© *The Journal of Risk and Insurance*, 2009, Vol. 76, No. 3, 727-751
DOI: 10.1111/j.1539-6975.2009.01313.x

Modeling Mortality With Jumps: Applications to Mortality Securitization

Hua Chen
Samuel H. Cox

Abstract

In this article, we incorporate a jump process into the original Lee–Carter model, and use it to forecast mortality rates and analyze mortality securitization. We explore alternative models with transitory versus permanent jump effects and find that modeling mortality via transitory jump effects may be more appropriate in mortality securitization. We use the Swiss Re mortality bond in 2003 as an example to show how to apply our model together with the distortion measure approach to value mortality-linked securities. Pricing the Swiss Re mortality bond is challenging because the mortality index is correlated across countries and over time. Cox, Lin, and Wang (2006) employ the normalized multivariate exponential tilting to take into account correlations across countries, but the problem of correlation over time remains unsolved. We show in this article how to account for the correlations of the mortality index over time by simulating the mortality index and changing the measure on paths.

Introduction

Mortality risk management is fundamental to life insurance and pension industries. Mortality models make it available to quantify the risks and provide the basis of pricing and reserving. Traditionally, reinsurance, and more recently, securitization, renders a mean of transferring or hedging mortality risks. Naturally, mortality models are fundamental to these transactions.

Hua Chen is an Assistant Professor in the Department of Risk, Insurance, and Healthcare Management at Temple University, A625 Alter Hall, 1801 Liacouras Walk, Philadelphia, PA 19122. Samuel H. Cox, FSA, CFA, is Dr. L. A. H. Warren Chair Professor in the Warren Centre for Actuarial Studies and Research at University of Manitoba, 181 Freedman Crescent, Winnipeg, MB R3T 5V4, Canada. Dr. Chen can be contacted via e-mail: hchen@temple.edu. We presented a former version of this article at the 42nd Actuarial Research Conference in Pittsburgh, the SCOR—*Journal of Risk and Insurance* Conference on New Forms of Risk Sharing and Risk Engineering in Paris. We appreciate the helpful comments and suggestions from participants at these meetings. We benefited a lot from the discussions with Shaun Wang, Eric Ulm, Daniel Bauer, and Richard MacMinn. We are especially grateful to anonymous referees for their very helpful comments that have much improved this article. We also thank the editors.

Mortality securitizations differ from reinsurance transactions in several ways. Perhaps the most important is that investors in a securitization typically do not have the mortality expertise that a reinsurer has. Also, in order to avoid moral hazard problems, the basis of a securitization may be a public index rather than the actual lives insured by the ceding party. Therefore, it is important that a mortality model clearly conveys the nature of risk transfer to investors, reflects the characteristics of available data, and provides for scenario analysis.

A wide variety of stochastic models have been proposed for modeling the dynamics of mortality over time. Cairns, Blake, and Dowd (2006a) provide a detailed overview and categorization. Most of the literature in this field is in the framework of short-rate models, among which continuous time models focus on the spot force of mortality and discrete time models concentrate on the spot mortality rates. Continuous time models (e.g., Milevsky and Promislow, 2001; Dahl, 2004; Biffis, 2005; Dahl and Møller, 2006; Miltersen and Persson, 2005; Schrager, 2006) help us understand the evolution of mortality rates over time, but they are relatively intractable. We prefer discrete time models because mortality rates are measured at most once a year and discrete time models are relatively easy to implement in practice.

The Lee–Carter model is among the earliest discrete time models. Lee and Carter (1992) model the central mortality rates to be log-linearly correlated with a time-dependent mortality factor and adjust for age-specific effects using two sets of age-dependent coefficients. In this way, the model captures both the mortality trend overall and the age-specific changes on different age groups. Thus, it describes the development of the mortality curve over time quite well. The age adjustment is necessary because mortality improvement varies across age groups. Moreover, the adverse mortality shocks, such as the 1918 influenza pandemic, attack different groups with varied intensities. We will discuss these two points in detail in the data section. The Lee–Carter approach has been extended by Brouhns, Denuit, and Vermunt (2002), Renshaw and Harberman (2003), Denuit, Devolder, and Goderniaux (2007), and further revisited by Li and Chan (2007). Cairns, Blake, and Dowd (2006b) propose a two-factor model that is similar in the spirit of that of the Lee–Carter approach. The first factor indicates the mortality trend over time and equally affects mortality at all ages, whereas the second factor's effect on mortality is proportional to age. Based on their model setup, the mortality curve in a given year increases in ages, which does not reflect the fact that mortality rates of infants and children are much higher than those of middle ages. In addition, their model does not include mortality jumps.

The purpose of mortality securitization is to hedge extremely adverse mortality risks. Thus, it is appropriate to take into account adverse mortality jumps in mortality securitization modeling (Cox, Lin, and Wang, 2006). Nevertheless, most of the papers on this topic have ignored mortality jumps (e.g., Renshaw, Haberman, and Hatzoupoulos, 1996; Sithole, Haberman, and Verrall, 2000; Milevsky and Promislow, 2001; Olivieri and Pitacco, 2002; Dahl, 2004; Denuit, Devolder, and Goderniaux, 2007). Even though some of them recognize that short-term catastrophe shocks may cause adverse mortality jumps, they do not model the jumps explicitly. For example, Lee and Carter (1992) treat the 1918 influenza pandemic as a highly unusual event and employ an intervention model to remove its influence. Li and Chan (2007) also regard pandemic events as nonrepetitive exogenous intervention and implement outlier detection and adjustment to unveil the "true" model underlying the outlier-free mortality series.

To our knowledge, there are only a few papers considering mortality jumps in mortality securitization modeling. Biffis (2005) uses affine jump diffusions to address the risk analysis and market valuation of life insurance contracts in the continuous time framework. Luciano and Vigna (2005) find that a jump process is better than a diffusive component alone in stochastic mortality modeling. Cox, Lin, and Wang (2006) come to the same conclusion. They, however, model age-adjusted death rates instead of age-specific death rates. Thus, their model fails to represent age-specific changes of mortality rates. In addition, they model the jump process in a way such that it has permanent effects on mortality rates, although many adverse mortality jumps are caused by short-term catastrophic events and only have transitory effects.

In this article, we propose to incorporate a jump process into the Lee–Carter model, restricting mortality jumps to have one-period effects. We fit the model to the U.S. age-specific mortality rates and forecast the development of the mortality curve. We show that the model with jumps outperforms that without jumps, and the model with transitory jump effects is more appropriate in mortality securitization modeling. We then discuss the outlier-adjusted Lee–Carter model (Li and Chan, 2007) to further explore the source of mortality jumps. We find that the so-called "outliers" play an important role in mortality securitization modeling, and we should not delete these outliers from the time-series mortality data in order to establish a proper model for pricing mortality-linked securities. We use the Swiss Re mortality bond issued in 2003 as an example to estimate the implied market prices of risk and show that the model with transitory jump effects leads to better pricing.

In an incomplete insurance market, there are mainly two approaches for security valuation. One is to adapt the arbitrage-free pricing framework to the securitization of mortality risk. Cairns, Blake, and Dowd (2006a, b) provide a detailed discussion on this issue and give as an example the pricing of the European Investment Bank (EIB) longevity bond. The second method is to use a distortion operator to create an equivalent risk-adjusted distribution and obtain the fair value of the security under this risk-adjusted measure. Examples of this approach, based on the Wang transform (Wang, 2000, 2002), include Lin and Cox (2005, 2008), Dowd et al. (2006), and Denuit, Devolder, and Goderniaux (2007). The Swiss Re mortality bond issued in 2003 covers mortality risks across five countries and over three consecutive years. Previous research (Cox, Lin, and Wang, 2006) uses the normalized multivariate exponential tilting, which is a generalization of the Wang transform, to take into account correlations across countries. This research, however, modifies the contract terms by linking the principal repayment with the maximum of the mortality index in 3 years and ignores correlations over time. In this article, we use the Wang transform and make the first attempt to account for correlations of the mortality index over time. The basic idea is to forecast the mortality index on paths and change the measure on each path to get the risk-adjusted mortality index. We employ the Wang transform not only because it is easy to implement but also because the Wang transform can be easily extended in a multiple-period and multivariate-risk context.[1]

The body of this article is organized as follows. The second section describes data and historical facts for further motivation of the problem. The third section presents

[1]See Wang (2002) for extending the Wang transform to a multiple-period setting, and Wang (2007) for extending the Wang transform to a multivariate-risk setting.

