters are estimated from this artificially generated data set, and the corresponding k_t process is also fit and resimulated. As a result, we obtain a projected mortality table (see Section 3.3 for projection method). The process is repeated a large number of times (say, $N = 5000$) giving a collection of projected mortality tables. Approximate 90% prediction intervals for e_0 are given by the 5% percentile and 95% percentile of the e_0's derived from these tables. Better bootstrap prediction intervals have been proposed (see, e.g., [16]), but any improvement in accuracy will depend heavily on the validity of the Lee-Carter model assumption. See also D'Amato et al. [17] for a stratified bootstrap sampling method which reduces simulation error.

4.2. Block Bootstrap Method

The ordinary bootstrap is not valid in case of spatial dependence, since it is based on simple random sampling with replacement from the original sample. In order for the bootstrap sampling distribution of the statistic of interest to be a valid approximation of the true sampling distribution, it is thus necessary for the original sample to be a random sample from the underlying population. Spatial dependence represents a violation of this assumption; it can cause either a systematic overestimation or systematic underestimation of the amount of uncertainty in a parameter estimate, depending upon whether negative or positive dependence dominates.

Confidence and prediction intervals constructed from bootstrap samples in the presence of positive spatial autocorrelation will thus be too narrow so that their actual coverage probabilities will be lower than their nominal levels. In the presence of negative autocorrelation, actual coverage levels will systematically exceed nominal levels.

One remedy for this problem is the block bootstrap, which gives improved approximate prediction intervals, taking account of the spatial dependence, at least locally. The basic idea is to resample rectangular blocks of observations, instead of individual observations one at a time. A full bootstrap resample is obtained by concatenating randomly sampled blocks of observations. Thus, apart from the "seams" between the blocks, some of the dependence between neighboring observations will be retained in a bootstrap resample. If the blocks are taken large enough, then most of the dependence between observations located near each other will be retained in the bootstrap resample. Of course, block sizes should not be taken too large; otherwise, the resamples begin to resemble the original sample too closely, and this will again cause uncertainties to be underestimated. The method provides reasonable approximations (for large samples) in much the way that the m-dependent central limit theorem provides a normal distribution approximation for averages of locally dependent data. It should be noted that the method will fail without adjustment for data with long memory. A full account of the block bootstrap for time series is given by Davison and Hinkley [16], and the paper by Nordman et al. [18] describes the spatial block bootstrap we will employ. See Taylor and McGuire [19] for another actuarial application of bootstrap for dependent data.

To set up a new artificial set of residuals, we start with an array, which has the same dimensions as the original matrix of residuals. The empty array is then partitioned into smaller rectangular blocks. Each block is replaced by a block of the same size, which is
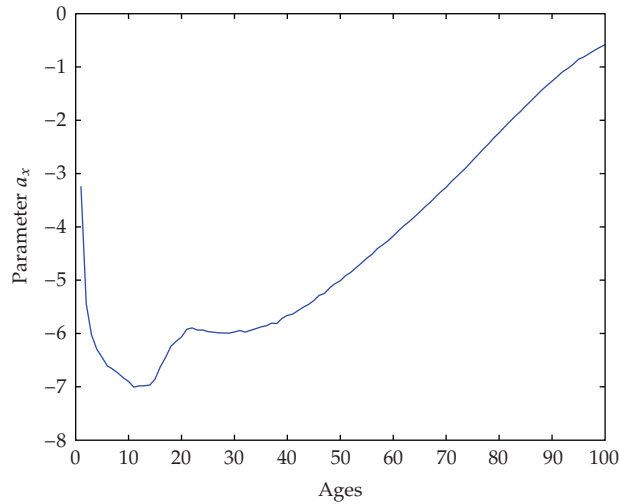
Figure 1: Parameter $\widehat{a}_x$ based on the Swedish male 1921–1960 mortality data.

randomly selected from the original matrix. These random blocks are constructed as follows. First, uniformly randomly select an element from the original matrix. The associated block consists of all residuals in the appropriately sized rectangle to the southeast of the chosen point. In cases where part of the rectangular block falls outside the original matrix, a periodic extension of the matrix is used to fill out the remaining cells of such a rectangle. As stated earlier, more details can be found, for example, in Nordman et al. [18].

