The Lee-Carter Model for Forecasting Mortality, Revisited

Siu-Hang Li* and Wai-Sum Chan†

Abstract

Interrupting phenomena are commonly encountered in time-series data analysis with the study of mortality trends being no exception. Nevertheless, previous demographic forecasts have paid little attention to the existence of such phenomena. In this study we use mortality data from Canada and the United States to perform time-series outlier analysis on the key component of the Lee-Carter model: the mortality index. We begin by employing a systematic outlier detection process to ascertain the timing, magnitude, and persistence of any outliers present in historical trends of the mortality index. We then try to match the identified outliers with important events that could possibly justify the vacillations in human mortality levels. At the same time, we adjust the effect of the outliers for model reestimation. The empirical results indicate that the outlier-adjusted model could achieve better fits and more efficient forecasts of variables such as the central rates of death and the life expectancies at birth. Finally, we conclude our study with possible extensions on the valuations of life annuities and the probabilistic distribution of the highest attained age, incorporating the effect of mortality improvement portrayed by the revised model.

1. Introduction

The Lee-Carter model (Lee and Carter 1992) is the current gold standard of mortality trend fitting and projection. Over the past 10 years, the model and its variants have been used by actuaries for a wide range of purposes, from the forecasting of mortality reduction factors (Renshaw and Haberman 2003) to the assessment of retirement income adequacy (Chia and Tsui 2003). Other applications in demographic science include population projections (Booth and Tickle 2003), the forecasting of sex differentials in mortality (Carter and Lee 1992), and the projection of mortality patterns for the ''oldest-old'' (Buettner 2002). Intrinsically, the model assumes that the dynamics of death rates over time are driven by a single time-varying parameter, namely, the mortality index. The mortality forecast relies on the extrapolation of this index under an appropriate statistical linear time-series model.

With no exception to the mortality index, longitudinal data are often contaminated with various forms of discrepant observations, which, in the statistical literature, are commonly referred to as *outliers*. Outliers have a variety of sources: they may arise from mere recording or typographical errors, or from nonrepetitive exogenous interventions, such as pandemics or hostilities in a mortality series. Thus, the analysis of outliers may reveal invaluable information about the external shocks that affect the series, which may then enable actuaries to predict how the series will respond if these or similar interruptive events recur. When adequate data are available, outlier analysis may possibly shed some light on recent discussions on the impact of the SARS crisis (see Panjer 2003) and the effect of terrorist attacks (see Rowley and Bishop 2001; Wolak et al. 2003) on future mortality assumptions.

In addition, outlier detection and adjustment are critical to both model estimation and forecasting. Tsay (1986) noted that the existence of

* Siu-Hang Li is a graduate student in the Department of Statistics and Actuarial Science at the University of Waterloo, Ontario, Canada, shli@math.uwaterloo.ca.

† Wai-Sum Chan, PhD, FSA, is a Professor in the Department of Finance, Chinese University of Hong Kong, Shatin, N. T., Hong Kong, P. R. China, chanws@cuhk.edu.hk.

outliers might cause serious biases in both the sample autocorrelation and partial autocorrelation functions, thus resulting in erroneous model specification. Hillmer (1984) and Ledolter (1989) found that outliers could have a disastrous effect when they occur near the forecast origin, in part because of the bias in model parameter estimation and in part the carryover effect of the outlier on the forecast. Li and Chan (2005) further demonstrated that this carryover effect is exceptionally severe in the original Lee-Carter model (see Lee and Carter 1992; Tuljapurkar, Li, and Boe 2000), in which the mortality index is modeled by a random walk with drift.

Yet the existence of outliers often has been overlooked in extrapolative mortality forecasts. Even when the presence of outliers was acknowledged, the identification process was exclusively based on prior knowledge, rather than a sound statistical detection method. This is exemplified by Lee and Carter (1992), who viewed subjectively the known influenza epidemic in 1918 as an anomaly and dealt with it by means of an intervention model with a dummy variable. Although not commonly used in mortality studies, outlier analysis is widely accepted by actuaries in stochastic investment modeling (see, e.g., Chan and Wang 1998; Chan 1998). It has also been applied in fields as diverse as macroeconomic analysis (Balke and Fomby 1994), data mining (Liu and Hudak 2001), and clinical studies (Thomas, Plu, and Thalabard 1992).

In this study we perform a systematic time-series outlier analysis of the mortality index encompassed in the Lee-Carter model, using mortality data of both the American and the Canadian populations. We begin our analysis by identifying exogenous events that might significantly affect the mortality dynamics. We do this via Chen and Liu's (1993) method to search iteratively for outliers present on the mortality index, and then we match the identified outliers with events that may have caused them. Noting that the masking effect of outliers might lead to biases in parameter estimates (Tsay 1986) and even erroneous model identification (Chan 1995), we also implement the additional iterative cycle proposed by Li and Chan (2005), aiming at unveiling the true model underlying the outlier-free series of the mortality index. Along the way, we scrutinize the efficiency of the outlier-adjusted model by comparing the interval forecasts of the mortality index and related variables such as the central rates of death, the life expectancies at birth, and the actuarial present values of a life annuity at pensionable age.

So far actuaries and demographers have concentrated on the estimation of population aggregates, such as life expectancies at birth or at some other ages, but have paid far less attention to extreme values. Nonetheless, the rapid emergence of centenarians and supercentenarians has highlighted the importance of the "tail" in different types of life tables (Fishman 2002) and has motivated some actuaries to look for appropriate ways for closing off the life tables instead of the prevailing practice of using the value of 1 in an arbitrarily chosen limiting age (Johansen 2002). Thatcher (1999) made the first attempt to derive the probabilistic distribution of the highest attained age, actuarially denoted as ω, assuming that the age pattern of mortality follows a logistic distribution. Thatcher's study, however, made no effort to consider the effect of future mortality improvement on the maximum life span. Following in Thatcher's footsteps, we attempt to compute the probabilistic distribution of ω for various cohorts of Canadians, incorporating the mortality dynamics portrayed by the outlier-adjusted Lee-Carter model.

2. THE DATA

To fit the Lee-Carter model and to perform the outlier analysis, we require both the central rates of death and the exposures to risk. Below we list, for each population to be investigated, the sources of data, sample period, and modifications made (if any).

Canada: From 1921 to 2000 the death count and the exposure to risk (midyear population estimate) for every single year of age up to 99 are obtained from the Human Mortality Database (2005).

The United States: The matrix of age-specific central rates of death from 1900 to 2000 is available from the National Center for Health Statistics (2004a, b). Unfortunately the mortality data are presented in an abridged form: that is, values of death rates are shown at age 0, age group 1–4, decadal age groups 5–14, 15–24, and so on up to 75–84, and the open age group 85 and over.

Such a layout is not sufficient for the computation of various monetary functions involving life contingencies, which requires probabilities of death for every single year of age. To overcome this problem, we apply the disaggregation method proposed by Pollard (1988) to derive full life tables from the mortality data tabulated in 10-year age groups, assuming that the force of mortality (μ_x) varies in an exponential manner and that the population is stable within each of the age intervals. The death rates are further extended to age 99 by the method suggested in Coale and Guo (1989, pp. 614–15).

The exposure to risk is approximated by the midyear population estimate, which is available from the U.S. Census Bureau (2004). From 1990 to 2000 the estimates can be retrieved online, and from 1900 to 1989 the estimates are obtained by written request.

3. THE LEE-CARTER MODEL

The Lee-Carter model essentially describes the logarithmically transformed age-specific central rate of death ($m_{x,t}$) as a sum of an age-specific component that is independent of time (a_x), and the product of a time-varying parameter (k_t, also known as the mortality index) that summarizes the general level of mortality and an additional age-specific component (b_x) that represents how rapidly or slowly mortality at each age varies when the mortality index changes. Mathematically,

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}. \qquad (3.1)$$

The final term, $\varepsilon_{x,t}$, is the error term, which reflects the age-specific influences not captured by the model. Mortality forecasting is carried out using the model of the mortality index time series, on which the outlier analysis in this study is performed.

The equation underpinning the Lee-Carter model is known to be overparameterized. To stipulate a unique solution, we take a_x as the arithmetic mean of the $\ln(m_{x,t})$ over time, while the sums of b_x and k_t are normalized to unity and zero, respectively. As all parameters on the right-hand side of equation (3.1) are unobservable, fitting the model by the ordinary least squares method is impossible. To overcome the situation, we employ Lee and Carter's (1992) two-stage estimation procedure, which gives exact solutions. In the first stage, singular value decomposition (SVD) is applied to the matrix of $\{\ln(m_{x,t})—a_x\}$ to obtain estimates of b_x and k_t. The SVD procedure can be implemented by using various standard mathematical/statistical packages such as GENSTAT (Genstat 5 Committee 1993), MATLAB (Brose 1997), and IMSL MATH/ LIBRARY (IMSL 1997). In the second stage, the time series of k_t is reestimated by solving for k_t such that

$$D_t = \sum_x \{\exp(a_x + b_x k_t) N_{x,t}\}, \qquad (3.2)$$

where D_t is the total number of deaths in time t, and $N_{x,t}$ is the exposure to risk of age x in time t. This is to ensure that the mortality schedules fitted over the sample years will reconcile the total number of deaths and the population age distributions. An autoregressive integrated moving-average (ARIMA) model is then used to model the dynamics of k_t. The Box and Jenkins (1976) approach often is employed to obtain a fitted ARIMA model from the empirical k_t data.

In addition to the two-stage estimation procedure, many alternative methods have been proposed in the literature. For example, Wilmoth (1993) and Brouhns, Denuit, and Vermunt (2002) considered a regression-type model under the Poisson assumption on the number of deaths (D_t). In this way standard maximum likelihood estimates for the a_x's, b_x's, and k_t's can be obtained. This parametric approach has excellent statistical properties if the assumed Poisson density is the correct one, but could lead to grossly incorrect inferences if the D_t (or k_t) series is contaminated with outliers. To alleviate the situation, the penalized likelihood approach may be used. A penalty function $P(k_t)$ is subtracted from the original log-likelihood function, where $P(k_t) \geq 0$ is a roughness penalty that decreases as k_t becomes smoother. Maximum penalized likelihood estimates can be computed by maximizing the modified log-likelihood function. This method has the advantage of smoothing the estimated k_t's and allowing for an extrapolation directly from the penalty structure. However, it requires a subjective choice of $P(k_t)$, and different choices of the penalty function will yield different estimators. Furthermore, it may smooth out useful information concerning occasional disturbances to the general mortality evolution pattern (i.e., outliers in

k_t). Therefore, we shall concentrate only on the original fitting methodology in this article.

In Appendix A Figures A.1 and A.2, respectively, plot the fitted values of k_t based on the Canadian and American mortality (they also plot forecasts, which should be ignored for now). The trends are roughly linear despite the existence of several aberrant observations, which will be discussed in later sections. For completeness, corresponding fitted values of a_x and b_x for both countries are shown in Table 1.

4. OUTLIER ANALYSIS

Having fitted the Lee-Carter model, we are now ready to proceed to the outlier analysis of the mortality index, k_t. Outlier analysis of time-series data comprises two key issues:

- To search for the location and type of outliers in a contaminated time series (known as the outlier detection problem in the time-series literature).
- To obtain better estimates of parameters in the underlying time-series model through the incorporation of outlier effects within a model (known as the outlier adjustment problem).

Time-series outlier analysis was first studied by Fox (1972), who employed the likelihood ratio test for outlier detection, but it has also been considered by many other researchers. Chang, Tiao, and Chen (1988) developed an iterative procedure for outlier detection and employed intervention models to incorporate outlier effects. Chen and Liu (1993) proposed an augmented iterative procedure for the joint estimation of model parameters and outlier effects. Other references on this topic include Abraham and Chuang (1989), Ljung (1993), Muirhead (1986), Tsay (1986, 1988), Vaage (2000), and Wei (1990).