TABLE 1
Mortality Improvement by Different Age Groups, From 1900 to 2003

Age Groups	1900	2003	Ratio	Age Groups	1900	2003	Ratio
All	2,518.0	832.7	0.331	35–44	1,023.1	201.6	0.197
< 1	16,244.8	700.0	0.043	45–54	1,495.4	433.2	0.290
1–4	1,983.8	31.5	0.016	55–64	2,723.6	940.9	0.345
5–14	385.9	17.0	0.044	65–74	5,636.1	2,255.0	0.400
15–24	585.5	81.5	0.139	75–84	12,330.0	5,463.1	0.443
25–34	819.8	103.6	0.126	≥ 85	26,088.2	14,593.3	0.559

Note: The "All" row is the age-adjusted death rate by all causes per 100,000, from NCHS reports HIST293 and GMWK293R. The other rows are the age-specific death rates per 100,000, from NCHS reports HIST290 and GMWK290R. The mortality improvement ratio is calculated by the authors.

the mortality models with permanent versus transitory jump effects and discusses their applications in different scenarios. Numerical examples of pricing mortality-linked securities are demonstrated in the fourth section. Concluding remarks and discussions are provided in the last section.

DATA DESCRIPTIONS AND HISTORICAL FACTS: FURTHER MOTIVATION

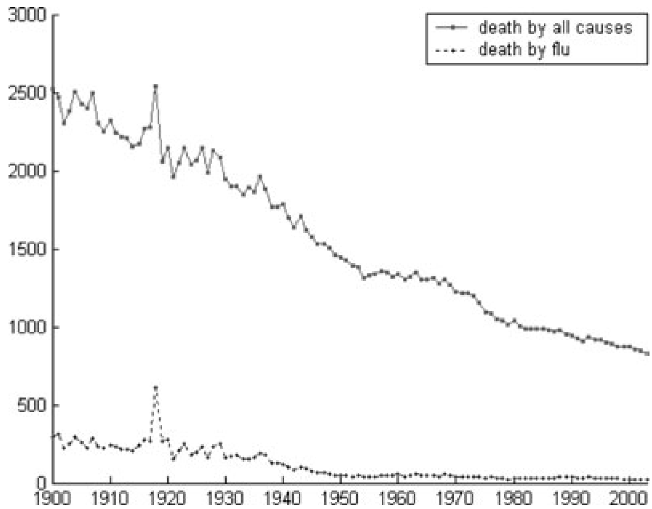
Our mortality data are from the National Center for Health Statistics (NCHS). The NCHS reports the age-adjusted death rate and age-specific death rate per 100,000 population (2000 standard) for selected causes of death from 1900 to 2003.[2] Age-adjusted death rates are used to compare relative mortality risks across groups and over time; they are indices rather than direct measures. The age-specific death rates are tabulated for age 0, age group 1–4, then 10-year groups such as 5–14, 15–24, up to 75–84, and age group 85 and over. Selected causes include heart disease, cancer, stroke, influenza, and pneumonia.

Table 1 provides evidence of mortality improvement. Overall, the age-adjusted death rate by all causes decreased to 832.7 per 100,000 in 2003, which is 33.1 percent of the 1900s level (2,518 per 100,000). However, the improving mortality has variant effects across age groups. The mortality rates for age group 1–4 fell to 1.6 percent of its initial value, but that for age group 85 and over only dropped to 55.9 percent of its initial value. These proportions differ by a factor of 35 at the extremes. A proper mortality model should capture this age-specific effect of mortality improvement on all ages.

The trend of mortality improvement is further demonstrated in Figure 1. Furthermore, Figure 1 compares the dynamics of age-adjusted death rates by all causes to that by influenza and pneumonia from 1900 to 2003. Although the death rates caused by influenza and pneumonia become small after 1950 (less than 0.00005), which makes the comparison difficult to visualize, we can observe the similar fluctuation patterns of death rates by all causes and by flu in the first half of the past century. The two curves even jump at the same time, which is remarkably evidenced in the year 1918.

[2]Source: http://www.cdc.gov/nchs/datawh/statab/unpubd/mortabs.htm.

FIGURE 1
U.S. Age-Adjusted Death Rates (per 100,000) by All Causes and by Influenza and Pneumonia, From 1900 to 2003



Deaths caused by flu account for 9.4 percent of all deaths before 1950 on average, 3.4 percent from 1950 to 2003, and 6.3 percent for the whole period examined. In 1918, this proportion reaches its historic peak at 24.1 percent. The similar patterns of these two curves and the high portions of deaths caused by flu suggest that we should not exclude flu events when modeling the mortality.

We also calculate the correlation coefficient between death rates by all causes except flu and death rates by flu across different age groups for each year from 1900 to 2003, which is shown in Figure 2. The correlation is above 0.95 for most of the time with only a few exceptions, and it averages to 0.9753. Interestingly, the correlation falls to the lowest value (0.8339) in 1918, which indicates the 1918 Spanish flu has different effects on the death rates of different age groups. The age-specific effect of this flu attack is revealed in detail in Table 2. The 1918 influenza pandemic raised the mortality rate by 30 percent overall. It affected the age groups 15–24 and 25–34 the most, whereas for individuals aged 55 and over, the death rates decreased a little bit. A proper mortality model should reflect the age-specific effect of short-term catastrophic shocks on mortality.

MORTALITY MODELING: THE GENERALIZED LEE–CARTER MODEL

The Lee–Carter Model Overview

Ever since Lee and Carter presented their original work in 1992, the Lee–Carter model has been widely used in mortality trend fitting and projection. The Census Bureau population forecast has used it as a benchmark for the long-run forecast of U.S. life expectancy. The two most recent Social Security Technical Advisory Panels have suggested the Trustees adopt this method or other methods consistent with it (Lee and Miller, 2001).

FIGURE 2
Correlation Coefficients Between Age-Specific Death Rates by All Causes Except Flu and Those by Flu, From 1900 to 2003

TABLE 2
Mortality Deterioration for Each Age Group, From 1917 to 1918

Age Groups	1917	1918	Ratio	Age Groups	1917	1918	Ratio
All	1,397.1	1,810	1.296	35–44	900.8	1,339.3	1.487
< 1	10,457.2	11,167.2	1.068	45–54	1,385.6	1,524.1	1.1
1–4	1,066	1,573.5	1.476	55–64	2,678.6	2,648.1	0.989
5–14	256	412.8	1.613	65–74	5,728.4	5,505	0.961
15–24	468.9	1,070.6	2.283	75–84	12,386.2	11,295.7	0.912
25–34	649.1	1,643.5	2.532	≥85	24,593.6	22,213.5	0.903

Let $m_{x,t}$ denote the central death rate for age x at time t. The model decomposes this time series of age-specific death rates into two sets of age-specific constants a_x and b_x, and a time-varying factor k_t, (Lee and Carter referred to k_t as mortality index. In order to distinguish k_t and the mortality index defined in the securitization contract, we refer to k_t as mortality factor thereafter.) Mathematically, the Lee–Carter model can be represented as follows:

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

where a_x represents the age pattern of death rates, b_x represents age-specific reactions to the time-varying factor, and $e_{x,t}$ is the error term that captures the age-specific effects not reflected in the model.

The Lee–Carter model cannot be fitted by the ordinary least square approach because all variables on the right-hand side of the model are unobservable. Moreover, this model is obviously overparameterized. To obtain a unique solution, a normalization conditions is imposed such that the b_x terms sum to unity and the k_t terms sum to zero, that is,[3]

$$\sum_x b_x = 1 \quad \text{and} \quad \sum_t k_t = 0. \tag{2}$$

Then a_x becomes the average value of ln $(m_{x,t})$ over time, that is,

$$a_x = \frac{1}{T}\sum_{t=1}^{T} \ln(m_{x,t}), \tag{3}$$

where T is the length of the time series of mortality data.

Lee and Carter suggest a two-stage procedure to solve this problem. In the first stage, the singular value decomposition (SVD) method is applied to the matrix of $\ln(m_{x,t}) - a_x$ to obtain estimates of b_x and k_t. In the second stage, the k_t factors are reestimated by iteration, given the values of a_x and b_x obtained in the first step, such that the implied number of deaths equals to the actual number of deaths.

$$D_t = \sum_x (Pop_{x,t} \exp(a_x + b_x k_t)), \tag{4}$$

where D_t is the actual total number of deaths at time t, and $Pop_{x,t}$ is the population in age group x at time t.