Nordman et al. [18] give advice on choice of (square) block size for situations where the parameter of interest is a smooth function of the mean. The optimal block size is based on a nonparametric plug-in principle and requires 2 initial guesses (see the appendix). Our parameters of interest do not completely satisfy their assumptions, but we have experimented with their recommendations. In the absence of firm theoretical guidance, we have found it useful to plot a correlogram of the original raw residuals and compare with the resampled residuals. The correlogram is a plot of the spatial autocorrelation against distance. See Venables and Ripley [20] for more information. If the correlograms match reasonably well, this gives us confidence in our block choice.

5. Applications to Data

5.1. Fitting Performance on Swedish Males 1921 to 1960

We fit the Lee-Carter model to the Swedish male data. The estimated parameters $\widehat{a}_x$, $\widehat{b}_x$, and $\widehat{k}_t$ are displayed in Figures 1, 2, and 3. We then computed the deviance residuals. Contour maps are plotted in Figure 4, while the correlogram plots are given in Figure 5. For comparison purposes, we put together the results based on the original deviance residuals and the resampled residuals of different block sizes.

Both plots are highly suggestive of spatial dependence, because of the occurrence of large patches of large positive and large negative residuals, that is, clustering in Figure 4(a). Evidence of short range dependence is also seen in the correlogram of the raw residuals.
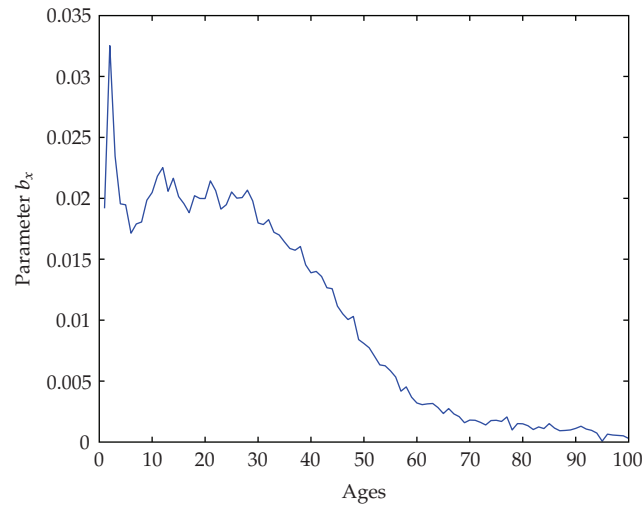
Figure 2: Parameter $\hat{b}_x$ based on the Swedish male 1921–1960 mortality data.
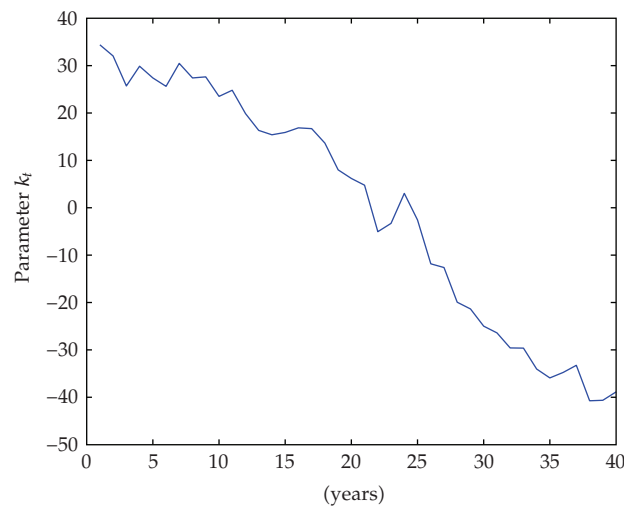

Figure 3: Parameter $\hat{k}_t$ based on the Swedish male 1921–1960 mortality data.

To see if such clustering could be due purely to chance, we simulated from the (Poisson) Lee-Carter model (three times) and fit the model to the simulated data and plotted the deviance residuals. The results are shown in Figures 6 and 7. These plots show no sign of spatial dependence; clusters appear to be much smaller than for the real data. We conclude that the ordinary bootstrap will not give valid prediction intervals. Thus, we employed a block bootstrap as described in the previous section.

