Modeling Mortality with Jumps: Transitory Effects and Pricing Implication to Mortality Securitization

Hua Chen and Samuel H. Cox

May, 2007

(Please do not cite or quote without permission)

Hua Chen (Corresponding author)
Department of Risk Management and Insurance
Georgia State University
P.O. Box 4036
Atlanta, GA 30302-4036
Phone: (405) 370-5899
Fax: (404) 413-7499
Email: hchen10@student.gsu.edu

Samuel H. Cox
Warren Centre for Actuarial Studies and Research
University of Manitoba
181 Freedman Crescent
Winnipeg, Manitoba R3T 5V4, Canada
Phone: (204) 474-7426
Fax: (204) 474-7545
Email: sam_cox@umanitoba.ca

Modeling Mortality with Jumps: Transitory Effects and Pricing Implication to Mortality Securitization

Abstract:

In this paper, we incorporate a jump-diffusion process into the original Lee-Carter model, and use it to forecast mortality rates and analyze mortality securitization. The outlier-adjusted Lee-Carter model is examined to provide further evidence of mortality jumps. We also explore alternative models with transitory versus permanent jump effects. We find that modeling the mortality via permanent jump effects induces errors in parameter estimation and distortion in mortality securitization consequently. We use the Swiss Re mortality bond as an example to show how to apply our model and the distortion measure approach to value mortality-linked securities. Pricing the Swiss Re mortality bond is difficult because the mortality index is correlated across countries and over time. Cox, Lin and Wang (2006) employ normalized multivariate exponential tilting to take into account correlations across countries. We show in this paper how to account for correlations of the mortality index over time by simulating the mortality index and changing the measure on paths.

Key words:

Lee-Carter model, outlier-adjusted Lee-Carter model, transitory jump effects, permanent jump effects, market price of mortality risk.

1. Introduction

Mortality risk management is fundamental to the life insurance and pension industries. Mortality models are crucial as a means of quantifying these risks and providing the basis of pricing and reserving. Traditionally reinsurance, and more recently, securitization, provide a means of transferring or hedging mortality risks. Naturally mortality models are fundamental to these transactions.

Mortality securitizations differ from reinsurance transactions in several ways. Perhaps the most important is that the investor in a securitization typically does 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; Dahl and Møller 2005; Miltersen and Persson 2005; Biffs 2005; Schrager 2006) help us understand the evolution of mortality rates over time, but are relatively intractable at the present time. We prefer discrete time models because they are easy to be implemented 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 index, 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 change 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 short-term mortality shocks, such as the 1918 influenza pandemic, attack different groups with different 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 Haberman (2003), Denuit, Devolder and Goderniaux (2007), and further revisited by Li and Chan (2007). Recently, Cairns, Blake and Dowd (2006b) propose a two-factor model for mortality modeling and morality-linked security pricing. The first factor equally affects mortality at all ages, whereas the second factor's effect on mortality is proportional to age. However, the mortality curve is increasing in ages in their model setup, which does not reflect the fact that mortality rates of infants and children are much higher than those at middle ages. Moreover, their model does not allow mortality jumps. We show that this may lead to a pricing error in morality securitizations.

Mortality jumps must be taken into account in mortality securitization modeling, because the rationale behind selling or buying mortality securities is to hedge mortality risks (Cox, Lin and Wang 2006). Nevertheless, most of papers on this topic, as in Cairns, Blake and Dowd (2006b), ignore mortality jumps (see Renshaw, Haberman, and Hatzoupoulos, 1996; Sithole, Haberman, and Verrall 2000; Milevsky and Promislow, 2001; Olivieri and Pitacco, 2002; Dahl, 2003; Denuit, Devolder and Goderniaux, 2007). Even if they recognize that short-term catastrophe shocks may cause mortality jumps, they do not model mortality 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) regard pandemic events as non-repetitive exogenous intervention too, 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. Cox, Lin and Wang (2006) find that mortality jumps do have a significant effect on mortality modeling. They, however, model the age-adjusted death rates instead of the mortality curve. Thus their model fails to represent age-specific changes of mortality rates. Furthermore, they model the jump process in a way that mortality jumps have permanent effects on mortality rates, although many mortality jumps are caused by short-term catastrophic events and have transitory effects only.

In this paper we propose to incorporate a jump-diffusion process into the Lee-Carter model, restricting mortality jumps to have one-period effects. We fit the model to US 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 permanent jump effects induces big deviations in the parameter estimation compared with that with transitory jump effects. We then discuss the outlier-adjusted Lee-Carter model presented by Li and Chen (2007) to further explore the source of mortality jumps. We find that the so-called “outliers” are actually very important to our morality securitization modeling, and we cannot delete the outliers from our time-series data in order to establish a proper model for pricing mortality securities. We use the Swiss Re mortality bond (2003) as an example of pricing the mortality-linked securities, and illustrate that the model with permanent jump effects results in large pricing distortions.

In an incomplete insurance market, there are mainly two approaches for security valuation.

One way is to adapt the arbitrage-free pricing framework of interest-rate derivatives to the valuation and securitization of mortality risk. Cairns, Blake and Dowd have a detailed discussion on this issue and give as an example the pricing of the EIB longevity bond (see Cairn, Blake and Dowd 2006a, 2006b). 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-neutral measure. Examples of this approach, based on the Wang transform (Wang 2000, 2002), include Lin and Cox (2005), Dowd, Blake, Cairns and Dawson (2006), Denuit, Devolder and Goderniaux (2007). The Swiss Re mortality bond (2003) covers mortality risks across countries and over time, which makes the valuation problem very difficult. Previous research (e.g. Cox, Lin and Wang, 2006) employ the normalized multivariate exponential tilting, which is a generalization of the Wang transform, to take into account correlations across countries. They, however, modify the contract terms by linking the principal repayment with the maximum of the mortality index in three years, and ignore correlations over time. In this article, we employ 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.

The remaining of this article proceeds as follows. In section 2 we describe the data and demonstrate historical facts for further motivation of the problem. The mortality jump caused by the 1918 Spanish flu is clearly evidenced. The high correlation between death by flu and death by all causes suggests we cannot ignore flu events which may cause a huge jump in mortality. In section 3 and 4 we briefly review the Lee-Carter model, propose a jump-diffusion process with temporary jump effects to fit the mortality index, and use conditional maximum likelihood estimation (CMLE) to calibrate the parameters. In Section 5 we discuss the outlier-adjusted

Lee-Carter model, further explore the source of mortality jumps and justify the necessity of modeling mortality with jumps. In Section 6 we take the Swiss Re mortality bond (2003) as an example to illustrate how to use our mortality model and the Wang transform to price the Swiss Re mortality bond with the mortality index weighted by ages and correlated over time. Concluding remarks and discussions are provided in Section 7.

2. Data descriptions and historical facts: further motivation

Our 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.[1] 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 5-14, 15-24, up to 75-84, and the 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% of the 1990's level (2518.0 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% of its initial value, but that for age group 85 and over only dropped to 55.9% 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 improving on all ages.