Based on the U.S. mortality data for different age groups from 1900 to 2003, we implement this two-stage procedure, report the fitted values of a_x, b_x for 11 age groups in Table 3, and plot the final estimates of the mortality factor k_t in Figure 3. Generally, the mortality rates of young age groups respond more rapidly when the mortality factor k_t changes. The mortality factor k_t is decreasing over time, which shows the trend of mortality improvement. The big jump around 1918 indicates the severe influenza pandemic in that year.

Model k_t With Jumps: Permanent Versus Transitory Effects?

To make forecasts of the future distribution of k_t, we need to choose a suitable model to fit the k_t series. Cox, Lin, and Wang (2006) combine a geometric Brownian motion and a compound Poisson process to model mortality rates with permanent jump effects. We present their model in a discrete-time setting and then propose our model with transitory jump effects.

Denote N_t, the number of jumps occurring in year t. For the sake of simplicity, we suppose there is at most one jump event in each year, with the probability of jump

[3]The normalization conditions can be chosen differently, but the choice of conditions will not affect the estimation of parameters, a_x, b_x, and k_t, up to a linear transformation (see Lee and Carter, 1992). It is necessary to point out that these constraints are used only once for the identification purpose in this step, and there are no such constraints when we choose a proper model to fit and forecast the k_t series.

TABLE 3
Fitted Values of a_x and b_x From the Lee–Carter Model

Age Group	a_x	b_x
Under 1	−3.3935	0.14501
1–4	−6.2072	0.19673
5–14	−7.1833	0.14942
15–24	−6.2877	0.10037
25–34	−5.9837	0.10531
35–44	−5.4745	0.08580
45–54	−4.7734	0.06044
55–64	−4.0071	0.04602
65–74	−3.2289	0.04193
75–84	−2.4146	0.04035
85 over	−1.6084	0.02862

FIGURE 3
Dynamics of the Mortality Factor k_t From the Lee–Carter Model, From 1990 to 2003

p.[4] That is,

$$N = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } 1 - p \end{cases}. \tag{5}$$

[4]More generally, we can use a Poisson random variable to model the jump frequency N. However, it does not add new insights as to the comparison between permanent jump effects and transitory jump effects, and only complicate the mathematics.

We assume the jump severity variable Y, is identically independently distributed normal variables with mean m and standard deviation s, and Y is independent of the jump frequency variable N.

In the spirit of Cox, Lin, and Wang (2006), a model with permanent jump effects describing the evolution of the mortality factor, k_t, can be written as

$$k_{t+1} = \begin{cases} k_t + \mu - pm + \sigma Z_{t+1}, & \text{if } N_{t+1} = 0 \\ k_t + \mu - pm + \sigma Z_{t+1} + Y_{t+1}, & \text{if } N_{t+1} = 1 \end{cases}, \tag{6}$$

where μ and σ are constants, and Z_t is a standard normal random variable that is independent of Y and N.

Equation (6) shows clearly that if a jump event occurs in year $t+1$, the magnitude of the jump, Y_{t+1}, is included in the mortality factor k_{t+1}, and this jump effect persists forever. However, most of the adverse mortality jumps, like the 1918 Spanish flu and the 2004 earthquake and tsunami, are caused by short-term catastrophic events and have merely transitory effects on mortality rates. We can observe mortality rates pop up when the catastrophic events happen, and fall back to the normal level when the events end. Therefore, we believe that the model with permanent jump effects is inappropriate for adverse mortality risk modeling. We propose our model with transitory jump effects in the following.

Let $\tilde{k}_t$ represent the mortality factor when there is no jump event. We model $\tilde{k}_t$ as a random walk with drift

$$\tilde{k}_{t+1} = \tilde{k}_t + \mu + \sigma Z_{t+1}, \tag{7}$$

If a jump event occurs in year $t+1$, that is, $N_{t+1} = 1$, The jump Y_{t+1} makes the actual mortality factor k_{t+1} change from $\tilde{k}_{t+1}$ to $\tilde{k}_{t+1} + Y_{t+1}$, that is,

$$k_{t+1} = \tilde{k}_{t+1} + Y_{t+1}. \tag{8}$$

If there is no jump in year $t+1$, that is, $N_{t+1} = 0$, we know that

$$k_{t+1} = \tilde{k}_{t+1}. \tag{9}$$

Equations (8) and (9) can be combined and written in one equation

$$k_{t+1} = \tilde{k}_{t+1} + Y_{t+1} N_{t+1}. \tag{10}$$

Therefore, the dynamics of the mortality factor k_t can be completely expressed as

$$\begin{cases} \tilde{k}_{t+1} = \tilde{k}_t + \mu + \sigma Z_{t+1} \\ k_{t+1} = \tilde{k}_{t+1} + Y_{t+1} N_{t+1} \end{cases}. \tag{11}$$

Simple algebra obtains

$$k_{t+1} = k_t + \mu + \sigma Z_{t+1} + Y_{t+1} N_{t+1} - Y_t N_t. \tag{12}$$

TABLE 4
Parameter Estimates for Different Model Selections

Parameter	Estimate	Parameter	Estimate
	Model with jumps—transitory effect: $\ln(L) = -62.52$		
μ	-0.2173	σ	0.3733
m	0.8393	s	1.4316
p	0.0436	LRT statistics	63.49
	Model with jumps—permanent effect: $\ln(L) = -65.47$		
μ	-0.2172	σ	0.3872
m	-0.3062	s	2.3133
p	0.0396	LRT statistics	57.60
	Model without jumps: $\ln(L) = -94.27$		
μ	-0.2172	σ	0.6043

Note: The critical value for the chi-square distribution ($df = 3$, alpha $= 0.01$) is 11.34. Therefore, our likelihood ratio test rejects the model without jump at the 1% significance level.

Let $z_t = k_{t+1} - k_t$. If we have a time series of K observations of k_t, there will be $K - 1$ observations of z's values. z_t and z_{t+1} can be expressed as

$$z_t = \mu + \sigma Z_{t+1} + Y_{t+1}N_{t+1} - Y_t N_t, \tag{13}$$

$$z_{t+1} = \mu + \sigma Z_{t+2} + Y_{t+2}N_{t+2} - Y_{t+1}N_{t+1}. \tag{14}$$

If $N_{t+1} = 0$, then z_t is independent on z_{t+1}. If $N_{t+1} = 1$, then z_t is correlated with z_{t+1} because of the Y_{t+1} part. It is noteworthy that we should use the conditional maximum likelihood estimation (CMLE) to calibrate the parameters because the values of z-series are not independent.[5]

The upper panel of Table 4 reports the parameter estimates for the model with transitory jump effects. The expected rate of change of the mortality factor μ, which is equal to -0.2173, implies that the mortality factor k_t decreases by 0.2173 per year on average. The negative sign of μ is consistent with the fact that the U.S. population mortality improves over time. The probability that there is a jump in a given year is equal to 0.0436.

We also estimate the parameters for the model with permanent jump effects for the purpose of comparison,[6] the results of which are shown in the middle panel of Table 4. The variables μ and σ are roughly the same as before, whereas there are significant

[5]Lin and Cox (2008) combine a geometric Brownian motion with a Markov chain to capture the transitory effect of mortality jumps. However, they argue that because the probability of having a catastrophic mortality event is very low, the correlations of data are negligible. Therefore, they use independence as an approximation and do not take into account the correlations of the data. If we do not consider the correlations of the data, the parameter estimates for the transitory jump model are $\mu = -0.2172$, $\sigma = 0.4018$, $m = -3.2391$, $s = 0$, $p = 0.0098$.

[6]See Appendix A for the derivation of the log-likelihood function of the model with permanent jump effects.

differences in the mean and variance of the jump severity distribution between two models. The frequency of jumps decreases from 0.0436 to 0.0396.

The estimation results for the model without jumps are in the lower panel of Table 4. The variable μ is unchanged, while σ increases substantially to 0.6043. This is because the model without jumps incorporates the variations caused by the jump process into the volatility term. We report the values of the log-likelihood functions for different models, and perform the likelihood ratio test. The test result rejects the model without jumps at the 1 percent significance level.

Evidence From the Outlier-Adjusted Lee–Carter Model: Do Outliers Matter?

We have already shown that the model with jumps fits the mortality factor k_t better than the model without jumps. The next question is where the mortality jumps come from. Before answering this question, let us briefly review the work by Li and Chan (2005, 2007). They argue that mortality series are often contaminated with discrepant observations, which may result from recording or typographical errors, or from non-repetitive exogenous interventions, such as pandemics or hostilities. In order to reveal the "true" mortality trend, they perform a systematic time-series outlier analysis for the mortality data in United States and Canada, and fit the adjusted outlier-free mortality series to the Lee–Carter model. For the U.S. data from 1900 to 2000, they find seven outliers, which occurred in years 1916, 1918, 1921, 1928, 1936, 1954, and 1975, respectively. Most of these outliers result from influenza epidemics according to their explanations, except for the data in 1954 and 1975.

