The Dissertation Committee for Shuo-Li Chuang Certifies that this is the approved version of the following dissertation:

The Stochastic Mortality Modeling and the Pricing of Mortality/Longevity Linked Derivatives

Committee:

Patrick L. Brockett, Supervisor

Linda Golden

Richard D. MacMinn

Kumar Muthuraman

Thomas W. Sager

The Stochastic Mortality Modeling and the Pricing of Mortality/Longevity Linked Derivatives

by

Shuo-Li Chuang, B.A.; M.B.A.; M.S.I.R.O.M

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2013

The Stochastic Mortality Modeling and the Pricing of Mortality/Longevity Linked Derivatives

Shuo-Li Chuang, Ph.D
The University of Texas at Austin, 2013

Supervisor: Patrick L. Brockett

The Lee-Carter mortality model provides the very first model for modeling the mortality rate with stochastic time and age mortality dynamics. The model is constructed modeling the mortality rate to incorporate both an age effect and a period effect. The Lee-Carter model provides the fundamental set up currently used in most modern mortality modeling. Various extensions of the Lee-Carter model include either adding an extra term for a cohort effect or imposing a stochastic process for mortality dynamics. Although both of these extensions can provide good estimation results for the mortality rate, applying them for the pricing of the mortality/longevity linked derivatives is not easy. While the current stochastic mortality models are too complicated to be explained and to be implemented, transforming the cohort effect into a stochastic process for the pricing purpose is very difficult. Furthermore, the cohort effect itself sometimes may not be significant.

We propose using a new modified Lee-Carter model with a Normal Inverse Gaussian (NIG) Lévy process along with the Esscher transform for the pricing of mortality/longevity linked derivatives. The modified Lee-Carter model, which applies the Lee-Carter model on the growth rate of mortality rates rather than the level of mortality rates themselves, performs better than the current mortality rate models shown in Mitchell et al (2013). We show that the modified Lee-Carter model also retains a similar stochastic structure to the Lee-Carter model, so it is easy to demonstrate the implication of the model. We proposed the additional NIG Lévy process with Esscher transform assumption that can improve the fit and prediction results by adapting the mortality improvement rate. The resulting mortality rate matches the observed pattern that the mortality rate has been improving due to the advancing development of technology and improvements in the medical care system. The resulting mortality rate is also developed under a martingale measure so it is ready for the direct application of pricing the mortality/longevity linked derivatives, such as q-forward, longevity bond, and mortality catastrophe bond. We also apply our proposed model along with an information theoretic optimization method to construct the pricing procedures for a life settlement. While our proposed model can improve the mortality rate estimation, the application of information theory allows us to incorporate the private health information of a specific policy holder and hence customize the distribution of the death year distribution for the policy holder so as to price the life settlement. The resulting risk premium is close to the practical understanding in the life settlement market.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Life expectancy at an age x for a member of a specified population is defined as the average number of years remaining at age x for the specific population under study. It is a population specific average and deviation from this average at the individual level is affected by quality of health care, environmental influences, diet, wars, epidemics, and natural disasters. Because of improved public health and other factors, life expectancy has risen throughout the developed world. Table 1.1 shows the life expectancy at age 65 for people in twelve developed countries. Life expectancy continues to increase; in 2009 there were 455,000 Americans over 100 years old. This number is expected to increase 5.5% per year thereby doubling the centenarian population every 13 years (United Nations, 2009).

The mortality rate at a given age is the number of people who die at that age divided by the number of people who are alive at the beginning of that age interval. Thus, the mortality rate at age 65 is the number of people who die at last age 65 divided by the number of people who reach age 65. Increased life span has finan-

	Male		Female	
	1980	2006	1980	2006
United Kingdom	12.6	17.3	16.6	20.1
Germany	12.8	17.2	16.3	20.5
Italy	13.3	17.8	17.1	21.6
France	13.6	18.0	18.2	22.3
Norway	14.3	17.7	18.2	20.9
Greece	14.6	17.4	16.8	19.6
Spain	14.6	17.9	17.8	22.0
Ireland	12.6	16.8	15.7	20.2
Sweden	14.3	17.6	17.9	20.8
Switzerland	14.3	18.5	18.2	22.1
United States	14.1	17.4	18.3	20.3
Japan	14.6	18.5	17.7	23.4

Source: Credit Suisse, OECD

Table 1.1: Life expectancy at age 65

cial consequences for the individual, community, and society, and accordingly, with decreasing mortality rates (for all ages) there is a need for improved capital management strategies in order that annuity providers, insurers, pension plan fiduciaries, and individuals are able to handle the additional financial burden. Current mortality forecasting models have consistently underestimated mortality rates resulting in pension and annuity providers taking on increasing risk and liabilities (Biffis and Blake, 2009).

There are many ways to mitigate the financial consequences of mortality/longevity risks. Insurance and reinsurance are traditional approaches, but these markets are not sufficiently large in financial assets to support these risks by themselves (e.g., they lack sufficient capacity and liquidity) hence they are unable to support the gigantic global financial exposure currently estimated to be about 20 trillion dollars (Loeys et al., 2007; Biffis and Blake, 2009). Pension and annuity providers are eager to participate in mitigating their mortality/longevity risks but they need the assistance of the much larger capital markets to be able to do this. The first step in creating a capital market instrument that can be used to mitigate mortality/longevity risks is to develop a mortality model that reflects current mortality rates and has the ability to project future mortality rates accurately. Next, they need to develop a way to use the model to create financial instruments whose pay off is based on actual and projected mortality rates so as to transfer mortality/longevity risks to counter parties in a way that is both priced competitively and accepted within the marketplace.

Modeling mortality began with Graunt (1662) who examined the London Bills of Mortality who showed that the life span of individuals was predictable in the aggregate, and developed a life table to describe this structured process of death. Edmond Halley (1693) had better data and was the first to actually show how to construct a rigorous (essentially modern) mortality table from empirical data and how to price life annuities using this table. De Moivre (1725) first postulated a functional form mortality table and showed how one could do annuity table calculations using this mathematical model. He postulated a uniform distribution of deaths model, and showed this simplified annuity calculation methods. Instead of taking a strictly mathematical formulation as a postulate, Gompertz (1825) took a biological approach and let the mathematical form arise as a consequence of the biological considerations. Specifically, he postulated that the mortality rate at age

x, $m(x)$, reflected the body's inclination to succumb to death at time t and that $w(x) = 1/m(x)$ was the body's vitality or ability to resist death at age x. A further postulate that changes change in the body's vitality at age x is proportional to the vitality (ability it has to withstand death) $w(t)$ that it has to start with led to the differential equation $w(x) = cw(x)$, whose solution is an exponential function. Reciprocating this formula for $w(x)$ to obtain a mathematical model for the mortality rate $m(x)$, yielded a mortality curve known as the Gompertz curve. It fits mortality data quite well from about age 30 to age 102-105 (Gavrilov and Gavrilova, 2011). The above mortality models are static in nature, i.e., they involve the age x of the person at a fixed point (year) in time but ignore the evolution of mortality over time, and ignore random variation in mortality (i.e., the given mortality rates are the expected rates at a frozen point in time). Recent mortality modeling which attempts to incorporate both the age at death as well as the temporal dimension is now mostly based on the two dimensional (age and time) structure supplied by Lee and Carter (1992). The Lee-Carter model assumes two factors, an age effect and a period or time effect, have an impact on mortality modeling. The temporal effect reflects changes in mortality rates across all ages as time and medicine progress, and the age component reflects that mortality increases with increasing age at any point in time.

Although the Lee-Carter model provides a fundamental set-up for mortality modeling, it does not adequately take into account rare events such as war, epidemics or natural disasters. When such an event occurs, it creates a sudden abnormal increase in the mortality rate. This is referred to as a mortality jump. To account for these jumps, extensions to the Lee-Carter model have been proposed[1]. Another extension added to the Lee-Carter model was to incorporate another term which accounts for the cohort effect found by Renshaw and Haberman (2006). This cohort effect extension attempts to reflect the observed phenomenon that in some times and places people born in the period of years (say 1946) will experience a similar mortality rate pattern which differs from those of their immediate predecessors (1945) or those born in the following year (1947).

While the Lee-Carter model and each extension serve to improve mortality forecasting, an unintended consequence is that the mortality models so constructed can become so complex that it is almost impossible to use them when attempting to create financial instruments. To adequately price a mortality/longevity linked financial derivative, a stochastic mortality model is needed along with a martingale measure[2]. The mortality model must be easy to explain and not compromise the accuracy of mortality forecasting. The Lee-Carter extensions without stochastic process specifications cannot be used for martingale pricing. And, as noted earlier, even the extensions with stochastic process specifications tend to be complicated; and it is not intuitive to link them to martingale pricing techniques common in the financial literature. Therefore, a new model is needed that ensures accurate mortality forecasting and at the same time enables martingale pricing.

Lévy processes are used within finance and risk management disciplines but there

[1]See Ballotta and Haberman(2006), Deng et al. (2012), Milevsky and Promislow (2001), Renshaw et al. (1996), and Sithole et al. (2000).

[2]A martingale measure is desired here (in order to serve as a risk-neutral measure) for arbitrage free asset pricing in an incomplete market (such as that governing mortality and longevity derivatives which do not have an underlying tradable underlying security upon which they are based).

is limited literature discussing the use of employing Lévy processes on mortality rate modeling. The major reason is because Lévy processes are stationary stochastic processes while the process of mortality rates is not. To take into account the non-stationary status of mortality rates within the Lee-Carter framework, Hainuat and Devolder (2008) introduce an Ornstein-Uhlenbeck process with tempered α-stable subordinators. They combine this with a Lévy process to account for small and large jumps. To arrive at a martingale price they then use the Esscher transform (Gerber and Shiu, 1994). Although using a Lévy-Ornstein-Uhlenbeck process solves the non-stationary mortality issue and enables martingale pricing, it is a complicated process involving many steps.

By looking at the growth rate of mortality rate at a fixed age x instead of level of mortality as was done by Lee and Carter, Mitchell et al. (2013) (hereafter the modified Lee-Carter model) structurally modified the Lee-Carter model. In so doing, they are able to fit and forecast mortality rates better than the Lee-Carter model and all the various extensions discussed previously.

The growth rate model of log mortality used by the modified Lee-Carter model has similar properties to certain models of asset returns in financial theory. Since modeling asset returns has developed substantially and since modeling the growth rate of mortality behaves similar to asset returns we will first examine utilizing the techniques developed in finance to accurately forecast mortality rates. Then we shall adapt the the modified Lee-Carter model to include a Lévy process that will enable us to obtain an equivalent martingale measure that can be used for pricing and create a fair price for mortality/longevity derivatives. Finally, this model will be used for pricing of mortality/longevity linked securities, such as mortality swaps, q-forward, and mortality bonds.

1.1 Structure of the Dissertation

Chapter 2 introduces Lévy processes and the Esscher transform. We first define a Lévy process and characterize a Lévy process via the corresponding Lévy triplet, on the characteristic function of the stochastic process which delineates the various distributions that correspond to the Lévy process. For our application, we select a Normal Inverse Gaussian (NIG) distribution to generate a Lévy process and then we apply the Esscher transform and show how this can result in arbitrage free (risk neutral or martingale) pricing. We show that the corresponding parameters for the Esscher transformed equivalent martingale stochastic process have a closed form for the corresponding parameter enabling a simplified calculation.

Chapter 3 discusses the Lee-Carter model and its various extensions, including changing the distributional assumptions of error terms, the Lee-Carter model with the cohort effect, and stochastic modeling with application of reduction factor model and Affine process. These extensions are widely applied, and we will describe their general structures. The advantages and disadvantages will also be presented.

Chapter 4 first introduces the modified Lee-Carter model. We propose an additional assumption to the modified Lee-Carter model, which is the the modified Lee-Carter model with an NIG Lévy driving process and transformed using the Esscher transform. The proposed model has a structure similar in many ways to the Lee-Carter model so is easy to use and explain. The application of the Esscher transform enables a closed form martingale pricing formula for the mortality/longevity linked derivatives.

Chapter 5 presents the estimation results from our proposed model. Using mortality data from the United States and the United Kingdom (England and Wales) we also compare our estimation results with the Lee-Carter model and the modified Lee-Carter model.

Chapter 6 is concerned with another financial product that is sensitive to longevity and mortality risk, namely life settlements which are newly created secondary markets in life insurance products. We first introduce the deterministic and probabilistic approaches to life settlement pricing. Then we apply information theory to "customize" the probability distribution of the death year for a policyholder when his/her private information (such as a projected life expectancy) is available. We apply our proposed model for mortality rate prediction and calculate the prices for a life settlement. We also propose an approach to calculate the implied risk premium for a life settlement. An empirical comparison is made with Murphy (2006).

In Chapter 7 we price the mortality/longevity linked derivatives, including q-forward, mortality catastrophe bonds, and the EIB bond. We take into account mortality improvement rates using the modified Lee-Carter model; therefore, our proposed model produces lower mortality rate estimations. Hence, the prices and premiums generated from our model are lower, which may stimulate market interest in hedging the mortality or longevity risks. Chapter 8 summarizes our contributions and findings. Future study and research questions are also discussed.

Chapter 2

Lévy Processes and the Esscher Transform

2.1 Introduction

Stochastic processes have been used in financial modeling for over a century. Bachelier is credited with being the first person to mathematically model the Brownian Motion stochastic process in his 1900 PhD dissertation at the University of Paris (the Sorbonne), five years ahead of Einsteins development for physics. Bacheliers dissertation concerned the stochastic process of speculative prices in the securities market, but this path was not commonly pursued until much later (especially in the late 1960s or early 1970s). Financial markets are complex, however, using stochastic processes, it is possible to create models that accurately represent complicated statistical correlations. In so doing, we are able to make sense of various financial markets and make predictions regarding rates of return and pricing structures. The main characteristics of the stochastic processes used in financial modeling include the drift term, the diffusion term (if any), and a potential jump change process. Merton (1976) first used as a stochastic processes model in financial modeling. Merton suggested that the dynamics of a stock return can be captured by a jump-diffusion model with a mixture of independent Brownian motions and Poisson processes. In financial modeling, a jump-diffusion model adequately describes the dynamics of many asset returns. This model has a drift term, a diffusion term, and a jump term. A drift term describes a long-term trend pattern. A diffusion term is often assumed to have a Brownian motion structure and captures smooth but random fluctuations of the asset return around the long term drift. The jump term captures the potential large changes that may be caused by rare events and may be modeled by a compound Poisson process. Diffusion together with jumps provide the random resultant variations of asset returns. A jump-diffusion model implies that the size of the small changes has a Gaussian distribution, but asset returns are recognized to have non-Gaussian behaviors that exhibit fat tails and excess kurtosis because of compound Poisson jump process. This realization means that another model is needed to account for these non-Gaussian behaviors.

A general Lévy process (a stationary process with independent increments) can be used to account for the non-Gaussian distribution of assets. It can be used to generate a distribution that can have fatter tails than the normal distribution, have excess kurtosis, and exhibit many other non-Gaussian behaviors (Wu, 2012). Lévy processes have been recognized as valuable tools in finance and risk management throughout recent decades. They are valuable because they can be used with data that appear to have components that conform to both Gaussian and non-Gaussian distributions. A Lévy process is a continuous time stochastic process with stationary independent increments. Due to the stationary independent increments assumption, each time-varying random variable at a fixed time in a Lévy process can be expressed as a sum of independent identically distributed random variables. Therefore, both large jumps and small jumps are characterized by additional distributional assumptions. Hence, a Lévy process can combine greater flexibility with analytical tractability. The potential distributions possible for Lévy process are delineated by their characteristic functions (Fourier Transform) to have a representation given by Lévy and Khinchine (Sato, 2000; Schoutens, 2003).