The first phase of an outlier detection procedure is to specify the underlying stochastic structure, that is, the appropriate ARIMA model, for the outlier-free series. However, the true underlying stochastic structure is seldom known a priori in practice, and this brings about the following problems, as noted by Chen and Liu (1993):

- The existence of outliers may result in erroneous model selection. Chan (1995) showed that outliers create spurious autocorrelations, which can lead to erroneous model specification. This cannot be avoided even when the sample size is enlarged.
- Even if the model is correctly specified, outliers may lead to biases in the estimation of parameters. This affects the detection of outliers and ultimately obscures the reestimation of parameters; the whole process repeats itself indefinitely.

In this study we employ both the Chen and Liu (1993) method and the additional iterative cycle proposed by Li and Chan (2005) to ameliorate the problem of model selection. For brevity we will restrict our discussion to points necessary for describing the applications in this paper.

Table 1

Fitted Values of a_x and b_x at Selected Ages

Age x	Canada a_x	Canada b_x	United States a_x	United States b_x
0	−3.6316	0.0241	−3.3311	0.0178
5	−7.1821	0.0225	−6.4166	0.0229
10	−7.6024	0.0204	−7.1241	0.0186
15	−7.1422	0.0151	−6.6891	0.0157
20	−6.5346	0.0119	−6.2542	0.0128
25	−6.5758	0.0138	−6.1025	0.0131
30	−6.4769	0.0141	−5.9509	0.0134
35	−6.2490	0.0132	−5.7021	0.0122
40	−5.9287	0.0116	−5.4534	0.0110
45	−5.5456	0.0098	−5.1038	0.0093
50	−5.0952	0.0082	−4.7543	0.0076
55	−4.6900	0.0065	−4.3690	0.0066
60	−4.2118	0.0064	−3.9837	0.0057
65	−3.7520	0.0063	−3.5889	0.0054
70	−3.3336	0.0061	−3.1941	0.0052
75	−2.8779	0.0061	−2.7685	0.0052
80	−2.4355	0.0053	−2.3428	0.0052
85	−1.9761	0.0048	−1.9178	0.0052
90	−1.5384	0.0042	−1.5042	0.0049
95	−1.1960	0.0020	−1.1056	0.0042

4.1 Outlier Models

We assume that the outlier-free time series Z_t follows an autoregressive integrated moving average, ARIMA(p, d, q) model,

$$\varphi(B)(1 - B)^d Z_t = \theta(B)a_t, \qquad (4.1)$$

where B is the backshift operator such that $B^s Z_t, = Z_{t-s}$,

$$\varphi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p,$$

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q,$$

and $\{a_t\}$ is a sequence of white noise random variables, iid with mean 0 and constant variance σ^2.

By definition, outliers are nonrepetitive exogenous interventions. Quantitatively, an outlier-contaminated time series Y_t consists of an outlier-free time series Z_t plus an exogenous intervention effect, denoted as $\Delta_t(T, \omega)$, that is,

$$Y_t = Z_t + \Delta_t(T, \omega), \tag{4.2}$$

where T is the locality of the outlier and ω is the magnitude of the outlier.

We consider four common types of outliers, namely, an additive outlier (AO), innovational outlier (IO), level shift (LS), and temporary change (TC). Let $D_t^{(T)}$ be an indicator variable that equals 1 when $t = T$ and 0 otherwise:

- An AO affects only the level of a single observation, that is,

$$\Delta_t(T, \omega) = \omega D_t^{(T)}. \tag{4.3}$$

- An IO affects all observations beyond T through the memory of the underlying outlier-free process, that is,

$$\Delta_t(T, \omega) = \frac{\theta(B)}{\varphi(B)(1 - B)^d} \omega D_t^{(T)}. \tag{4.4}$$

- A LS affects a series at a given time, and its effect is permanent, that is,

$$\Delta_t(T, \omega) = \frac{\omega}{1 - B} D_t^{(T)}. \tag{4.5}$$

- A TC affects a series at a given time, and its effect decays exponentially according to a dampening factor, say, δ, that is,

$$\Delta_t(T, \omega) = \frac{\omega}{1 - \delta B} D_t^{(T)}. \tag{4.6}$$

In practice, the value of δ often lies between 0.6 and 0.8 (Liu and Hudak 1994, p. 76). In our study we take $\delta = 0.7$ as recommended by Chen and Liu (1993).

A graphical illustration of each type of outlier is given in Figure 1. For more detailed discussions of these types of outliers, see Chen and Tiao (1990), Fox (1972), and Tsay (1988). In general, a time series may contain more than one, say, m, outliers, and we have the general time-series outlier model

$$Y_t = Z_t + \sum_{i=1}^{m} \Delta_t(T_i, \omega_i). \tag{4.7}$$

4.2 Detection and Adjustment

The detection process proposed by Chen and Liu (1993) is primarily based on the effect of outliers on the estimated residuals. To begin, we define a polynomial $\pi(B)$ as

$$\pi(B) = \frac{\varphi(B)(1 - B)^d}{\theta(B)} = 1 - \pi_1 B - \pi_2 B^2 - \cdots. \tag{4.8}$$

Then equation (4.1) can be rewritten as

$$\pi(B) Z_t = a_t, \tag{4.9}$$

and the fitted residuals $\hat{e}_t$, which may be contaminated with outliers, can be readily obtained and expressed as

$$\hat{e}_t = \pi(B) Y_t. \tag{4.10}$$

So, for each type of outlier, we have the following:

$$\text{AO:} \quad \hat{e}_t = \omega D_t^{(T)} + a_t, \tag{4.11}$$

$$\text{IO:} \quad \hat{e}_t = \omega \pi(B) D_t^{(T)} + a_t, \tag{4.12}$$

$$\text{TC:} \quad \hat{e}_t = \omega \frac{\pi(B)}{(1 - \delta B)} D_t^{(T)} + a_t, \tag{4.13}$$

$$\text{LS:} \quad \hat{e}_t = \omega \frac{\pi(B)}{1 - B} D_t^{(T)} + a_t. \tag{4.14}$$

Alternatively, we can rewrite equations (4.11)–(4.14) as a general time-series regression,

$$\hat{e}_t = \omega d(j, t) + a_t, \tag{4.15}$$

where $j \in J = \{\text{AO, IO, TC, LS}\}$; $d(j, t) = 0$ for all j and $t < T$; $d(j, T) = 1$ for all j; and for all $k \geq 1$,

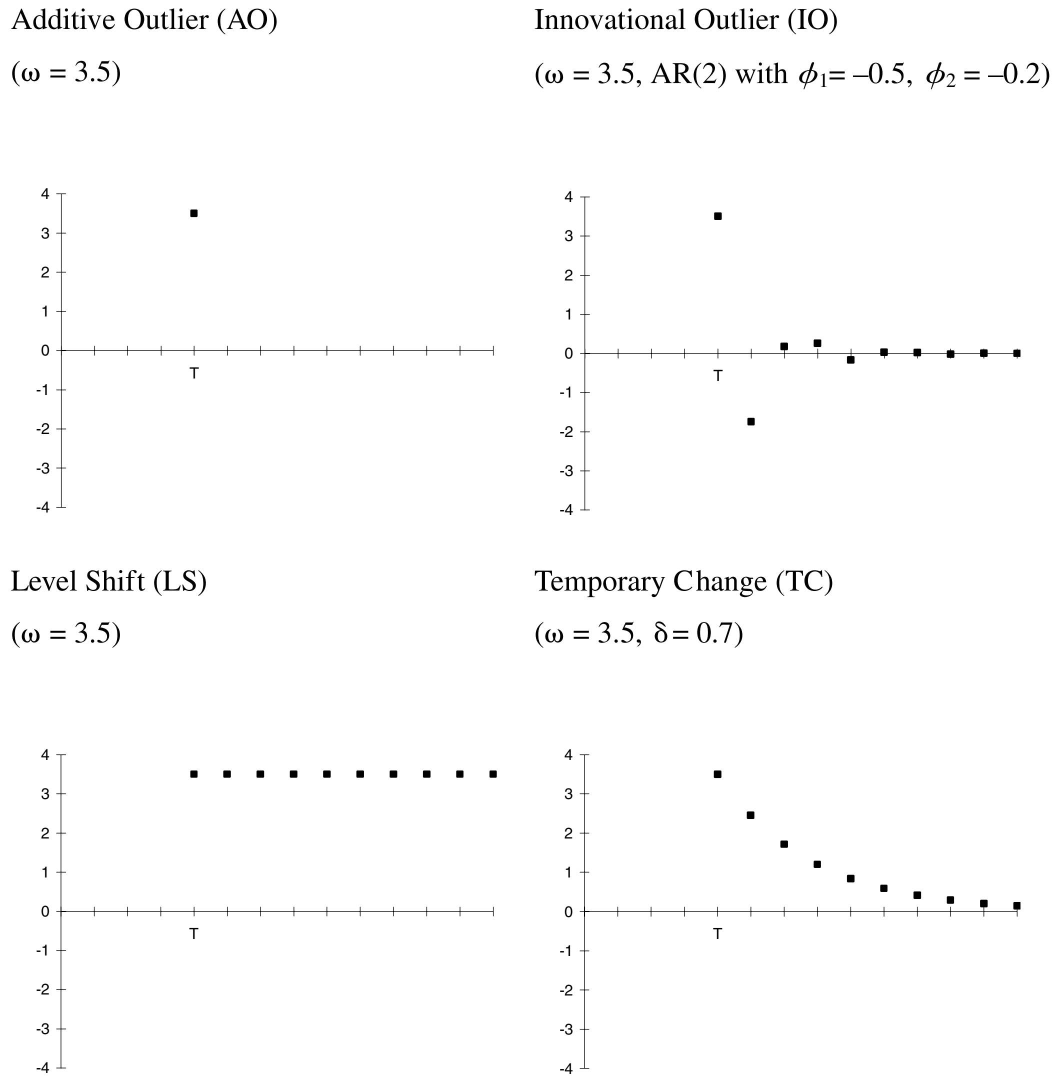
$$d(\text{AO}, T + k) = 0,$$

$$d(\text{IO}, T + k) = -\pi_k,$$

$$d(\text{TC}, T + k) = \delta_k - \sum_{i=1}^{k-1} \delta^{k-j} \pi_j - \pi_k,$$

$$d(\text{LS}, T + k) = 1 - \sum_{i=1}^{k} \pi_k.$$

Figure 1
Different Types of Time-Series Outliers



Hence, for a given T (suspected locality of the outlier) and j (suspected type of outlier), the standardized t-statistic $\tau(j, T)$ for the effect of the outlier, that is, the slope parameter ω, can be readily computed using the principle of least squares. The final test statistic is the maximum value of this t-statistic over all possible T and j, that is,

$$\mathcal{T} = \max_{1 \le T \le n} \max_{j \in J} \{\tau(j, T)\}. \qquad (4.16)$$

For a given j, the test statistics follow approximately a normal distribution. An outlier of type j is detected if the final test statistic is greater than a critical value of C. We employ $C = 2.5$ in this paper as recommended by Liu and Hudak (1994) for a reasonable level of sensitivity. With the type and the location of an outlier, we can jointly reestimate the model parameters and the outlier effect. After the estimation is complete, we can adjust the original series for the effect of the detected outlier. The process is repeated until no other outlier is found.

The outlier detection-estimation-adjustment process can be implemented by using various standard statistical software packages for time-series analysis, such as AUTOBOX (Ord and Lowe 1996), SAS/ETS (SAS Institute 2004), and SCA (Liu and Hudak 1994).