Do the mortality jumps in our model mainly come from the flu events? Do these outliers really stand outside the mortality trend? We delete the outliers found by Li and Chan from the original mortality data, estimate the mortality factor k_t again (Figure 4), and compare the results of the model with transitory jump effects and that without jumps (Table 5). We find that after eliminating the outliers the mortality factor k_t declines more smoothly and do not show significant jumps in the evolution process. In addition, when we fit the mortality factor using the model with transitory jump effects, the probability of a jump in a given year p becomes zero, which actually makes the model with jumps equivalent to the model without jumps. We therefore infer that the mortality jumps arise from these so-call "outliers," which are mostly caused by flu epidemics.

Is it appropriate to exclude the outliers when we model the death rates for the purpose of pricing mortality-linked securities? As Chan (2002) elucidates, "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 important (such as pricing catastrophe risks ...), then a model which is sympathetic to outliers in the data ought to be used."[7] We have observed that flu plays an important role in all the causes of deaths in the second section and detected that mortality jumps mostly result from influenza epidemics. We believe that outliers should not be neglected and that adverse mortality jumps should be explicitly modeled in mortality securitization because the rationale behind selling or buying mortality securities is to hedge mortality risks. Besides, historic data

[7]See Chan (2002, pp. 559-560).

FIGURE 4
Dynamics of the Mortality Factor k_t From the Outlier-Adjusted Lee–Carter Model

Note: Mortality data in years 1916, 1918, 1921, 1928, 1936, 1954, and 1975 are deleted according to the outlier analysis by Li and Chan (2007).

TABLE 5
Parameter Estimates Using the Outlier-Free Mortality Data From 1900 to 2003

Parameter	Estimate	Parameter	Estimate
Model with jumps—transitory effect: $\ln(L) = -48.38$			
μ	−0.2317	σ	0.4005
m	−0.3774	s	0.9316
p	0		
Model without jumps: $\ln(L) = -48.38$			
μ	−0.2317	σ	0.4005
Likelihood ratio test (LRT) statistic = 0			

Note: Based on the likelihood ratio test, we cannot reject the model without jumps at the 1% significance level.

show that influenza pandemics happen with frightening regularity, occurring every 30 to 50 years (Knapp, 2006). We should not view the pandemic as a one-time event that will never happen again.

Discussions on Model Application to Securitizations

In general, there are two types of mortality risks we need to consider. The first is the adverse mortality risk we discussed in this article, which is usually caused by catastrophic events and result in much higher mortality rates than would normally be experienced. The 1918 Spanish flu killed up to 50 million people worldwide and 500,000 in the United States (Rasmussen, 2005). The earthquake and tsunami in 2004 resulted in 300,000 dead and missing across southern Asia and eastern Africa (Cox, Lin, and Wang, 2006). The Swiss Re mortality bond issued by the Swiss Reinsurance company in 2003 is an example of a securitization designated as a hedge for life insurers. It expanded Swiss Re's capacity to pay catastrophe mortality losses.

The second is longevity risk, which is referred to as the uncertainty in the future improvement in mortality rates. If the realized mortality rate is much lower than the assumed mortality rates in the premium pricing and reserve calculating, life annuity providers will incur large losses. The EIB longevity bond offered in November 2004 was designed as a hedge for pension plans and other annuity providers. Although it failed to generate sufficient demand to be launched, it did attract public attention and provides an instructive case study.

Modeling adverse mortality jumps seems more important for securities linked to short-term catastrophic shocks. Ignoring these jumps may lead to underestimating the probability of having a catastrophic event and overestimating the variation of the mortality factor (see Table 4), which may bring large errors in pricing mortality-linked securities and calculating the risk premium. We have made clear arguments that the model with transitory jump effects is more appropriate than that with permanent jump effects in mortality securitization. We, therefore, take the Swiss Re mortality bond issued in 2003 as an example to show how to apply the model with transitory jump effects to calculate the implied market price of risk.

It is noteworthy that by no means is the model with permanent jump effects not applicable elsewhere. Actually, the model with permanent jump effects should work properly in the context of longevity securitization because the main concern of investors in the longevity market is the unexpected mortality improvement in the duration of longevity bonds. For example, heart disease is one of the leading causes of death in the United States. It accounted for 28.5 percent of the total death in 2002 according to the report of National Center of Health Statistics (Vol. 53, No. 5, October 2004). If there were an unexpected breakthrough of medicine protecting people from heart attack, the mortality rates would decrease substantially and the effects of mortality improvement would last for a long period, even forever.

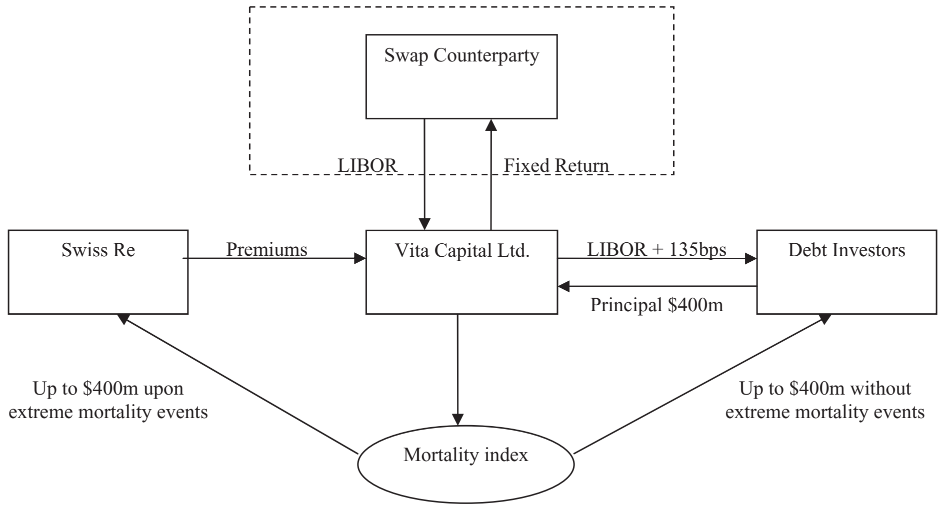
MORTALITY-LINKED SECURITIES PRICING

The Swiss Re Mortality Bond

The Swiss Re transaction is diagramed in Figure 5. The Swiss Reinsurance company is the largest reinsurance company in the world. In order to reduce its exposure to catastrophic mortality risks, it issued its first pure mortality bond transaction through a special purpose vehicle (SPV) called Vita Capital. The issue size was $400 million. It is a 3-year deal: the bond was issued in December 2003 and matured on January 1, 2007. Coupons are paid quarterly at a rate of 3-month U.S. dollar LIBOR plus a spread of 135 basis points. Vita Capital executed a swap transaction to swap Swiss Re's fixed premium payment for LIBOR. The principal repayment is at risk and depends on a mortality index weighted over five countries.[8] If the mortality index does not exceed 1.3 times the 2002 base level during any of the 3 years, the principal is fully

[8] The five countries are United States, United Kingdom, France, Italy, and Switzerland. The weights assigned to each country are: United States 70 percent, United Kingdom 15 percent, France 7.5 percent, Italy 5 percent, and Switzerland 2.5 percent.

FIGURE 5
The Contract Design of the Swiss Re Mortality Bond Issued in 2003



Source: Revised from Blake, Cairns, and Dowd (2006).

repayable. Otherwise, the investors will receive a reduced principal repayment if the mortality index exceeds this threshold, and will get nothing back if the index is above 1.5 times the base level. As discussed by Cowley and Cummins (2005), there is always a trade-off between the moral hazard problem and the basic risk. Basing the payoff of mortality bonds on a population index has the advantage of reducing investors' concern about moral hazard problems. However, it also introduces the basis risk, in that the insurer's mortality experience could deteriorate significantly than the index. For this reason, mortality bonds are more likely to appeal to large, diversified insurers or reinsurers.