Constructing an exponential Lévy process that is a generalized version of a jump diffusion model enables processing both small jumps and large jumps when proper distributional properties are assumed. While the most studies of jump-diffusion models that require a diffusion term to fill the gaps in between the arrival of jumps modeled by a compound Poisson process, a Lévy process has infinite number of jumps within any finite interval to process all jumps without adding a diffusion term and a jump process (Carr et al., 2002; Carr and Wu, 2003; Wu, 2005). A specified distribution that possesses high kurtosis (frequent small jumps) and fat tails (several infrequent large jumps) can deal with the non-Gaussian nature of asset returns. In this research, a Normal Inverse Gaussian (NIG) distribution will be specified for the stochastic innovation term in the dynamics of the asset returns. An NIG distribution characterizes the distribution of the jump. Since there are many small jumps and few large jumps, the distribution of jump is high-kurtosis and fat-tailed, and the

NIG distribution fits this requirement. An NIG distribution has four parameters to specify location, degree of asymmetry, kurtosis, and skewness, so it provides an excellent fit for the distribution of jumps. Combining the exponential Lévy process with this NIG distribution will produce better estimation results for pricing assets.

Another aspect that must be considered within this research is calculating martingale prices. Martingale pricing is used to quantify asset prices within a risk-free or risk-neutral environment. These risk neutral distributions are particularly useful for determining arbitrage free prices in an incomplete market (such as that exhibited by derived instruments based on mortality rates or other non-traded underlying indices). There are many ways to calculate martingale prices. Here we will use the Esscher transform. We use this method for two reasons. First, it is easier to use than other methods, such as the application of the Radon-Nykodym derivative. Second, the parameters estimated from an NIG Lévy process can be directly input into the Esscher transform to get a closed form solution for the process thus eliminating the need for numerical calculations.

When pricing a financial asset in an incomplete market, it is necessary to convert the physical (or observed empirical) original measure of the asset to an equivalent martingale measure to ensure that the pricing reflects an arbitrage free pricing that allows for calculations in a risk-neutral environment so that discounted present values can be calculated using the risk-free discount rate. The Radon-Nykodym derivative, which determines the likelihood ratio of two measures, is often used to calculate a martingale measure (Konstantopoulos et al., 2011). However, the Radon-Nykodym derivative is difficult to use and sometimes martingale measure cannot be found.

The Esscher transform is a special case of the Radon-Nykodym derivative that Gerber and Shiu discussed and used for asset pricing in 1994. They found, that when using an exponential Lévy process, it is possible to sidestep the calculation of the Radon-Nykodym derivative by applying the Esscher transform. Moreover, a closed form solution for the parameter of the Esscher transform can be found in these cases (Cawston and Vostrikova, 2009; Önalan, 2009).

In the rest of this chapter we introduce the Lévy processes that are commonly applied in finance and risk management. Then we show how the Esscher transform can be applied to these Lévy processes to find equivalent martingale measures. We choose an NIG Lévy process as an example to illustrate how we can find the parameter of the Esscher transform, as we shall utilize this process subsequently.

2.2 Lévy Processes

Lévy processes were named after the French mathematician Paul Lévy (1886-1971) who introduced the processes that eventually become the crucial component in the modern theory of stochastic processes as applied to asset pricing. The definition of a Lévy process is presented below.

Definition 2.1. A stochastic process $(Y_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ having values in $\mathbb{R}^d$ is said to be a Lévy process if it possesses the following properties:

1. Independent increments: For $0 \leq t_1 < t_2 \leq t_3 < t_4$, $(Y_{t_2} - Y_{t_1})$ and $(Y_{t_4} - Y_{t_3})$ are independent.

2. Stationary increments: For any $h > 0$, $(Y_{t+h} - Y_t)$ does not depend on t.

3. Right continuous: For all $\epsilon > 0$ and $h > 0$,

$$\lim_{h \to 0} \mathbb{P}(|Y_{t+h} - Y_t| > \epsilon) = 0.$$

The definition implies the most important property of a Lévy process, namely at any point in time the distribution of Y_t is infinitely divisibility (Protter, 2002; Schoutenes, 2003). This means that Y_t at any fixed time can be expressed as the sum of independent and identically distributed (i.i.d.) random variables. This characteristic is particularly useful when the sequence exhibits large jumps, small jumps, or both. It is not surprising that the Brownian motions that are frequently represented as diffusion processes and the Poisson processes that are often shown as jump processes are both Lévy processes.

Since the distribution of a random variable X can be characterized by its characteristic function obtained from the Fourier transform. The law (or probability distribution) of a Lévy process can thus be determined by the same characteristic function. Consequently, we next define the characteristic function of the Lévy process so we may utilize it subsequently.

Definition 2.2. Let $(Y_t)_{t \geq 0}$ be a Lévy process on $\mathbb{R}^d$, then there exists a continuous function $\psi(u)$ called characteristic exponent of Y_t, such that

$$E(e^{iuY_t}) = e^{t\psi(u)}.$$

In the Definition 2.2, the law (or distribution) of Y_t is determined by the function $\psi(.)$. From Definition 2.2, we observe the distribution of Y_1, and by stationarity the distribution of the entire process is determined. Therefore, to characterize the distribution of a Lévy process $(Y_t)_{t\geq 0}$, we can just specify the $\psi(.)$ for Y_1. The exact specification is determined by the Lévy-Khintchine representation which completely specifies the characteristic function of stationary processes with independent increments.

Theorem 2.3 (Lévy-Khintchine representation)**.** *Let $(Y_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$, then*

$$E(e^{iuY_t}) = e^{t\psi(u)},$$

where

$$\psi(u) = \gamma ui - \frac{1}{2}\sigma_\nu^2 u^2 + \int_{-\infty}^{\infty} \left(\exp(iuy) - 1 - iuy\mathbb{I}_{\{|y|<1\}}\right) \nu(dy).$$

The $(\sigma_\nu^2, \nu, \gamma)$ is called Lévy triplet that uniquely characterizes the Lévy process.

The Lévy-Khintchine representation shows that a Lévy process can possess distributional properties that can be specified through the parameters of the Lévy triplet. Moreover, when the moment generating function of a random variable exists, it can be determined from the characteristic function. The function $\psi(-iu)$ is the logarithm of the moment generating function and is used frequently for calculations involving moments. It is usually called cumulant generating function, and we define it as $\kappa(u)$; that is,

$$\kappa(u) := \psi(-iu) = \gamma u + \frac{1}{2}\sigma_\nu^2 u^2 + \int_{-\infty}^{\infty} \left(\exp(uy) - 1 + uy\mathbb{I}_{\{|y|<1\}}\right) \nu(dy).$$

We then have a proposition for the moment generation function.

Proposition 2.4. Let $(Y_t)_{t\geq 0}$ be a Lévy process with the triplet $(\sigma_\nu^2, \nu, \gamma)$. The moment generating function $M_Y(u) = \mathbb{E}(e^{uY_t})$, $\xi \in \mathbb{R}$, is finite if and only if

$$\int_{|x|\geq 1} e^{uy}\nu(dy) < \infty.$$

Then

$$E(e^{uY_t}) = e^{t\psi(-iu)} = e^{\kappa(u)}$$

where ψ is the characteristic exponent of $(Y_t)_{t\geq 0}$.

The intriguing part of the proposition 2.4 is that it looks similar to an *exponential Lévy process*, also called a *geometric Lévy process*. The definition of an exponential Lévy process is defined below.

Definition 2.5. An exponential Lévy process $(X_t)_{t\geq 0}$ is given as follows

$$X_t = X_0 e^{Y_t},$$

where $(Y_t)_{t\geq 0}$ is a Lévy process.

An exponential Lévy process is often seen in the financial modeling. For example, the asset return dynamics are often assumed to follow an exponential Lévy process

$$S_t = S_0 \exp(Y_t),$$

where S_t is the stock price at time t, and Y_t is the stock price dynamics which is assumed to be a Lévy process. Since the market is incomplete, pricing the asset with physical (or real) measure has risk concerns. In pricing an exponential Lévy process in an risk-free environment, the goal is to make the discounted stock price as a martingale; that is,

$$\tilde{S}_t = S_0 \exp(-r_f t + Y_t) \tag{2.1}$$

into a martingale under some corresponding martingale measure. This martingale measure yields the risk neutral measure capable of evaluating expected discounted present values using the risk-free discount rate. Without the martingale measure, some other discount rate is needed to price. The assumption that the process to be priced follows an exponential Lévy process has some benefits of calculating the risk neutral price for financial derivatives. We will show this in the next section.

2.3 Esscher Transform

The Esscher transform is named in honor of the Swedish actuary Fredrick Esscher who introduced this transformation for a special case in Esscher (1932). The transformation is also known as an exponential tilting in the statistical literature. The Esscher transform has been a popular approach in financial modeling, particularly asset pricing. To achieve the pricing under the risk neutral environment, an equivalent martingale measure must be selected. The minimal entropy martingale which minimizes the entropy difference between the physical measure and the risk neutral measure is an intuitive choice. However, calculation may be complicated. Prause (1999), Esche and Schweizer (2005) and Hubalek and Sgarra (2006) show that for an exponential Lévy process applying Esscher transform process can obtain the martingale having the property of preserving the Lévy structure of the model as well as approximating the minimal entropy martingale measure well.

The definition of the Esscher transform is described as follows.

Definition 2.6. Let Y be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\theta \in \mathbb{R}$. The Esscher transform $\mathbb{P}^\theta$ corresponding to Y and a parameter θ is defined by its likelihood ratio

$$\frac{d\mathbb{P}^\theta}{d\mathbb{P}} = \frac{e^{\theta Y}}{E(e^{\theta Y})},$$

provided that $E(e^{\theta Y})$ exists.

By Proposition 2.4, the Esscher transform of a Lévy process is

$$\frac{d\mathbb{P}^\theta}{d\mathbb{P}} = \frac{e^{Y_t\theta}}{E(e^{Y_t\theta})} = e^{\theta Y_t - t\kappa(\theta)}. \tag{2.2}$$

In financial modeling, a pricing process can be constructed by utilizing the martingale measure obtained from the Radon-Nikodym derivative. This can be viewed as the ratio of two likelihoods and is often applied when change of measures is needed to obtain a risk neutral martingale measure related to the original physical measure. However, the Radon-Nikodym sometimes may be complicated. Utilizing the Esscher transform can simplify the process, because the Esscher transform provides a measure change and is a special case of the Radon-Nikodym derivative by obtained from the ratio of two measures.

Theorem 2.7 (Radon-Nikodym theorem)**.** *Let* λ_1 *and* λ_2 *be two measures on* $(\Omega, \mathcal{F})$ *and* λ_2 *be* $\sigma-$*finite. If* $\lambda_1 \ll \lambda_2$*, then there exists a random variable* Z *on* $(\Omega, \mathcal{F}, \lambda_1)$

such that $Z \geq 0$ almost surely, and

$$\lambda_1(A) = \int_A Z\lambda_2 \tag{2.3}$$

holds for $A \in \mathcal{F}$.

The (2.3) can be re-written as

$$\lambda_1(A) = \int_A d\lambda_1 = \int_A \frac{d\lambda_1}{d\lambda_2} d\lambda_2 = \int_A Z\lambda_2, \tag{2.4}$$

and $Z = d\lambda_1/d\lambda_2$ is called the Radon-Nikodym derivative of λ_1 with respect to λ_2. The Radon-Nikodym can be viewed as the ratio of two likelihoods. Comparing (2.2) and (2.4), we observe that the Esscher transform is a special case of the Radon-Nikodym derivative, wherein assume the ratio has a special form. Therefore, the transformed measure $\mathbb{P}^\theta$ is actually the equivalent measure with respect to the original (physical) measure $\mathbb{P}$. The issue in using the Esscher transform is to select the parameter θ such that the resulting changed measure is a martingale.

To obtain the parameter θ for the corresponding martingale measure of the Esscher transform in the Lévy process case, the following theorem may be used to simplify the calculation.

Theorem 2.8. *Let $(Y_t)_{t\geq 0}$ be a Lévy process, then the Esscher transformed process $\mathbb{P}^\theta$ is also a Lévy process and*

$$\kappa(u) = \kappa(u + \theta) - \kappa(\theta).$$

The theorem 2.8 can also be demonstrated by Bayes's theorem. For all $0 \leq s \leq t \leq T$, we have

$$\begin{aligned} E^{\mathbb{P}_T^{\theta}}(e^{u(Y_t - Y_s)}|\mathcal{F}_s) &= \frac{E^{\mathbb{P}}(e^{u(Y_t-Y_s)}Y_T^{\theta}|\mathcal{F}_s)}{E^{\mathbb{P}}(Y_T^{\theta}|\mathcal{F}_s)} \\ &= \frac{1}{Y_s^{\theta}}E^{\mathbb{P}}[E^{\mathbb{P}}(e^{u(Y_t-Y_s)}Y_T^{\theta}|\mathcal{F}_t)|\mathcal{F}_s] \\ &= \frac{1}{Y_s^{\theta}}E^{\mathbb{P}}(e^{u(Y_t-Y_s)}Y_T^{\theta}|\mathcal{F}_s) \\ &= E^{\mathbb{P}}(e^{(u+\theta)(Y_t-Y_s)-(t-s)\kappa(\theta)}|\mathcal{F}_s) \\ &= e^{(t-s)(\kappa(u+\theta)-\kappa(\theta))}. \end{aligned}$$

Therefore, Theorem 2.8 holds. In most cases, including financial modeling and our case, the u is one. Therefore, we have

$$\kappa(1) = \kappa(1+\theta) - \kappa(\theta).$$

In (2.1), the Esscher transform can make it as a martingale with the corresponding martingale measure, and $\kappa(1)$ is r_f (Cawston and Vostrikova, 2009; Önalan, 2009). Therefore, the θ^* for the Esscher transform martingale measure is the solution to

$$r_f = \kappa(1+\theta^*) - \kappa(\theta^*). \tag{2.5}$$

2.4 NIG Lévy Processes

If a process is an exponential Lévy process, then by Lévy-Khintchine representation, we can characterize this Lévy process by the Lévy triplet. We specifically choose an

NIG distribution for the further modeling.

2.4.1 NIG Distribution

An NIG distribution is a mixture of Normal distribution and inverse Gaussian (IG) distribution. Let a random variable $Z > 0$ be an inverse Gaussian random variable with density function

$$f_{IG}(z; \alpha, \beta) = \frac{\alpha}{\sqrt{2\pi\beta}} z^{-\frac{3}{2}} \exp\left(\frac{(\alpha - \beta z)^2}{2\beta z}\right).$$

To construct an NIG random variable, let a random variable Z be an IG random variable

$$Z \sim IG(\delta\sqrt{\alpha^2 - \beta^2}, \sqrt{\alpha^2 - \beta^2}).$$

The random variable Y has an NIG distribution if its conditional distribution is a normal distribution

$$Y|Z \sim N(\mu + \beta z, z),$$

and the upon unconditional density function of Y is

$$f_{NIG}(y; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp(\delta\sqrt{\alpha^2 - \beta^2} + \beta(y - \mu)) \frac{K_1(\alpha\delta\sqrt{1 + \left(\frac{y-\mu}{\delta}\right)^2}}{\sqrt{1 + \left(\frac{y-\mu}{\delta}\right)^2}}, \qquad (2.6)$$

where $\alpha > 0$, $\beta < |\alpha|$, $\delta > 0$, and $K_1(.)$ is the modified Bessel function of the third kind with order 1 (see Appendix A). To generate an NIG random variable, let X be a random variable drawn from a standard Normal distribution. Then

$$Y = \beta Z + \sqrt{Z} X + \mu$$

where $Y \sim NIG(\alpha, \beta, \mu, \delta)$. There are four parameters in the NIG density function to control shape and location. Figure 2.1 shows how these parameters affect the shape and location of an NIG distribution. The parameter α controls the tail heaviness. A smaller α introduces heavier tail. The parameter β is to control how symmetric the shape can be. When β is zero, the shape is symmetric. The parameter μ is called the location parameter and designates the center of the distribution. The δ is called scale parameter and controls the spread of the distribution. A lager δ results in a high kurtosis shape.