5.2. Prediction Intervals for e_0 Based on Block Bootstrap

We considered a number of block sizes. The theoretical block sizes (suggested by Nordman et al. [18]) depend on the actual parameter being estimated, and they also depend
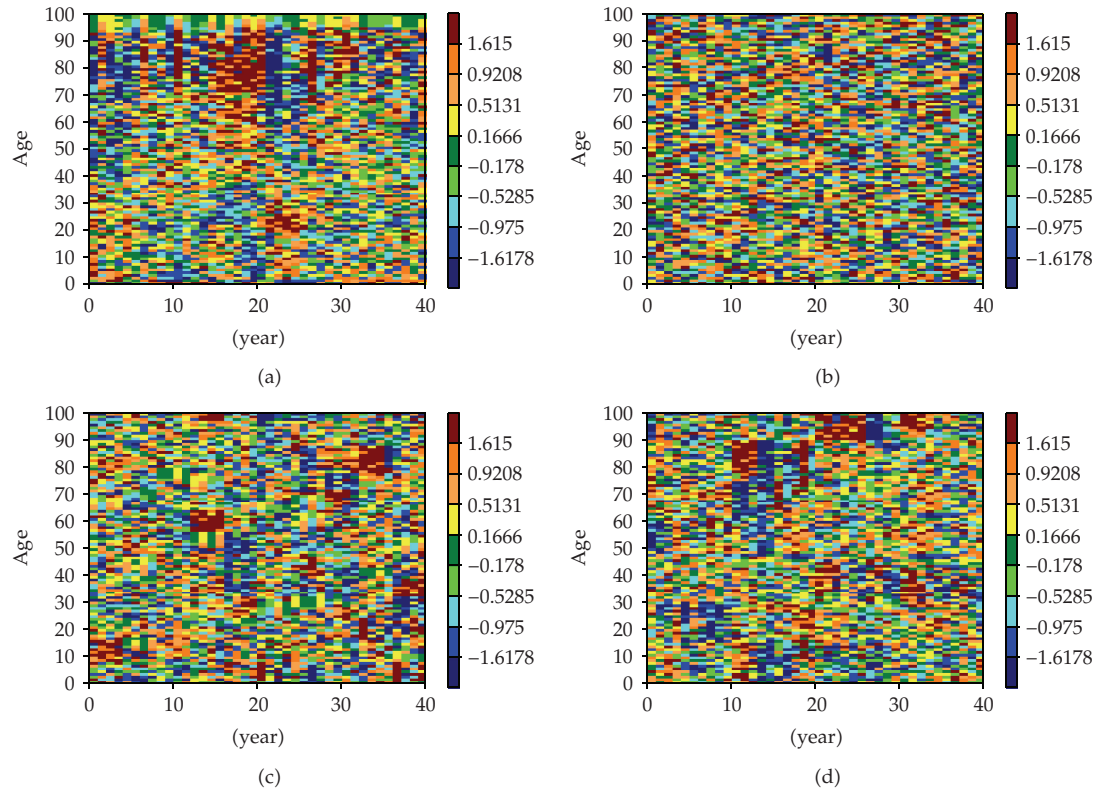
Figure 4: Contour maps for deviance residuals: (a) raw, (b) 1×1, (c) 8×4, and (d) 15×10.

Table 1: The theoretical block sizes for the forecast horizons 1 to 46 years.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
m_1	5	11	11	10	5	6	9	13	12	14	13	12	11	6	5	3
m_2	20	19	13	15	17	18	17	18	11	15	18	21	19	17	20	18
t	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
m_1	6	8	10	11	13	6	2	3	1	9	10	11	10	9	7	9
m_2	16	18	22	21	21	19	17	14	14	11	8	2	11	7	10	15
t	33	34	35	36	37	38	39	40	41	42	43	44	45	46		
m_1	9	7	1	4	1	7	8	8	12	11	10	11	10	10		
m_2	16	16	11	4	0	10	10	12	8	12	15	8	9	14		

on the initial guesses. The optimal sizes estimated depend on different forecast horizons. Table 1 gives two sets of block sizes: one (marked as m_1) is based on initial guesses of 5×5 and 10×10, and the other (marked as m_2) on 10×10 and 20×20.