Let q_t represent the weighted average mortality index at year t, and q_0 the 2002 base level of the mortality index. The principal repayment of this bond as a fraction of the issue size shows as follows:

$$\text{Principal Repayment } = \max\left(1 - \sum_{t=t_0}^{T} \text{loss}_t, 0\right), \tag{15}$$

where $t_0 = 2004$, $T = 2006$, and the loss ratio in year t, loss $_t$, is defined as a call option spread with the upper strike $U = 1.5q_0$ and the lower strike $M = 1.3q_0$

$$\text{loss}_t = \frac{\text{Max}(q_t - M{,}0) - \text{Max}(q_t - U{,}0)}{U - M}. \tag{16}$$

There are two difficulties we need to deal with in order to value the Swiss Re mortality bond properly. First, the mortality index defined in the contract is a weighted average across five countries. Mortality rates in these five countries may be correlated with each other because, for example, pandemics may attack the population of these countries at the same time. Cox, Lin, and Wang (2006) adopt the normalized multivariate exponential tilting to take into account the correlation across countries. Second, the principal repayment of the Swiss Re mortality bond, \$400 million $\times$ max $(1 - \sum_{t=2004}^{2006} \text{loss}_t, 0)$, is based on the loss ratios (and hence, the mortality index) in three consecutive years. The mortality index evolves over time and depends on the last period of experience (see Equation (11)). The correlation of mortality index over time makes the pricing problem more difficult. In order to simplify the problem, Cox, Lin, and Wang modify the contract term such that

$$\text{Principal Repayment} = 1 - \frac{\text{Max}(q - M, 0) - \text{Max}(q - U, 0)}{U - M}, \tag{17}$$

where $q = \max(q_{2004}, q_{2005}, q_{2006})$.

As shown above, they take the maximum of the mortality index in these 3 years and link the principal repayment to this maximum value. In this way, they actually ignore the correlation over time and change the multiple-period problem into a single-period problem.

In this article, we adopt the Wang transform and take into account the correlations of the mortality index over time. For simplicity, we assume the mortality index only depends on the U.S. mortality data. However, our methodology can be easily extended to the case where the reference index is a weighted average across several countries by using the normalized multivariate exponential tilting.

Change Measures via the Wang Transform

As mentioned in the introduction section, the Wang transform has been widely used as a universal framework for pricing financial and insurance risks. For a given asset (or loss) variable X with a cumulative distribution function (CDF), $F(x)$, the Wang transform produces a risk-adjusted CDF, $F^*(x)$, such that

$$F^*(x) = \Phi[\Phi^{-1}(F(x)) - \lambda], \tag{18}$$

where Φ is the standard normal cumulative distribution and λ represents the market price of risk. After we obtain the risk-adjusted distribution $F^*(x)$, we can calculate the expectation of X under $F^*(x)$, which is denoted by $E^*[X]$. Further discounting this value back to time zero using the risk-free interest rate, we can get the fair value of the asset (or loss) X.

One important feature of the Wang transform is that it preserves the normal and lognormal distribution, which enables it to replicate the capital asset pricing model (CAPM) if the return of an underlying asset has a normal distribution and recover the Black–Scholes formula if the return of the underlying asset is lognormally distributed. Specifically, if X has a normal (μ, σ^2) distribution under the physical measure P, then after the Wang transform X is still normally distributed with $\mu^* = \mu + \lambda\sigma$ and $\sigma^* = \sigma$ under the risk-adjusted measure Q. If X has a lognormal (μ, σ^2) distribution under P, then X is still a lognormal variable with $\mu^* = \mu + \lambda\sigma$ and $\sigma^* = \sigma$ under Q.

Recall that in the model section, we work with the following dynamics under the physical probability measure P

$$\begin{cases} \tilde{k}_{t+1} = \tilde{k}_t + \mu + \sigma Z_{t+1} \\ k_{t+1} = \tilde{k}_{t+1} + Y_{t+1} N_{t+1} \end{cases}, \tag{19}$$

where $Z_t \sim N(0, 1)$, $Y_t \sim N(m, s^2)$ and $N_t = \{ \begin{smallmatrix} 1, & \text{with probability } p \\ 0, & \text{with probability } 1-p \end{smallmatrix}$ under P.

Under the assumption that the standard normal variable Z, the jump severity variable Y, and the jump frequency variable N are independent of each other, we apply the Wang transform to Z, Y, and N, respectively.[9,10] Let λ_1, λ_2, and λ_3 represent the market prices of risk associated with each of the variables Z, Y, and N. The dynamics of the mortality factor under the risk-adjusted measure Q becomes

$$\begin{cases} \tilde{k}^*_{t+1} = \tilde{k}^*_t + \mu + \sigma Z^*_{t+1} \\ k^*_{t+1} = \tilde{k}^*_{t+1} + Y^*_{t+1} N^*_{t+1} \end{cases}, \tag{20}$$

where $Z^*_t \sim N(\lambda_1, 1)$, $Y^*_t \sim N(m + \lambda_2 s, s^2)$, $N_t = \{ \begin{smallmatrix} 1, & \text{with probability } p^* \\ 0, & \text{with probability } 1-p^* \end{smallmatrix}$, and $p^* = 1 - \Phi [\Phi^{-1} (1 - p) - \lambda_3]$.[11]

Pricing Procedure and Results

We use the following procedure to calculate the market prices of risk of the Swiss Re mortality bond issued in 2003 based on the Wang transform.

- Because the mortality factor k_t is correlated over time, we cannot simulate k_t for each year independently. We need to simulate k_t (t = 2004, 2005, and 2006) on a path. We do simulations 10,000 times, using the model we have proposed and the parameter estimates shown in the upper panel of Table 4.

[9]We can use normalized multivariate exponential tilting when measuring and pricing multivariate risks. Because we assume that Z, Y, and N are independent of each other, applying the Wang transform to each variable individually is equivalent to applying the normalized multivariate exponential tilting. In addition, the independence among Z, Y, and N preserves after we apply the Wang transform. See Wang (2007) for details.

[10]The other reason we apply the Wang transform to each of the variables Z, Y, and N, instead of applying the Wang transform to the mortality index q is to take into account the correlations of mortality index over time. As we mentioned before, the principal repayment is based on the mortality index in three consecutive years, which are correlated with each other. Therefore, if we intend to apply the Wang transform to q_{2004}, q_{2005}, and q_{2006}, we need to figure out the correlation matrix of these three variables first, which is difficult to achieve right now. In addition, the estimated market prices of risk w.r.t. q_{2004}, q_{2005}, and q_{2006} are restricted to particular years, and cannot be used to price new bond issues.

[11]We denote the variables after the Wang transform with superscript *, which indicates that the distribution of the variable changes in the risk-adjusted measure. We DO NOT mean that we get a new variable.

- We use the Wang transform to change from the physical measure P to the risk-adjusted measure Q, and calculate the values of k_t^* on each path under Q using Equation (20), given initial values of the market prices of risk λ_1, λ_2, and λ_3.
- We calculate the mortality rates for different age groups by formula $m_{x,t}^* = \exp(a_x + k_t^* b_x)$ under Q. The year 2000 standard population and corresponding weights are used to compute the weighted average mortality index q_t^* under Q for each year.[12]
- We calculate the loss ratios, loss $_t^*$, under Q by Equation (16), and compute the risk-adjusted expected value of the principal repayment as of the period T.

$$E_T^*[\text{repayment}] = 400 \text{ million} \times \left(\frac{1}{10{,}000} \sum_{i=1}^{10{,}000} \left[\max \left(1 - \sum_t \text{loss}_{i,t}^*, 0 \right) \right] \right). \quad (21)$$

- We discount the coupon payments each period and the principal repayment back to the beginning of year 2004, using the risk-free interest rate.[13] Equating the discounted expected payoff to the issue size of the Swiss Re mortality bond $400 million, we can obtain the implied market prices of risk, λ_1, λ_2, and λ_3, via the numerical iteration such as the quasi-Newton method.

Note that we have one mortality bond price and need to estimate three market prices of risk. Therefore, we cannot solve λ_1, λ_2, and λ_3 simultaneously. As Cairns, Blake, and Dowd (2006b) demonstrate, we can estimate λ_1, λ_2, and λ_3 by fixing two of them and then solving for the third. We can also assume the market prices of risk are equal, that is, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, and solve for λ consequently.[14] Based on the par spread of the Swiss Re bond 1.35 percent, the implied market prices of risk for different models are shown in Table 6.[15]

Under the model with transitory jump effects, if we assume that the risk associated with the jump process is diversifiable, that is, $\lambda_2 = \lambda_3 = 0$, the market price of risk

[12]The year 2000 standard population and corresponding weights can be obtained from the technique notes of the NCHS report GMWK293R. The weight is 0.013818 for age under 1 year, 0.055317 for ages 1–4, 0.145565 for ages 5–14, 0.138646 for ages 15–24, 0.135573 for ages 25–34, 0.162613 for ages 35–44, 0.134834 for ages 45–54, 0.087247 for ages 55–64, 0.066037 for ages 65–74, 0.044842 for ages 75–84, and 0.015508 for age 85 and over.