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

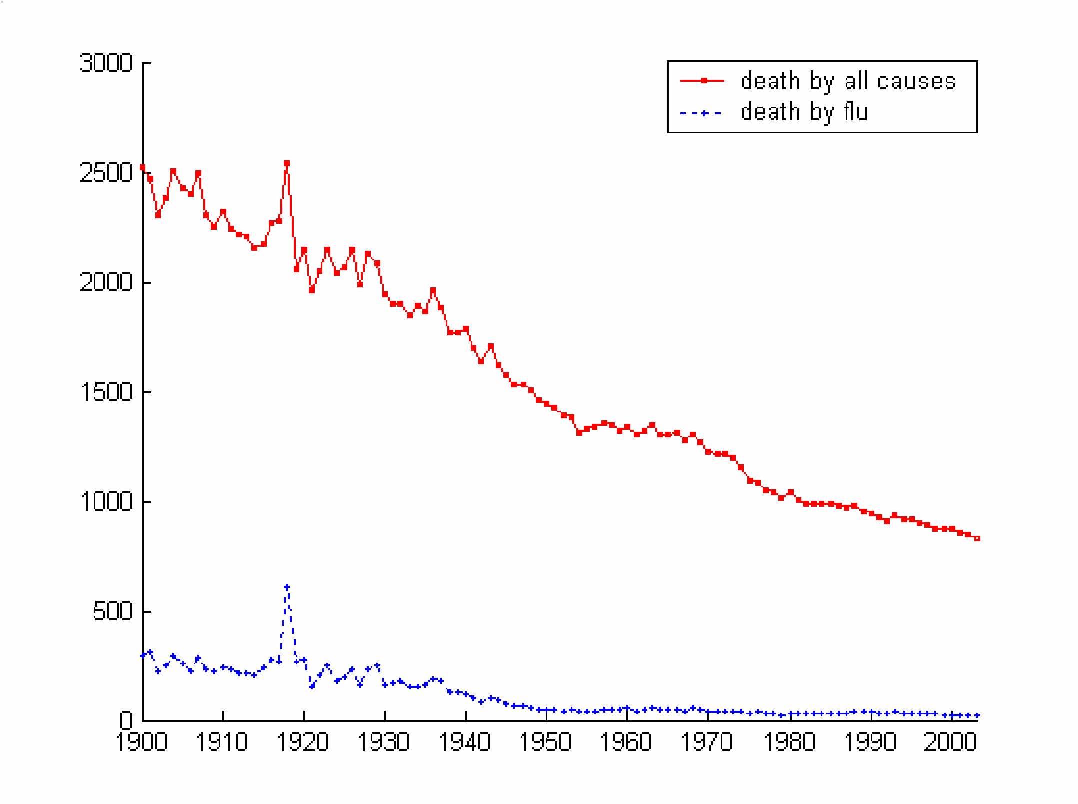
Table 1: The mortality improvement by different age-groups, from 1990 to 2003

Age groups	1900	2003	Ratio	Age groups	1900	2003	Ratio
All	2518.0	832.7	0.331	35-44	1023.1	201.6	0.197
<1	16244.8	700.0	0.043	45-54	1495.4	433.2	0.290
1-4	1983.8	31.5	0.016	55-64	2723.6	940.9	0.345
5-14	385.9	17.0	0.044	65-74	5636.1	2255.0	0.400
15-24	585.5	81.5	0.139	75-84	12330.0	5463.1	0.443
25-34	819.8	103.6	0.126	>=85	26088.2	14593.3	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 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 still observe the similar pattern of fluctuations of death rates by all causes and by flu in the first half of the past century. The two graphs even jump at the same time, which is remarkably evidenced in year 1918. The correlation coefficient between the two trends is 0.9116, which also indicates a close correspondence between flu-caused deaths and all deaths. By our calculation, deaths caused by flu account for 9.4% of all deaths before 1950 in average, 3.4% from 1950 to 2003, and 6.3% for the whole period examined. In 1918, this proportion reaches its historic peak at 24.1%. The high correlation between the two graphs and high portion of deaths caused by flu suggest that we should not exclude flu events when modeling the mortality.

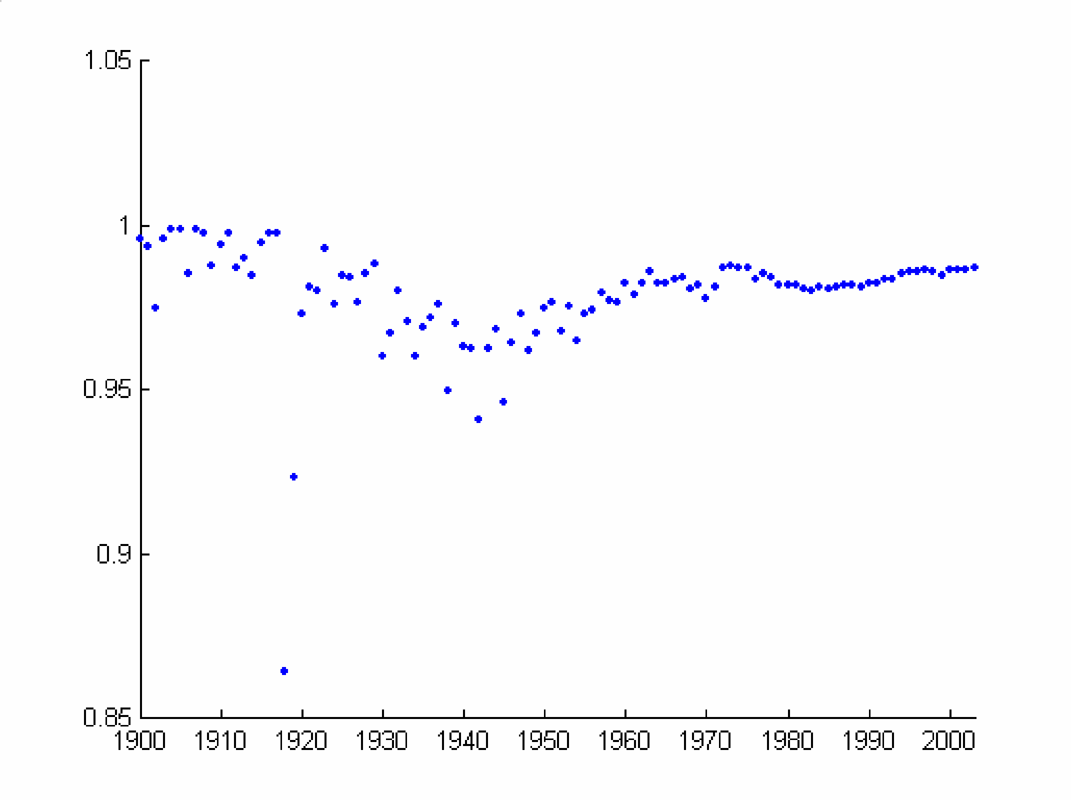
Figure 1: US age-adjusted death rate per 100,000 by all causes and by influenza and Pneumonia, from 1900 to 2003.



Note: Data are from NCHS reports HIST293 and GMWK293R

The high correlation between deaths by flu and all deaths is more evident if we examine age-specific death rate data. We can calculate the correlation coefficient between death rates by all causes 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 0.9787. Interestingly, the correlation falls to the lowest value of 0.86 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 on death rates is revealed in more details in Table 2. The 1918 influenza pandemic raised the mortality rate by 30% 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.

Figure 2: Correlation coefficients between the age-specific death rates per 100,000 by all causes and that by Influenza and Pneumonia, each year from 1900-2003



Note: Age-specific death rates are from NCHS reports HIST290 and GMWK290R. The correlation coefficients are calculated by the authors.