To prevent problems with potential erroneous model selection, we also employ the following external iteration cycle proposed by Li and Chan (2005) to incorporate outlier effects in model specification.

Step 1 Tentative Model Identification

Use the Box and Jenkins approach (Box and Jenkins 1976), or any other appropriate method, to identify tentatively the order of the underlying outlier-free ARIMA(p, d, q) model.

Step 2 Outlier Detection and Adjustment

Use Chen and Liu's (1993) iterative procedures for a joint estimation of outlier effects and ARIMA model parameters. After the incorporation of outlier effects, an outlier-adjusted data series is obtained.

Step 3 Reidentification of the Model

Using the method employed in Step 1, identify the ARIMA model underlying the adjusted data series. If the reidentification makes a difference in p, d, and/or q, go back to Step 2 using the original unadjusted data series under the reidentified order (p, d, q). Otherwise, terminate the iteration cycle, and the ultimate estimates of outliers and ARIMA model parameters are those obtained in the immediately previous Step 2.

5. EMPIRICAL RESULTS

5.1 Detected Outliers

Table 2 displays all the outliers found in the mortality indices derived from the American and Canadian data. Note that a negative outlier stands for improvement in mortality, whereas a positive one means deterioration. The outliers mostly are found in the first half of the century, and they can be roughly classified as pandemic-related and war-related.

World War I best explains the mild positive level shift in 1916. After a long stretch of isolationism, President Woodrow Wilson requested that the U.S. Congress declare war, which it did on April 6, 1917. The American contribution to the war was substantive: total deaths of the American armed forces amounted to 213,799[1] (Tucker 2005). This deadly war ended in 1918, and the fact that young men were no longer being killed could possibly account for the negative temporary change in 1921.

Table 2
Summary of Outliers Detected in the Mortality Indices

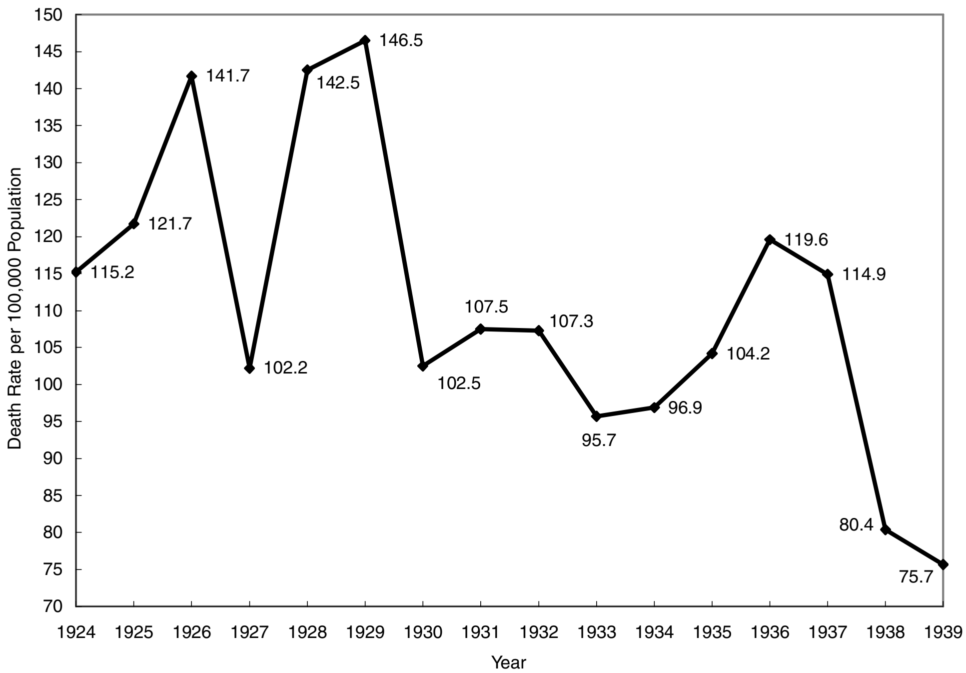
Year	Size	t-Value	Type
Canada			
1926	7.119	3.49	LS
1937	6.131	4.25	AO
United States			
1916	7.120	3.55	LS
1918	24.473	13.06	AO
1921	−9.012	−4.40	TC
1928	8.117	3.89	TC
1936	9.465	4.64	TC
1954	−6.112	−3.00	TC
1975	−7.880	−3.94	LS

The positive additive outlier in 1918 is most likely a consequence of the "Spanish flu" pandemic, which particularly affected healthy young adults. This pandemic of unprecedented virulence spread around the world and killed approximately 479,000 Americans (Crosby 2003) and between 20 and 40 million people worldwide over the next six months (Holmes 2004); it was thought to have been the most deadly pandemic in human history. It is noteworthy that this gigantic outlier may have masked the potential dampening effect of the previous outlier, and this offers an explanation to why the outlier in 1916 is identified as a level shift instead of a temporary change as one may expect.

Figure 2 shows the death rates from influenza and pneumonia in the United States during this time of turmoil—the Great Depression and war. Abrupt jumps are observed in 1926, 1928–29, and 1936–37. This suggests that some small-scale and probably not widely reported influenza epidemic may have occurred during this time of economic insecurity, which may have given rise to

[1] This includes deaths from battlefield wounds and diseases. Readers should note that uncertainty always exists in the determination of the number of combatant deaths. Moreover, scholars are seldom in agreement on the inclusion of deaths from indirect causes, e.g., the flu epidemic.

Figure 2
Death Rates from Influenza and Pneumonia, United States, 1924–1939



Source: National Center for Health Statistics (2004a).

outliers in the late 1920s and 1930s. Note that because of the lack of information, the type and magnitude of the outlier may not be reliable if it lies within the first or last few observations. It is therefore possible that the type of the outlier in 1926 is misidentified.

In 1950 North and South Korea went to war. The Korean War killed about 54,000 U.S. servicemen (American Battle Monuments Commission 2006) and temporarily reduced the population of healthy and robust Americans. The cessation of young men being killed after the war's end might account for the negative temporary change in 1954.

Finally, the negative level shift in 1975 may be associated with better nutrition, the use of antibiotics, or some other improvements in social conditions.

5.2 The Outlier-Adjusted Model

Below we evaluate the model performance in each mortality series under investigation. The identified ARIMA order does not make a difference in Li and Chan's (2005) iterative cycle: (0,1,0) fits both indices well throughout the entire process. Ljung-Box's Portmanteau statistics (Ljung and Box 1978) are computed up to the tenth lags for testing serial correlation of the residuals. The results (not shown) do not suggest any inadequacy of the fitted models, either with or without outlier adjustment.

To provide a tentative impression of the fitting performance, we report the residual standard errors in Table 3. Upon outlier detection and adjustment, values of residual standard error are significantly reduced. However, under the principle of parsimony, one should employ the least possible number of parameters for adequate representations, and it is therefore premature to make conclusions solely from the reduction in residual standard errors as the total number of model parameters is increased, as additional intervention components for outlier effects are included. To compare the fitting performance of models with a different number of parameters, Akaike (1974) introduced the Akaike Information Criterion (AIC), which is defined as

$$\text{AIC} = n \ln(\hat{\sigma}^2) + 2M, \qquad (5.1)$$

where n is the number of effective observations,

Table 3
Fitting Diagnostics, ARIMA Model for the Mortality Indices

	Residual Standard Error		AIC	
	Before Adjustment	After Adjustment	Before Adjustment	After Adjustment
Canada	2.3675	2.0401	139.89	120.08
United States	4.3190	2.2047	294.60	174.12

$\hat{\sigma}$ denotes the residual standard deviation, and M represents the number of parameters in the model. The criterion rewards the reduction of the residual standard deviation but at the same time harshly penalizes the increased number of parameters. In other words, a smaller AIC value is more favorable. In Table 3 we also report the AIC values both before and after outlier adjustment. The results support the view that the outlier-adjusted models provide a better fit.

We also conducted a simple within-sample test of the performance of the models. We fitted both the original and the outlier-adjusted model to mortality data prior to year 1990. Based on the refitted models, we constructed forecasts of the life expectancy at birth to year 2000 and compared these forecasts with the actual values. The results, shown in Table 4, are supportive to the performance of the models. In all forecasts, the actual values lie within the 95% pointwise confidence intervals. Also, the absolute percentage errors (APEs) are at an acceptable level of less than 1.09% for the original Lee-Carter model and 1.04% for the outlier-adjusted version.

Chatfield (1993) pointed out that it is important to provide interval forecasts so that forecasts obtained by different methods may be compared more thoroughly. In Table 5 we report the interval forecast for the central rates of death at selected ages; and in Appendix A, Figures A.1–A.6 we plot both the mean and the interval forecasts of the mortality index, the life expectancy at birth, and the actuarial present value (APV) of a life annuity at age 65. The confidence intervals are substantially narrower after the adjustment for outliers; they are coherent with the empirical results of Ledolter (1989), which found that estimated prediction intervals are sensitive to outliers, as they strongly inflate the estimate of the innovation variance. For mathematical derivations of the interval forecast for the APV of life annuities, refer to Appendix B.

Readers should note the interpretation of the reduction in width of the interval forecasts. There are a number of reasons for the increased narrowness. First, as mentioned earlier, the detection and adjustment process effectively mitigates the problem of biases in parameter estimation due to outliers. Second, if an outlier occurs at the forecast origin, an assumption on the type of that outlier is required, and the uncertainty of the type of the outlier is not reflected in the interval forecast. Third, the narrowness essentially reflects the optimistic nature of the outlier-adjusted forecasts. Under the outlier-adjusted models we presume that the same event will not recur in the future, and so the question arises which model a practitioner should pursue during the forecasting exercise. Chan (2002, pp. 559–60) offers the following suggestion for stochastic investment modeling:

> Whether or not it is appropriate to adjust the data for outliers depends on the purpose to which the model so derived will be used. If the model will be used in an application for which extreme stochastic fluctuations are less important (e.g. to ensure that premiums are adequate in most, but not extreme, scenarios), then it may be preferable to use a model based on outlier-adjusted data. If, however, the model will be used in an application for which extreme stochastic fluctuations are important (such as pricing catastrophe risks or ensuring that investment guarantee reserves are sufficient to keep an insurance company solvent in all but the most extreme scenarios), then a model which is sympathetic to outliers in the data ought to be used.