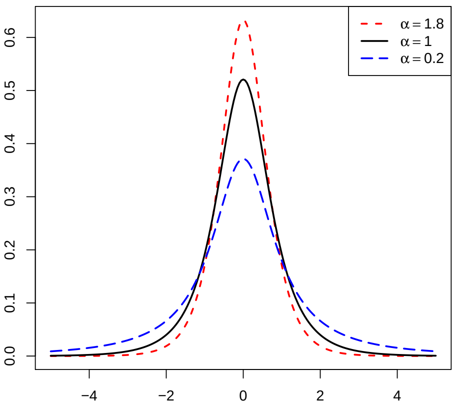
An NIG distribution have nice convolution property and scaling property, making an NIG distribution easy for many analytical applications (Blasild 1981; Kalemanova et al., 2005). .

Proposition 2.9 (Convolution Property)**.** Let Y_1 and Y_2 be NIG random variables with

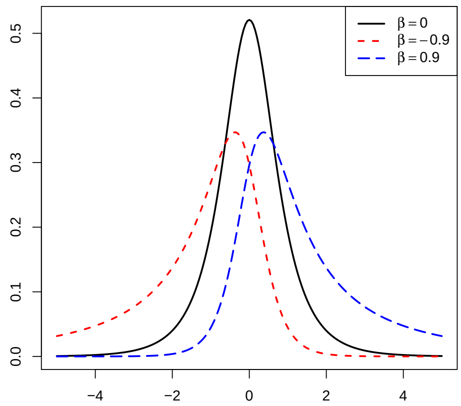
$$Y_1 \sim NIG(\alpha, \beta, \mu_1, \delta_1)$$
$$Y_2 \sim NIG(\alpha, \beta, \mu_2, \delta_2).$$

Then $Z = Y_1 + Y_2$ is also an NIG random variable where

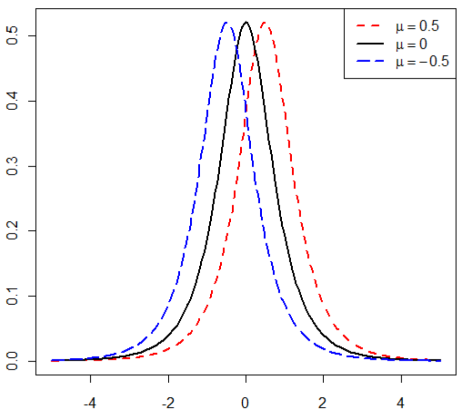
$$Z \sim NIG(\alpha, \beta, \mu_1 + \mu_1, \delta_1 + \delta_2)$$



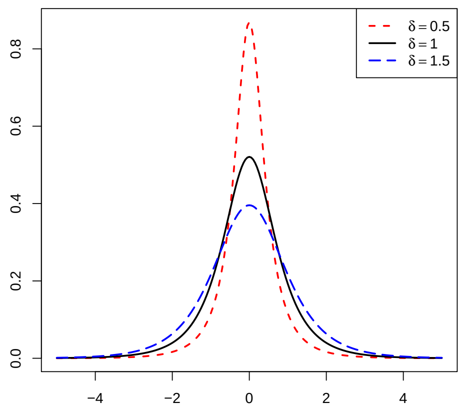
(a) α variable with $(\beta, \mu, \delta) = (0, 1, 0)$



(b) β variable with $(\alpha, \mu, \delta) = (1, 1, 0)$



(c) μ variable with $(\alpha, \beta, \delta) = (1, 1, 0)$



(d) δ variable with $(\alpha, \beta, \mu) = (1, 0, 1)$

Figure 2.1: Densities of an NIG distribution

Proposition 2.10 (Scaling Property)**.** Let Y be an NIG random variable where

$$Y \sim NIG(\alpha, \beta, \mu, \delta).$$

Given a non-negative constant c, then cY is also an NIG random variable where

$$cY \sim NIG\left(\frac{\alpha}{c}, \frac{\beta}{c}, c\mu, c\delta\right).$$

2.4.2 Estimation

There are two ways to estimate the parameters of the NIG. The first approach is the method of moments.

Let $\gamma = \sqrt{\alpha^2 - \beta^2}$, and the central moments of an NIG random variable (Kalemanova et al., 2005) are

$$\begin{aligned} E(y) &= \mu + \frac{\delta\beta}{\gamma} \\ Var(Y) &= \frac{\delta\alpha^2}{\gamma^3} \\ \text{skewness} &= \frac{3\beta}{\alpha\sqrt{\delta\gamma}} \\ \text{kurtosis} &= \frac{3(1 + 4\beta^2/\alpha^2)}{\delta\gamma} \end{aligned} \tag{2.7}$$

The sample moments are

$$
\begin{aligned}
\text{mean} &= \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \\
\text{variance} &= \hat{\sigma}_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\
\text{skewness} &= \hat{S}_Y = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\hat{\sigma}_Y}\right)^3 \\
\text{kurtosis} &= \hat{K}_Y = -3 + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\hat{\sigma}_Y}\right)^4
\end{aligned}
\tag{2.8}
$$

Equate (2.7) and (2.8), and we first can find the estimation for γ shown as follows

$$
\hat{\gamma} = \frac{3}{\hat{\sigma}_Y\sqrt{3\hat{K}_Y - 5\hat{S}_Y^2}},
$$

and then we can find the estimation of all parameters shown as follows

$$
\begin{aligned}
\hat{\delta} &= \frac{9}{\hat{\gamma}(3\hat{K}_Y - 4\hat{S}_Y^2)} \\
\hat{\alpha} &= \hat{\sigma}_Y\hat{\gamma}\sqrt{\frac{\hat{\gamma}}{\hat{\delta}}} \\
\hat{\beta} &= \sqrt{\hat{\alpha}^2 - \hat{\gamma}^2} \\
\hat{\mu} &= \bar{x} - \frac{\hat{\delta}\hat{\beta}}{\hat{\gamma}},
\end{aligned}
$$

provided the condition $3\hat{K}_Y > 5\hat{S}_Y^2$ is satisfied for the existence of the moments.

The second approach is the maximum likelihood method. Prause (1999) shows the log-likelihood function for a generalized hyperbolic distribution. Since the density function of NIG distribution is (2.6) and an NIG distribution is one of the special cases of generalized hyperbolic distributions, we can have a log-likelihood shown as follows

$$
\begin{aligned}
L_{NIG}(y|\alpha, \beta, \mu, \delta) =& \log a(\alpha, \beta, \delta) - \frac{1}{2}\sum_{i=1}^{n} \log\left(\delta^2 + (x_i - \mu)^2\right) \\
&+ \sum_{i=1}^{n}\left(\log K_{-1}\left(\alpha\sqrt{\delta^2 + (x_i-\mu)^2}\right) + \beta(x_i - \mu)\right)
\end{aligned}
$$

where

$$
a(\alpha, \beta, \delta) = \frac{\alpha\delta}{(2\pi)^{1/2}(\alpha^2 - \beta^2)^{1/4} K_{-1/2}\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}.
$$

The partial derivatives of the log-likelihood are shown as follows

$$
\begin{aligned}
\frac{d}{d\alpha}L_{NIG} =& n\frac{\delta\alpha}{\sqrt{\alpha^2 - \beta^2}} \\
&- \sum_{i=1}^{n}\sqrt{\delta^2 + (x_i-\mu)^2}R_{-1}(a(\alpha,\beta,\delta)\sqrt{\delta^2 + (x_i-\mu)^2}) \\
\frac{d}{d\beta}L_{NIG} =& n\left(-\frac{\alpha\beta}{\sqrt{\alpha^2-\beta^2}} - \mu\right) + \sum_{i=1}^{n} x_i \\
\frac{d}{d\delta}L_{NIG} =& n\left(\frac{1}{\delta} + \sqrt{\alpha^2-\beta^2}\right) \\
&+ \sum_{i=1}^{n}\left(\frac{-2\delta}{\delta^2 + (x_i-\mu)^2} - \frac{\alpha\delta}{\sqrt{\delta^2 + (x_i-\mu)^2}}\right) \\
\frac{d}{d\mu}L_{NIG} =& n\beta\sum_{i=1}^{n}\frac{x_i - \mu}{\sqrt{\delta^2 + (x_i-\mu)^2}} \\
&\times\left(\frac{2}{\sqrt{\delta^2 + (x_i-\mu)^2}} + \alpha R_{-1}\left(\alpha\sqrt{\delta^2 + (x_i-\mu)^2}\right)\right)
\end{aligned}
$$

where

$$R_{-1} = \frac{1}{\pi} \int_0^\infty \frac{xdy}{y\left(J_1(\sqrt(y)) + Y_1(\sqrt{y})\right)(y + x^2)},$$

$J_1(.)$ and $Y_1(.)$ are Bessel function of the first order with degree 1 and Bessel function of the second order with degree 1 (See Appendix B). We can furthermore derive direct solutions for μ and β shown as follows

$$\begin{aligned} \hat{\mu} =& -\frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} + \frac{1}{n}\sum_{i=1}^{n} x_i \\ \hat{\beta} =& \frac{1}{n}\sum_{i=1}^{n} \frac{x_i - \mu}{\sqrt{\delta^2 + (x_i - \mu)^2}} \\ & \times \left(\frac{2}{\sqrt{\delta^2 + (x_i - mu)^2}} + \alpha R_{-1}\left(\alpha\sqrt{\delta^2 + (x_i - \mu)^2}\right)\right) \end{aligned}$$

These two approaches sometimes yield different results. The method of moment is calculated from the sample moments, and it may have better visual approximation of the shape. The method of moment also has the advantage of availability of the close forms, although the solutions do not always exist. The maximum likelihood method produces the estimation that is more likely to happen. It has useful asymptotic properties for statistical inference. In practice, one can calculate the estimations from the method of the moment first, and then use them as the starting points for the itereatively solving for maximum likelihood estimates.

2.5 NIG Lévy Process

Let $(Y_t)_{t\geq}0$ be an NIG Lévy process. By the Lévy-Khintchine representation, we can use the Lévy triplet (A, ν, γ) to characterize the process as follows

$$\begin{aligned}
\sigma_\nu^2 &= 0 \\
\nu(dy) &= \frac{\alpha\delta}{\pi}\frac{\exp(\beta y) K_1(\alpha|y|)}{|y|}dy \\
\gamma &= \frac{2\alpha\delta}{\pi}\int_0^1 \sinh(\beta y) K_1(\alpha y)\, dy.
\end{aligned}$$

We would like to determine the parameter of the Esscher transformation needed to obtain the martingale equivalent measure of the exponential NIG Lévy process, i.e., to apply (2.5) when we have (2.1). The moment generation function of an NIG random variable Y is

$$M_Y(\theta) = \exp(\mu\theta + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta+\theta)^2})) = \exp(\kappa(\theta)), \tag{2.9}$$

We can find the function $\kappa(.)$ from (2.9), and plug it in (2.5). The solution θ^* can be formulated as follows

$$\theta^* = -\beta - \frac{1}{2} + \sqrt{\frac{(\mu - r_f)^2}{\delta^2 + (\mu - r_f)^2}\alpha^2 - \frac{(\mu - r_f)^2}{4\delta^2}}. \tag{2.10}$$

The transformed measure with respect to θ^* is $NIG(\alpha, \beta + \theta^*, \mu, \delta)$.

Chapter 3

Lee-Carter Model

3.1 Introduction

Between 1900 and 2012 life expectancy in the developed countries has increased approximately 40 years. Table 3.1 shows the life expectancy at birth for populations in twelve countries between 1900 and 2012. Technology has enabled vast improvements in diverse industries such as medical care, public works, and nutrition thereby increasing the overall survival rate of children and senior citizens. Consequently, the mortality rate continues to decline.

Before the 1990s, forecasting mortality rates involved subjective judgments and opinions of experts often leading to under-estimating the rate of mortality decline (Lee and Miller, 2000). Expert opinions were based on demographic features including medical, behavioral, and social influences. These demographic features created a static sequence mortality model that did not account for random acts such as world wars or epidemics. The inability of this forecasting to account for these real world

	Male		Female	
	1900	2012	1900	2012
United Kingdom	46.4	78.1	50.1	82.4
Germany	43.8	77.9	46.6	82.6
Italy	42.9	79.2	43.2	84.6
France	45.3	78.4	48.7	84.3
Norway	52.3	77.7	55.8	83.1
Greece	38.1	77.5	39.7	82.8
Spain	33.9	78.3	35.7	84.5
Ireland	47.8	78.1	49.3	82.7
Sweden	52.8	78.9	55.3	83.6
Switzerland	45.7	78.3	48.5	84.2
United States	48.3	76.1	51.1	81.1
Japan	42.8	80.6	44.3	87.4

Source: CIA World Factbook and Kinsella (1992).

Table 3.1: Life expectancy at birth

occurrences means that this model does not truly reflect our society and therefore cannot be used to accurately describe mortality rates.

The Lee and Carter (1992) removes the subjective aspects of previous static forecasting methods by introducing a period effect and age effect. This means that mortality rates are shown to change with time as well as age group. Many institutions, including the U.S. Census Bureau, the United Nations Population Division, and governments and actuaries in Canada, Mexico and some European countries employ the model when forecasting (Sullivan, 2001; Booth et al., 2002).

Lee and Carter discovered a pattern and created an index that shows a decrease in mortality rates over time. The problem with this forecasted index of mortality rates is that it assumes a Gaussian (normal) distribution. However, the change of mortality rate is actually not a well-behaved Gaussian distribution because the occurrence of a disaster or an epidemic can cause a dramatic increase in mortality rates. These non-Gaussian aspects along with some other effects related to mortality rates require extensions to the Lee-Carter model. In the following paragraphs we will describe these extensions.

The first extension includes non-Guassian distributions such as Normal Inverse Gaussian (NIG) distribution, Student's t-distribution, and other generalized hyperbolic distributions to account for a change in mortality rate due to a disaster or epidemic. The error terms in the Lee-Carter model thus are assumed to be the aforementioned distributions.

The second extension adds a reduction factor that can be a number or a function to exhibit the downward trend of mortality rates. Using this extension, the future mortality rate of an age or an age group is modeled using the current mortality rate with a reduction factor since mortality rates are decreasing due to technological improvements. The reduction factor is modeled as an exponential function with a stochastic jump-diffusion process that has a drift term as a general pattern and a diffusion term to exhibit the randomness. Financial research has successfully used this stochastic process to model the non-Gaussian distributed stock return time series. Since the mortality rate also displays similar non-Gaussian behavior, applying a stochastic process to mortality rate modeling is analogous to applying the stochastic process in financial modeling.

As stated earlier, mortality rates have been decreasing since the turn of the last century and these rates seem to trend similarly for similar populations. Yet, in the United Kingdom, research published by the Continuous Mortality Investigation (CMI) Bureau shows that people born in the United Kingdom between 1925 and 1945 (centered on the generation born in 1931) have experienced more rapid improvement in mortality than generations born on either side of this period. Taking into account this cohort effect, Renshaw and Haberman (2006) introduce the second extension to the Lee-Carter model. In this model, an extra term is applied to capture the observation that people within the same birth cohort should have similar mortality experiences. Although there are several other models with the cohort effect, Renshaw and Habermans (2006) is the most popular one and is widely applied in actuarial science.