Table 2: The change of death rates per 100,000 for each age group, from 1917 to 1919

Age groups	1917	1919	Ratio	Age groups	1917	1919	Ratio
All	1397.1	1810	1.296	35-44	900.8	1339.3	1.487
<1	10457.2	11167.2	1.068	45-54	1385.6	1524.1	1.1
1-4	1066	1573.5	1.476	55-64	2678.6	2648.1	0.989
5-14	256	412.8	1.613	65-74	5728.4	5505	0.961
15-24	468.9	1070.6	2.283	75-84	12386.2	11295.7	0.912
25-34	649.1	1643.5	2.532	>=85	24593.6	22213.5	0.903

Note: Data are from NCHS report HIST290 and GMWK290R. The mortality improvement ratio is calculated by the authors.

3. Mortality modeling: the classical Lee-Carter Model

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 their long-run forecast of US life expectancy. The two most recent Social Security Technical Advisory Panels have recommended the adoption of the method, or forecasts consistent with it, by the Trustees.

Let $m_{x,t}$ be the central death rate for age x at time t, then 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 index k_t. 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 index, and $e_{x,t}$ is the error term which captures the age-specific effect not reflected in the model.

The Lee-Carter model cannot be fitted by the ordinary least square approach, because all variables on the right side of the model are unobservable. Moreover, this model is obviously under-identified. To obtain a unique solution, we impose the normalization conditions such that the b_x terms sum to unity and the k_t terms sum to zero, i.e.,

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

Then a_x becomes the average value of $\ln(m_{x,t})$ over time, i.e.,

$$a_x = \sum_t \ln(m_{x,t}) \tag{3}$$

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 re-estimated by iteration, given the values of a_x and b_x in the first step, such that the implied number of deaths equals to the actual number of deaths.

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

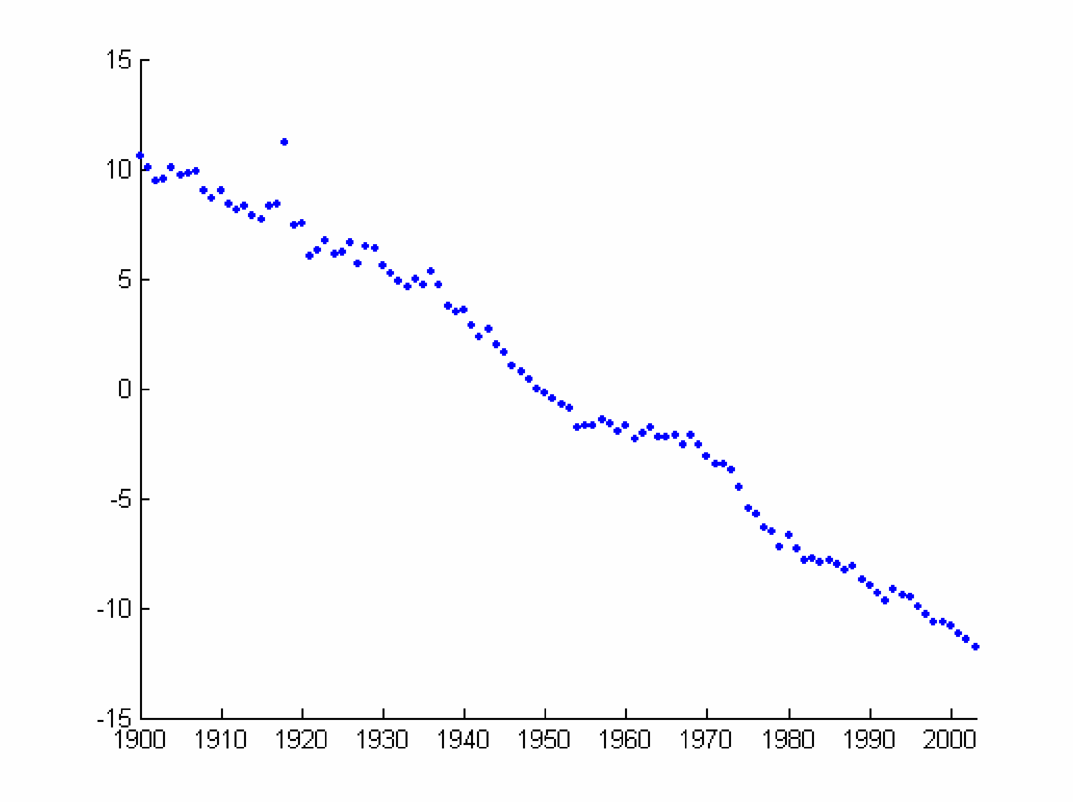
where D_t is the actual total number of deaths at time t, and $Pop_{x,t}$ is the population in age group x at time t.

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 index k_t in Figure 3. We can see that generally the mortality rates of young age groups respond more rapidly when the mortality index changes. As we expect, the mortality index 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.

Table 3: Fitted value of a_x and b_x (SVD) for the Lee-Carter model, from 1990 to 2003

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	0.085801
45-54	-4.7734	0.060443
55-64	-4.0071	0.046024
65-74	-3.2289	0.041927
75-84	-2.4146	0.040345
85 over	-1.6084	0.028619

Figure 3: Dynamics of the mortality index k_t for the Lee-Carter model, from 1990 to 2003



4. Model k_t with a jump-diffusion process: permanent versus transitory effect?

To make forecast of the future distribution of k_t, we need to choose a suitable model to fit k_t. Cox, Lin and Wang (2006) combine a geometric Brownian motion and a compound Poisson process to model the age-adjusted mortality rates for US and UK. We cannot apply their model here. First, the mortality index k_t decreases from positive to negative values. This property restrains us from modeling it with a geometric Brownian motion, because a geometric Brownian motion will never become negative when starting from a positive value. Instead, a process driven by a standard Brownian motion may be applicable. Second, Cox, Lin and Wang (2006) include jumps into the stochastic differential equation, which makes the jumps have permanent effects on mortality rates. However, most of the 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. We believe that a model with permanent jump effects is inappropriate for mortality modeling, especially for mortality securitization modeling. Therefore, in this section we model the mortality index k_t with a standard Brownian motion and a discrete Markov chain with jumps which only have transitory effects. For the purpose of comparison, we also present the model with permanent jump effects in the Appendix A.

Let N_t be the total number of jumps during the time interval $(0,t)$. Suppose there is at most one jump event in each of time period $(t-h,t)$, then N_t can be expressed as a discrete Markov chain with $N_0 = 0$ and transition path:

$$N_{t+h} = \begin{cases} N_t + 1, & prob = p \\ N_t, & prob = 1 - p \end{cases} \tag{5}$$

Let $N_{[t-h,t]} = N_t - N_{t-h}$ be the number of jumps occurring in the period $(t-h,t)$, then $N_{[t-h,t]}$ is a Bernoulli random variable with probability of jump p.

Let $\tilde{k}(t)$ denote the mortality index when there are no jump events. It can be driven by a standard Brownian motion:

$$d\tilde{k}(t) = udt + \sigma dW_t \tag{6}$$

where u and σ are the instantaneous rate of change and the instantaneous volatility of the mortality index when there are no jumps, and W_t is a standard Brownian motion with mean 0 and variance t.