This rationale also applies to mortality forecasting. The outlier-adjusted version may be more preferred if the model is used in an application where the major mortality trend is the primary focus, for example, in forecasting cohort life tables for the assessment of pecuniary loss in per-

Table 4
Within Sample Forecast of Life Expectancy at Birth (e_0)

Year	Actual Value	Without Outlier Adjustment			With Outlier Adjustment		
		Central Estimate	95% Interval	APE	Central Estimate	95% Interval	APE
Canada							
1991	77.08	77.33	(76.92, 77.73)	0.31%	77.33	(76.93, 77.67)	0.32%
1992	77.34	77.46	(76.89, 78.03)	0.16	77.48	(76.91, 77.95)	0.18
1993	77.22	77.60	(76.91, 78.29)	0.49	77.62	(76.93, 78.19)	0.51
1994	77.41	77.73	(76.94, 78.52)	0.42	77.76	(76.97, 78.41)	0.45
1995	77.53	77.86	(76.99, 78.73)	0.43	77.89	(77.03, 78.61)	0.47
1996	77.79	77.99	(77.05, 78.93)	0.26	78.03	(77.09, 78.81)	0.31
1997	77.99	78.12	(77.11, 79.12)	0.16	78.16	(77.16, 79.00)	0.23
1998	78.20	78.24	(77.18, 79.31)	0.05	78.30	(77.23, 79.18)	0.12
1999	78.42	78.37	(77.25, 79.49)	0.07	78.43	(77.31, 79.35)	0.01
2000	78.77	78.49	(77.33, 79.66)	0.35	78.56	(77.40, 79.52)	0.27
United States							
1991	75.54	76.07	(75.15, 77.00)	0.71	76.07	(75.14, 76.56)	0.70
1992	75.82	76.23	(74.93, 77.53)	0.54	76.21	(74.91, 76.91)	0.52
1993	75.57	76.38	(74.81, 77.96)	1.07	76.36	(74.48, 77.20)	1.04
1994	75.74	76.53	(74.73, 78.33)	1.05	76.50	(74.70, 77.46)	1.01
1995	75.85	76.68	(74.69, 78.68)	1.09	76.65	(74.65, 77.71)	1.04
1996	76.20	76.83	(74.67, 79.00)	0.84	76.79	(74.62, 77.94)	0.78
1997	76.56	76.98	(74.67, 79.29)	0.55	76.93	(74.61, 78.16)	0.48
1998	76.75	77.13	(74.68, 79.58)	0.50	77.07	(74.62, 78.38)	0.42
1999	76.76	77.27	(74.70, 79.85)	0.67	77.20	(74.63, 78.58)	0.58
2000	76.87	77.41	(74.73, 80.10)	0.71	77.34	(74.65, 78.78)	0.61

sonal injury litigations (Sarony, Chan, and Chan 2003). Some other measures that allow for the recurrence of outliers should be undertaken if the model is used in an application where extreme fluctuations are important, for example, in predicting the highest attained age. Further discussions on this issue are provided in the next section.

6. EXTENSION—PREDICTING THE HIGHEST ATTAINED AGE OF CANADIANS

Suppose that N members of a birth cohort survive to age x_0; then when all of these N members have died, there will be a highest value, say, ω_N, among their ages at death, and prior to realization, ω_N can be regarded as a random variable with the distribution function

$$\Pr(\omega_N < x) = (1 - {}_{x-x_0}p_{x_0})^N. \qquad (6.1)$$

Based on equation (6.1), Thatcher (1999) showed that if human lifetime follows a logistic distribution, then the distribution of ω_N has a closed form, given by

$$\Pr(\omega_N < x) = \left[1 - \exp\left(\gamma(x_0 - x) + \beta^{-1}\ln\left(\frac{1 + \alpha\exp(\beta x_0)}{1 + \alpha\exp(\beta x)}\right)\right)\right]^N, \qquad (6.2)$$

where α, β, and γ are the parameters in the logistic mortality model, defined by

$$\mu_x = \frac{\alpha\exp(\beta x)}{1 + \alpha\exp(\beta x)} + \gamma. \qquad (6.3)$$

Thatcher also demonstrated that the distribution function is insensitive to both the starting age x_0 and the size of cohort N. In practice, we may let $x_0 = 70$ and approximate N by the midyear population estimate at age x_0.

Thatcher's methodology, however, has not yet considered the possible changes of human mortality levels in the future. Here we explore the impact of future mortality improvement on the highest attained age by integrating the model specified in equation (6.1) with the outlier-adjusted Lee-Carter model. To illustrate the integration of the models, let us suppose that we are interested in estimating the distribution of

Table 5
95% Interval Forecast of Central Death Rates at 2025

Age	Without Outlier Adjustment		With Outlier Adjustment		Percentage Reduction in Width of Interval
	Lower Limit	Upper Limit	Lower Limit	Upper Limit	
Canada					
1	0.000022	0.000094	0.000022	0.000081	18.43%
20	0.000226	0.000405	0.000228	0.000383	14.00
60	0.005439	0.007468	0.005468	0.007239	12.70
United States					
1	0.000038	0.000169	0.000053	0.000125	44.81
20	0.000239	0.000489	0.000279	0.000423	42.38
60	0.007288	0.010055	0.007805	0.009422	41.57

the highest attained age of the cohort born in year 1920, and that we let $x_0 = 70$ in equation (6.1). Then the distribution of ω_N requires the death probabilities, q_x, for $x \geq 70$. For $x = 70, \ldots, 80$, q_xs are available in the historical data, and for $x > 80$, q_xs can be obtained from the Lee-Carter mortality projection.

The proposed analysis is heavily dependent on the appropriateness of the Lee-Carter model in projecting old-age mortality. To justify our methodology, we perform another within-sample test, similar to that in Section 5. We first estimate the Lee-Carter model parameters by the historical data from year 1921 to 1990, beginning at age 70. Then we forecast the life expectancy at age 85 from year 1991 to 2000 and compare the forecast with the actual life expectancies. The results are presented in Table 6. The average absolute percentage error is approximately 3.5%, which is slightly higher than that in the within-sample forecast of e_0. In all the forecasts, the actual values lie within the 95% pointwise confidence intervals. Overall, the results indicate a reasonable fitness of the Lee-Carter model at the older ages.

Table 6
Within-Sample Forecast of the Life Expectancy at Age 85 (e_{85}), Canada

Year	Actual Value	Without Outlier Adjustment			With Outlier Adjustment		
		Central Estimate	95% Interval	APE	Central Estimate	95% Interval	APE
Male							
1991	4.66	4.91	(4.66, 5.16)	5.30%	4.91	(4.66, 5.17)	5.38%
1992	4.82	4.94	(4.58, 5.31)	2.50	4.96	(4.59, 5.32)	2.77
1993	4.70	4.98	(4.54, 5.42)	5.99	5.00	(4.55, 5.45)	6.46
1994	4.86	5.02	(4.50, 5.53)	3.29	5.05	(4.53, 5.57)	3.93
1995	4.88	5.05	(4.47, 5.63)	3.56	5.09	(4.51, 5.68)	4.39
1996	5.00	5.09	(4.45, 5.72)	1.75	5.14	(4.49, 5.78)	2.74
1997	5.05	5.12	(4.43, 5.81)	1.34	5.18	(4.48, 5.88)	2.51
1998	5.08	5.16	(4.42, 5.90)	1.58	5.23	(4.48, 5.98)	2.94
1999	5.21	5.19	(4.41, 5.98)	0.32	5.27	(4.47, 6.07)	1.18
2000	5.73	5.23	(4.40, 6.06)	8.75	5.32	(4.47, 6.17)	7.21
Female							
1991	6.08	6.07	(5.85, 6.29)	0.20	6.08	(5.90, 6.25)	0.12
1992	6.26	6.14	(5.82, 6.45)	2.06	6.15	(5.91, 6.40)	1.80
1993	6.05	6.20	(5.81, 6.59)	2.55	6.23	(5.93, 6.53)	3.01
1994	6.10	6.27	(5.82, 6.72)	2.66	6.31	(5.96, 6.66)	3.30
1995	6.12	6.33	(5.82, 6.84)	3.51	6.38	(5.99, 6.78)	4.34
1996	6.16	6.40	(5.84, 6.96)	3.92	6.46	(6.03, 6.89)	4.93
1997	6.11	6.46	(5.86, 7.07)	5.75	6.54	(6.07, 7.01)	6.95
1998	6.21	6.53	(5.88, 7.18)	5.15	6.61	(6.11, 7.12)	6.52
1999	6.33	6.60	(5.90, 7.29)	4.17	6.69	(6.15, 7.23)	5.70
2000	6.53	6.66	(5.93, 7.40)	2.09	6.77	(6.20, 7.34)	3.76

Accurate crude data are also crucial to the analysis. Nevertheless, it is neither possible nor desirable to rely solely on crude data for several reasons: (1) Raw death counts beyond age 99 prior to calendar year 1950 are not available. This prohibits us from fitting the Lee-Carter model and consequently deriving the distribution of ω_N by relying entirely on crude data. (2) The deaths and number of persons at risk dwindle toward the advanced ages, leading to enormous sampling error and highly volatile crude death rates. (3) Age exaggeration is common among supercentenarians, leading to a possible upward bias in the estimated probability distribution of the highest attained age. Although the problem of age exaggeration may be resolved by using validated data, the number of validated cases of supercentenarians in Canada as of September 2001 is only 25 (see Robine and Vaupel 2002), which is far too scanty to produce useful statistical inferences of any kind. Based on verified ages at death in Canada, Bourbeau and Desjardins (2002) concluded that, for the time being, official statistics are not to be counted on to provide a conclusive picture of patterns of mortality at the highest ages, and mathematical techniques must still be counted on to establish the later years of the life tables. The use of extrapolative methods, either mathematical or statistical, is therefore inevitable in the analysis of the highest attained age (see, e.g., Han 2005; Robertson, Dupuis, and Jones 2005; Thatcher 1999).

On the other hand, it is inappropriate to base too heavily on extrapolation, since the pace of mortality improvement at the extreme ages is usually much less. For instance, the ratio of b_{85} to b_{95} in Table 1 suggests that the speed of mortality improvement at age 95 is only about half of that at age 85. This means if we begin the extrapolation at age 85, the projected death rate at the extreme ages could be seriously overstated. Having considered the availability and the quality of the available data, we use crude data up to age 99 and resort to extrapolative estimates afterwards.

We extend the projected cohort death rates to age 100 and above by the relational mortality model (Himes, Preston, and Condran 1994), which requires no subjective assumption on the limiting age. The model, recommended by the Working Group on Projecting Old-Age Mortality and Its Consequences of the Population Division in the United Nations (United Nations 1997), consists of a standard age pattern of mortality by sex and by single year of age from 45 to 99, calibrated from 82 different mortality schedules (for each sex) observed in a variety of low-mortality countries. (The standard mortality schedule from age 45 to 99 can be found in Himes et al., p. 273.) The standard is then made useful at higher ages by fitting a straight line to the logits of the age-specific death rates, beginning at age 80, as recommended by Himes et al. Mathematically,

$$\text{logit}(m_x^s) = \alpha + \beta x, \tag{6.4}$$

where $\text{logit}(m_x^s) = \ln(m_x^s/(1 - m_x^s))$, and m_x^s denotes the standard central death rate at age x. Equation (6.4) is then used to produce standard age-specific death rates from age 100 and up. Finally, we can relate the extended standard schedule to the projected cohort life tables by regressing on their logit transformations, that is,

$$\text{logit } m_x = \delta + \gamma \text{ logit } m_x^s, \tag{6.5}$$

where m_x's are the death rates in the projected cohort life table. The regression coefficients, δ and γ, are estimated by m_x and m_x^s for $x = 85 \ldots 99$. The projected cohort life table can be extended to age 100 and above based on the estimates of δ and γ.

It is noteworthy to see the limiting behavior of this extension. Equation (6.4) implies that

$$\lim_{x\to\infty} m_x^s = \lim_{x\to\infty} \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}} = 1, \tag{6.6}$$

for our estimates of α and β, which satisfy the condition of $\alpha < 0$ and $\beta > 0$. Then it follows that

$$\lim_{x\to\infty} m_x = \frac{e^{\delta+\gamma}}{1 + e^{\delta+\gamma}}. \tag{6.7}$$

In other words, as x tends to infinity, the value of m_x tends to a certain limit that is less than one. This is consistent with the observed leveling off of mortality rates at the extreme ages (see Bourbeau and Disjardins 2002; Robine and Vaupel 2002).