[13]We use the U.S. Treasury yield rates on December 30, 2003 as the risk-free interest rates. We calculate the coupon payment by assuming it is paid annually, because we do not have quarterly data for the U.S. Treasury yield.

[14]The purpose of this article is to propose an appropriate mortality model and develop a valuation strategy to account for the pricing difficulties of the Swiss Re mortality bond. For simplicity, we only consider one transaction in 2003. Our model and pricing strategy can be easily extended to multiple-transaction situations. We can find the optimal λ_1, λ_2, and λ_3 by minimizing the target function $\sum [\hat{P}_i(\lambda_1,\lambda_2,\lambda_3) - P_i]^2$, where $\hat{P}_i(\lambda_1,\lambda_2,\lambda_3)$ is the modeled price of the ith issue depending on the parameters λ_1, λ_2, and λ_3, and P_i is the actual market quote for the ith transaction.

[15]We do not report the case where there is only systematic risk associated with the jump frequency. For the model with transitory jumps, our estimation results show that it is not sufficient to adjust the distribution of the jump frequency only.

TABLE 6
Implied Market Prices of Risk for Different Models

Model with transitory jumps	λ_1	5.1449	0	1.5000
	λ_2	0	3.4808	1.5000
	λ_3	0	0	1.5000
Model with permanent jumps	λ_1	4.6408	0	0.8072
	λ_2	0	2.0006	0.8072
	λ_3	0	0	0.8072
Model without jumps	λ_1	2.9921		

associated with Z is 5.1449. If there is no systematic risk with respect to Z and N, that is, $\lambda_1 = \lambda_3 = 0$, the market price of risk associated with Y is 3.4808. Of course, these can be regarded as the extreme cases. If we assume $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, we can solve for $\lambda = 1.5$.

We also notice that when switching to the model with permanent jump effects, the implied market prices of risk drop dramatically in each case. This can be explained by the difference in the intrinsic model setups and the difference in the volatility of the jump severity distribution. First, if we model mortality jumps with permanent effects, the jump magnitudes accumulate over time. The forecasted mortality rates are thereby more inclined to reach the predetermined threshold level, which indicates the risk on the principal repayment is higher. Second, from Table 4 the volatility of the jump severity is 2.3133 in the model with permanent jump effects, whereas it is 1.4316 in the model with transitory jump effects. The higher volatility of the jumps in the former model leads to higher risk. When the par spread of the Swiss Re is fixed at 135 basis points, the higher risk imposed on the principal repayment, the lower market price of risk we estimate. Therefore, modeling mortality with permanent jump effects underestimates the market prices of risk.

We come to the model without jumps at last. The estimated market price of risk associated with the random variation Z is 2.9921, which is much lower than that for the model with jumps, whether the jumps have permanent effects or transitory effects. The model without jumps overestimates the variation of the mortality factor whereas underestimating the probability of catastrophic events. We suspect that the effect of overestimating the variation dominates the effect of underestimating the catastrophic probability, which brings down the market price of risk associated with Z.

It is natural to compare our results with the estimations in other studies. For example, Cox, Lin, and Wang (2006) estimate the market price of risk for the Swiss Re deal is 0.83 when they assume λs are equal, whereas our estimation is 1.5 when λs are equal. Why do our estimates have higher values? One reason is that they apply the model with permanent jump effects, which underestimates the market price of risk. The other reason is that the estimates are based on different reference variables. Cox, Lin, and Wang (2006) base their estimation on the mortality rates q_t, whereas our estimation is based on the mortality factor k_t. Because the mortality factor k_t has a much larger scale than the true mortality rates (k_t changes from 10.6 in 1900 to -11.8

TABLE 7
Premium Spreads of a 4-Year Hypothetic Mortality Bond With Different Tranches

	Tranche A	Tranche B	Tranche C
Tranche size	\$100m	\$100m	\$100m
Upper strike U	1.5	1.45	1.40
Lower strike M	1.3	1.25	1.20
Premium spread (bps)	56	181	434

in 2003, and the age-adjusted mortality rate changes from 0.02518 in 1990 to 0.00833 in 2003), it is not surprising to see that the market prices of risk estimated in our article have higher values.

Applications in Premium Calculating

What investors of mortality-linked securities care about is how much premiums they can obtain from the transaction. Now we are ready to calculate the premium spreads of a hypothetic mortality bond. Suppose a 4-year mortality bond was issued in December 2004 (i.e., $t_0 = 2005$ and $T = 2008$). It is written on a mortality index q_t with the base level in year 2003 ($q_0 = q_{2003}$). The mortality index q_t only depends on the U.S. death rates and is a weighted average across different age groups

$$q_t = \sum_x \omega_x m_{x,t}, \tag{22}$$

where ω_x denotes the weight assigned to age group x based on the year 2000 standard population.

In order to attract different types of investors, the bonds are structured to several tranches with different lower strike M and upper strike U, as shown in Table 7. The principal repayment of each tranche at maturity follows Equations (15) and (16). We calculate the premium spreads of each tranche based on the mortality model with transitory jump effects and our pricing methodology, given the market prices of risk $\lambda_1 = \lambda_2 = \lambda_3 = 1.5$.

It is noteworthy to point out that the upper strike and lower strike are decreasing from Tranche A to Tranche C. Tranche C has the lowest upper strike and lower strike, which means investors in Tranche C are exposed to the highest risk of losing part or all of the principals. Naturally, Tranche C has the highest premium spread.

CONCLUSIONS AND DISCUSSION

In this article, we provide a deep discussion of mortality modeling and mortality-linked security pricing. We make the first attempt to incorporate a jump process into the Lee–Carter model and discuss in detail how to model the mortality factor k_t with transitory jump effects. We point out that transitory jump effects introduce data dependence and ignoring the dependence of data may cause inaccurate parameter estimation, which is the problem in previous studies (e.g., Lin and Cox, 2008). We compare the model with permanent jump effects to the model with transitory jump effects and discuss how to apply different models in different scenarios. Finally, we contribute to the existing literature by showing how to account for correlations of the mortality index over time when pricing the Swiss Re mortality bond issued in 2003. The basic idea is to simulate the paths of the mortality index and change measures on each path.

Our proposed model has the advantage of estimating the mortality trend and jump process simultaneously. However, a simple model is always preferred by investors and risk modeling firm, as it is easy to understand and implement. One line of future research may focus on the improvement of mortality modeling. A regime-switching model can be one option. In order to capture more extreme price movements in the stock returns and stochastic variability in the volatility parameter, Hardy (2001) proposes a regime-switching lognormal process. The basic idea is to assume that volatility takes one of K discrete values, switching between these values randomly. The further research can apply the regime switching model to mortality data so as to capture the extreme mortality risk.

Another approach we may consider is the extreme value theory (EVT). The reason we incorporate the jump part into our model is to capture the extreme mortality risk. However, the potential values of a risk have a probability distribution that we will never observe *exactly* although past losses may provide *partial* information about that distribution. Traditional parametric models are ill suited to extreme probabilities, because parametric modeling produces a good fit in the central regions at the expense of good fit in the tail (Diebold, Schuermann, and Stroughair, 1998). Fortunately, EVT provides us with guidance to achieve reliable estimates of extreme risks. The key idea in the EVT is to "estimate extreme probabilities by fitting a model . . . using only the extreme event data rather than all the data, thereby fitting the tail, and only the tail"(Sanders, 2005). However, this advantage can also be viewed as a disadvantage of the EVT approach: it focuses on the tail distribution, thus ignoring the central regions of the data. An ideal mortality model should have a good fit on the historical data and also extrapolate more extreme, out of sample data.

Another line of future research may focus on how to decide an "optimal" transform in an incomplete market. As suggested by Cox, Lin, and Wang (2006), although the change of measures is not unique in an incomplete market, we can try to apply the minimum martingale transform to find a strategy that minimizes the variance of the payoff risk. Second, we ignore the issue of parameter uncertainty in this article. We simulate the mortality index using estimates of the parameters, assuming these parameter estimates are true values without ambiguity. We may relax this assumption and consider the valuation and hedging of mortality securities under parameter uncertainty.

APPENDIX A

Log-Likelihood Function of the Model With Permanent Jump Effects

The model with permanent jump effects can be simplified to

$$k_{t+1} = k_t + \mu - pm + \sigma Z_{t+1} + Y_{t+1} N_{t+1}. \tag{A1}$$

Let

$$z_t = k_{t+1} - k_t = \mu - pm + \sigma Z_{t+1} + Y_{t+1} N_{t+1}. \tag{A2}$$

If $N_{t+1} = 0$, the variable $z_t \mid (N_{t+1} = 0)$ will be normally distributed with mean $M_n = \mu - pm$, and variance $S_n^2 = \sigma^2$.