If a jump occurs in the interval $(t-h,t)$, i.e., $N_{[t-h,t]} = 1$, we assume the jump size $Y_{[t-h,t]}$

are identically independently distributed normal variables with mean m and standard deviation s, and $Y_{[t-h,t]}$ is independent of the Brownian motion W_t. The jump $Y_{[t-h,t]}$ makes the actual mortality index $k(t)$ change from $\tilde{k}(t)$ to $\tilde{k}(t)+Y_{[t-h,t]}$. That is,

$$k(t)=\tilde{k}(t)+Y_{[t-h,t]} \tag{7}$$

If there is no jump in the interval $(t-h,t)$, i.e., $N_{[t-h,t]}=0$, we know that:

$$k(t)=\tilde{k}(t) \tag{8}$$

(7) and (8) can be combined and written in one equation:

$$k(t)=\tilde{k}(t)+Y_{[t-h,t]}N_{[t-h,t]} \tag{9}$$

Therefore, the dynamics of the mortality index $k(t)$ can be completely expressed as:

$$\begin{cases} d\tilde{k}(t)=udt+\sigma dW_t \\ k(t)=\tilde{k}(t)+Y_{[t-h,t]}N_{[t-h,t]} \end{cases} \tag{10}$$

By integrating the first equation in (10) from t to $t+h$, we get

$$\tilde{k}(t+h)=\tilde{k}(t)+uh+\sigma[W_{t+h}-W_t], \tag{11}$$

From the second equation in (10), we can derive

$$\begin{aligned} k(t+h) &= \tilde{k}(t+h)+Y_{[t,t+h]}N_{[t,t+h]} \\ &= \tilde{k}(t)+uh+\sigma[W_{t+h}-W_t]+Y_{[t,t+h]}N_{[t,t+h]} \\ &= k(t)-Y_{[t-h,t]}N_{[t-h,t]}+uh+\sigma[W_{t+h}-W_t]+Y_{[t,t+h]}N_{[t,t+h]} \end{aligned} \tag{12}$$

Let $z_t=k(t+h)-k(t)$. If we have a time series of K observations of $k(t)$, there will be $K-1$ observations of z's values with time interval equal to h=1. z_t and z_{t+h} can be expressed as:

$$z_t=uh+\sigma[W_{t+h}-W_t]+Y_{[t,t+h]}N_{[t,t+h]}-Y_{[t-h,t]}N_{[t-h,t]}, \tag{13}$$

$$z_{t+h}=uh+\sigma[W_{t+2h}-W_{t+h}]+Y_{[t+h,t+2h]}N_{[t+h,t+2h]}-Y_{[t,t+h]}N_{[t,t+h]}, \tag{14}$$

If $N_{[t,t+h]}=0$, then z_t is independent on z_{t+h}. If $N_{[t,t+h]}=1$, then z_t is correlated with

z_{t+h} because of the $Y_{[t,t+h]}$ part. We cannot use the traditional maximum likelihood estimation to calibrate the parameters when the data are not independent. [2] Instead, we should use conditional probability to derive the log-likelihood function, which is so-called Conditional Maximum Likelihood Estimation (CMLE). Detailed derivation of the log-likelihood function is included in Appendix B.

Table 4: Parameter estimates via Maximum Likelihood Estimation, using mortality data from 1900 to 2003

Model with jumps-transitory effect: Ln(likelihood) = -62.52			
Parameter	Estimate	Parameter	Estimate
u	-0.2173	σ	0.3733
m	0.8393	s	1.4316
p	0.0436		
Model with jumps:-permanent effect			
Parameter	Estimate	Parameter	Estimate
u	-0.2172	σ	0.3872
m	-0.3062	s	2.3133
p	0.0396		
Model without jumps: Ln(likelihood) = -94.27			
Parameter	Estimate	Parameter	Estimate
u	-0.2172	σ	0.6043
Likelihood Ratio Test (LRT) statistics = 63.49			

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

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 index, u, is -0.2173, which implies

[2] Lin and Cox (2006) try to combine a geometric Brownian motion with a Markov chain to capture the transitory effect of mortality jumps. However, they don't take into account the correlations of the data, which may bring big errors in their maximum likelihood estimation and cause the estimation results to deviate from the true values. If we don't consider the correlations of the data, then the parameter estimates are u = -0.2172, σ = 0.4018, m = -3.2391, s = 0, p = 0.0098.

the mortality index k_t decreases by -0.2173 per year on average. The negative sign of u is consistent with the fact that the U.S population mortality improves over time. The instantaneous volatility is equal to 0.3733. 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[3], the results of which are shown in the middle panel of Table 4. The instantaneous mean and volatility of the mortality index are roughly the same as before, while there are significant 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. We will show later that modeling jumps to have permanent effects brings a large pricing error in the mortality securitization.

The estimation results for the model without jumps are in the lower panel of Table 4. The instantaneous mean of the mortality index is unchanged, while the instantaneous volatility increases to 0.6043, which is an increment by 61 percent, 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 the model with transitory jump effects and for the model without jumps in Table 4, and perform the likelihood ratio test. Our test result rejects the model without jumps at the significance level of 1%.

5. Evidence from the outlier-adjusted Lee-Carter Model: Do outliers matter?

We have already shown that a jump-diffusion process fits the mortality index better than the model without jumps. Our next question is where the mortality jumps come from. Before answering this question, we need to briefly review the work done by Li and Chan (2005, 2007).

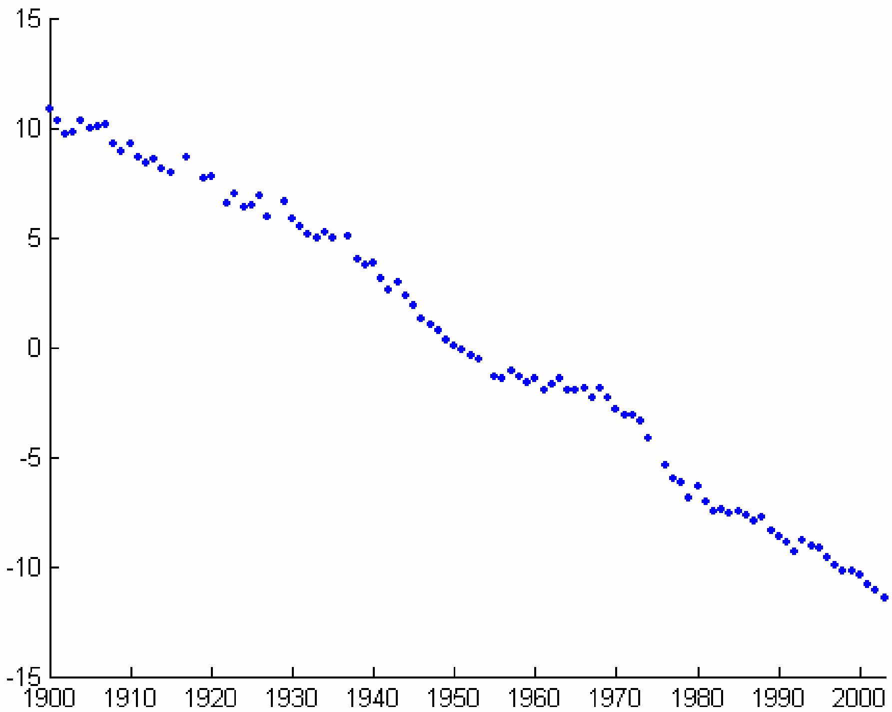
---

[3] See Appendix A for the model setup of the model with permanent jump effects.

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 US and Canada, and fit the adjusted outlier-free mortality series to the Lee-Carter model. For the US data from 1900 to 2000, they find 7 outliers, which occurred in year 1916, 1918, 1921, 1928, 1936, 1954 and 1975, respectively. These outliers are closely related to or resulted from influenza epidemics according to their explanations, except for the data in 1954 and 1975.