As noted earlier, it is important to allow for the recurrence of outliers in the analysis of the highest attained age in which extreme fluctuations of mortality rates are important. Here we consider

Table 7
Predicted Highest Attained Age for Selected Cohorts of Canadians

	Year of Birth (Attained Age in Year 2006)				
	1891 (115)	1901 (105)	1911 (95)	1921 (85)	1931 (75)
	Male, without Outlier Adjustment				
Mode	110	110	111	111	111
99th percentile	116	116	116	116	117
	Male, with Outlier Adjustment				
Mode	110	110	111	111	111
99th percentile	116	116	116	116	117
	Female, without Outlier Adjustment				
Mode	112	113	114	114	115
99th percentile	118	119	120	120	120
	Female, with Outlier Adjustment				
Mode	112	113	114	115	115
99th percentile	118	119	120	120	120

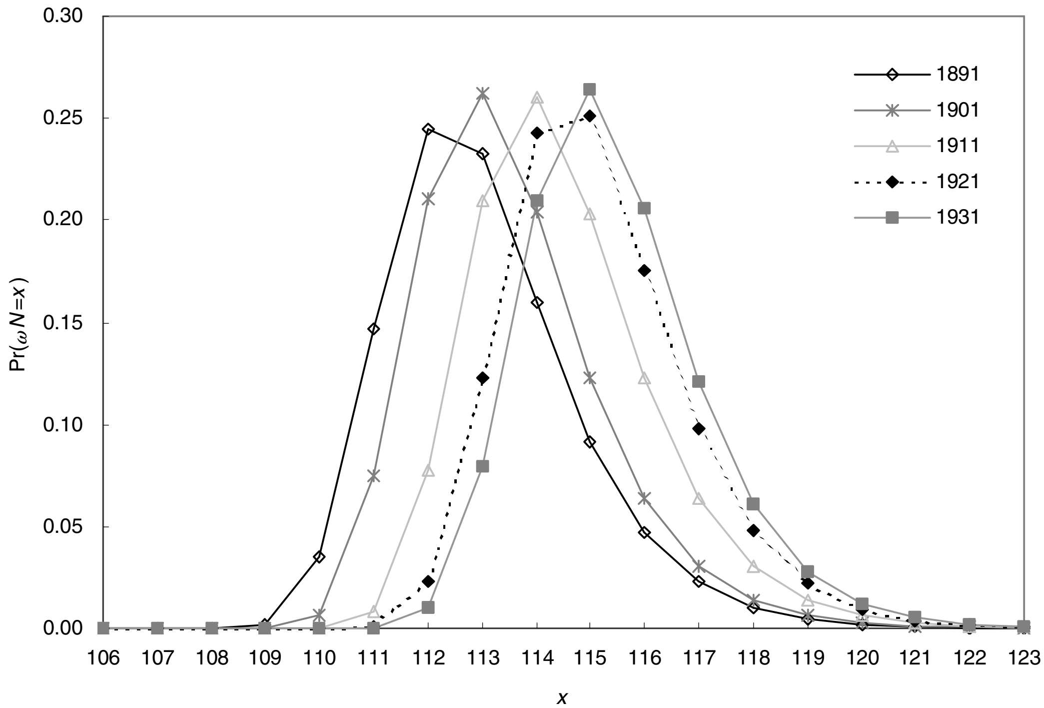
the possibility of outlier recurrence by using a bootstrapping procedure: Outliers in the post-sample forecasts are obtained by sampling with replacement from the pool of detected outliers in the historical series of k_t, separately for each sex. The outliers drawn (if any) are then superimposed to the sample path of the future k_ts.

Overall the procedures in estimating the probability distribution of ω_N can be summarized as follows:

1. Estimate the Lee-Carter parameters a_x, b_x, and k_t, for $x = 70, \ldots, 99$ and $t = 1921, \ldots, 2000$.
2. Based on the time series of k_t obtained in Step 1, obtain the estimates of outliers and the parameters p, d, and q of the outlier-adjusted time-series model, using iterative procedures mentioned in Section 5.
3. To incorporate the possible recurrence of outliers, we generate 1000 sample paths of k_t, where $t > 2000$, based on the ARIMA parameters obtained in Step 2. For each of the sample paths:
   a. Sample with replacement from the pool of detected outliers in the historical series of k_t. We assume that each of the detected outliers has an equal recurrence probability of 1/80. Then superimpose the sampled outliers to the sample path of k_t.
   b. Obtain a projection of future death rates using the estimates of a_x and b_x obtained in Step 1, and the sample path perturbed with outliers.
   c. Obtain the cohort death rates for different year of births. For instance, for the cohort born in 1920, take $q_{70}, \ldots, q_{80}$ from the historical data, and take $q_{81}, \ldots, q_{99}$ from the projection.
   d. Extend the cohort death rates to age 100 and above by the relational model.
   e. Compute the distribution function specified by equation (6.1) using the extended cohort death rates.
4. Let $\Pr^{(i)}(\omega_N < x)$ be the estimate of the distribution function of ω_N in the ith sample path. We take the average of $\Pr^{(i)}(\omega_N < x)$, $i = 1, \ldots, 1000$, as the ultimate estimate of $\Pr(\omega_N < x)$.

Table 7 summarizes the predicted distributions of the highest attained age for cohorts born in year 1891, 1901, 1911, 1921, and 1931. Figure 3 provides a graphical representation of the probability mass functions. Note that these cohorts attained age $x_0 = 70$ in year 1961, 1971, 1981, 1991, and 2001, respectively. For readers' information, results derived from the original Lee-Carter model are also presented in tandem. The results in the first column may be compared with the actual situation: As of May 2006, the oldest validated living person in Canada is Julie Winnefred Bertrand, female, born in year 1891, 114 years old in 2006 (see Gerontology Research

Figure 3
Predicted Probability Mass Function of the Highest Attained Age for Selected Cohorts of Female Canadians



Group 2005). The current age of Julie is fairly close to the mode of the predicted distribution and is well accommodated by the 99th percentile. Across the row of the table, we can observe that the improvement of mortality has a positive effect on the highest attained age. Over the period of the analysis, the average increase of the mode of the highest attained age is approximately 0.25 year and 0.75 year in a 10-year period for male and female, respectively.

The predicted trends of highest attained age, however, are not identical to the results in other similar studies. For instance, based on the limiting distribution of the peaks over threshold, Han (2005) predicted that the average increase of the highest attained age in every 10-year period would be 0.8 year for male and 1.0 year for female. The differences lie on the assumptions used in the derivations. Han relied solely on the asymptotic results in the extreme value theory, and hence their results are valid only if the underlying survival distribution belongs to the peaks over threshold domain of attraction. Similarly, as we relied on the relational model in extending the cohort death rates, the validity of our results are heavily dependent on the accuracy of the relational model. Unfortunately these conditions can never be justified.

Finally we would like to sound a few cautionary notes in the interpretation of the results. First, the 99th percentile of the highest attained age is completely different from that of the lifetime: The former is the extreme of the *maximum* life span, whereas the latter is merely the extreme of the life span. Although there may be some extent of positive correlation, these two values are significantly different in magnitude and are by no means comparable with each other. Second, the probabilistic distribution of the highest attained age is conditional on the mortality experience of the specific cohort from which the distribution is derived, and thus it is not applicable to any other cohort whose mortality experience is different. Third, the variance of the distribution takes no account of the parameter uncertainty in the Lee-Carter forecast of future death rates, and hence the actual tail of the distribution may be even thicker.

7. CONCLUDING REMARKS

In this paper we performed a systematic outlier analysis of the mortality index built in the Lee-Carter model. Through outlier detection we found that mortality levels in the United States and Canada are vulnerable to events like pandemics and wars. By incorporating the effect of outliers, we created an outlier-adjusted Lee-Carter model. Given the better fit and the enhancement in forecast efficiency, the outlier-adjusted version seems to be an attractive alternative to the original model in predicting the long-term trend of mortality rates.

The appeal of the Lee-Carter model is the long-term linearity of the its time-series component, the mortality index k_t. The linearity holds not only in the mortality indices of Canada and the United States, but also in those of many other developed countries (see, e.g., Li and Chan 2005; Tuljapurkar, Li, and Boe 2000). While deviation from linearity is possible, the statistical evidence should not be overruled by mere speculations or unproven theories. The persistent linearity in k_t suggests that the ARIMA model is a suitable candidate for the modeling of k_t. However, in future mortality analyses, researchers may make use of a more general class of nonlinear time-series model for a more rigorous examination of k_t (see Tong 1983).

The availability of high-quality data is a critical factor in the analysis of the highest attained age. Because of the absence of actual death counts beyond age 85, we are not able to produce reliable prediction of the highest attained age for the United States population. A possible solution to this problem is using the mortality date from the Social Security Administration (OASDI). It is also warranted to replicate the analysis using the mortality data from the Canadian Pension Plan/Quebec Pension Plan, and from the International Database of Longevity when the number of validated supercentenarians is sufficiently large.

Finally, as pointed out by Smith (1999) and Ledford and Robinson (1999), the prediction of the highest attained age might be improved by considering some modern statistical methods based on exceedances over a high threshold (see, e.g., Davison and Smith 1990; Pickands 1975). It would be interesting to integrate the outlier-adjusted Lee-Carter model and the results in the theory of exceedances over threshold in further research on the probabilistic distribution of the highest attained age.

8. ACKNOWLEDGMENTS

The authors would like to thank the editor and an anonymous referee for their invaluable comments on an earlier version of the article. They would also like to acknowledge stimulating discussions with the discussant (Prof. Robert Brown) and the participants in ‘‘Living to 100 and Beyond: Survival at Higher Ages Symposium, 2005,’’ organized by the Society of Actuaries.

The authors are grateful to the U.S. Census Bureau, the National Center for Health Statistics, the Human Mortality Database, and the Gerontology Research Group for providing the data used in this study. This work was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Competitive Earmarked Research Grant Project No. CUHK7413/05H).

APPENDIX A

Figure A.1

Mean Forecast and 95% Pointwise Confidence Interval for Mortality Index, k_t, Canada

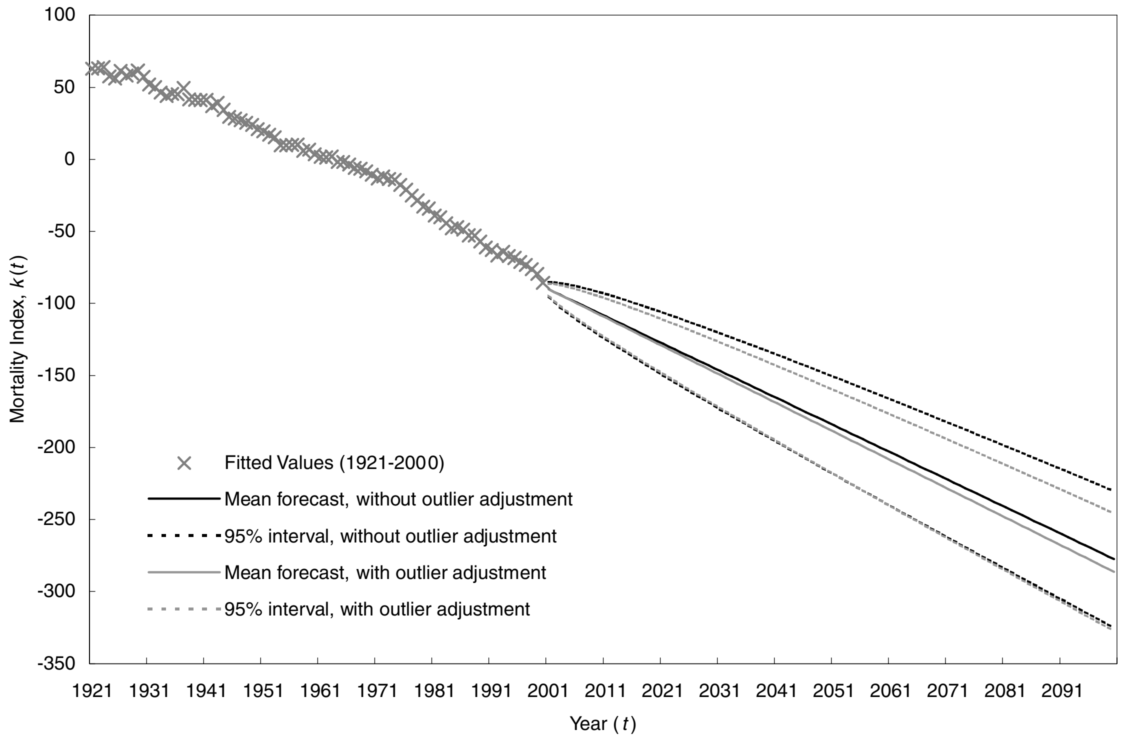


Figure A.2

Mean Forecast and 95% Pointwise Confidence Interval for Mortality Index, k_t, United States