Note that the results vary a lot, depending on the initial guess as well as the forecast horizon. Thus, it is not clear whether these block sizes are really the best, and we checked correlograms for some different block sizes including rectangular ones. Figure 5 shows correlograms for some of the better cases (raw residuals, and 8×4, 10×5, 15×10, and 20×12 block sizes). We have chosen to work with 15×10 blocks, since that case matches the raw residuals most closely.
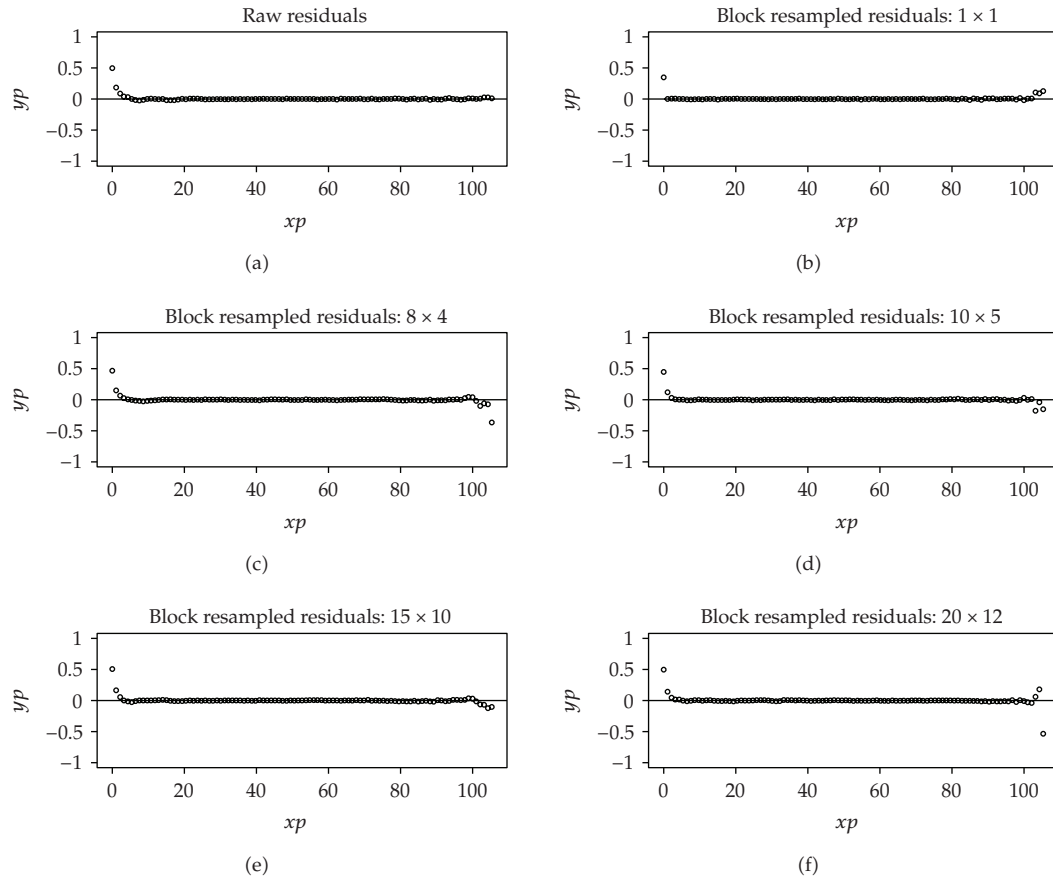
Figure 5: Correlograms for resampled deviance residuals based on different block sizes.

As further confirmation, revisit the residual plots in Figure 4. These correspond to (a) raw residuals, (b) 1×1, (c) 8×4, and (d) 15×10 block sizes. Quantitatively, one looks for the same kind of "patchiness" in the bootstrap residuals as can be seen in the raw residuals. There are obvious discrepancies when the block size is too small; there are no large patches of large positive or large negative residuals at the 1×1 block size, for the larger block sizes, we see patterns which are more like the patterns seen in the raw residuals.

Prediction intervals are displayed in Figure 8. We have plotted the 90% pointwise prediction bands based on the Denuit quantile formula, the ordinary bootstrap and the 15×10 block bootstrap. The observed life expectancy e_0 is also shown. We comment on these results in the next section.