Finally, there are other models that use age effect and period effect for forecasting mortality rates. Instead of modeling the mortality rate directly, these models apply the Lee-Carter model on mortality odds that is the ratio of the mortality rate and the survival rate. The advantage to modeling the mortality odds is that Generalized Linear Models (GLMs) are applied. GLMs provide simple linear form, are easy to estimate, and many distributional assumptions are available. Cairns et al. (2006) and Cairns (2007) discuss how to implement a GLM on the mortality odds. Although models for mortality odds are also popular for actuarial modeling, we will not discuss this type of model since the response variable is different from the Lee-Carter model.

In the following pages we will introduce the original Lee-Carter model and the extensions. We will show how to estimate the model parameters by the singular value decomposition (SVD) method. A literature review discussing non-Gaussian modifications to the original Lee-Carter model is included. Then we show how to apply a stochastic process on the Lee-Carter model and finally we show how to apply the cohort effect on the Lee-Carter model.

3.2 Lee-Carter Model

The Lee-Carter model proposed by Lee and Carter (1992) is described as follows

$$m_{x,t} = \exp(a_x + b_x k_t + \epsilon_{x,t}) \tag{3.1}$$

where $m_{x,t}$ is the mortality rate for the age group x in year t, k_t is a time-varying mortality index, a_x describes the general shape of mortality rate, b_x is the response of each group to the mortality rate, and $\epsilon_{x,t}$ is the error term for the age group x representing the information that is not captured by the model. The error term $\epsilon_{x,t}$ is assumed to have a Gaussian distribution with mean zero and variance σ_x^2 that is the variance of the log mortality rate within the age group x. The (3.1) can be re-written as

$$\log(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t}, \tag{3.2}$$

which is interpreted as that the logarithm of the mortality rate is linear in the mortality index with age-specific components a_x and b_x.

The a_x is mostly negative, corresponding to the fact that the pattern of decline of mortality rate. The intriguing design of this model involves both b_x and k_t components with both age effect and period effect. The b_x describes how the logarithm of the mortality rate changes when k_t changes. If b_x is large, the mortality rate varies a lot when general pattern k_t changes; if b_x is small, the mortality rate varies little when k_t changes. Using the influenza epidemic of 1918 as an example, we can observe the change of value of b_x. The disease raised death rates for people between the ages of 25 to 34 years old. Thus, the estimated b_x of age group 25 to 34 years old people is higher than those b_x's for older age groups. The model seems to imply that the all age-specific mortality rates move with the same direction, yet the b_x is not necessary to have the same sign for all age groups. A negative b_x means that the mortality rate at the age group tends to rise while other age group show decreasing patterns, and vice versa.

3.2.1 Estimation for Lee-Carter Model

From (3.2), besides for a_x and b_x, the k_t is also unknown. Therefore, there is no regressor and thus the regression techniques cannot be used here. However, the SVD technique is applicable for the Lee-Carter model. The SVD technique is the special case of the principal component method that is to express the data in such a way as to highlight their similarities and differences of the data. That is, we apply the SVD technique to extract the general pattern of mortality rates across all age groups and the specific factor for the mortality rate of each age group.

The a_x can be viewed as an average response over the sample period, so the estimation for a_x is shown as follows

$$\hat{a}_x = \frac{1}{T}\sum_{t=1}^{T} \log(m_{x,t}).$$

Then subtract $\hat{a}_x$ from the original logarithm of the mortality rate, and denote it as

$Y_{x,t}$ shown as follows

$$Y_{x,t} = \log(m_{x,t}) - \hat{a}_x.$$

Apply the SVD technique on $Y_{x,t}$ to find b_x and k_t. Assume that there are n age groups. The SVD technique is shown as follows

$$SVD(Y_{x,t}) = \sum_{i=1}^{n} \rho_i U_{x,i} V_{t,i}, \tag{3.3}$$

where $U_{n\times n}$ is a unitary matrix, $V_{t\times t}$ is a unitary matrix, and ρ_i is the singular value of $Y_{x,t}$. From (3.3), b_x and k_t can be estimated by

$$SVD(Y_{x,t}) = \sum_{i=1}^{n} \rho_i U_{x,i} V_{t,i} = \sum_{i=1}^{n} b_x^i k_t^i,$$

where b_x^i is the i^{th} element of b_x, and k_t^i is the i^{th} element of k_t. The b_x and k_t can be arbitrary, since one of these two elements can be multiplied by a constant while the other is divided by the constant without changing the estimated values by the model. Lee and Carter (1992) suggest two constraints shown as follows to specify the estimation

$$\sum_x b_x = 1$$
$$\sum_t k_t = 0.$$

Therefore, the estimated results are $\hat{b}_x = U_{x,1}$ and $k_t = \rho_1 V_{t,1}$.

3.2.2 Specification for the Error Term

The mortality rate is declining year by year, so k_t is expected to trend downward. Therefore, k_t is a non-stationary time series and it cannot be applied on any forecasting models directly. To forecast the mortality rate, the k_t is needed to be de-trended. Moreover, the k_t may show some jumps, such as the phenomenon that the mortality rate of the United States shows jumps during the influenza epidemic of 1918. Lee and Carter (1992) assume that the k_t has a lag term with Gaussian innovations; that is,

$$k_t = k_{t-1} + c + e_t \tag{3.4}$$

where c is a constant, and e_t is an error term to exhibit the uncertainty in k_t. The error term has a Gaussian distribution with mean zero and variance t. From (3.4), we can observe that the change of the mortality index is assumed to be Gaussian. In response to the event that can have significant impacts on k_t, such as the influenza epidemic of 1918, Lee and Carter (1992) suggest replace c with a dummy variable, so that the impact of the event can be disclosed more obviously.

3.3 Extensions of the Lee-Carter Model

3.3.1 Application of Non-Gaussian Error Terms

Lee and Carter (1992) assume a structure with Gaussian innovations for mortality index meaning that the distribution of the difference of mortality index is Gaussian and the distribution of error terms are Gaussian. In fact, the difference of mortality index is high-kurtosis, fat tail, and asymmetric, due to some extreme events such as wars and epidemics. Therefore, some researchers propose to modify the distributional assumption to incorporate the non-Gaussian innovations and mortality index change.

In financial modeling, non-Gaussian innovations have been researched for decades. Bølv and Benth (2000), Eberlein and Keller (1995), Lillestøl (2000), Prause(1997), and Rydberg (1997) suggest that a normal inverse Gaussian (NIG) involving a Lévy process can provide a good fit for the distribution with high kurtosis and fat tails. Prause (1999), Barndorff-Nielson and Shephard (2011), and Mencia and Sentana (2004) show that Student's t-distribution and its skew extensions, such as the generalized hyperbolic skew Student's t-distribution can fit the distribution with asymmetry well.

Wang et al. (2011) apply these non-Gaussian distributions, including NIG, Student's t-distribution, and generalized hyperbolic skew Student's t-distribution, to the error terms (both ϵ_t and e_t) in the Lee-Carter model. Their research results are based on the data from six countries, including the United States, France, and several other countries. They first apply normality test proposed by Jarque and Berra (1980) on the error terms of the Lee-Carter model. They find that applying the assumption of Gaussian distributions on the error terms is not appropriate, because error terms have a non-Gaussian shape. They then apply non-Gaussian distributions on the error terms and find that the prediction results are much better than those of the original Lee-Carter model under different criteria, including Akaike information criterion (AIC), Bayesian information criterion (BIC), and other criteria.

It seems reasonable to assume non-Gaussian error terms. However, this assumption means that the abnormal jumps are from error terms or simply noises. It does not solve the issue that the mortality rate time series itself has abnormal jumps. We should looking for the technique that can model the mortality rate (mortality index) itself as a process with abnormal jumps.

3.3.2 Application of Stochastic Processes on the Lee-Carter Model

The innovation of the Lee-Carter model is motivated by the decreasing pattern of the mortality rate over time. The time-varying component in the Lee-Carter model, k_t is estimated based on past information for the mortality rate shown in (3.4). In other words, from (3.1) and (3.4), we can derive the equation shown as follows

$$m_{x,t} = m_{x,t-1} \exp\left(b_x c + b_x \varepsilon_t + e_{x,t}\right),$$

For forecasting, we can further derive the equation shown as follows

$$m_{x,t} = m_{x,0} \exp(a_x t + b_x \sum_{i=1}^{t} k_1 + \sum_{i=1}^{t} \epsilon_1),$$

where $m_{x,0}$ is the current mortality rate or the initial mortality rate for age group x. Therefore, when forecasting the mortality rate, it is intuitive to formulate the mortality model as the current mortality rate times some factor that becomes smaller over time, which is called the reduction factor; that is, the mortality rate model can have the general form shown as follows

$$m_{x,t} = m_{x,0} RF_{x,t}, \tag{3.5}$$

where $RF_{x,t}$ is the reduction factor. The models with the form shown in (3.5) are categorized as the reduction factor method.

The Lee-Carter model allows introduction of the reduction factor method, and enables observation that the variation of the mortality dynamics is captured by the reduction factor. The reduction factor method is thus an extension of the Lee-Carter model. However, the Lee-Carter model uses the central mortality rate for modeling purposes. The reduction factor model although derived from the Lee-Carter model, for the most part, does not use the central mortality rate for modeling. Therefore, $l_{x,t}$ is the number of people who are age x and exposed to risk at time t and $d_{x,t}$ is the number of deaths at time t for people age x. The central mortality rate is calculated as follows

$$m_{x,t} = \frac{d_{x,t}}{\frac{l_{x,t}+l_{x+1,t}}{2}}.$$

Originally the reduction factor method was used to compute annuity pricing involving life contingencies. However, central mortality rate data are published yearly by governments and thus are insufficient for computing purposes. More elaborate mortality measures must be used that have a continuous estimation. Two measures other than the central mortality rate are often used: the initial mortality rate and the force of mortality rate. The initial mortality rate can be adjusted directly by population estimation, which provides the flexibility of incorporating population demographics. The initial morality rate is defined as

$$q_{x,t} = \frac{d_{x,t}}{l_{x,t}}. \tag{3.6}$$

The $q_{x,t}$ is the probability that individual dies exactly at age x, meaning that an individual lives x complete years. In reality, people are more likely to die in the year of age x and $x+1$, so we need another mortality rate to describe the death rate in the year of age x and $x+1$.

The force of mortality is the instantaneous mortality rate, meaning the probability that an individual age x dies in the next instant of time. Denote ${}_hq_x$ as the initial mortality rate that a person age x dies before reaching $x+h$ years. The force of mortality can be approximated by taking the limit as h approaches zero. Denote μ_x as the force of mortality rate, the μ_x is shown as follows

$$\mu_x = \lim_{h \to 0} \frac{{}_hq_{x,t}}{h} = -\frac{\frac{d}{dh} l_{x,t}}{l_{x,t}}.$$

We can observe that the force of mortality rate is the relative rate of decline in this group at age x. If we assume a constant force of mortality, the force of mortality rate can be estimated as the central mortality rate (Bravo, 2010; Bravo et al., 2010). The relationship between these three mortality measures are shown as follows

$$m_{x,t} \simeq -\log(1 - q_{x,t}) \simeq \mu_{x+1/2}. \tag{3.7}$$

The application of the reduction method on the initial mortality rate is adopted by the United Kingdom CMI Bureau and the United States Society of Actuaries as an extrapolative approach. The United Kingdom CMI Bureau published the reduction factor in 1999 shown as follows

$$RF_{x,t} = \alpha_x + (1 - \alpha_x)(1 - f_x)^{t/20}$$

where

$$\alpha_x = \begin{cases} 0.13, & x < 60 \\ 1 + 0.87\frac{x-110}{50}, & 60 \leq x \leq 110 \end{cases}$$

and

$$f_x = \begin{cases} 0.55 & x < 60 \\ \frac{0.55(110-x)+0.29(x-60)}{50}, & 60 \leq x \leq 110. \end{cases}$$

The United States Society of Actuaries recommends to use

$$q_{x,t} = q_{x,0}(1 - AA_x)^t$$

where AA_x is the age-dependent mortality improvement factor (Scale AA_x) and the time 0 is the year 1994. The reduction factors are different due to the different evolution pattern of the mortality rates in different countries.

Currently, the reduction method is built on the force of mortality rate and models the reduction factor as a stochastic process proposed by Ballotta and Haberman (2006), Renshaw et al. (1996), Sithole et al. (2000), and Milevsky and Promislow (2001). The reduction factor is shown as follows

$$RF(x,t) = \exp((a + b_x)t) + \sigma Y_t) \tag{3.8}$$

where a represents the general change of the reduction factor, b_x is the age-specific change of the reduction factor, $(Y_t)_{t\geq 0}$ is a stochastic process to describe the randomness of mortality rate with corresponding variation σ, and change of $(Y_t)_{t\geq 0}$ is usually modeled as an Ornstein-Uhlenbeck process to incorporate the fact that the change of the mortality rate is mean-reverse; that is, the $(Y_t)_{t\geq 0}$ may be modeled as

$$dY_t = -aY_t dt + dB_t,$$

where the first part is the mean-reverse drift term and the B_t is a standard Brownian motion. Milevsky and Promislow (2001) simplify (3.8) by modeling force of mortality rate for fixed cohort and the model is called Brownian Gompertz model.

Applying the Affine process on the reduction factor method regarding the force of mortality rate is the latest modeling development. Denote the ${}_hp_x$ as the survival probability of an individual age x and survives to age $x+h$. The (3.7) can be written as follows

$$\mu_x = -\frac{\frac{d}{dh}\,{}_hp_x}{{}_hp_x}$$

and we can further derive the survival probability as follows

$${}_hp_x = \exp\left(-\int_x^{x+h} \mu_{x,s} ds\right).$$

The conditional (spot) survival probability for the time h given the end of time T and the current age x can be assume to have the form shown as follows

$${}_{T-h}p_{x+h} = \mathbb{E}\left[\exp\left(-\int_h^n \mu_{x,s} ds\right) |\mathcal{F}_h\right] = \exp(A_x(h,T) + B_x(h,T)W_h) \quad (3.9)$$

where $\mathcal{F}_h$ is the information set up to time h and W_h is a stochastic process. The (3.9) is the special case of the Affine process.

Dahl (2004) first proposes an Affine mortality structure shows as follows

$$_{T-h}p_{x+h} = \exp(A_x(h,T) + B_x(h,T)\mu_{x+h}) \tag{3.10}$$

and the dynamics of the force of mortality is specified by the equations shown as follows

$$\begin{aligned} d\mu_{x+h} &= \alpha^{\mu}_{x,h}dt + \sigma^{\mu}_{x,h}dB^{\mu}_{h} \\ \alpha^{\mu}_{x,h} &= \delta^{\alpha}_{x,h}\mu_{x,h} + \xi^{\alpha}_{x,h} \\ \sigma^{\mu}_{x,h} &= \sqrt{\delta^{\sigma}_{x,h}\mu_{x,h} + \delta^{\sigma}_{x,h}}, \end{aligned}$$

where many parameters in the stochastic components are needed to be estimated. The detailed descriptions are shown in Dahl (2004). Dahl and Møller (2006) propose this structure on the reduction factor directly. They also apply the dynamic model for the short rate on the force of mortality rate. Bliff (2005) applies a two dimensional Affine process on the force of mortality rate. The first component is the random intensity of force of mortality and the second component describes the dynamics of the stochastic drift.

The Affine mortality structure has been applied in the pricing of the mortality/longevity linked derivatives. Although the Affine mortality structure provides details for the continuous mortality rate movement, it is complicated and computation-intensive. Moreover, it may make the dynamics of mortality movement too complicated to be explained, especially when the martingale pricing technique is required for pricing related derivatives.