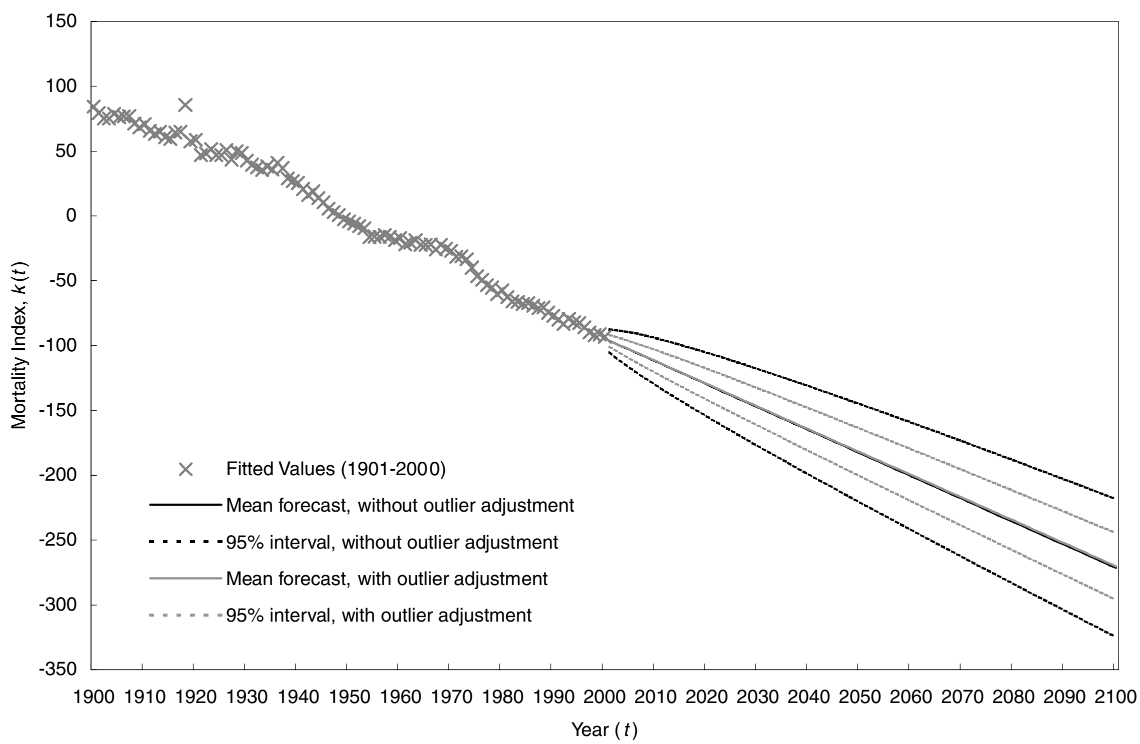


Figure A.3

Mean Forecast and 95% Pointwise Confidence Interval for Complete Life Expectancy at Birth, Canada

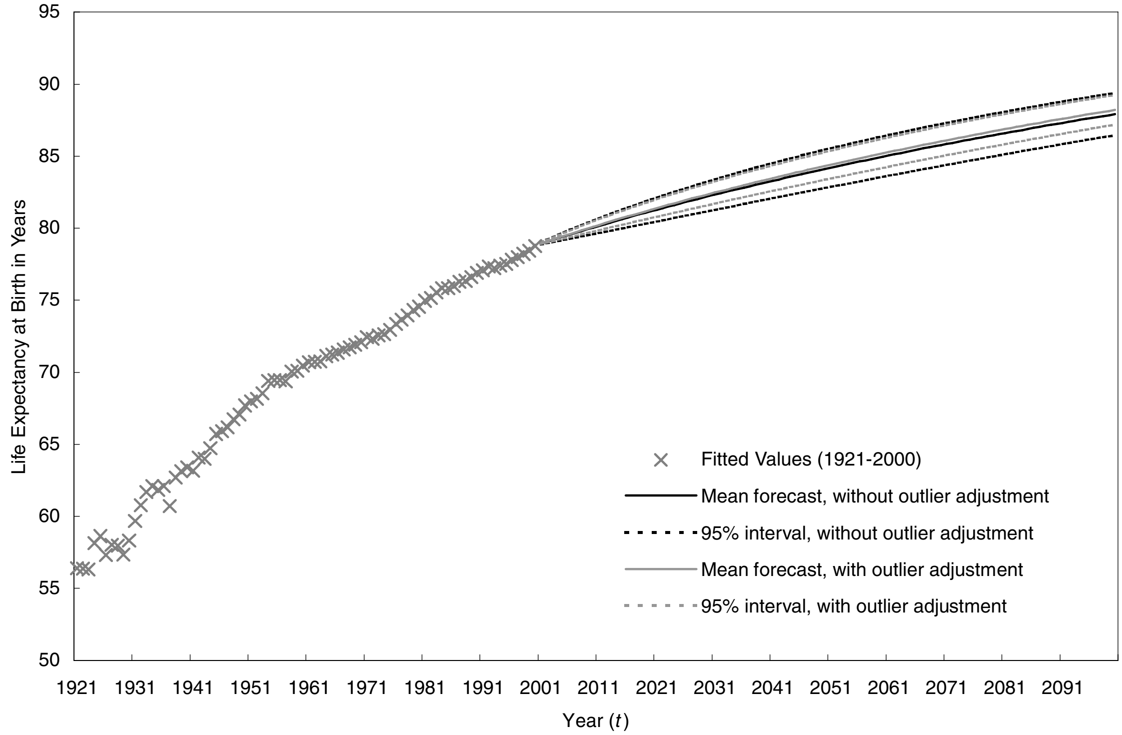


Figure A.4

Mean Forecast and 95% Pointwise Confidence Interval for Complete Life Expectancy at Birth, United States

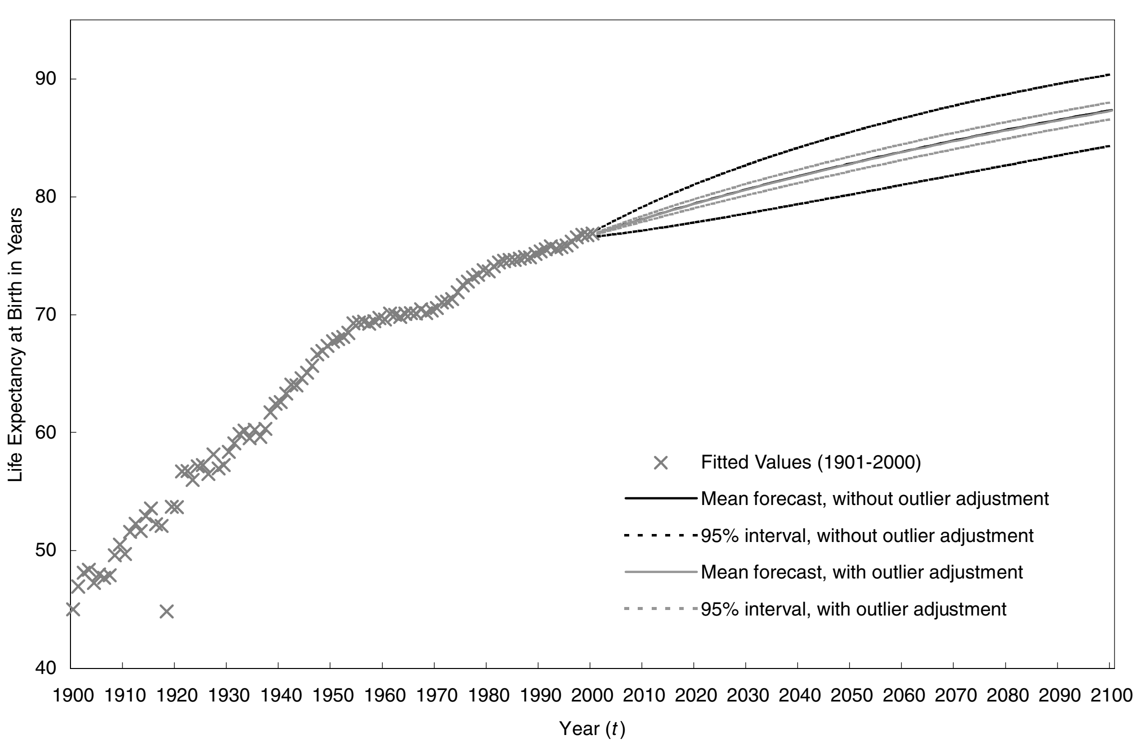


Figure A.5
Mean Forecast and 95% Pointwise Confidence Interval for Actuarial Present Value of Life Annuity at Age 65, *i* = 2.5%, Canada