If $N_{t+1} = 1$, the variable $z_t \mid (N_{t+1} = 1)$ will be normally distributed with mean $M_y = \mu + (1 - p)m$, and variance $S_y^2 = \sigma^2 + s^2$.

The density function of z_t, denoted by $f(z_t)$, can be written in terms of conditional probabilities

$$\begin{aligned} f(z_t) &= f(z_t \mid N_{t+1} = 0)\Pr(N_{t+1} = 0) + f(z_t \mid N_{t+1} = 1)\Pr(N_{t+1} = 1) \\ &= \frac{1}{\sqrt{2\pi}S_n}\exp\left(-\frac{(z_t - M_n)^2}{2S_n^2}\right) \times (1 - p) + \frac{1}{\sqrt{2\pi}S_y}\exp\left(-\frac{(z_t - M_y)^2}{2S_y^2}\right) \times p. \end{aligned} \tag{A3}$$

If we have a time series of K observations of k_t, there will be $K - 1$ observations of z's values. The log-likelihood function can be expressed as follows:

$$\begin{aligned} \sum_{i-1}^{K-1} \ln f(z_i) = \sum_{i=1}^{K-1} \ln\Bigg(&\frac{1}{\sqrt{2\pi}S_n}\exp\left(-\frac{(z_t - M_n)^2}{2S_n^2}\right)(1 - p) \\ &+ \frac{1}{\sqrt{2\pi}S_y}\exp\left(-\frac{(z_t - M_y)^2}{2S_y^2}\right)p. \end{aligned} \tag{A4}$$

APPENDIX B

Log-Likelihood Function of the Model With Transitory Jump Effects

Recall in the third section, second part, we defined

$$z_t = \mu + \sigma Z_{t+1} + Y_{t+1}N_{t+1} - Y_t N_t, \tag{B1}$$

$$z_{t+1} = \mu + \sigma Z_{t+2} + Y_{t+2}N_{t+2} - Y_{t+1}N_{t+1}. \tag{B2}$$

If $N_{t+1} = 0$, then z_t is independent on z_{t+1}. If $N_{t+1} = 1$, then z_t is correlated with z_{t+1} because of the Y_{t+1} part. Under the CMLE, the likelihood function can be obtained as follows:

$$\begin{aligned} f(z_1,z_2,\ldots z_{K-1}) &= f(z_{K-1} \mid z_1,z_2,\ldots z_{K-2})f(z_1,z_2,\ldots,z_{K-2}) \\ &= f(z_{K-1} \mid z_{K-2})f(z_{K-2} \mid z_1,z_2,\ldots z_{K-3})f(z_1,z_2,\ldots z_{K-3}) \\ &= f(z_{K-1} \mid z_{K-2})f(z_{K-2} \mid z_{K-3})\ldots f(z_2 \mid z_1)f(z_1). \end{aligned} \tag{B3}$$

Therefore, the log-likelihood function is

$$\begin{aligned} \ln f(z_1,z_2,\ldots z_{K-1}) = {} & \ln f(z_{K-1} \mid z_{K-2}) + \ln f(z_{K-2} \mid z_{K-3}) \\ & + \cdots + \ln f(z_2 \mid z_1) + \ln f(z_1). \end{aligned} \tag{B4}$$

Next, we will derive $f(z_{t+1} \mid z_t.)$ and $f(z_1)$, respectively.

If $N_{t+1} = 0$, then $\quad z_{t+1} = \mu + \sigma Z_{t+2} + Y_{t+2}N_{t+2}$ (B5)

The variable $z_{t+1} \mid (N_{t+1} = 0, N_{t+2} = 0)$ will be normally distributed with mean $M_{nn} = \mu$, and variance $S_{nn}^2 = \sigma^2$.

The variable $z_{t+1} \mid (N_{t+1} = 0, N_{t+2} = 1)$ will be normally distributed with mean $M_{ny} = \mu + m$, and variance $S_{ny}^2 = \sigma^2 + s^2$.

If $N_{t+1} = 1$, we add Equations (B1) and (B2) and simplify to get

$$z_{t+1} = -z_t + 2\mu + \sigma(Z_{t+1} + Z_{t+2}) + Y_{t+2}N_{t+2} - Y_t N_t. \tag{B6}$$

If no mortality jump event occurs in year t and $t+2$, the variable $z_{t+1} \mid (z_t, N_t = 0, N_{t+1} = 1, N_{t+2} = 0)$ will be normally distributed with mean $M_{nyn} = -z_t + 2\mu$ and variance $S_{nyn}^2 = 2\sigma^2$.

Similarly, $z_{t+1} \mid (z_t, N_t = 1, N_{t+1} = 1, N_{t+2} = 0)$ will be normally distributed with mean $M_{yyn} = -z_t + 2\mu - m$ and variance $S_{yyn}^2 = 2\sigma^2 + s^2$.

The variable $z_{t+1} \mid (z_t, N_t = 0, N_{t+1} = 1, N_{t+2} = 1)$ will be normally distributed with mean $M_{nyy} = -z_t + 2\mu + m$ and variance $S_{nyy}^2 = 2\sigma^2 + s^2$.

The variable $z_{t+1} \mid (z_t, N_t = 1, N_{t+1} = 1, N_{t+2} = 1)$ will be normally distributed with mean $M_{yy} = -z_t + 2\mu$ and variance $S_{yy}^2 = 2\sigma^2 + 2s^2$.

The conditional density function of $z_{t+1} \mid z_t$ denoted by $f(z_{t+1} \mid z_t)$ can be written as

$$\begin{aligned}
f(z_{t+1} \mid z_t) = {} & f(z_{t+1} \mid N_{t+1} = 0, N_{t+2} = 0)\,\Pr(N_{t+1} = 0, N_{t+2} = 0) \\
& + f(z_{t+1} \mid N_{t+1} = 0, N_{t+2} = 1)\,\Pr(N_{t+1} = 0, N_{t+2} = 1) \\
& + f(z_{t+1} \mid z_t, N_t = 0, N_{t+1} = 1, N_{t+2} = 0)\,\Pr(N_t = 0, N_{t+1} = 1, N_{t+2} = 0) \\
& + f(z_{t+1} \mid z_t, N_t = 1, N_{t+1} = 1, N_{t+2} = 0)\,\Pr(N_t = 1, N_{t+1} = 1, N_{t+2} = 0) \\
& + f(z_{t+1} \mid z_t, N_t = 0, N_{t+1} = 1, N_{t+2} = 1)\,\Pr(N_t = 0, N_{t+1} = 1, N_{t+2} = 1) \\
& + f(z_{t+1} \mid z_t, N_t = 1, N_{t+1} = 1, N_{t+2} = 1)\,\Pr(N_t = 1, N_{t+1} = 1, N_{t+2} = 1) \\
= {} & \frac{1}{\sqrt{2\pi}\,S_{nn}} \exp\left(-\frac{(z_{t+1} - M_{nn})^2}{2S_{nn}^2}\right)(1-p)^2 \\
& + \frac{1}{\sqrt{2\pi}\,S_{ny}} \exp\left(-\frac{(z_{t+1} - M_{ny})^2}{2S_{ny}^2}\right)p(1-p) \\
& + \frac{1}{\sqrt{2\pi}\,S_{nyn}} \exp\left(-\frac{(z_{t+1} - M_{nyn})^2}{2S_{nyn}^2}\right)p(1-p)^2 \\
& + \frac{1}{\sqrt{2\pi}\,S_{yyn}} \exp\left(-\frac{(z_{t+1} - M_{yyn})^2}{2S_{yyn}^2}\right)p^2(1-p) \\
& + \frac{1}{\sqrt{2\pi}\,S_{nyy}} \exp\left(-\frac{(z_{t+1} - M_{nyy})^2}{2S_{nyy}^2}\right)p^2(1-p) \\
& + \frac{1}{\sqrt{2\pi}\,S_{yyy}} \exp\left(-\frac{(z_{t+1} - M_{yyy})^2}{2S_{yyy}^2}\right)p^3.
\end{aligned} \tag{B7}$$

The variable $z_1 \mid (N_1 = 0, N_2 = 0)$ will be normally distributed with mean $\hat{M}_{nn} = \mu$, and variance $\hat{S}^2_{nn} = \sigma^2$.