So 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 our original mortality data, estimate the mortality index k_t again (Figure 4), and compare the results of the model with transitory jump effects and that without jumps (Table 5). We can see that after eliminating the outliers the mortality index k_t declines more smoothly and doesn't show significant jumps in the evolution process. In addition, when we fit the mortality index 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 indifferent with the model without jumps. We therefore infer that the mortality jumps in our model arise from these so-call "outliers", which are mostly caused by flu epidemics.

Figure 4: Dynamics of the mortality index k_t for the outlier adjusted Lee-Carter Model, from 1990 to 2003



Note: Mortality data in year 1916, 1918, 1921, 1928, 1936, 1954, and 1975 are deleted according to the outliers analysis of Li and Chan (2007)

Table 5: Parameter estimates via Maximum Likelihood Estimation, using the adjusted outlier-free mortality data from 1900 to 2003

Model with jumps-transitory effect: Ln(likelihood) = -48.38			
Parameter	Estimate	Parameter	Estimate
u	-0.2317	σ	0.4005
m	-0.3774	s	0.9316
p	0		
Model without jumps: Ln(likelihood) = -48.38			
Parameter	Estimate	Parameter	Estimate
u	-0.2317	σ	0.4005
Likelihood Ratio Test (LRT) statistics =0			

Note: Based on our Likelihood ratio test, we cannot reject the model without jumps at the 99% significance level.

Actually, for the model with transitory jumps effects the estimated values of m, s, and p are not very stable if we choose different initial values. However, the value of the log-likelihood function remains unchanged.

Is it appropriate to exclude the outliers when we model the death rates for pricing mortality-linked securities? As Chan (2002) recognizes, "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."[4] We have witnessed the high correlation between deaths caused by flu and deaths by all reasons in data section, and recognized that mortality jumps mostly result from influenza epidemics. We conclude that outliers should not be neglected and that mortality jumps should be explicitly modeled in mortality securitization, since the rationale behind selling or buying mortality securities is to hedge mortality risks. Besides, historic data shows 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 which will never happy again.

6. Example of pricing mortality securities: the Swiss Re mortality bond

In general, there are two types of mortality risks we need to consider. The first is longevity risk, which refers to the uncertainty in 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/BNP 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

[4] See Chan (2002), page 559-560

provided an instructive case study.

The second mortality risk is short-term catastrophe shocks, which are 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.

Modeling mortality jumps seems more important for securities linked to short-term catastrophic risks. Ignoring mortality jumps may cause us to underestimate the probability of having a catastrophic event, and overestimate the variation of the mortality index (see Table 4) at the same time, which may bring large errors in pricing mortality-linked securities and calculating the risk premium. We, therefore, take the Swiss Re mortality bond (2003) as an example to show how to apply the model here to mortality securitization.

The Swiss Re mortality bond and the pricing difficulties

The Swiss Reinsurance company is the second largest reinsurance company in the world. In order to reduce its exposure to catastrophic mortality risks, it issued its first pure mortality bond of $400 million in December 2003. The bond matured on 1 January 2007. Coupons are paid quarterly at a rate of three-month U.S. dollar LIBOR plus a spread of 135 basis points. However, the principal is at risk and depends on a weighted average of the mortality index across five countries and different age-groups.[5] If the mortality index does not exceed 1.3 times the 2002

[5] The five countries are U.S., U.K., France, Italy and Switzerland. The weights assigned to each country are: U.S. 70%, U.K. 15%, France 7.5%, Italy 5%, Switzerland 2.5%.

base level during any of the three years, the principal is fully 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. Let q_t denote the weighted average mortality index at year t (t =2004, 2005, 2006), and q_0 be the 2002 base level of the mortality index, then the payoff schedule of this bond is shown as follows:

$$f(t) = \begin{cases} LIBOR + spread, & \text{t} = 1,2,...\text{T - 1} \\ LIBOR + spread + \max\left(1 - \sum_{t=2004}^{2006} loss_t, 0\right), & \text{t = T} \end{cases} \tag{15}$$

where the loss percentage in year t is defined as

$$loss_t = \begin{cases} 0, & q_t \le 1.3q_0 \\ \dfrac{q_t - 1.3q_0}{0.2q_0}, & 1.3q_0 < q_t \le 1.5q_0 \\ 1, & q_t > 1.5q_0 \end{cases} \tag{16}$$

There are two difficulties we need to deal with in order to value the Swiss Re mortality bond. First, the mortality index defined in the contract is a weighted average across five countries. The correlation of mortality risks across countries makes the pricing problem difficult. Cox, Lin and Wang (2006) solve this problem by adopting the normalized multivariate exponential tilting to take into account correlations across countries. Second, the principal repayment of the Swiss Re mortality bond is based on the experience of the mortality index in three consecutive years. The correlation of the mortality index over time makes the problem even more difficult. Cox, Lin and Wang (2006) take the maximum of the mortality index in three 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 paper, we adopt the Wang transform and find a way to take into account correlations of the mortality index over time. For simplicity, we assume the mortality index only depends on the US mortality rate which is a weighted average of the age-specific mortality rates, by the year 2000 standard population. However, our derivation can be extended to the mortality index weighted by countries and age-groups, which can be done by the multivariate exponential tilting, for example.

Pricing the Swiss Re via the Wang transform on paths

As mentioned in the introduction section, the Wang transform has been widely used as a universal framework for pricing financial and insurance risks. Assumed that the true underlying probability distribution is known without ambiguity, for a given asset X with cdf $F(x)$, the Wang transform will produce a risk-adjusted cdf $F^*(x)$:

$$F^*(x) = \Phi[\Phi^{-1}(F(x)) - \lambda], \qquad (17)$$

where Φ is the standard normal cumulative distribution and the parameter λ is the market price of risk, reflecting the level of systematic 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 X.

One important feature of the Wang transform is that it preserves the normal and lognormal distribution, which enables it to replicate the CAPM if the return for an underlying asset has a normal distribution and recover the Black-Scholes formula if the return for 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 also 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:

$$\begin{cases} d\tilde{k}(t) = udt + \sigma dW_t \\ k(t) = \tilde{k}(t) + Y_{[t-h,t]} N_{[t-h,t]} \end{cases} \tag{18}$$

where $W_t \sim N(0,t)$, $Y_{[t-h,t]} \sim N(m,s^2)$ and $N_{[t-k,t]} = \begin{cases} 0, & prob = 1-p \\ 1, & prob = p \end{cases}$ under P.

Assuming the Brownian motion W_t, the jump severity Y, and the jump frequency N are independent with each other, we can apply the Wang transform to W_t, Y, and N respectively. Under the risk-adjusted measure Q, W_t^* is normally distributed with mean $\lambda_1 t$ and variance t, $Y_{[t-h,t]}^*$ is normally distributed with mean $m + \lambda_2 s$ and variance s^2, and $N_{[t-h,t]}^*$ is a Bernoulli random variable with the probability of jumps p^*, where $p^* = 1 - \Phi[\Phi^{-1}(1-p) - \lambda_3]$.[6] [7]

Mathematically, the dynamics of the mortality index under Q becomes:

$$\begin{cases} d\tilde{k}^*(t) = udt + \sigma dW_t^* \\ k^*(t) = \tilde{k}^*(t) + Y_{[t-h,t]}^* N_{[t-h,t]}^* \end{cases} \tag{19}$$

where $W_t^* \sim N(\lambda_1 t, t)$, $Y_{[t-h,t]}^* \sim N(m + \lambda_2 s, s^2)$, and $N_{[t-h,t]}^* = \begin{cases} 0, & prob = 1-p^* \\ 1, & prob = p^* \end{cases}$

Here the parameter λ_1, λ_2 and λ_3 represent the market prices of risk associated with the Brownian motion, the jump severity and the jump frequency, respectively. Because we have an incomplete market for mortality-linked securities, the values of λ_1, λ_2 and λ_3, and thus the choice of the risk-adjusted measure Q, are not unique.