6. Discussion

In Figure 8, we note that the theoretical prediction intervals calculated from the Denuit quantile formula is the narrowest. This makes sense since those intervals have only accounted for the uncertainty in forecasting k_t's future random path. With bootstrap methods, sampling errors in the parameter estimates and partial model uncertainty have also been considered.
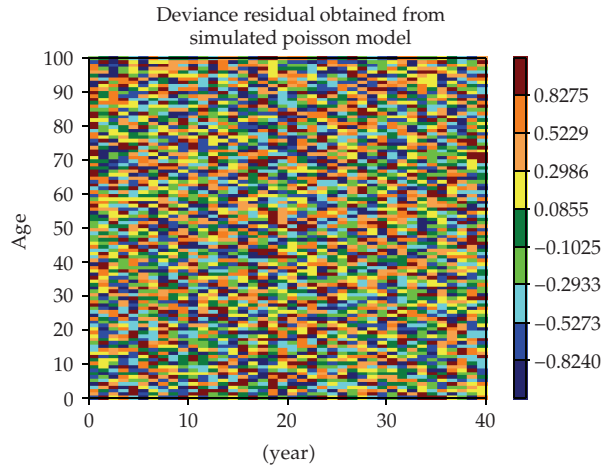
Deviance residual obtained from
simulated poisson model

(year)

Figure 6: Contour map for simulated deviance residuals.
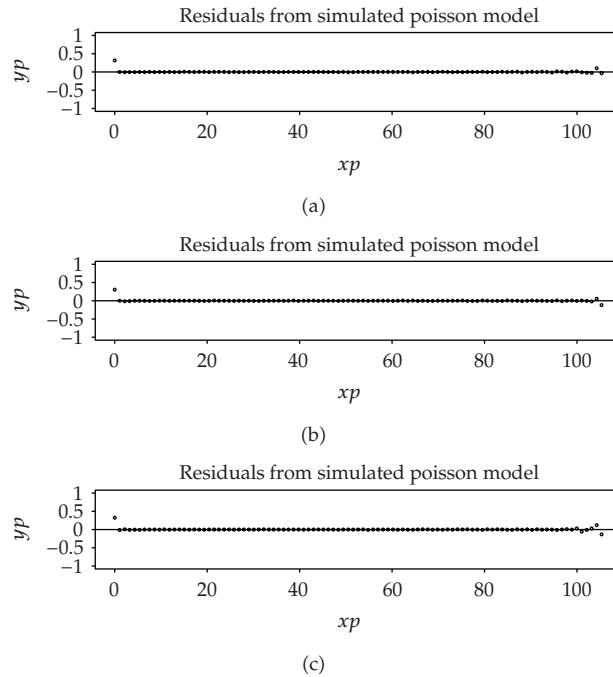

(a)

(b)

(c)

Figure 7: Correlograms for simulated deviance residuals.

Therefore, the prediction intervals from simulations are expected to be wider than the theoretical ones, which has been confirmed by our results as well.

Secondly, while the prediction intervals from the ordinary bootstrap are only marginally wider than their theoretical counterparts, we note that the prediction intervals from the block bootstrap make a substantial difference. This is due to the fact that the deviance residuals are not pattern-free random noise. When this is the case, the residuals may carry important information about the correlations among the estimated parameters,
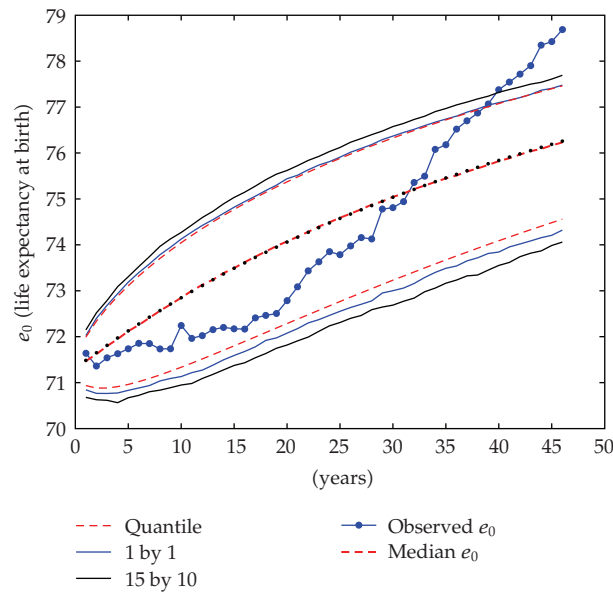
Figure 8: Forecasted e_0 with prediction intervals from different methods (1921–1960).

and/or may indicate some type of model risk. Ordinary bootstrap resampling with replacement then destroys this information and thereby diminishes the variability. Therefore, the block bootstrap is needed to obtain accurate prediction intervals. By taking into account the other types of risk properly, we obtain wider prediction intervals which serve as a better representation of the uncertainty associated with mortality predictions.