3.3.3 The Lee-Carter Model and the Cohort Effect

The cohort effect is generally referred as the birth cohort effect, meaning that people born in the same year or same period of years should have similar mortality improvement experience. Renshaw and Haberman (2006) describe a cohort as follows

$$\text{cohort} = \text{period} - \text{age}.$$

For example, the cohort 1962 means people who are 50 years old in 2012 were born in 1962. Considerable research related to the cohort effect has been conducted in the United Kingdom. Those born between 1925 and 1945 show a negative cohort effect; that is, a higher mortality improvement rate than those born after 1945 (Willets, 2004; Willet et al., 2004; Renshaw and Haberman, 2006; Booth and Tickle, 2010). Table 3.2 and 3.3 show the average annual rate of mortality improvement for England and Wales, stratified by age group, mortality decade, and sex. The mortality improvement rate of the cohort of males who were born between 1925 and 1935 is higher than the other cohorts . Table 3.4 shows the average mortality improvement by birth cohort and gender for England and Wales between 1961 and 2001. Notice that the 1925 to 1944 birth cohort has a higher mortality improvement rate.

Renshaw and Haberman (2006) propose an age-period-cohort model (APC)

Age Group	1960's	1970's	1980's	1990's
25-29	**1.5**	0.1	0.4	-0.9
30-34	**1.7**	1.4	-0.6	-0.8
35-39	**1.7**	1.0	0.3	0.9
40-44	0.1	**2.1**	2.1	0.5
45-49	-0.2	**1.8**	**2.3**	1.3
50-54	0.2	0.9	**3.1**	2.3
55-59	1.0	0.9	**3.1**	**2.4**
60-64	1.0	1.0	2.0	**3.2**
65-69	0.1	1.4	1.6	**3.1**
70-74	0.1	1.2	1.7	2.2

Source: Willet (2003) and Gustafsson (2011).

Table 3.2: The average annual rate of mortality improvement (in percentage) for the England and Wales male population, stratified by age group and mortality decade

Age Group	1960's	1970's	1980's	1990's
25-29	**1.6**	0.7	2.6	0.4
30-34	**2.5**	1.2	0.8	0.6
35-39	**1.7**	**1.5**	1.4	0.7
40-44	0.5	**2.0**	1.7	0.3
45-49	0.4	**1.8**	**2.2**	1.2
50-54	0.0	0.6	**2.7**	1.4
55-59	0.3	0.3	**2.0**	**2.0**
60-64	1.1	0.3	0.9	**2.7**
65-69	0.2	0.9	0.6	**2.4**
70-74	0.4	1.3	1.0	1.2

Source: Willet (2003) and Gustafsson (2011).

Table 3.3: The average annual rate of mortality improvement (in percentage) for the England and Wales female population, stratified by age group and mortality decade

Birth cohort	Male	Female
1900-1924	1.2	0.8
1925-1944	2.2	2.0
1945-1959	0.4	1.0

Source: Willet (2003) .

Table 3.4: Average mortality improvement by birth cohort and gender for the population of England and Wales over the period 1961-2001.

known as M2 model shown as follows

$$m_{x,t} = \exp(a_x + b_x k_t + \beta_x \gamma_{t-x} + \epsilon_{x,t}),$$

where γ_{t-x} is the cohort effect. The APC model successfully capture the cohort effect additional to a period effect and an age effect in the mortality rate of the United Kingdom. The APC model shows significant improvement over the Lee-Carter model. Currie (2006) introduces the simpler version of Renshaw and Haberman (2006) model known as M3 model shown as follows

$$m_{x,t} = \exp(a_x + \frac{1}{n_a} k_t + \frac{1}{n_a} \gamma_{t-x} + \epsilon_{x,t}),$$

where n_a is the number of ages (or age groups) in the data.

The cohort effect generally has a smaller impact than the period effect, and sometimes it is not even significant (Booth and Tickle, 2010). The cohort effect may be included in the period effect and the age effect, so sometimes the cohort effect may not be shown. Besides, the cohort effect is different from country to country and long-term data is needed for further estimation. It is also difficult to convert the model with cohort effect to the model with martingale pricing technique.

3.4 Summary

The Lee-Carter model provides the fundamental framework for contemporary mortality modeling meaning it includes a period effect and an age effect. Originally, Gaussian distributions were assumed for error terms and the change of mortality rate. The first extension of the Lee-Carter model is to relax the assumption of the error terms to have non-Gaussian distributions. Although changing the assumption of distributions can provide a better fit, it does not solve the issue that the change of mortality rate shows non-Gaussian behaviors, such as fat tails and high kurtosis.

The Lee-Carter model was originally applied on the central mortality rate and can be converted into a reduction factor method. The reduction factor method provides a general pattern for the mortality improvement. Because the purpose of mortality modeling is primarily used for the calculation of products involving life contingencies, the initial mortality rate and the force of mortality are used in the reduction factor method. Now that the reduction factor method uses the Affine process it has become very complicated. Although the reduction factor method was originally designed to capture mortality dynamics for pricing purposes, it may be difficult to explain what those parameters mean in the reduction factor method, especially when the martingale pricing technique is involved.

Although the Lee-Carter model has an age effect and a period effect, some observed mortality dynamics are still not captured. The cohort effect means that people born in the same year or during the same period of years should have a similar mortality experience. The age-period-cohort model adds an additional cohort effect to the Lee-Carter model. The additional cohort effect in the Lee-Carter model may improve the fit and provides a practical explanation of the mortality dynamics, but the cohort effect may be too insignificant to be noticed due to the fact that some part of the cohort effect is already included in the age effect and period effect. It is also difficult to convert the model with cohort effect to the model with martingale pricing technique.

Chapter 4

Modified Lee-Cater Model with a Lévy Process

4.1 Introduction

The purpose of mortality modeling is to accurately estimate mortality dynamics. If an accurate model can be created, it is assumed that the results can be used to calculate mortality/longevity derivatives and annuities. As stated previously, the Lee-Carter model takes into account that mortality rates change with time, as well as age group thus making it more reliable than subjective, expert opinions. Also, the Lee-Carter model is simple and elegant making it easy to use and easy to interpret its results. Adding the cohort effect to the model and renaming it the age-period-cohort (APC) model takes into account that people born in the same year or range of years exhibit similar mortality rates thus providing a richer explanation of mortality evolution. And other extensions to the model have taken into account the impact events such as war, famine and disease have on mortality rates by applying stochastic processes. However, these models do not work well or are not easy to interpret/implement when trying to calculate derivatives and annuities.

All mortality forecasting models have drawbacks and limitations. The most extensions of the Lee-Carter model are related to stochastic mortality modeling. As mentioned previously, they require data manipulation and intensive computation. The complexity of the models are to improve the performance of the forecasting, so the settings of the models are not intuitive. There are many parameter needed to be estimated and those variables are difficult to explain. Moreover, it is even more complicated to involve the martingale pricing technique for pricing the related derivatives.

The APC model may perform better than the Lee-Carter model for both in-sample estimation and out-of-sample forecasting, although the additional cohort effect may be very small or insignificant (Booth and Tickle, 2010). Moreover, although the APC model has a regression form, similar to the Lee-Carter model, it is not a regression model. Some statistical techniques are needed to extract variables from the data. Also, the Lee-Carter model only requires central mortality rate inputs and estimations are arrived at using the Singular Value Decomposition (SVD) while the APC model requires additional information such as number of deaths and number of people at risk for all ages in each calendar year, which is not readily available. A data extrapolation process is needed before further implementing the APC model. The extrapolation process smooths the curves of number of deaths and number of people at risk so that the annual number of deaths and people at risk for all ages can be found. The process generates numbers to “fill up” the missing values of numbers of deaths and the numbers of people at risk for all ages in each calendar year due to the fact that annual data for all ages are not available.

There are other obstacles when using the APC model. Namely, APC model calculation requires intensive computing resources. Moreover, calculating the cohort effect often requires extensive, historical data which sometimes are not available. Additionally and most importantly for this research, the APC model does not lend itself to stochastic modeling that is first needed to arrive at a martingale price which then enables accurate annuity pricing and creation of valuable derivatives. The abundance of data required by these mortality models combined with accounting for rare events and the cohort effect push the complexity factors of these models beyond the point of comprehension. It becomes impossible to understand and explain the results each model produces. Therefore, to create fair-pricing for derivatives and annuities a new model is necessary that is able to accurately model mortality rates, accepts stochastic processing, takes into account jumps and ensures that the results are easy to interpret.

We propose a model based on the simple and elegant Lee-Carter model. However, instead of using the Lee-Carter data structure based on the level of mortality we incorporate the Mitchell et al. (2013) data structure that uses the growth rate of mortality. By doing this we overcome the complexity issues that the extensions introduced to care of de-trending issues and jumps simultaneously. We identify the Mitchell et al. (2013) model as the modified Lee-Carter model. Analysis shows that the modified Lee-Carter model performs better than the Lee-Carter model, its extensions and the APC model. Using data from the United States, United Kingdom, Canada, Sweden, Australia, and six other countries they apply different criteria such as root sum squared errors and the unexplained variance to arrive at their conclusion.

There are other distinguishing features of the proposed model. It requires no prior data manipulation and simply uses data readily available from governmental websites. Additionally, we keep the original form of the Lee-Carter model and only change assumptions regarding jump characteristics in mortality rates.

The modified Lee-Carter model also performs better than the Lee-Carter model, its extensions and the APC model for out-of-sample prediction purposes if the proper distributional assumptions of the mortality index are considered. Remember that the Lee-Carter model does not account for jumps therefore its results produce a Gaussian distribution; however, epidemics and natural disasters happen and must be accounted for in a mortality model. The modified Lee-Carter model takes jumps into account by introducing a Normal Inverse Gaussian (NIG) distribution. The NIG distributional assumption enables the modified Lee-Carter model to adapt the fact that the distribution of the mortality rate change is high kurtosis and fat tailed. The predicted mortality rates under the NIG distributional assumptions have a narrower confidence interval and yield more stable prediction results.

Mortality modeling has made significant progress since the inception of the Lee-Carter model. Our proposed model intends to build on the modified Lee-Carter model by combining a stochastic process with an NIG distribution to model the mortality index and arrive at fair-pricing of derivatives. We propose an NIG Lévy process for modeling mortality dynamics. A Lévy process allows for an infinite number of jumps therefore large rare jumps and small normal jumps are captured. The additional NIG distribution specification is assumed for the non-Gaussian behavior of the mortality index.

The modified Lee-Carter model and the NIG Lévy process specification are well-suited for applying the Esscher transform to achieve martingale pricing inspired by Milevsky and Promislow (2001). Once we have applied the Esscher transform we can re-formulate the modified Lee-Carter model so that the resulting mortality rate is under a martingale measure. In Chapter 5 we show improved results in estimated and predicted mortality rates compared with the Lee-Carter model and the modified Lee-Carter model. These results enable us to price mortality/longevity linked derivatives, such as a q-forward, a mortality catastrophe bond and a longevity bond.

In the following sections we first show that we can formulate the modified Lee-Carter model to look similar to the Lee-Carter model. We then introduce an NIG Lévy process. Using an analogy to martingale pricing in financial modeling, we explain how mortality rates under a martingale measure can be achieved. The Esscher transform on the modified Lee-Carter model is then presented.

4.2 Modified Lee-Carter Model

In the Lee-Carter model the mortality rate is an exponential function with an age effect and a period effect shown in (3.1). The change of mortality index is modeled with a lag term shown in (3.4). Therefore, from(3.1) and (3.4), we can re-write it as follows

$$m_{x,t} = m_{x,t-1} \exp\left(b_x c + b_x e_t + \epsilon_{x,t} - \epsilon_{x,t-1}\right).$$

We can find that part of mortality dynamics depends on the information from the previous period, which is natural due to the fact that the mortality rate is a time series. Denote $(\epsilon_{x,t} - \epsilon_{x,t-s})$ as $\zeta^s_{x,t}$, alternatively we can have the equation shown as follows

$$\log(m_{x,t}) - \log(m_{x,t-1}) = b_x c + b_x e_t + \zeta^1_{x,t}. \tag{4.1}$$

From (4.1), we observe that the age effect appears on all components of the model, which is a feature of the Lee-Carter model. Besides, because the components of the Lee-Carter model are all tangled with age effect and period effect, the additional cohort effect may improve the fit. However, the improvement depends on the strength of the age effect and period effect. Besides the age effect, we also find that the difference of the logarithm of mortality rate can be modeled as the combination of a drift term, a random variable, and an error term that depends on not just the time t but also the past information s (s is 1 in (4.1)). The (4.1) shows that taking the difference of the the logarithm of mortality rate for the first step essentially does the same thing as modeling on the logarithm of the mortality rate and then taking the difference of the mortality index. To predict future mortality rates, we further derive a form based on the current mortality rate by the iterations shown as follows

$$m_{x,t} = m_{x,0} \exp\left(b_x ct + b_x \sum_{i=0}^{t-1} e_{t-i} + \zeta^t_{x,t}\right). \tag{4.2}$$

The (4.2) shows that the future mortality rate at time t can be obtained by multiplying the current mortality rate with an exponential function that delivers the dynamics of the mortality rate changes.

The mortality rate change of the Lee-Carter model has a drift term and some random components that are accumulated with time. If the random components $\sum_{i=0}^{t-1} e_{t-i}$ can have both the scaling property and the convolution property, it is appropriate to convert the original Lee-Carter model with some stochastic components. Thus, e_t is assumed to have an independently identical distributed (i.i.d.) Gaussian distribution with mean zero and variance t, and $\zeta_{x,t}^s$ is assumed to have i.i.d. Gaussian distribution with mean zero and variance $\sigma_x^2 s$ where σ_x^2 is the variance of the mortality rate change for age group x. The common argument lies on the assumption of e_t. Because the mortality rate change performs non-Gaussian behavior due to the rare events, the realized distribution of the e_t is non-Gaussian (Wang et al., 2011).

Because the mortality index of the Lee-Carter model is needed to be de-trended before further modeling can take place, it makes sense to take the difference of the logarithm of mortality rate at the the first step. By doing so, the k_t will be stationary and ready for modeling. Mitchell et al. (2013) suggest that constructing a model on the difference of the logarithm of mortality, rather than on the level and then taking difference, can provide a better fit. They propose the modified Lee-Carter model shown as follows

$$m_{x,t} = m_{x,t-1} \exp(a_x + b_x k_t + \varepsilon_{x,t})$$

or alternatively,

$$\log(m_{x,t}) - \log(m_{x,t-1}) = a_x + b_x k_t + \varepsilon_{x,t}. \tag{4.3}$$

By comparing (4.1) and (4.3), we see they look similar; the only difference is the modified Lee-Carter model is less restrictive on parametrization. Mitchell et al. (2013) apply the model on data from eleven countries and compare the results with the Lee-Carter model and its extensions such as the Lee-Carter model with cohort effect proposed by Cairns et al. (2007), Haberman and Renshaw (2011), and Renshaw and Haberman (2006). The modified Lee-Carter model performs better for in-sample fitting under different criteria, such as BIC, root sum of squared errors, and unexplained variance. The modified Lee-Carter also performs better than the APC model.

To forecast the future mortality rate, the modified Lee-Carter model can be re-written as follows by iterations

$$m_{x,t} = m_{x,0} \exp(a_x t + b_x \sum_{i=1}^{t} k_i + \sum_{i=1}^{t} \epsilon_{x,i}).$$

which is again similar to (4.2). The error term is usually assumed to have a Gaussian distribution, so $\sum_{i=1}^{t} \epsilon_{x,i}$ has a Gaussian distribution. Mitchell et al. (2013) choose an NIG distribution by taking the advantage of the fact that an NIG can fit the distribution with the high kurtosis and the fat tails well. The out-of-sample forecasting shows more accurate predicted values and the narrower confidence interval.

Since the k_t series is derived from log mortality change, each element represents a change for the time point. This similar fashion applies on the error term ϵ_t series.