Pricing procedures and results

---

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

[7] Note that the market price of risk increases with the time horizon, i.e., $\lambda = \lambda_1 \sqrt{t}$, where λ_1 represents the market price of risk per annum. $W_t^* \sim N(\lambda * \sqrt{t}, t) \sim N(\lambda_1 \sqrt{t} * \sqrt{t}, t) \sim N(\lambda_1 t, t)$. See Wang (2002) for more detailed discussion on this issue.

We follow the procedure below to calculate the market prices of risk of the Swiss Re mortality bond (2003) based on the Wang transform.

1. Because the mortality index $k(t)$ is correlated over time, we cannot simulate $k(t)$ for each year independently. We need to simulate the path of the mortality index $k(t)$ ($t =$ 2004, 2005, and 2006). We do this by 10,000 times, using the jump-diffusion process (18) and the parameter estimates shown in the upper panel of Table 4.

2. We use the Wang transform, change from the physical probability measure P to the risk-adjusted probability measure Q, and calculate the values of $k^*(t)$ on each path under Q using equation (14), given initial values of the market prices of risk λ_1, λ_2 and λ_3.

3. We calculate the mortality rates for different age groups by the 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. [8]

4. We calculate the loss percentage, $loss^*_t$, under Q by the equation (16), and compute the risk-adjusted expected value of the principal repayment of the Swiss Re mortality bond as of the period T.

$$E^*_T[repayment] = 400{,}000{,}000 \times \left(\frac{1}{10{,}000} \sum_{i=1}^{10{,}000} [\max(1 - \sum_t loss^*_{i,t}, 0)] \right), \tag{20}$$

5. We discount the coupon payments each period and the principal repayment back to the beginning of year 2004, using the risk-free rate[9]. Letting the discounted expected payoff equal to

[8] 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 age 1-4, 0.145565 for age 5-14, 0.138646 for age 15-24, 0.135573 for age 25-34, 0.162613 for age 35-44, 0.134834 for age 45-54, 0.087247 for age 55-64, 0.066037 for age 65-74, 0.044842 for age 75-84, 0.015508 for age 85 and over.

[9] We use the US Treasury yield rates on December 30, 2003 as the risk-free rates. We calculate the coupon payment by assuming it is paid annually, because we don't have quarterly data for US Treasury yield rates.

the issue size of the Swiss Re mortality bond $400,000,000, we can obtain the 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, i.e., $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, and solve for λ consequently, which is the method used by Cox, Lin and Wang (2006). [10]

Table 6: Estimated 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		

Based on the par spread of the Swiss Re bond 1.35%, the estimated market prices of risk for different models are shown in Table 7.[11] What we estimate here are the market prices of risk associated with the Brownian motion, the jump severity, and the jump frequency for the

[10] The purpose of this paper 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 happened in 2003. Actually, the Swiss Reinsurance company issued another 3 transactions of mortality bonds in 2006. Our model and pricing strategy can be easily extended to multiple-transaction situations. We can find 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 i th issue depending on the parameters λ_1, λ_2 and λ_3, and P_i is the actual market quotes for the i th transaction.

[11] We don't 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 enough to adjust the distribution of the jump frequency.

mortality index $k(t)$, instead of the true mortality rates. Because the mortality index $k(t)$ has a much larger scale than the true mortality rates ($k(t)$ changes from 10.6 in 1900 to -11.8 in 2003, and the age-adjusted mortality rate changes from 0.02518 in 1990 to 0.00833 in 2003), it is not surprised to see that the market prices of risk estimated in our paper have such high values.

Under the model with transitory jump effects, if we assume that the risk associated with the jump process is diversifiable, i.e., $\lambda_2 = \lambda_3 = 0$, the market price of risk associated with the Brownian motion is 5.1449. If there is no systematic risk of the Brownian motion and the jump frequency, i.e., $\lambda_1 = \lambda_3 = 0$, the market price of risk associated with the jump severity is 3.5004. 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 we switch to the model with permanent jump effects, the estimated market prices of risk drop dramatically in each case. This can be explained by the large difference in the volatility of the jump severity distribution and the difference in the intrinsic model setup. First, if we model mortality jumps to have permanent effects, the jump effects 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, while it is 1.4316 in the model with transitory jump effects. The higher volatility of the jumps in the former model raises the risk further. When the par spread of the Swiss Re is fixed at 1.35%, the higher risk imposed on the principle repayment, the lower market price of risk we estimate. Therefore, modeling the mortality with permanent jump effects leads to a large distortion of the market prices of risk.

We come to the model without jumps at last. The estimated market price of risk associated with the Brownian motion 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 index while underestimating the probability of catastrophic events. We suspect that the effect of overestimating the variation predominates the effect of underestimating the catastrophic probability, which brings down the market price of risk associated with the Brownian motion.

7. Conclusion and discussion

In this paper, we have a deep discussion in mortality modeling and mortality-linked security pricing. A good stochastic mortality model for pricing mortality securities should meet the following criteria, while none of the previous research addresses all these problems.

1. The model should capture both the mortality trend over time and the age-specific changes for different age groups. Modeling the age-adjusted mortality rates is not enough, because the payoffs of mortality securities are sometimes linked to a mortality index based on different age groups.

2. The model should incorporate a mortality jump process explicitly. Mortality jumps caused by short-term catastrophic events, such as the 1918 Spanish Flu, cannot be ignored because the rationale of mortality securitization is to hedge mortality risks.

3. Mortality jumps should have transitory effects on mortality rates. It is inappropriate to model mortality jumps having permanent effects, especially when we value mortality bonds, because most of mortality jumps are caused by short-term catastrophic events and the effects should fade away after one or several periods.

4. The model with transitory jump effects introduces correlations of the data. When

estimating the parameters in the model, we cannot simply assume the data are independent and use the traditional maximum likelihood estimation.

We extend the work of Cox, Lin and Wang (2006), and address all problems mentioned above in this paper. We make the first attempt to incorporate a jump process into the Lee-Carter model and discuss in detail how to model the mortality index $k(t)$ with transitory jump effects. Secondly, we derive the conditional log-likelihood function and estimate the parameters via the approach of Conditional Maximum Likelihood Estimation (CMLE), provided that the data are correlated over time. We show that big estimation errors occur if we assume data are independent, which is the problem in Lin and Cox (2005). Thirdly, we compare the model with permanent jump effects to that with transitory jump effects, and present how the difference will cause a large pricing distortion. Finally, our article contributes to the existing literature by showing how to account for correlations of the mortality index over time when pricing the Swiss Re mortality bond (2003). The basic idea is to simulate the paths of the mortality index and change measures on paths.

A 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. Secondly, we ignore the issue of parameter uncertainty in this paper. We simulate the mortality index using estimates of the parameters, assuming these parameter estimates are true values without ambiguity. Another line of further research is to relax this assumption and consider the valuation and hedging of mortality securities under parameter uncertainty.