Another observation concerns the importance of the k_t as a component of the variability. Previously, it has been claimed that the uncertainty from k_t dominated all other errors. This is again an argument that is only true under restricted conditions (see Section 3.3 for more details). Figure 9 presents k_t's share in the total forecast error and shows that other sources of variability are not negligible even when the forecast horizon becomes very long.

Finally, we note that most of the observed life expectancies lie within our prediction intervals, but some of them (mainly in the later years) fall outside, indicating that there may still be factors influencing mortality that are not explained by the Lee-Carter model, even with dependence accounted for. This is certainly a cause for concern: model misspecification is a problem when we use the Lee-Carter model for long-term mortality prediction. Sometimes introducing more variables may improve model fitting and remove correlations in the residuals to some extent. Examples of those models include Renshaw and Haberman [21] that allows for the cohort period effect and Renshaw and Haberman [22] that adds a second or even third bilinear term to the Lee-Carter for better fit. However, as shown in Dowd et al. [23] and Cairns et al. [24], those models may suffer from nonrobustness in parameter estimation and sometimes generate implausible or unstable predictions. It is our belief that simpler models are preferable for long-term prediction, since they run less risk of being overfitted.

However, we remark that the block bootstrap is not able to solve the fundamental problem of model misspecification, which is not the purpose of the paper nor a problem that can be easily fixed. This paper tackles the problem from a different perspective—we adopt the Lee-Carter model framework but work on deriving more accurate measures of

Figure 9: The ratio of k_t's variability over all errors, 1 by 1 versus 15 by 10 (1921–1960).

the uncertainty degree in mortality prediction. There are other approaches using a similar idea—for example, the Bayesian method and MCMC simulation as discussed by Cairns et al. [25]. The advantage of our proposed block bootstrap method is in that it is distribution-free. Without knowing the exact form of correlation among errors, we are able to provide more accurate prediction intervals for mortality predictions.

Appendix

A. Block Size Selection

For completeness, we describe the block selection strategy suggested by Nordman et al. [18]. That approach uses a version of spatial bootstrap based on nonoverlapping blocks, and it is designed to minimize the asymptotic mean-squared error in the estimation of the variance of the parameter estimators.

Accordingly, the asymptotically optimal block size is given by

$$b_{\text{OL}} = b_{\text{NOL}}\sqrt{1.5}, \tag{A.1}$$

where b_{NOL} is the asymptotically optimal block size for the bootstrap with nonoverlapping blocks.

Let b_1 and b_2 be initial guesses at block sizes which are chosen so that $b_2 \propto b_1^{3/4}$. Then the optimal nonoverlapping block size is given by

$$b_{\text{NOL}} = \left(\frac{B_0^2 \, \text{row}(r_D)\text{col}(r_D)}{2\sigma^2(b_1)} \right)^{1/4}, \tag{A.2}$$

where $\sigma^2(b)$ is the sample variance of the bootstrap estimates of the parameter estimates based on a $b \times b$ block, and

$$B_0 = 2b_2\left(\sigma^2(b_2) - \sigma^2(2b_2)\right). \tag{A.3}$$

Acknowledgments

References

[1] R. Willets, "Mortality in the next millenium," in *Proceedings of the meeting of the Staple Inn Actuarial Society*, December 1999.

[2] GAD, "National population projections: review of methodology for projecting mortality," Tech. Rep. 8, Government Actuary's Department, 2001, National Statistics Quality Review Series.

[3] E. Pitacco, "Survival models in a dynamic context: a survey," *Insurance: Mathematics & Economics*, vol. 35, no. 2, pp. 279–298, 2004.