Assume that each element in these two series is an i.i.d. random variable from some distribution, we can re-write the equation as follows

$$m_{x,t} = m_{x,0}\exp(a_x t + b_x \sum_{i=1}^{t} k_1 + \sum_{i=1}^{t} \epsilon_1).$$

If the k_1 can have a distribution that have the convolution property and the scaling property, $b_x \sum_{i=1}^{t} k_1$ will also have the same distribution. An NIG is an excellent choice, and it provides nice fit for the distribution of k_1 that has high kurtosis, asymmetry, and fat tails.

Let k_1 be a $NIG(\alpha, \beta, \mu, \delta)$ random variable. By the convolution property and the scaling property of an NIG distribution, we have

$$N_{x,t} := b_x \sum_{i=1}^{t} k_1 \sim NIG(\alpha', \beta', \mu' t, \delta' t)$$

where

$$\alpha' = \alpha/|b_x|, \ \beta' = \beta/|b_x|, \ \mu' = b_x\mu, \text{ and } \delta' = |b_x|\delta.$$

The convolution property and the scaling property also make the conversion of the modified model to the model with Lévy process possible.

Assume $(N_{x,t})_{t>0}$ be an NIG Lévy process, the modified Lee-Carter model becomes

$$m_{x,t} = m_{x,0}\exp(a_x t + N_{x,t}), \tag{4.4}$$

which is an exponential Lévy process. We should note that there is no error term. If an error term is included in the model, it will be a diffusion term. Carr et al.(2002), Carr and Wu (2003), Cont and Tankov (2004), Schoutens (2003), and Wu (2005) mention that the diffusion term can be dropped since a Lévy process can incorporate both large jumps and small jumps well. The addition of a diffusion term will not have significant impacts on the estimation.

4.3 Mortality Rate Under Martingale Measure

4.3.1 Concept

In financial modeling, when the market is complete, all state prices can be observed. The expected price can be calculated according to a measure that describes how the price is going to change. However, in reality, the market is incomplete, so not all state prices are ever available. When one invests in a financial product with risk, there is a risk premium that compensates his/her risk-taking behavior. However, the risk appetite of each investor is usually unknown. Such circumstance means that it impossible to price under the original physical measure. Therefore, a risk neutral pricing methodology must be constructed to create fair pricing of a financial instrument.

Moreover, when one invests in a financial product with risk, the risk premium is considered in the price. The risk is the volatility of the return of the financial product. For example, The expected return of a financial product should be at least better than the return of a risk-free bond.

To arrive at a fair price a martingale pricing technique is used. The expected rate of return under martingale pricing is the risk-free rate, which is the expected

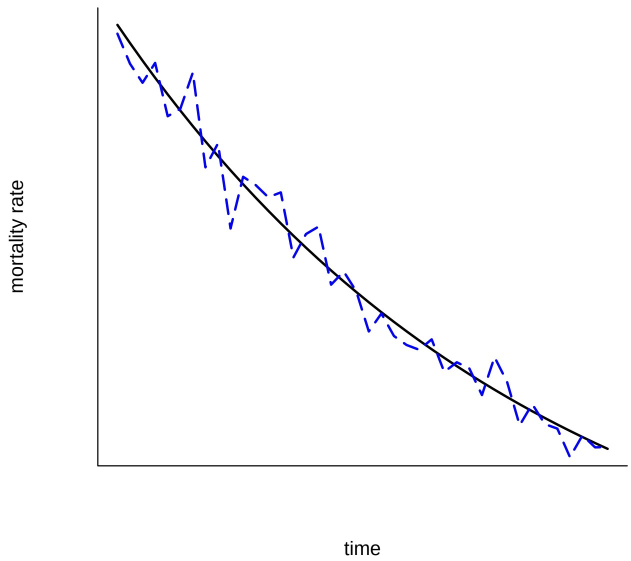


Figure 4.1: The illustration of the mortality rate under the martingale measure

rate of return of the risk-free bond. Under a martingale measure we do not need to know all possible state prices and we do not need to know the risk appetite for each investor. When a pricing procedure is under a martingale measure, the probability of each state is adjusted with the volatility of the financial instrument, meaning that the expected variance of the return change is zero. Therefore, the price under the martingale measure is risk-free or risk-neutral, so the expected return is simply the risk-free rate.

A similar technique to a martingale measure can be applied when pricing mortality/longevity linked derivatives. Mortality rates and asset returns have patterns and the movement along the curves, whether up or down, show similarities. While the mortality improvement rate represents the expected rate of change of the mortality rate, the dynamics of the mortality index show the volatility of mortality rates. Based on the mortality improvement rate we can construct martingale pricing for mortality/longevity derivatives similar to the martingale pricing based on the risk-free rate in financial modeling. The mortality rate possesses some stochastic jumps similar to the rates of return of most financial derivatives. Imagine that we can estimate the mortality improvement rate from historical data and apply it for prediction purposes. The prediction results are used for the mortality/longevity linked products. Since there is a volatility risk in the mortality rate, there is a risk premium for it. Different hedgers require different risk premiums, and we do not know the risk appetite for each of them. The mortality rate under a martingale measure means that the volatility of the mortality rate change is considered, so that the expected mortality rate change is zero. In this sense, the expected mortality rate under the martingale measure is simply calculated from the mortality improvement rate.

Figure 4.1 illustrates the mortality rate under the martingale measure. The black solid line is the mortality rate with only the estimation of the mortality improvement rate. Since there is no mortality variation, the mortality rate is as expected and shows no variance. It is the adjusted mortality rate under a martingale measure. The blue dashed line is one of the possible mortality rates with variations and mortality improvement. It does have the mortality risk generated from the variance of the mortality rate under the real measure. For pricing purposes, we need to adjust the blue dashed line to the black solid line, producing a risk-free mortality rate for further pricing.

Mortality risk derivatives are related to the mortality rate and longevity risks derivatives are related to the survival rate. Since we know mortality rates under a martingale measure, we can use it to price mortality linked derivatives, such as q-forward, mortality swap, and mortality bond. However, the pricing of a longevity linked derivative and annuities requires knowledge of survival rates which are the exponential function of negative mortality rates. To calculate the survival rate under a martingale measure, first researchers adjust the parameters of the stochastic processes to mimic the mortality dynamics. Then, they calculate the mortality rate under the martingale measure and derive the survival rate. Another approach is to assume a special exponential form and a constant force of mortality. Milevsky and Promislow (2001) assume that the hazard rate has a Gompertz expectation with a squared root volatility and they include an approximation formula for the survival rate. We will use the mortality rate under the martingale measure for survival rate calculation in later chapters.

4.3.2 Application of Esscher transform

In the previous chapter, we introduce the Esscher transform as a short cut for the martingale measure calculation when we have an NIG Lévy process. The modified Lee-Carter model with NIG Lévy process shown in (4.4) is indeed an exponential Lévy process. We would like to apply Theorem 2.8 and (2.5) to calculate θ for the corresponding Esscher transform.

We previously discuss that the mortality improvement rate is similar to the risk-free rate in financial modeling. Compare (4.4) and (2.1), we can re-write (2.5) as follows

$$\kappa(\theta^* + 1) - \kappa(\theta^*) = -a_x.$$

Since we already assume an NIG distribution for the Lévy process, we can apply formula (2.10) to calculate θ^* as follows

$$\theta^* = -\beta' - \frac{1}{2} + \sqrt{\frac{(\mu' + a_x)^2}{\delta'^2 + (\mu' + a_x)^2}\alpha'^2 - \frac{(\mu' + a_x)^2}{4\delta'^2}}.$$

After applying the martingale measure, the transformed mortality index $N^{\theta}_{x,t}$ has an NIG distribution; that is,

$$N^{\theta}_{x,t} \sim NIG(\alpha', \beta' + \theta, \mu' t, \delta' t).$$

Therefore, the mortality rate shown as follows is under a martingale measure obtained from the Esscher transform

$$m^{\theta}_{x,t} = m_{x,0} \exp(a_x t + N^{\theta}_{x,t}).$$

The transformed mortality rate $m^{\theta}_{x,t}$ has included the information of the stochastic change of mortality rate change, and this change depends on the parameter for the age group x.

4.4 Summary

The modified Lee-Carter model performs better than the Lee-Carter model, its extensions, and the APC model. Mitchell et al. (2013) show that the modified Lee-Carter model with an NIG distributional specification for the mortality index has more accurate predicted values and a narrower confidence interval. We propose an NIG Lévy process on the modified Lee-Carter model to not only improve performance but to also provide access to the pricing of mortality/longevity linked derivatives.

The convenience of the NIG Lévy process is not only for simplifying the calculation process of the martingale measure, but also for simplifying forecasting procedures, which is particularly useful for pricing purposes. When pricing mortality/longevity linked derivatives, we need to forecast by using a martingale measure. Although many mortality modeling approaches are available, they either only provide excellent mortality rate forecasting ability or are able to perform annuity calculation in some complicated way. By assuming mortality rate dynamics as an exponential NIG Lévy process shown in (4.4), we can produce accurate predictions enabling easy access to martingale pricing through the Esscher transform.

Chapter 5

Mortality Modeling

5.1 Introduction

Using two data sources, we present estimation results from the Lee-Carter model, the modified Lee-Carter model, and from our proposed model, modified Lee-Carter model with Normal Inverse Gaussian (NIG) Lévy process and the Esscher transform (hereafter proposed model)[1]. The data sources include United States mortality rate data from HIST290 of the National Center for Health Statistics and mortality rate data for England and Wales from the Human Mortality database. The data can be found respectively at

`http://www.cdc.gov/nchs/datawh/statab/unpubd/mortabs.htm`

and

[1] We will not have estimation results of the age-period-cohort model (APC) and the Lee-Carter model with the extensions, since Mitchell et al. (2013) have already shown that the modified Lee-Carter model performs better than these models. Also, we will not estimate any models related to the Affine process because such models are applied on the force of mortality rate and the National Center for Health Statistics and Human Mortality Database contain central mortality rate data.

http://www.mortality.org/.

The data sets from the National Center for Health Statistics include death rates per 100,000 people from 1900 to 2006. The data are broken into eleven age groups: (<1), (1-4), (5-14), (15-24), then every ten years until (75-84), and finally (>85). The data sets from the Human Mortality Database include death rates from 1930 to 2009. The data is broken into 24 age groups: (<1), (1-4), (5-9), (10-14), then every five years until (105-109), and finally (>110). Some data sets are missing data from the age groups over 100, therefore we only use data up to the 95-99 age group. The two data sets also include tables listing gender specific mortality rates. Furthermore, first we focus on the general population of the United States, then the female population of the United States. Lastly, we analyze the male populations of England and Wales. In this analysis we will show the in-sample estimations and the out-of-sample forecasting of our proposed model, the modified Lee-Carter model, and the Lee-Carter model. The parameter estimates for all three selected populations are also presented. In later chapters, we will utilize these estimation results to price mortality/longevity linked derivatives including a q-forward, a mortality catastrophe bond, an EIB bond, and a life settlement.

	Modified Lee-Carter	Lee-Carter
United States (general)	0.1200	0.1232
England and Wales (male)	0.2818	0.3097
United States (female)	0.0312	0.1246

Table 5.1: RSSE for the estimaiton results

5.2 Comparison of the Lee-Carter Model and Modified Lee-Carter Model

5.2.1 Comparison of the Estimation Results

We observe that mortality estimation results regarding our three primary populations are similar when using the Lee-Carter model and the modified Lee-Carter model. Figure 5.1 shows the estimation results of the Lee-Carter model and the modified Lee-Carter model for United States general population 65-74 age group from 1900 to 2006. Figure 5.2 shows the estimation results of the Lee-Carter model and the modified Lee-Carter model for the United States female population 65-74 age group from 1900 to 2006. Figure 5.3 shows the estimation results of the Lee-Carter model and the modified Lee-Carter model for the England and Wales male population 65-69 age group in from 1960 to 2009[2]. We can observe that the Lee-Carter model and the modified Lee-Carter model seem to show the similar es-

[2]Although the data of the male populations of England and Wales are available prior to 1960, we only use data after 1960. This follows current research trends that use data after 1960 to model mortality of England and Wales. Secondly, mortality rates for the male population in England and Wales do not have severe jumps. This characteristic can be used to examine whether the NIG distributional assumption can produce the robust estimation results when the realized distribution tends to be Gaussian.

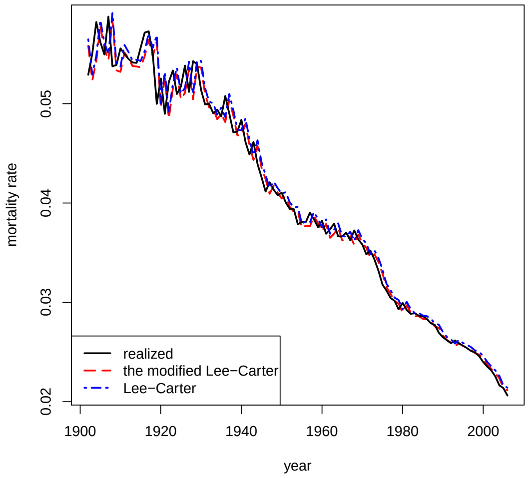


Figure 5.1: The estimation results of the Lee-Carter model and the modified Lee-Carter model for United States general population 65-74 age group from 1900 to 2006.6

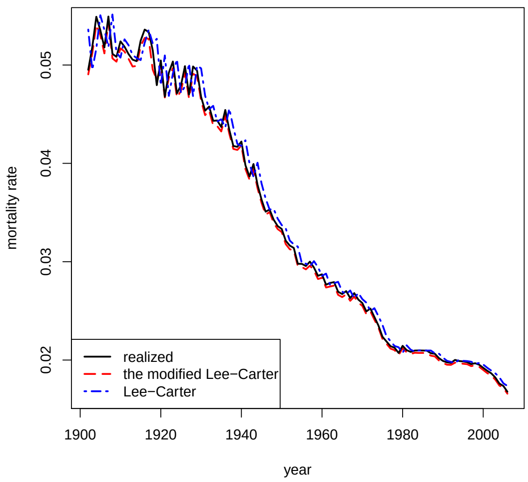


Figure 5.2: The estimation results of the Lee-Carter model and the modified Lee-Carter model for the United States female population 65-74 age group from 1900 to 2006

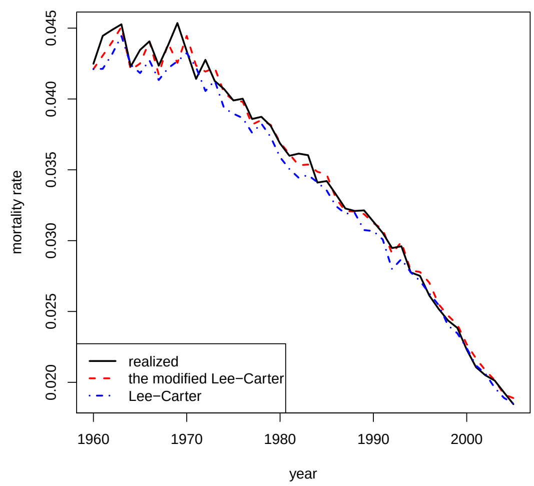


Figure 5.3: The estimation results of the Lee-Carter model and the modified Lee-Carter model for the England and Wales male population 65-69 age group in from 1960 to 2009

timation results for three populations. Table 5.1 shows the root sum of squared errors (RSSE) for the estimation results of three distinct populations. The modified Lee-Carter model has a smaller RSSE value for all three populations signifying that it performs better than the Lee-Carter model.