The variable $z_1 \mid (N_1 = 1, N_2 = 0)$ will be normally distributed with mean $\hat{M}_{yn} = \mu - m$, and variance $\hat{S}^2_{yn} = \sigma^2 + s^2$.

The variable $z_1 \mid (N_1 = 0, N_2 = 1)$ will be normally distributed with mean $\hat{M}_{ny} = \mu + m$, and variance $\hat{S}^2_{ny} = \sigma^2 + s^2$.

The variable $z_1 \mid (N_1 = 1, N_2 = 1)$ will be normally distributed with mean $\hat{M}_{yy} = \mu$, and variance $\hat{S}^2_{yy} = \sigma^2 + 2s^2$.

The density function of z_1, which is denoted by $f(z_1)$, can be written as

$$\begin{aligned}
f(z_1) = {} & f(z_1 \mid N_1 = 0, N_2 = 0)\Pr(N_1 = 0, N_2 = 0) \\
& + f(z_1 \mid N_1 = 1, N_2 = 0)\Pr(N_1 = 1, N_2 = 0) \\
& + f(z_1 \mid N_1 = 0, N_2 = 1)\Pr(N_1 = 0, N_2 = 1) \\
& + f(z_1 \mid N_1 = 1, N_2 = 1)\Pr(N_1 = 1, N_2 = 1) \\
= {} & \frac{1}{\sqrt{2\pi}\,\hat{S}_{nn}} \exp\left(-\frac{(z_1 - \hat{M}_{nn})^2}{2\hat{S}^2_{nn}}\right)(1-p)^2 + \frac{1}{\sqrt{2\pi}\,\hat{S}_{yn}} \exp\left(-\frac{(z_1 - \hat{M}_{yn})^2}{2\hat{S}^2_{yn}}\right)p(1-p) \\
& + \frac{1}{\sqrt{2\pi}\,\hat{S}_{ny}} \exp\left(-\frac{(z_1 - \hat{M}_{ny})^2}{2\hat{S}^2_{ny}}\right)(1-p)p + \frac{1}{\sqrt{2\pi}\,\hat{S}_{yy}} \exp\left(-\frac{(z_1 - \hat{M}_{yy})^2}{2\hat{S}^2_{yy}}\right)p^2.
\end{aligned} \tag{B8}$$

Substituting the formulas of $f(z_{t+1} \mid z_t.)$ and $f(z_1)$ into the log-likelihood function (B4), we can calculate the log-likelihood function numerically.

REFERENCES

Biffis, E., 2005, Affine Processes for Dynamic Mortality and Actuarial Valuations, *Insurance: Mathematics and Economics*, 37(3): 443-468.

Blake, D., A. J. G. Cairns, and K. Dowd, 2006, Living with Mortality: Longevity Bonds and Other Mortality-Linked Securities, *British Actuarial Journal*, 12(1): 153-197.

Brouhns, N., M. Denuit, and J. K. Vermunt, 2002, A Poisson Log-Bilinear Regression Approach to the Construction of Projected Lifetables, *Insurance: Mathematics and Economics*, 31(3): 373-393.

Cowley, A., and J. D. Cummins, 2005, Securitization of Life Insurance Assets, *Journal of Risk and Insurance*, 72(2): 193-226.

Chan, W. S., 2002, Stochastic Investment Modeling: A Multiple Time-Series Approach, *British Actuarial Journal*, 8(3): 545-591.

Cairns, A. J. G., D. Blake, and K. Dowd, 2006a, Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk, *ASTIN Bulletin*, 36(1): 79-120.

Cairns, A. J. G., D. Blake, and K. Dowd, 2006b, A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration, *Journal of Risk and Insurance*, 73(4): 687-718.

Cox, S. H., Y. Lin, and S. S. Wang, 2006, Multivariate Exponential Tilting and Pricing Implications for Mortality Securitization, *Journal of Risk and Insurance*, 73(4): 719-736.

Dahl, M. H., 2004, Stochastic Mortality in Life Insurance: Market Reserves and Mortality-Linked Insurance Contracts, *Insurance: Mathematics and Economics*, 35(1): 113-136.

Dahl, M., and T. Møller, 2005, Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk, *Insurance: Mathematics and Economics*, 39(2): 193-217.

Denuit, M., P. Devolder, and A. Goderniaux, 2007, Securitization of Longevity Risk: Pricing Survivor Bonds with Wang Transform in the Lee-Carter Framework, *Journal of Risk and Insurance*, 74(1): 87-113.

Diebold, F., T. Schuermann, and J. Stroughair, 1999, Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management, in: A.-P. N. Refenes, J. D. Moody and A. N. Burgess, eds., *Advances in Computational Finance* (Amsterdam: Kluwer Academic Publishers), pp. 3-12. Reprinted in *Journal of Risk Finance*, 1(Winter 2000): 30-36.

Dowd, K., D. Blake, A. J. G. Cairns, and P. Dawson, 2006, Survivor Swaps, *Journal of Risk and Insurance*, 73(1): 1-17.

Hardy, M., 2001, A Regime-Switching Model of Long-Term Stock Returns, *North American Actuarial Journal*, 5(2): 41-53.

Knapp, D., 2006, Avian Flu: Bracing for a Pandemic, *Risk Management Magazine*, *July 2006*, 42-49.

Lee, R. D., and L. Carter, 1992, Modelling and Forecasting the Time Series of US Mortality, *Journal of the American Statistical Association*, 87(419): 659-671.

Lee, R., and T. Miller, 2001, Evaluating the Performance of Lee-Carter Mortality Forecasts, *Demography*, 38(4): 537-549.

Li, S. H., and W. S. Chan, 2005, Outlier Analysis and Mortality Forecasting: The United Kingdom and Scandinavian Countries, *Scandinavian Actuarial Journal*, 2005 (3): 187-211.

Li, S. H., and W. S. Chan, 2007, The Lee-Carter Model for Forecasting Mortality, Revisited, *North American Actuarial Journal*, 11(1): 68-69.

Lin, Y., and S. H. Cox, 2005, Securitization of Mortality Risks in Life Annuities, *Journal of Risk and Insurance*, 72(2): 227-252.

Lin, Y., and S. H. Cox, 2008, Securitization of Catastrophe Mortality Risks, *Insurance: Mathematics and Economics*, 42(2): 628-637.

Luciano, E., and E. Vigna, 2005, Non Mean Reverting Affine Processes for Stochastic Mortality. ICER Applied Mathematics Working Paper No. 4. Available at SSRN: http://ssrn.com/abstract=724706.

Milevsky, M. A., and S. D. Promislow, 2001, Mortality Derivatives and the Option to Annuitize, *Insurance: Mathematics and Economics*, 29(3): 299-318.

Miltersen, K. R., and S.-A. Persson, 2005, Is Mortality Dead? Stochastic Forward Force of Mortality Determined by No Arbitrage, Working Paper, University of Bergen.

Olivieri, A., and E. Pitacco, 2002, *Inference About Mortality Improvements in Life Annuity Portfolios*, Transactions of the 27th International Congress of Actuaries, Cancun, Mexico.

Rasmussen, E., 2005, Flu Pandemic: A Real Threat, *Bests Review*, March 2005.

Renshaw, A. E., and S. Harberman, 2003, On the Forecasting of Mortality Reduction Factors, *Insurance: Mathematics and Economics*, 32(3): 379-401.

Renshaw, A. E., S. Haberman, and P. Hatzoupoulos, 1996, The Modeling of Recent Mortality Trends in United Kingdom Male Assured Lives, *British Actuarial Journal*, 2(2): 449-477.

Sanders, D. E. A., 2005, The Modeling of Extreme Events. *British Actuarial Journal*, 11(3): 519-572.

Schrager, D. F., 2006, Affine Stochastic Mortality, *Insurance: Mathematics and Economics*, 38(1): 81-97.

Sithole, T. Z., S. Haberman, and R. J. Verrall, 2000, An Investigation into Parametric Models for Mortality Projections, with Applications to Immediate Annuitants' and Life Office Pensioners' Data, *Insurance: Mathematics and Economics*, 27(3): 285-312.

Wang, S. S., 2000, A Class of Distortion Operators for Pricing Financial and Insurance Risks, *Journal of Risk and Insurance*, 67(1):15-36.

Wang, S. S., 2002, A Universal Framework for Pricing Financial and Insurance Risks, *ASTIN Bulletin*, 32(2): 213-234.

Wang, S. S., 2007, Normalized Exponential Tilting: Pricing and Measuring Multivariate Risks, *North American Actuarial Journal*, 11(3): 89-99.