[4] A. J. G. Cairns, D. Blake, and K. Dowd, "Pricing death: Frameworks for the valuation and securitization of mortality risk," *ASTIN Bulletin*, vol. 36, no. 1, pp. 79–120, 2006.

[5] A. J. G. Cairns, "A discussion of parameter and model uncertainty in insurance," *Insurance: Mathematics and Economics*, vol. 27, no. 3, pp. 313–330, 2000.

[6] R. D. Lee and L. R. Carter, "Modeling and forecasting U.S. mortality," *Journal of the American Stotistical Association*, vol. 87, pp. 659–675, 1992.

[7] N. Brouhns, M. Denuit, and J. Vermunt, "Measuring the longevity risk in mortality projections," *Bulletin of the Swiss Association of Actuaries*, vol. 2, pp. 105–103, 2002.

[8] N. Brouhns, M. Denuit, and I. Keilegom, "Bootstrapping the Poisson log-bilinear model for mortality forecasting," *Scandinavian Actuarial Journal*, vol. 3, pp. 212–224, 2005.

[9] M.-C. Koissi, A. F. Shapiro, and G. Högnäs, "Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval," *Insurance: Mathematics and Economics*, vol. 38, no. 1, pp. 1–20, 2006.

[10] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, vol. 57 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, New York, NY, USA, 1993.

[11] Human Mortality Database, Berkley: University of California and Rostock: Max Planck Institute for Demographic Research, 2008, http://www.mortality.org/.

[12] N. Brouhns, M. Denuit, and J. K. Vermunt, "A Poisson log-bilinear regression approach to the construction of projected lifetables," *Insurance: Mathematics and Economics*, vol. 31, no. 3, pp. 373–393, 2002.

[13] M. Denuit, "Distribution of the random future life expectancies in log-bilinear mortality projection models," *Lifetime Data Analysis*, vol. 13, no. 3, pp. 381–397, 2007.

[14] R. Lee and T. Miller, "Evaluating the performance of the Lee-Carter method for forecasting mortality," *Demography*, vol. 38, no. 4, pp. 537–549, 2001.

[15] J. H. Maindonald, *Statistical Computation*, John Wiley & Sons, New York, NY, USA, 1984.

[16] A. C. Davison and D. V. Hinkley, *Bootstrap Methods and Their Application*, vol. 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge, UK, 1997.

[17] V. D'Amato, S. Haberman, and M. Russolillo, "Efficient bootstrap applied to the Poisson Log-Bilinear Lee Carter model," in *Proceedings of the Applied Stochastic Models and Data Analysis (ASMDA '09)*, L. Sakalauskas, C. Skiadas, and E. K. Zavadskas, Eds., pp. 374–377, 2009.

[18] D. J. Nordman, S. N. Lahiri, and B. L. Fridley, "Optimal block size for variance estimation by a spatial block bootstrap method," *The Indian Journal of Statistics*, vol. 69, no. 3, pp. 468–493, 2007.

[19] G. Taylor and G. McGuire, "A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error," *North American Actuarial Journal*, vol. 11, no. 3, pp. 70–88, 2007.

[20] W. Venables and B. Ripley, *Modern Applied Statistics Using S*, Springer, New York, NY, USA, 4th edition, 2002.

[21] A. E. Renshaw and S. Haberman, "A cohort-based extension to the Lee-Carter model for mortality reduction factors," *Insurance: Mathematics and Economics*, vol. 38, pp. 556–570, 2006.

[22] A. E. Renshaw and S. Haberman, "Lee-Carter mortality forecasting with age-specific enhancement," *Insurance: Mathematics & Economics*, vol. 33, no. 2, pp. 255–272, 2003.

[23] K. Dowd, A. J. Cairns, D. Blake, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah, "Backtesting Stochastic Mortality Models: An Ex-Post Evaluation of Multi-Period-Ahead Density Forecasts," Pensions Institute Discussion Paper PI-0803.

[24] A. J. G. Cairns, B. David, K. Dowd et al., "A quantitative comparison of stochastic mortality models using data from England and wales and the United States," *North American Actuarial Journal*, vol. 13, no. 1, pp. 1–35, 2009.

[25] 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*, vol. 73, no. 4, pp. 687–718, 2006.