5.2.2 Estimation of the distributions

In order to predict the future mortality rate, we need to fit some distribution onto the mortality index k_t. We first draw a time series plot of the k_t to check the frequency and the amplitude of the jumps. Figure 5.4 shows the time series plot of the mortality index k_t for the general population of the United States from 1900 to 2006. Figure 5.5 shows the time series plot of the mortality index k_t for the

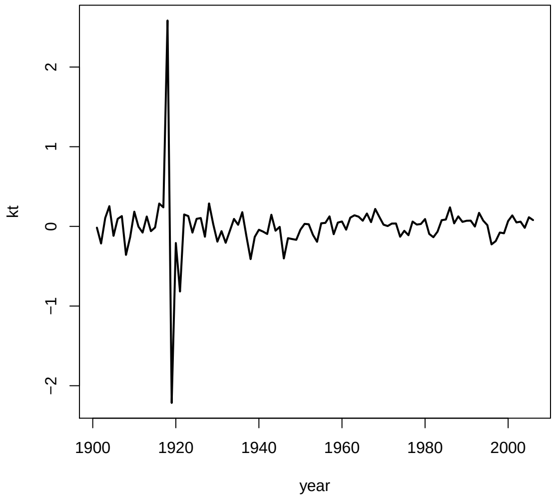


Figure 5.4: The time series plot of the mortality index of the general population of United States from 1900 to 2006

female population of the United States from 1900 to 2006. We can observe that there are some large jumps around 1918 due to the influenza epidemic of 1918. The World War I and World War II also cause some minor jumps around 1915 and 1950. Figure 5.6 shows the series plot of the mortality index k_t for the general population of the England and Wales from 1950 to 2009. The jumps of the mortality index are minor and the amplitude of a jump decreases by time, because the mortality rate was generally improved after 1960's in the United Kingdom.

We then would like to fit the distribution onto the mortality index k_t. We select a Gaussian distribution and an NIG distribution. When the time series plot of the k_t does not show significant large jumps, such as the k_t from the male population of England and Wales, a Gaussian distribution should fit well. However, if the time series plot show some abnormally large jumps, such as the k_t from the general

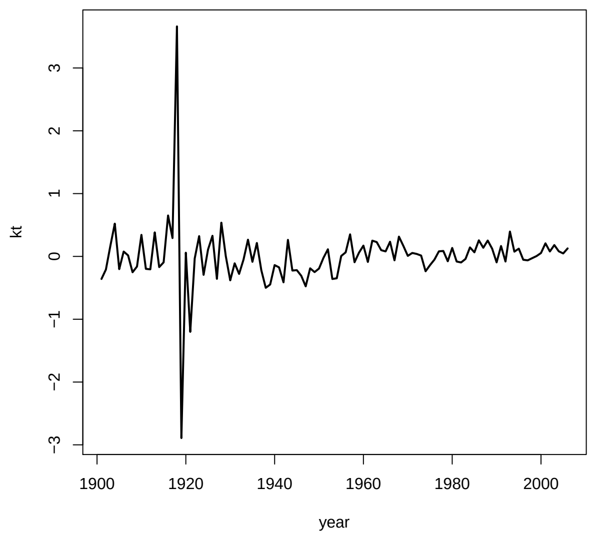


Figure 5.5: The time series plot of the mortality index of the female population of the United States from 1900 to 2006

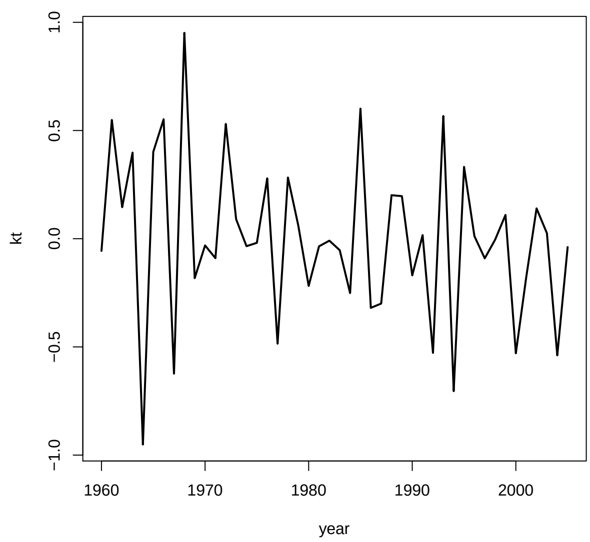


Figure 5.6: The time series of the mortality index of the male population of England and Wales from 1960 to 2009

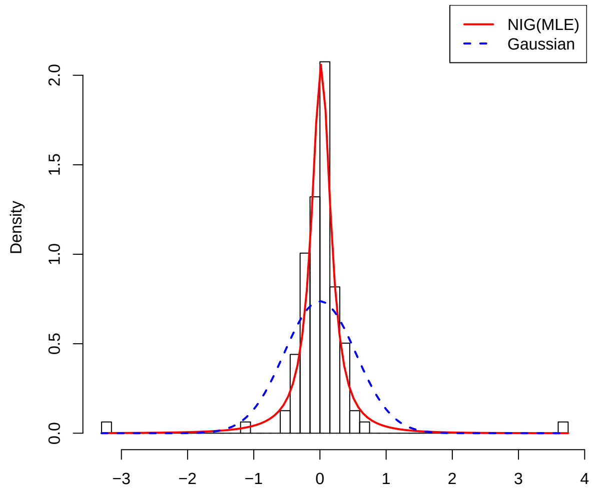


Figure 5.7: The fitted distribution of mortality index of the general population of United States from 1900 to 2006

population of the United States or the female population of the United States, an NIG distribution will fit better. The parameters of an NIG distribution can characterize a distribution with high kurtosis and fat tails, which is particularly suitable for fitting a distribution with abnormal jumps.

We fit an NIG distribution and a Gaussian distribution and compare the results. Figure 5.7 shows the fitted distribution of the mortality index of the general population of the United States from 1900 to 2006. Figure 5.8 shows the fitted distribution of the mortality index of the female population of the United States from 1900 to 2006. The red solid lines in Figure 5.7 and 5.8 are fitted by an NIG distribution. An NIG can indeed capture both the peak and tails of the distribution. The blue dashed lines in both figure fit a Gaussian distribution. A Gaussian distribution

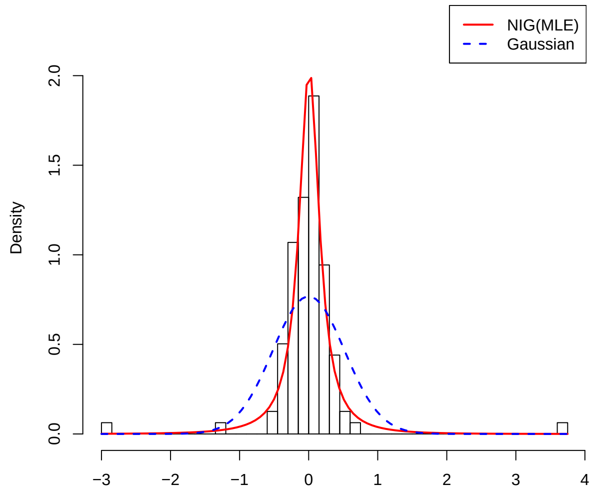


Figure 5.8: The fitted distribution of mortality index of the female population of the United States from 1900 to 2006

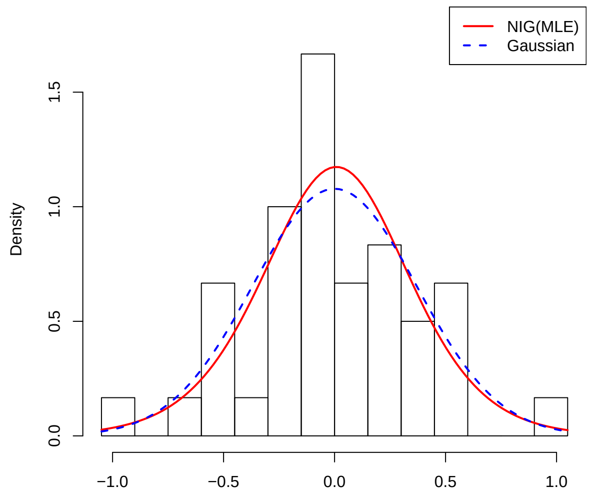


Figure 5.9: The fitted distribution of mortality index of the male population of England and Wales from 1960 to 2009

clearly misses the fit for high kurtosis and long tails. Figure 5.9 shows the fitted distribution of the mortality index of the male population of England and Wales from 1960 to 2009. Since there is no significant jump when we examine the time series plot of the k_t of England and Wales, the NIG distribution and Gaussian distribution produce similar results.

Next we include estimation results for our three specific populations. In Chapters 6 and 7 we use these results to price mortality/longevity linked derivatives including a q-forward, a mortality catastrophe bond, an EIB bond, and a life settlement. The estimation results of parameters for the Esscher transform are also calculated. Table 5.2 shows the parameter estimates for our proposed model when using United States general population mortality data from 1900 to 2002. These results will be used for pricing the mortality catastrophe bond. Table 5.3 shows the parameter estimates for our proposed model when using United States female population mortality data from 1900 to 2006. These results will be used for pricing the life settlement. Table 5.4 shows the parameter estimates for our proposed model when using England and Wales' male mortality data from 1960 to 2005. These results will be used for pricing a q-forward.

5.2.3 Comparison of Forecasting Results

Finally, we compare the forecasting abilities of the Lee-Carter model, the modified Lee-Carter model and our proposed model. We omit the last twenty-five-years of data from the United States' mortality rate data and the last ten-year data from the England and Wales' mortality rate data and compare predicted mortality rates

age group	a_x	b_x	θ_x
<1	-0.029785	0.049203	18.820900
1-4	-0.039399	0.146032	5.808083
5-14	-0.030146	0.129953	6.594559
15-24	-0.018681	0.211881	3.496684
25-34	-0.019956	0.234177	3.046362
35-44	-0.015863	0.122966	6.918658
45-54	-0.012140	0.051462	17.899149
55-64	-0.010599	0.026487	35.486287
65-74	-0.008750	0.018142	52.108885
75-84	-0.007842	0.008247	115.345985
>85	-0.005424	0.001450	658.591948

Table 5.2: The parameter estimates of the proposed model by using mortality data of the general population in the United States 1900-2002

age group	a_x	b_x	θ_x
<1	-0.029734	0.056170	17.206336
1-4	-0.040439	0.163779	4.675276
5-14	-0.032172	0.152349	4.732417
15-24	-0.024550	0.194816	2.556976
25-34	-0.023961	0.213318	2.070498
35-44	-0.018204	0.108566	6.055575
45-54	-0.014110	0.050465	16.990611
55-64	-0.012468	0.026635	36.315323
65-74	-0.010954	0.018690	53.222012
75-84	-0.009393	0.009130	112.756512
>85	-0.006539	0.006082	169.727237

Table 5.3: The parameter estimates of the proposed model by using mortality data of the female population in the United States 1900-2006

age group	a_x	b_x	θ_x
<1	-0.032198	0.023333	-178.133699
1-4	-0.031987	0.111591	-13.495506
5-9	-0.034720	0.021364	-201.019097
10-14	-0.021513	0.010027	-444.440563
15-19	-0.017207	0.016941	-223.883720
20-24	-0.012190	0.015223	-224.593224
25-29	-0.007775	-0.004471	-974.714395
30-34	-0.006461	-0.009347	-354.663995
35-39	-0.009081	0.009287	-402.117895
40-44	-0.011161	0.026122	-84.264573
45-49	-0.013071	0.024732	-106.291354
50-54	-0.016521	0.042925	-46.705424
55-59	-0.019693	0.045441	-49.273238
60-64	-0.019310	0.041984	-56.129553
65-69	-0.018400	0.058065	-28.450457
70-74	-0.016900	0.061690	-22.797928
75-79	-0.014691	0.083243	-9.666672
80-84	-0.011796	0.084675	-6.573880
85-89	-0.011457	0.106949	-3.249793
90-94	-0.007503	0.118011	-0.425454
95-99	-0.006616	0.112216	-0.133826

Table 5.4: The parameter estimates of the proposed model by using mortality data of the male population in England and Wales 1960-2005

and realized mortality rates. The predicted value is the average of 100,000 simulations and a 95 % confidence level is also provided. Figure 5.10 shows the predicted mortality rate of the general population of the United States. Figure 5.11 shows the predicted mortality rate of the female population of the United States. We observe that predictions from the modified Lee-Carter model and our proposed model are more accurate than the Lee-Carter model, and the confidence interval of our proposed model is narrower. Because our proposed model has the tightest confidence interval it provides the most stable prediction results. Figure 5.12 shows the predicted mortality rates of England and Wales' male population. With less data than our other populations and no abnormal jumps, the Lee-Carter model and the modified Lee-Carter cannot effectively predict the mortality rates because the realized mortality rates are out of the confidence intervals. Our proposed model produces better predicted values and an appropriate confidence interval that includes the realized values.

5.3 Summary

A comparison of estimation results from the Lee-Carter model, the modified Lee-Carter model, and from our proposed model suggest that the modified Lee-Carter model performs better than the Lee-Carter model. Although it uses the same structure as the Lee-Carter model by using the mortality growth rate instead of the level of mortality rate, the modified Lee-Carter model produces better results. Furthermore, the NIG distributional specification fits the mortality index better than a Gaussian distribution because the mortality rate distribution may be high-kurtosis

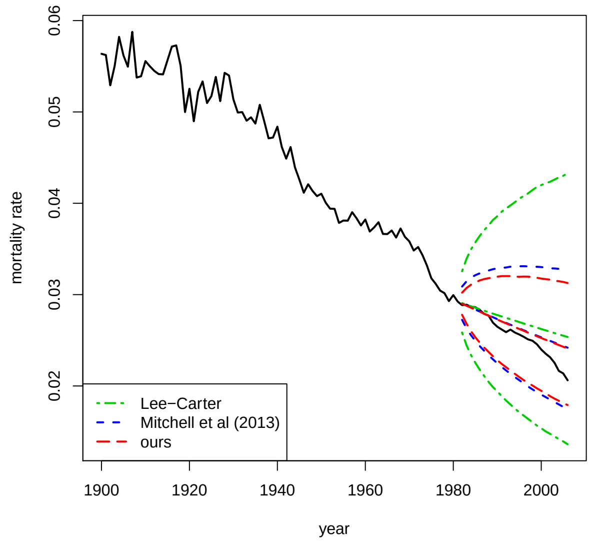


Figure 5.10: The predicted mortality rate of the general population of United States

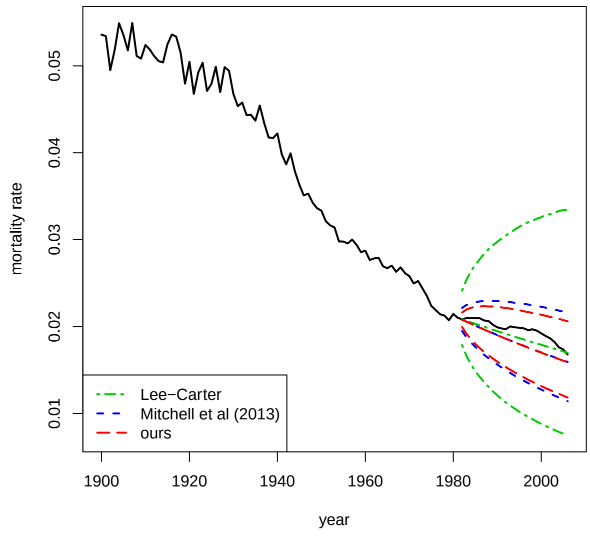


Figure 5.11: The predicted mortality rate of the female population of the United States

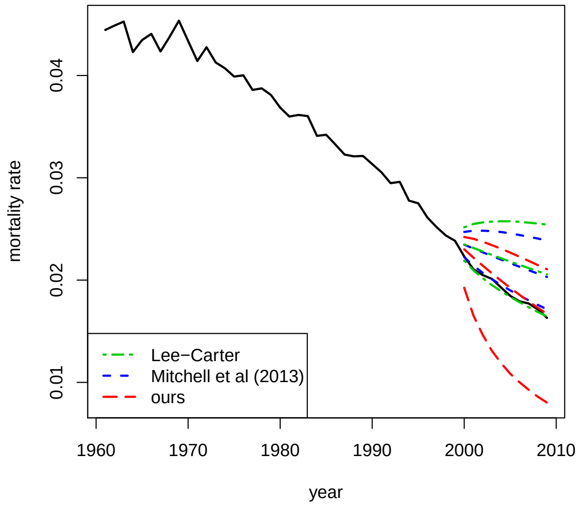


Figure 5.12: The predicted mortality rate of the male population of England and Wales

and fat-tailed due to the mortality jumps. If the mortality index does not show significant jumps, such as the mortality index of England and Wales, fitting with a Gaussian distribution or an NIG distribution will have similar results. A comparison of prediction results from the Lee-Carter model and our proposed model shows more accurate predicted values of the mortality rate, as well as more stable prediction results with a narrower confidence interval.