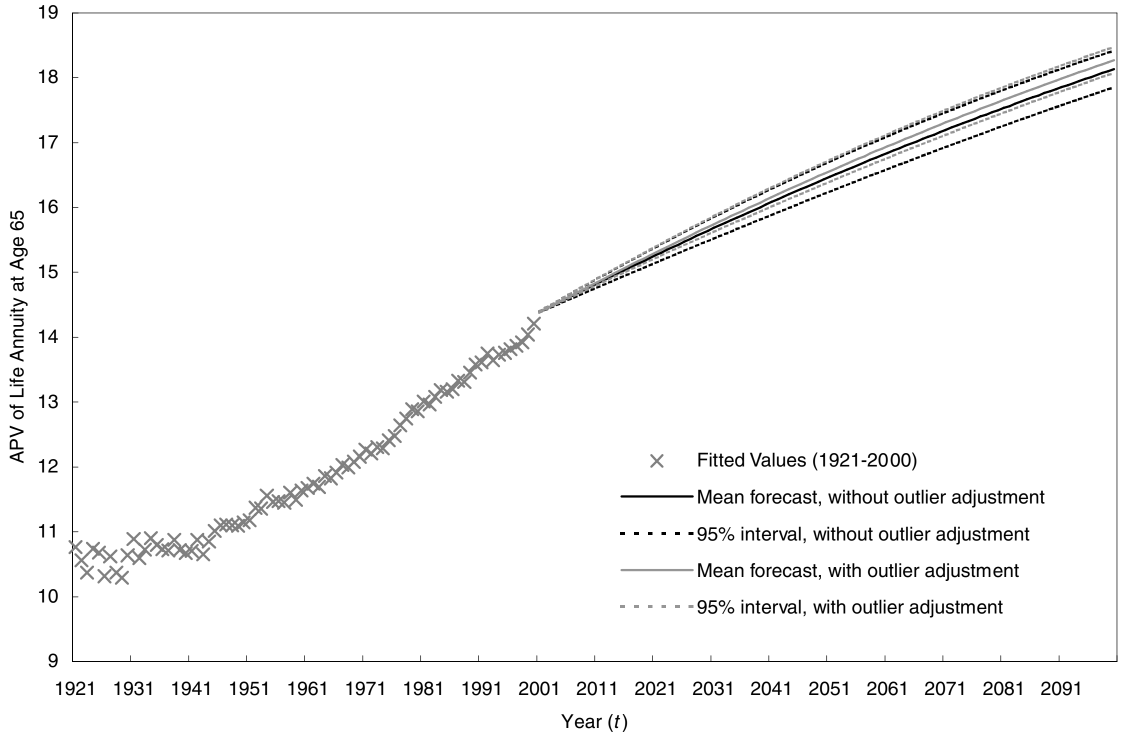
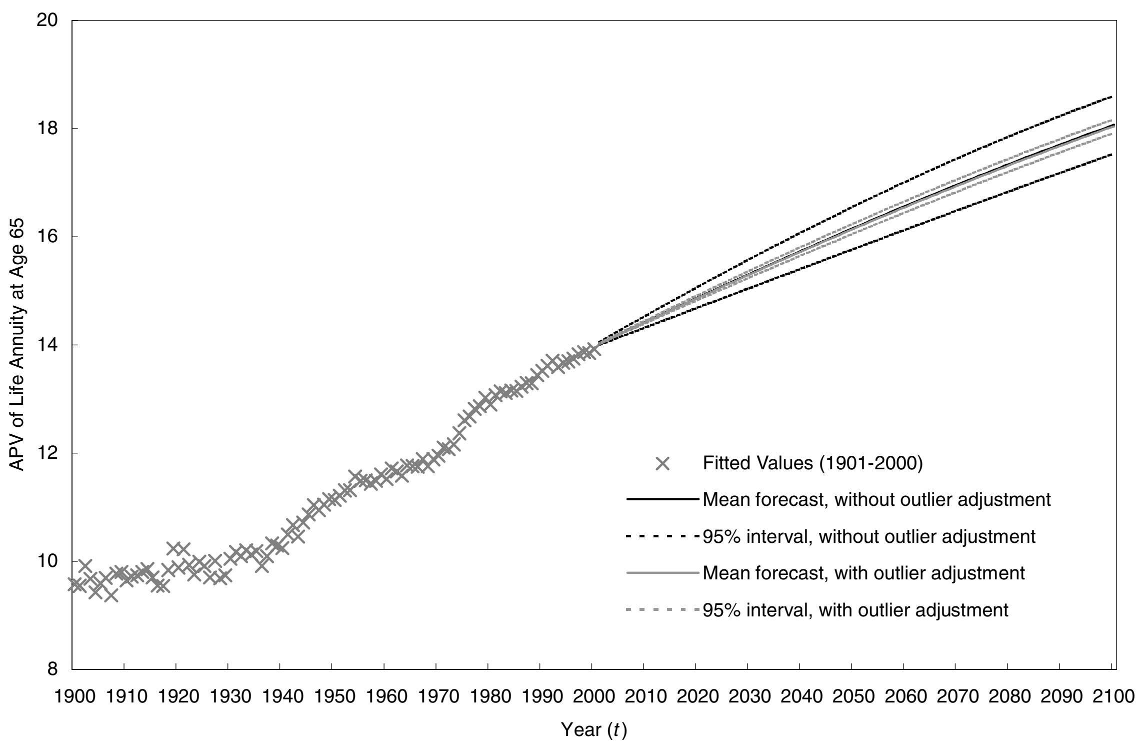


Figure A.6
Mean Forecast and 95% Pointwise Confidence Interval for Actuarial Present Value of Life Annuity at Age 65, *i* = 2.5%, United States



APPENDIX B

DERIVATION OF THE INTERVAL FORECAST OF THE ACTUARIAL PRESENT VALUE OF A LIFE ANNUITY UNDER THE LEE-CARTER MODEL

The s-period ahead forecast of the logarithm of each age-specific death rate is specified by

$$\widehat{\ln}(m_{x,t+s}) = \hat{a}_x + \hat{k}_{t+s}\,\hat{b}_x, \qquad \text{(B.1)}$$

where the carat indicates an estimate (for a_x and b_x) or a forecast (for k_t), and t denotes the base period. Assuming the model specification is correct, the true value of $\ln(m_{x,t+s})$ is given by

$$\ln(m_{x,t+s}) = (\hat{a}_x + \alpha_x) + (\hat{k}_{t+s} + u_{t+s})(\hat{b}_x + \beta_x) + \varepsilon_{x,t+s}, \qquad \text{(B.2)}$$

where α_x and β_x are the errors in estimating a_x and b_x, u_{t+s} is the error in the s-period ahead forecast of k_t, and ε_{t+s} is the error term that reflects all remaining age-specific influences not captured by the model. On the whole, the forecast error for the logarithm of the age-specific death rate can be expressed as

$$E_{x,t+s} = \ln(m_{x,t+s}) - \widehat{\ln}(m_{x,t+s}) = \alpha_x + \varepsilon_{x,t+s} + (\hat{b}_x + \beta_x)u_{t+s} + \beta_x\hat{k}_{t+s}. \qquad \text{(B.3)}$$

For a given x, the error terms are *assumed* to be independent of one another. In addition, following the suggestion by Lee and Carter (1992), $E_{x,t+s}$ can be decomposed into two parts, namely, $\varphi_{x,t+s} = u_{t+s}\,\hat{b}_x$ and $\xi_{x,t+s} = \alpha_x + \varepsilon_{x,t+s} + \beta_x\,\hat{k}_{x,t+s} + \beta_x u_{t+s}$, such that $\varphi_{x,t+s}$ is perfectly correlated across x, and $\xi_{x,t+s}$ is *assumed* independent across x. Both $\varphi_{x,t+s}$ and $\xi_{x,t+s}$ are *assumed* to have mean 0. On the ground of $E_{x,t+s}$, the forecast error for the APV of a life annuity can be readily derived using the following three arguments. Actuarial notations here follow Bowers et al. (1997).

Claim 1

The forecast error in $m_{x,t+s}$ is approximately $E_{x,t+s}\,\hat{m}_{x,t+s}$.

PROOF OF CLAIM 1

$$\ln(m_{x,t+s}) = \widehat{\ln}(m_{x,t+s}) + E_{x,t+s}$$

$$\Rightarrow m_{x,t+s} = \hat{m}_{x,t+s}\exp(E_{x,t+s}) \approx \hat{m}_{x,t+s}(1 + E_{x,t+s})$$

$$\Rightarrow m_{x,t+s} - \hat{m}_{x,t+s} \approx E_{x,t+s}\,\hat{m}_{x,t+s}.$$

Claim 2

On a small change of m_x, say, Δm_x, q_x changes by $p_x\Delta m_x$.

PROOF OF CLAIM 2

Assuming uniform distribution of death within each year of age, $m_x = q_x/(1 - 1/2q_x)$. This implies that $\partial m_x/\partial q_x = (1 - (q_x/2))^2$.

Then, using a first-order approximation,

$$\Delta q_x = \left(1 - \frac{q_x}{2}\right)^2 \Delta m_x \approx (1 - q_x)\Delta m_x \quad \text{(for small values of } \Delta m_x\text{)} = p_x\Delta m_x.$$

Claim 3

On a small increase in m_{x+y}, say, Δm_{x+y}, $\xi y \in \{0, 1, 2, \ldots\}$, a_x decreases by

$$\Delta m_{x+y}\ {}_{y+1|}\ddot{a}_x.$$

PROOF OF CLAIM 3

$$\ddot{a}_x = \ddot{a}_{x:\overline{x+y}|} + {}_{y+1}p_x v^{y+1}\,\ddot{a}_{x+y+1} = \ddot{a}_{x:\overline{x+y}|} + {}_{y}p_x(1 - q_{x+y})v^{y+1}\,\ddot{a}_{x+y+1}.$$

Suppose that m_{x+y} is increased by Δm_{x+y}; then by Claim 1, q_{x+y} will increase by $p_{x+y}\Delta m_{x+y}$. Let $\Delta\ddot{a}_x$ be the change of $\ddot{a}_x$ when m_{x+y} is increased by Δm_{x+y}; then

$$\ddot{a}_x + \Delta\ddot{a}_x = \ddot{a}_{x:\overline{x+y}|} + {}_{y}p_x(1 - q_{x+y} - p_{x+y}\Delta m_{x+y})v^{y+1}\,\ddot{a}_{x+y+1}$$
$$= \ddot{a}_x - \Delta m_{x+y}\ {}_{y+1}p_x v^{y+1}\,\ddot{a}_{x+y+1}$$
$$= \ddot{a}_x - \Delta m_{x+y}\ {}_{y+1|}\ddot{a}_x,$$
$$\therefore \Delta\ddot{a}_x = -\Delta m_{x+y}\ {}_{y+1|}\ddot{a}_x.$$

By Claim 3, the forecast error of $\ddot{a}_x$ for time $t + s$ can be written as $\sum_{y=0}^{\omega-x-1} - E_{x+y,t+s}m_{x+y,t+s}\ {}_{y+1|}\ddot{a}_y$. Denoting **E** as expectation and using equation (B.3), the approximate error variance can be written as

$$\mathrm{Var}(\hat{\ddot{a}}_x) = \mathbf{E}\left[\sum_{y=0}^{\omega-x-1} \hat{m}_{x+y,t+s}\,u_{t+s}\,\hat{b}_{x+y}\ {}_{y+1|}\ddot{a}_x\right]^2 + \mathbf{E}\left[\sum_{y=0}^{\omega-x-1} (\hat{m}_{x+y,t+s}\ {}_{y+1|}\ddot{a}_x)^2\right.$$

$$\times\ (\alpha_{x+y} + \varepsilon_{x+y,t+s} + \beta_{x+y}\hat{k}_{t+s} + \beta_{x+y}u_{t+s})^2\Bigg]. \tag{B.4}$$

Substituting variances, we get

$$\mathrm{Var}(\hat{\ddot{a}}_x) = \sigma^2_{u,t+s}\left[\sum_{y=0}^{\omega-x-1}\hat{m}_{x+y,t+s}\ \hat{b}_{x+y\ \ y+1|}\ddot{a}_x\right]^2 + \sum_{y=0}^{\omega-x-1}(\hat{m}_{x+y,t+s\ \ y+1|}\ddot{a}_x)^2 \times (\sigma^2_{a,x} + \sigma^2_{\varepsilon,x,t+s} + \sigma^2_{\beta,x}\hat{k}^2_{t+s} + \sigma^2_{\beta,x}\sigma^2_{u,t+s}). \tag{B.5}$$

Finally, the approximate 95% pointwise confidence interval is given by

$$\hat{\ddot{a}}_x \pm 2\sqrt{\mathrm{Var}(\hat{\ddot{a}}_x)}. \tag{B.6}$$

For technical details on the computations of $\sigma^2_{u,t+s}$, $\sigma^2_{a,x}$, $\sigma^2_{\beta,x}$, $\sigma^2_{\varepsilon,x,t+s}$, see Lee and Carter (1992, p. 669).

REFERENCES

ABRAHAM, BOVAS, AND ALICE CHUANG. 1989. Outlier Detection and Time Series Modelling. *Technometrics* 31: 241–48.

AKAIKE, HIROTUGU. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control* AC-19: 716–23.

AMERICAN BATTLE MONUMENTS COMMISSION. 2006. The American Battle Monuments Commission Database. www.abmc.gov/search/koreanwar.php. Accessed May 27, 2005.

BALKE, NATHAN S., AND THOMAS N. FOMBY. 1994. Large Shocks, Small Shocks, and Economic Fluctuations: Outliers in Macroeconomic Time-Series. *Journal of Applied Econometrics* 9: 181–200.

BOOTH, HEATHER, AND LEONIE TICKLE. 2003. The Future Aged: New Projections of Australia's Elderly Population. Department of Actuarial Studies Research Paper Series. Macquarie University, Sydney, Australia.