References:

The Actuary, 2004, Swiss Re obtains $400 Million of Mortality Risk Coverage, January/February, p. 16.

Bauer D. and Jochen Rub. 2006. Pricing Longevity Bonds using Implied Survival Probabilities. Working paper.

Biffis, E., 2005, Affine Processes for Dynamic Mortality and Actuarial Valuations, *Insurance: Mathematics and Economics*, 37: 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: 153-197

Blake, D., A. J. G.. Cairns, K. Dowd and R. MacMinn. 2006. Longevity Bonds: Financial Engineering, Valuation and Hedging. *Journal of Risk and Insurance*, 73: 647-672.

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: 373–93.

Buhlmann, H., 1980, An Economic Premium Principle, *ASTIN Bulletin*, 11(11): 52-60.

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: 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. *The 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. *The 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: 113-136

Dahl, M., and T. Møller, 2005, Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk, In the Proceedings of the 15th International AFIR Colloquium, Zurich, Available online at <http://www.afir2005.ch>.

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

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

Duffie, D., 1992, *Dynamic Asset Pricing Theory* (Princeton, New Jersey: Princeton University Press).

Grove, R.D., and A. M. Hetzel. 1968. Vital Statistics Rates in the United States 1940-1960, U.S. Department of Health, Education and Welfare, Public Health Service, National Center for Health Statistics, Washington, DC: U. S. Government Printing Office.

Karatzas, I., and S. E. Shreve. 1992. *Brownian Motion and Stochastic Calculus* (2nd edition) (New York: Springer-Verlag).

Kijima, M.. 2006. A Multivariate Extension of Equilibrium Pricing Transforms: The Multivariate Esscher and Wang Transforms for Pricing Financial and Insurance Risks,

ASTIN Bulletin, 36(1): 269-284.
Knapp, D. 2006. Avian Flu: Bracing for a Pandemic. *Risk Management Magazine, July 2006, 42-49*
Lee, R. D., 2000, The Lee-Carter Method of Forecasting Mortality, With Various Extensions and Applications, *North American Actuarial Journal*, 4(11): 80-93.
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.
Li, S. H., and W. S. Chan. 2005. Outlier Analysis and Mortality Forecasting: the United Kingdom and Scandinavian Countries. *Scandinavian Actuarial Journal.* (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)
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, 2006, Securitization of Catastrophe Mortality Risks, Working paper, Georgia State University.
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.
Moller, T., 2004, Stochastic Orders in Dynamic Reinsurance Markets, *Finance and Stochastic*, 8(4): 479-499.
Rasmussen, Erik. 2005. Flu Pandemic: A Real Threat. *Bests Review*, March 2005
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.
Renshaw, A. E., and S. Harberman. 2003. On the Forecasting of Mortality Reduction Factors. *Insurance: Mathematics and Economics* 32: 379–401.
Schrager, D. F., 2006, Affine Stochastic Mortality. *Insurance: Mathematics and Economics,* 38: 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. *The 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.

Appendix A: log-likelihood function of the model with permanent jump effects

If we assume that the jump events have permanent effects on mortality modeling, the dynamics of $k(t)$ can be expressed as:

$$dk(t) = \begin{cases} (u - pm)dt + \sigma dW_t \\ \quad \text{if the jump event does not occur at time t, i.e. } N_{[t,t+dt]} = 0 \\ (u - pm)dt + \sigma dW_t + Y_{[t,t+dt]} \\ \quad \text{if the jump event occurs at time t, i.e., } N_{[t,t+dt]} = 1 \end{cases} \tag{A.1}$$

It can be simplified to:

$$dk(t) = (u - pm)dt + \sigma dW_t + Y_{[t,t+dt]} N_{[t,t+dt]}, \tag{A.2}$$

By integrating both sides from t to $t+h$, we can get

$$k(t+h) = k(t) + (u - pm)h + \sigma[W_{t+h} - W_t] + Y_{[t,t+h]} N_{[t,t+h]}, \tag{A.3}$$

Let $z_t = k(t+h) - k(t) = (u - pm)h + \sigma[W_{t+h} - W_t] + Y_{[t,t+h]} N_{[t,t+h]}$, (A.4)

If $N_{[t,t+h]} = 0$ the variable $z_t \big| (N_{[t,t+h]} = 0)$ will be normally distributed with mean $M_n = (u - pm)h$, and variance $S_n^2 = \sigma^2 h$.

If $N_{[t,t+h]} = 1$ the variable $z_t \big| (N_{[t,t+h]} = 1)$ will be normally distributed with mean $M_y = (u - pm)h + m$, and variance $S_y^2 = \sigma^2 h + s^2$.

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

$$\begin{aligned} f(z_t) &= f(z_t \big| N_{[t,t+h]} = 0)\Pr(N_{[t,t+h]} = 0) + f(z \big| N_{[t,t+h]} = 1)\Pr(N_{[t,t+h]} = 1) \\ &= \frac{1}{\sqrt{2\pi} S_n} \exp\left(-\frac{(z_t - M_n)^2}{2S_n^2} \right) \times (1 - p) + \frac{1}{\sqrt{2\pi} S_y} \exp\left(-\frac{(z_t - M_y)^2}{2S_y^2} \right) \times p \end{aligned} \tag{A.5}$$

If we have a time series of K observations of $k(t)$, there will be $K-1$ observations of z's values with time interval equal to h=1. The log-likelihood function can be expressed as follows:

$$\sum_{i-1}^{K-1} \log f(z_i)$$

$$= \sum_{i=1}^{K-1} \ln\left(\frac{1}{\sqrt{2\pi} S_n} \exp\left(-\frac{(z_t - M_n)^2}{2S_n^2} \right)(1-p) + \frac{1}{\sqrt{2\pi} S_y} \exp\left(-\frac{(z_t - M_y)^2}{2S_y^2} \right) p \right), \tag{A.6}$$

Appendix B: log-likelihood function of the model with transitory jump effects

Recall in section 4, we have

$$z_t = uh + \sigma[W_{t+h} - W_t] + Y_{[t,t+h]} N_{[t,t+h]} - Y_{[t-h,t]} N_{[t-h,t]}, \tag{B.1}$$

$$z_{t+h} = uh + \sigma[W_{t+2h} - W_{t+h}] + Y_{[t+h,t+2h]} N_{[t+h,t+2h]} - Y_{[t,t+h]} N_{[t,t+h]}, \tag{B.2}$$

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

$$f(z_1, z_2, ...z_{K-1})$$

$$= f(z_{K-1} | z_1, z_2, ...z_{K-2}) f(z_1, z_2, ..., z_{K-2})$$

$$= f(z_{K-1} | z_{K-2}) f(z_{K-2} | z_1, z_2, ...z_{K-3}) f(z_1, z_2, ...z_{K-3})$$

$$= f(z_{K-1} | z_{K-2}) f(z_{K-2} | z_{K-3}) ... f(z_2 | z_1) f(z_1), \tag{B.3}$$

Therefore, the log-likelihood function is:

$$\ln f(z_1, z_2, ...z_{K-1})$$

$$= \ln f(z_{K-1} | z_{K-2}) + \ln f(z_{K-2} | z_{K-3}) + ... + \ln f(z_2 | z_1) + \ln f(z_1), \tag{B.4}$$

Next, we will derive $f(z_{t+h} | z_t)$ and $f(z_1)$ respectively.