BOURBEAU, ROBERT, AND BERTRAND DESJARDINS. 2002. Dealing with Problems in Data Quality for the Measurement of Mortality at Advanced Ages in Canada. *North American Actuarial Journal* 6(3): 1–13.

BOWERS, NEWTON L., JR., HANS U. GERBER, JAMES C. HICKMAN, DONALD A. JONES, AND CECIL J. NESBITT. 1997. *Actuarial Mathematics*. 2nd ed. Schaumburg, IL: Society of Actuaries.

BOX, GEORGE, AND GWILYM M. JENKINS. 1976. *Time Series Analysis Forecasting and Control*. 2nd ed. San Francisco: Holden-Day.

BROSE, GERALD J. 1997. *Numerical Methods with MATLAB: A Resource for Scientists and Engineers*. Boston: PWS Publishing.

BROUHNS, NATACHA, MICHEL DENUIT, AND JEROEN K. VERMUNT. 2002. A Poisson Log-Bilinear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics* 31: 373–93.

BUETTNER, THOMAS. 2002. Approaches and Experiences in Projecting Mortality Patterns for the Oldest-Old. *North American Actuarial Journal* 6(3): 14–29.

CARTER, LAWRENCE R., AND RONALD D. LEE. 1992. Modelling and Forecasting U.S. Sex Differentials in Mortality. *International Journal of Forecasting* 8: 393–411.

CHAN, WAI-SUM. 1995. Outliers and Financial Time Series Modelling: A Cautionary Note. *Mathematics and Computers in Simulation* 39: 425–30.

———. 1998. Outlier Analysis of Annual Retain Price Inflation: A Cross-Country Study. *Journal of Actuarial Practice* 6: 149–72.

———. 2002. Stochastic Investment Modelling: A Multiple Time-Series Approach. *British Actuarial Journal* 8: 545–91.

CHAN, WAI-SUM, AND SHAUN WANG. 1998. The Wilkie Model for Retail Price Inflation Revisited. *British Actuarial Journal* 4: 637–52.

CHANG, IH, GEORGE C. TIAO, AND CHUNG CHEN. 1988. Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics* 30: 190–204.

CHATFIELD, CHRIS. 1993. Calculating Interval Forecasts. *Journal of Business and Economic Statistics* 11: 121–35.

CHEN, CHUNG, AND LON-MU LIU. 1993. Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association* 88: 284–97.

CHEN, CHUNG, AND GEORGE C. TIAO. 1990. Random Level Shift Time Series Models, ARIMA Approximation, and Level Shift Detection. *Journal of Business and Economic Statistics* 8: 170–86.

CHIA, NGEE C., AND ALBERT K. C. TSUI. 2003. Life Annuities and Compulsory Savings and Income Adequacy of the Elderly in Singapore. *Journal of Pension Economics and Finance* 2: 41–65.

COALE, ANSLEY, AND G.UANG GUO. 1989. Revised Regional Model Life Tables at Very Low Mortality. *Population Index* 55: 613–43.

CROSBY, ALFRED W. 2003. *America's Forgotten Pandemic: The Influenza of 1918*. 2nd ed. Cambridge: Cambridge University Press.

DAVISON, ANTHONY C., AND RICHARD L. SMITH. 1990. Models for Exceedances over High Thresholds (with discussion). *Journal of the Royal Statistical Society Series B* 52: 393–442.

FISHMAN, MORRIS. 2002. Searching for Omega. *The Actuary* 36(5): 2.

FOX, ANTHONY J. 1972. Outliers in Time Series. *Journal of the Royal Statistical Society Series B* 34: 350–63.

GERONTOLOGY RESEARCH GROUP. 2005. Validated Living Supercentenarians, as of May 6, 2005. www.grg.org. Accessed May 6, 2005.

HAN, ZHONGXIAN. 2005. Living to 100 and Beyond: An Extreme Value Study. *Living to 100 and Beyond Symposium Monograph*. www.soa.org/ccm/content/research-publications/library-publications/monographs/life-monographs/living-to-100-and-beyond-monograph/.

HILLMER, STEVEN. 1984. Monitoring and Adjusting Forecasts in the Presence of Additive Outliers. *Journal of Forecasting* 3: 205–15.

HIMES, CHRISTINE L., SAMUEL H. PRESTON, AND GRETCHEN A. CONDRAN. 1994. A Relational Model of Mortality at Older Ages in Low Mortality Countries. *Population Studies* 48: 269–91.

HOLMES, EDWARD C. 2004. Enhanced: 1918 and All That. *Science* 303: 1787–88.

HUMAN MORTALITY DATABASE. 2005. University of California, Berkeley (U.S.), and Max Planck Institute for Demographic Research (Germany). www.mortality.org or www.humanmortality.de. Accessed March 15, 2005.

IMSL. 1997. *IMSL Math/Library: Fortran Subroutines for Mathematical Applications. Volume 1.* Houston: Vision Numerics.

JOHANSEN, ROBERT J. 2002. Advancing the Horizon for Centenarian Survival Studies. *North American Actuarial Journal* 6(3): Guest Editorial.

LEDFORD, ANTHONY W., AND MICHAEL E. ROBINSON. 1999. Discussion of A. R. Thatcher, The Long-Term Pattern of Adult Mortality and the Highest Attained Age. *Journal of the Royal Statistical Society Series A* 162: 39.

LEDOLTER, JOHANNES. 1989. The Effect of Additive Outliers on the Forecasts from ARIMA Models. *International Journal of Forecasting* 5: 231–40.

LEE, RONALD D., AND LAWRENCE R. CARTER. 1992. Modelling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* 87: 659–71.

LI, SIU-HANG., AND WAI-SUM CHAN. 2005. Outlier Analysis and Mortality Forecasting: The United Kingdom and Scandinavian Countries. *Scandinavian Actuarial Journal*, 3: 187–211.

LIU, LON-MU, AND GREGORY B. HUDAK. 1994. *Forecasting and Time Series Analysis Using the SCA Statistical System*. Chicago: Scientific Computing Associates.

LJUNG, GRETA M. 1993. On Outlier Detection in Time Series. *Journal of the Royal Statistical Society Series B* 55: 559–67.

LJUNG, GRETA M., AND GEORGE E. P. BOX. 1978. On a Measure of Lack of Fit in Time Series Models. *Biometrika* 65: 287–303.

MUIRHEAD, COLIN R. 1986. Distinguishing Outlier Types in Time Series. *Journal of the Royal Statistical Society Series B* 48: 39–47.

NATIONAL CENTER FOR HEALTH STATISTICS. 2004a. HIST290 Death Rates for Selected Causes by 10-Year Age Groups, Race, and Sex: Death Registration States, 1900–32, and United States, 1933–98. www.cdc.gov/nchs/datawh/statab/unpubd/mortabs/hist290.htm. Accessed January 25, 2004.

———. 2004b. GMWK23R Death Rates by 10-Year Age Groups: United States and Each State, 1999–2001. www.cdc.gov/nchs/datawh/statab/unpubd/mortabs.htm. Accessed January 25, 2004.

ORD, KEITH, AND SAM LOWE. 1996. Automatic Forecasting. *American Statistician* 50: 88–94.

PANJER, HARRY H. 2003. SARS and Actuarial Modeling. *The Actuary* 37(6): 27.

PAYNE, R. W., P. E. LANE, P. G. N. DIGBY, S. A. HARDING, P. K. LEECH, G. W. MORGAN, A. D. TODD, R. THOMPSON, G. TUNNICLIFFE WILSON, S. J. WELHAM, AND R. P. WHITE. 1993. *Genstat 5 Release 3 Reference Manual*. Oxford: Clarendon Press.

PICKANDS, JAMES. 1975. Statistical Inference Using Extreme Order Statistics. *Annals of Statistics* 3: 119–31.

POLLARD, JOHN, H. 1988. On the Derivation of a Full Life Table from Mortality Data Recorded in Five-Year Age Groups. *Mathematical Population Studies* 2: 1–44.

RENSHAW, ARTHUR E., AND STEVEN HABERMAN. 2003. On the Forecasting of Mortality Reduction Factors. *Insurance: Mathematics and Economics* 32: 379–401.

ROBERTSON, KATHRYN A., DEBBIE J. DUPUIS, AND BRUCE L. JONES. 2005. Living to 100 and Beyond: An Extreme Value Study. *Living to 100 and Beyond Symposium Monograph.* www.soa.org/ccm/content/research-publications/library-publications/monographs/life-monographs/living-to-100-and-beyond-monograph/.

ROBINE, JEAN-MARIE, AND JAMES W. VAUPEL. 2002. Emergence of Supercentenarians in Low-Mortality Countries. *North American Actuarial Journal* 6(3): 54–63.

ROWLEY, MARK C., AND PETER C. BISHOP. 2001. Impact of Terrorism on Future Mortality Assumptions. *Record of the Society of Actuaries* 28.

SARONY, NEVILLE, WAI-SUM CHAN, AND FELIX WAI-HON CHAN. 2003. *Personal Injury Tables Hong Kong: Tables for the Calculation of Damages*. Hong Kong: Sweet and Maxwell Asia.

SAS INSTITUTE. 2004. What's New in SAS/ETS 9? http://support.sas.com/software/91x/etsugwhatsnew900.htm. Accessed August 4, 2004.

SMITH, RICHARD L. 1999. Discussion of A. R. Thatcher, The Long-Term Pattern of Adult Mortality and the Highest Attained Age. *Journal of the Royal Statistical Society Series A* 162: 37.

THATCHER, ROGER A. 1999. The Long-Term Pattern of Adult Mortality and the Highest Attained Age. (with discussion). *Journal of the Royal Statistical Society Series A* 162: 5–43.

THOMAS, GUY, GENEVIEVE PLU, AND JEAN-CHRISTOPHE THALABARD. 1992. Identification of Pulses in Hormone Time Series Using Outlier Detection Methods. *Statistics in Medicine* 11: 2133–45.

TONG, HOWELL. 1983. *Threshold Models in Non-Linear Times Series Analysis*. Berlin: Springer.

TSAY, RUEY. S. 1986. Time Series Model Specification in the Presence of Outliers. *Journal of American Statistical Association* 81: 132–41.

———. 1988. Outliers, Level Shifts and Variance Changes in Time Series. *Journal of Forecasting* 7: 1–20.

TUCKER, SPENCER., ED. 2005. *The Encyclopedia of World War I: A Political, Social, and Military History*. Santa Barbara, CA: ABC-CLIO.

TULJAPURKAR, SHRIPAD, NAN LI, AND CARL BOE. 2000. A Universal Pattern of Mortality Decline in the G7 Countries. *Nature* 405: 789–92.

UNITED NATIONS. 1997. Projecting Old-Age Mortality and Its Consequences. Report on the working group, December 3–5, 1996. New York: United Nations.

U.S. CENSUS BUREAU. 2004. Monthly Population Estimates, 1990–2000. www.census.gov/popest/archives/1990s/nat_detail.html. Accessed January 25, 2004.

VAAGE, KJELL. 2000. Detection of Outliers and Level Shifts in Time Series: An Evaluation of Two Alternative Procedures. *Journal of Forecasting* 19: 23–37.

WEI, WILLIAM W. S. 1990. *Time Series Analysis: Univariate and Multivariate Methods*. Redwood City, CA: Addison-Wesley.

WILMOTH, JOHN R. 1993. Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality Change. Technical report. Department of Demography, University of California, Berkeley.

WOLAK, DANIEL, DARRELL KNAPP, CHARLES H. MEINTEL, AND KEVIN J. TRAPP. 2003. Pricing for Terrorism Risk in Group Life, Group Health and Group Disability Products. *Record of the Society of Actuaries* 29; 1–30.

Discussions on this paper can be submitted until July 1, 2007. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.