If $N_{[t,t+h]} = 0$, then $z_{t+h} = uh + \sigma[W_{t+2h} - W_{t+h}] + Y_{[t+h,t+2h]} N_{[t+h,t+2h]})$. (B.5)

$z_{t+h} | (N_{[t,t+h]} = 0, N_{[t+h,t+2h]} = 0)$ will be normally distributed with mean $M_{nn} = uh$, and variance $S_{nn}^2 = \sigma^2 h$. $z_{t+h} | (N_{[t,t+h]} = 0, N_{[t+h,t+2h]} = 1)$ will be normally distributed with mean $M_{ny} = uh + m$, and variance $S_{ny}^2 = \sigma^2 h + s^2$.

If $N_{[t,t+h]} = 1$, we add equations (B.1) and (B.2) together and simplify to get:

$$z_{t+h} = -z_t + 2uh + \sigma[W_{t+2h} - W_t] + Y_{[t+h,t+2h]}N_{[t+h,t+2h]} - Y_{[t-h,t]}N_{[t-k,t]}, \tag{B.6}$$

If no mortality jump event occurs during the period $(t-h,t)$ and $(t+h,t+2h)$, the variable $z_{t+h}\big|(z_t, N_{[t-h,t]} = 0, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 0)$ will be normally distributed with mean $M_{nyn} = -z_t + 2uh$ and variance $S_{nyn}^2 = 2\sigma^2 h$.

Similarly, $z_{t+h}\big|(z_t, N_{[t-h,t]} = 1, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 0)$ will be normally distributed with mean $M_{yyn} = -z_t + 2uh - m$ and variance $S_{yyn}^2 = 2\sigma^2 h + s^2$.

$z_{t+h}\big|(z_t, N_{[t-h,t]} = 0, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1)$ will be normally distributed with mean $M_{nyy} = -z_t + 2uh + m$ and variance $S_{nyy}^2 = 2\sigma^2 h + s^2$.

$z_{t+h}\big|(z_t, N_{[t-h,t]} = 1, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1)$ will be normally distributed with mean $M_{yy} = -z_t + 2uh$ and variance $S_{yy}^2 = 2\sigma^2 h + 2s^2$.

The conditional density function of $z_{t+h}\big|z_t$, denoted by $f(z_{t+h}\big|z_t)$, can be written as:

$$
\begin{aligned}
& f(z_{t+h}\big|z_t) \\
&= f(z_{t+h}\big|N_{[t,t+h]} = 0, N_{[t+h,t+2h]} = 0)\Pr(N_{[t,t+h]} = 0, N_{[t+h,t+2h]} = 0) \\
&+ f(z_{t+h}\big|N_{[t,t+h]} = 0, N_{[t+h,t+2h]} = 1)\Pr(N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1) \\
&+ f(z_{t+h}\big|z_t, N_{[t-h,t]} = 0, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 0)\Pr(N_{[t-h,t]} = 0, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 0) \\
&+ f(z_{t+h}\big|z_t, N_{[t-h,t]} = 1, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 0)\Pr(N_{[t-h,t]} = 1, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 0) \\
&+ f(z_{t+h}\big|z_t, N_{[t-h,t]} = 0, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1)\Pr(N_{[t-h,t]} = 0, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1) \\
&+ f(z_{t+h}\big|z_t, N_{[t-h,t]} = 1, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1)\Pr(N_{[t-h,t]} = 1, N_{[t,t+h]} = 1, N_{[t+h,t+2h]} = 1) \\
&= \frac{1}{\sqrt{2\pi}S_{nn}}\exp\left(-\frac{(z_{t+h} - M_{nn})^2}{2S_{nn}^2}\right)(1-p)^2 + \frac{1}{\sqrt{2\pi}S_{ny}}\exp\left(-\frac{(z_{t+h} - M_{ny})^2}{2S_{ny}^2}\right)p(1-p) \\
&+ \frac{1}{\sqrt{2\pi}S_{nyn}}\exp\left(-\frac{(z_{t+h} - M_{nyn})^2}{2S_{nyn}^2}\right)p(1-p)^2 + \frac{1}{\sqrt{2\pi}S_{yyn}}\exp\left(-\frac{(z_{t+h} - M_{yyn})^2}{2S_{yyn}^2}\right)p^2(1-p)
\end{aligned}
$$

$$+\frac{1}{\sqrt{2\pi}S_{nyy}}\exp\left(-\frac{(z_{t+h}-M_{nyy})^2}{2S_{nyy}^2}\right)p^2(1-p)+\frac{1}{\sqrt{2\pi}S_{yyy}}\exp\left(-\frac{(z_{t+h}-M_{yyy})^2}{2S_{yyy}^2}\right)p^3 \qquad \text{(B.7)}$$

The variable $z_1 \big| (N_{[0,1]}=0, N_{[1,2]}=0)$ will be normally distributed with mean $\hat{M}_{nn}=uh$, and variance $\hat{S}_{nn}^2=\sigma^2 h$.

The variable $z_1 \big| (N_{[0,1]}=1, N_{[1,2]}=0)$ will be normally distributed with mean $\hat{M}_{yn}=uh-m$, and variance $\hat{S}_{yn}^2=\sigma^2 h+s^2$

The variable $z_1 \big| (N_{[0,1]}=0, N_{[1,2]}=1)$ will be normally distributed with mean $\hat{M}_{ny}=uh+m$, and variance $\hat{S}_{ny}^2=\sigma^2 h+s^2$

The variable $z_1 \big| (N_{[0,1]}=1, N_{[1,2]}=1)$ will be normally distributed with mean $\hat{M}_{yy}=uh$, and variance $\hat{S}_{yy}^2=\sigma^2 h+2s^2$

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

$$\begin{aligned}
&f(z_1)\\
&= f(z_1 \big| N_{[0,1]}=0, N_{[1,2]}=0)\Pr(N_{[0,1]}=0, N_{[1,2]}=0)\\
&\quad + f(z_1 \big| N_{[0,1]}=1, N_{[1,2]}=0)\Pr(N_{[0,1]}=1, N_{[1,2]}=0)\\
&\quad + f(z_1 \big| N_{[0,1]}=0, N_{[1,2]}=1)\Pr(N_{[0,1]}=0, N_{[1,2]}=1)\\
&\quad + f(z_1 \big| N_{[0,1]}=1, N_{[1,2]}=1)\Pr(N_{[0,1]}=1, N_{[1,2]}=1)
\end{aligned}$$

$$=\frac{1}{\sqrt{2\pi}\hat{S}_{nn}}\exp\left(-\frac{(z_1-\hat{M}_{nn})^2}{2\hat{S}_{nn}^2}\right)(1-p)^2+\frac{1}{\sqrt{2\pi}\hat{S}_{yn}}\exp\left(-\frac{(z_1-\hat{M}_{yn})^2}{2\hat{S}_{yn}^2}\right)p(1-p)$$

$$+\frac{1}{\sqrt{2\pi}\hat{S}_{ny}}\exp\left(-\frac{(z_1-\hat{M}_{ny})^2}{2\hat{S}_{ny}^2}\right)(1-p)p+\frac{1}{\sqrt{2\pi}\hat{S}_{yy}}\exp\left(-\frac{(z_1-\hat{M}_{yy})^2}{2\hat{S}_{yy}^2}\right)p^2, \qquad \text{(B.8)}$$

Substituting the formulas of $f(z_{t+h} \big| z_t)$ and $f(z_1)$ into the log-likelihood function (B.4), we can calculate the log-likelihood function numerically.