Chapter 6

Life Settlement Pricing

6.1 Introduction

A life settlement, also called a viatical settlement, is a legal financial instrument that enables a life insurance policyholder to receive a cash value from the policy while still alive. A third party purchases the life insurance policy from the policyholder for a lump sum greater than its cash surrender value but less than the net death benefit value. The third party then takes ownership of the policy, pays the premiums and when the insured dies, receives the full death benefit. Often used by terminally ill or chronically ill people, a life settlement enables the insured to have access to the returns from the policy rather than leave it to a beneficiary.

The HIV-AIDS epidemic of the 1980s spurred interest in life settlements in the United States. At that time, mostly young, gay men were diagnosed with the HIV-AIDS virus. With high mortality rates and prohibitively expensive medications, many with life insurance policies opted to sell the policy and use the proceeds to cover their expenses. This created a secondary market for life insurance policies.

Although the impetus for creating a secondary market for life insurance policies was the HIV-AIDS epidemic, the appeal came from investors offering to buy the policy for more than the holder could obtain by surrendering the policy or simply letting it lapse. The cash value was calculated according to life tables created by reputable institutions, such as the Commissioners Standard Ordinary (CSO) and Society of Actuaries. The transaction created a win-win situation for both parties. Policyholders received a lump sum payment while still alive and investors obtained an acceptable rate of return. Given that HIV-AIDS medications did not exist in the 1980s, investors often saw rates of return up to 15% (Quinn, 2008; Siegert, 2010).

Following the financial success of life settlements, several companies began specializing in life settlement transactions. However, this new market soon collapsed because rates of return were diminishing. By 1996 new HIV-AIDS treatments were prolonging the lives of those infected with the virus making the investment less attractive to investors. The dramatic decrease in mortality rates caused investors to significantly decrease the offer price to policyholders infected with HIV-AIDS. Since the policies might take longer to mature, the price of a life settlement plummeted (Stone and Zissu, 2006).

Wanting to grow the secondary market for life insurance, life settlement specialists made senior citizens their new target market. Life insurance policyholders 55 years old or older often hold policies with death benefits worth $250,000 or more. Targeting those who no longer can afford the premium and others who simply no longer need the insurance and would prefer a lump sum payout while still alive, the settlement market is growing. In 2005 it was estimated to be worth \$10 billion and in 2008, \$21.5 billion (estimated by Conning & Company). Today, the market is estimated to be worth \$35 billion (estimated at the end of 2011 by Conning & Company). As stated in Chapter 1, in 2009 there were 455,000 Americans over 100 years old. This number is expected to increase 5.5% per year thereby doubling the centenarian population every 13 years. Life settlement specialists see this aging population as a market that will be worth approximately \$117 billion in 2017. Also, the Social Security Advisory Board predicts that approximately 1% of the general population will be centenarians in 2050 (Sisk, 2011).

Demographic and market conditions suggest that life settlement instruments are excellent investments. However, to make the instrument attractive to investors and policyholders, pricing life settlements must be optimized. Policyholders want a significant payout and investors want a high rate of return. If this can be accomplished, then the life settlement market will flourish.

In the following sections we explain the two currently used settlement pricing methodologies and then introduce a new pricing model that uses information theory to more accurately price the life settlement. Information theory takes into account private information about the policyholder. We include this information to calculate and create an adjusted life table. This adjusted life table enables an optimized strategy for pricing life settlements. Using our proposed model, we apply information theory to price a life settlement policy. We describe how to calculate a settlement price by applying the density of death years. We also propose ideas of how to calculate risk premiums for life settlements.

6.2 Deterministic and Probabilistic Life Settlement Pricing

There are two currently accepted life settlement pricing methodologies. The deterministic pricing methodology came before probabilistic pricing. To calculate settlement pricing, both methodologies use data from the Commissioners Standard Ordinary (CSO) and Valuation Basic Table (VBT) life tables. The deterministic methodology as its name suggests, assumes mortality occurs at the exact time listed in the mortality tables. This assumption was reasonable when life expectancies were shorter and easier to predict than today and during the HIV-AIDS epidemic. Longer life spans have necessitated that a probabilistic pricing methodology be used to price life settlements. In so doing, it does not assume mortality as occurring at an exact time but accounts for it occurring within a curve of life expectancy.

Let us use an example from the life tables. The CSO and VBT tables show that a 65-year-old woman in 2013 is expected to live another about 17 years. However, today, many senior citizens suffer from chronic illnesses such as COPD (chronic obstructive pulmonary disease), whose symptoms can be treated by modern medicine but whose cause cannot be cured. Chronic illnesses decrease life expectancy. Current mortality tables do not take this phenomenon into account but if they did, the accuracy of pricing of life settlements would improve.

Calculating deterministic life settlement pricing

Given a face value of an insurance policy B and the discounting factor $v = 1/(1+r)$, where r is the required rate of return, the present value of payoff of a life insurance policy is Bv^n where n is the number of year that the policy elapses. A policyholder pays a periodical premium. Let P denote the premium and $\ddot{a}_n$ denote the present value of an annuity due[1] with one dollar payable at the beginning of each year starting at time zero. The price calculated from the deterministic life settlement pricing method is

$$Bv^n - P\ddot{a}_n \tag{6.1}$$

when n is deterministic and given.

The problem with the deterministic life settlement pricing is that the year of settlement, which is the year that the policyholder will die, actually is unknown and hence not deterministic. Life expectancy is based on averages. Therefore our example of the 65 year old woman in 2013 living 17 more years could actually be shorter if she suffers from a chronic or debilitating illness. The woman's longevity is actually a random variable. Therefore, the price of a life settlement should be the expected value of (6.1) with respect to n, and this value is greater than or equal to the price from the deterministic life settlement pricing (Brockett et al., 2013).

Calculating probabilistic life settlement pricing

Concluding that a policyholder's actual length of life is a random variable, the probabilistic life settlement pricing considers the expected value of a life settlement shown as follows

$$\mathbb{E}(Bv^n - P\ddot{a}_n). \tag{6.2}$$

[1]The explicit form for $\ddot{a}_n$ is shown as follows

$$\ddot{a}_n = 1 + v + v^2 + \cdots + v^{n-1}.$$

To calculate (6.2), we only need to know the probability distribution of n (Bowers et al., 1986). The density of n can be obtained from mortality rate table. According to Forman (2010), most life settlement companies use the 2008 VBT table. Since some private information of a policyholder can be obtained, each company has its own way to modify the mortality rate table.

A popular adjustment approach is to start with a standard table and then multiply the mortality rate for each age by a selected constant to obtain a new table that takes account of the policyholder's private information. Forman (2010) gives an example report of life expectancy calculation results for an 84-year-old woman whose medical records led the underwriter to come up with an estimate of 9.2 years of mean life expectancy, a median life expectancy of 9.3 years and an 85% mortality value of 13 years. The suggested multiplier is 2.03. Although easy to implement, this method is limited because it does not take into account all the private information that the policyholder discloses[2].

6.3 Information Theory

6.3.1 Motivation

If it is possible to correctly price life settlements then investors and policyholders would be interested in these financial instruments. Using information theory we adjust original probabilities of mortality and create a policyholder specific "cus-

[2]The report states "Please note it is recommended that the information provided in this life expectancy evaluation be used in its entirety. If only subset of the data is used, you will be losing the interrelationships between the analytics."

tomized" life table that produces optimized pricing of life settlements.

Consider two densities $f(t)$ and $g(t)$, where t is a random number denoted the year of death. The $g(t)$ is the density of death year for general population in each year. The $f(t)$ is the density of death year for a specific individual in each year. The difference between $f(t)$ and $g(t)$ is that the $f(t)$ is derived from $g(t)$ with incorporating the information of the personal information of the individual. To make the calculation possible, we apply the minimum discrimination information estimation in the information theory.

6.3.2 Minimum Discrimination Information Estimation

Consider two densities $f(t)$ and $g(t)$ as mentioned above, the statistics $\ln(f(t)/g(t))$ is a sufficient statistics and is a log odds ratio. When we have a long sequence of observations, $(t_i)_{i\geq 0}$, the long-run odds ratio is shown as follows

$$E\left(\ln\frac{f(t)}{g(t)}\right)=\sum_i \ln\frac{f_{t_i}}{g_{t_i}} \tag{6.3}$$

which is the expected amount of information for discriminating between two densities f and g. Denote (6.3) as $I(f|g)$, and $I(f|g)$ is the quantity to measure how much divergence between densities f and g. It is not difficult to show that $I(f|g)$ is zero if and only if $f=g$ (Brockett et al. (2013)).

Let us rephrase the problem in a general setting. Assume that a density function g is given, and we wish to find another density f that is "as close as possible" to g with sanctification of $k+1$ number of generalized moment constraints including the expected value of some collection of functions. Let $a_j(t)$ denote one of such function, so we can have a collection of equations shown as follows

$$\begin{aligned}
\theta_0 &= 1 = \sum_{i}^{n} f(i) \\
\theta_1 &= \sum_{i}^{n} a_j(i) f(i) = \mathbb{E}(a_1(n)) \\
&\vdots \quad \vdots \\
\theta_k &= \sum_{i}^{n} a_j(i) f(i) = \mathbb{E}(a_k(n)).
\end{aligned}$$

The first constraint is to assure that f is a probability density function. The second to the $k+1$ constraint can be modified according to what the private information we can obtain from the policyholder. If we set $a_1(i) = i$ and $\theta_1 = m$, then the second constraint means the policyholder is expected to live m more years. We can also have the information of the distribution of length of life. For example, if the 85 percentile of length of life is 13 years, then we can modify the condition to the equation shown as follows

$$\sum_{i=1}^{13} f(i) = 0.85.$$

Such construction of the conditions is versatile and can be utilized whatever information we can obtain from the policyholder.

To find the f, we need to solve the problem presented as follows

$$\min_{f} I(f|g) \tag{6.4}$$

with the conditions derived from the personal information of the policyholder.

Brockett et al. (1980) show that the optimization (6.4) with linear conditions is a convex programming problem and that the dual mathematical programming problems are actually unconstrained and involve only exponential and linear terms (Brockett et al., 2013). The number of dual programming problems is equal to the number of constraints, and the unique solution proposed by Brockett et al. (1980) has the general form

$$f(t) = g(t) \exp(-(\beta_0 + 1) - \beta_1 a_1(t) - ... - \beta_k a_k(t)).$$

6.4 Adjusting Mortality with Known Information

6.4.1 Applying the Minimum Discrimination Information Estimation

In this section we apply the minimum discrimination information estimation. Suppose we can obtain mortality rate table for the general population and calculate the density g. For an individual who is willing to sell his/her life insurance policy, we know the number of years that he/she is expected to live is m . We would like to find a density function f with the information of his/her expected length of life and is "as close as possible" to g. The f should satisfy the conditions shown as follows

$$\begin{aligned} \sum_{i=1}^{k} f(i) &= 1 \\ \sum_{i=1}^{k} f(i)i &= m. \end{aligned} \tag{6.5}$$

Hence the optimization problem is shown as follows

$$\min I(f|g) = \min \sum_i f(i) \ln(f(i)/g(i))$$

subject to the constraints (6.5). Let β_0 and β_1 be the Lagrange multipliers and then the original optimization problem can be rephrased as to minimize the Lagrange function shown as follows

$$L(f, \beta_0, \beta_1) = \sum_i f(i) \ln(f(i)|g(i)) - \beta_0 \left(1 - \sum_i f(i)\right) - \beta_1 \left(m - \sum_i f(i)i\right)$$

with the conditions that $f(i) \geq 0$ for all i. The first order conditions are shown as follows

$$\begin{aligned}
&\ln(f(i)|g(i)) + 1 + \beta_0 + i\beta_1 = 0, \text{ for } i = 0, ...k, \\
&-1 + \sum_i f(i) = 0, \\
&-m + \sum_i f(i)i = 0.
\end{aligned}$$

The first condition gives the function shown as follows

$$f(i) = g(i) \exp(-1 - \beta_0 - i\beta_1)$$

for all i. To utilize the information of the last two conditions, Brockett et al. (1980) suggest to consider the function $\Phi(\beta) = \sum_i g(i) \exp(-i\beta)$. Since the sum of densities is one, we have

$$1 = \sum_i g(i) \exp(-1 - \beta_0 - k\beta_1) = \exp(1 - \beta_0)\Phi(\beta_1),$$

so we can derive β_0 shown as follows

$$\beta_0 = \ln \Phi(\beta_1) - 1.$$

To obtain β_1, note that $\Phi'(\beta_1) = -\sum_i g(i)i \exp(-i\beta_1)$, so we then has the equation shown as follows

$$\begin{aligned}
\Phi'(\beta_1) &= -\sum_i ig(i) \exp(-i\beta_1) \\
&= -\exp(1 + \beta_0) \sum_i ig(i) \exp(1 - \beta_0 - i\beta_1) \\
&= -\exp(1 + \beta_0) \sum_i if(i) \\
&= -\Phi(\beta_1)m
\end{aligned}$$

Hence, β_1 can be obtained by solving $\Phi'(\beta_1) = -\Phi(\beta_1)m$ numerically.

6.4.2 Example

For the purpose of illustration, assume that the distribution of deaths during the year of death is assumed to be a uniform distribution. Consider the 2008 VBT table as the mortality table for general population, and a 70-year-old female illustrated in the table is expected to live another 8.5 years estimated by the medical underwriter.

Her whole life insurance paid up at age 110 was issued on a female standard risk at age 40 based on the 2008 VBT. The face value (death benefit) of the policy is \$50,000 with annual premium P = \$565.5. Our goal is to apply the estimation described in the previous section and obtain a new probability of death in each year for this female based on the information of her 8.5 life expectancy. Then we will calculate the price of her life settlement.

Table 6.1 shows the estimation results with $\beta_0 = -2.4558$ and $\beta_1 = 0.1300$. Figure 6.1 compares the original mortality rate curve and the adjusted mortality rate curve for the 70 year old female. The adjusted mortality rate increases about six percent at age 70 and continues until death. The adjusted density of death year puts more probabilities for the early years to disclose the information that the policyholder has the shorter life expectancy.

6.5 Comparison of Life Settlement Pricing

When we utilize the additional private information of the policyholder, the resulting value of the life settlement will be higher than the value without knowing the information. The additional private information has its own value. Table 6.2 shows the comparison of the life settlement prices calculated without adjustment of information and the price from deterministic life settlement pricing by using four mortality tables: disabled retirees, healthy annuitants[3], the CSO 2001, and the 2008 VBT. We can find that no matter what mortality table we are using, the deterministic life

[3]The mortality rates of disabled retirees and healthy annuitants are from the mortality table RP-2000 and it is available at
`www.soa.org/files/research/exp-study/rp00_mortalitytables.pdf`

age	year	mortality rate (2008 VBT)	probability of death in year (2008 VBT)	adjusted probability of death in year	adjusted mortality rate
70	0	0.019540	0.019540	0.083785	0.083785
71	1	0.021490	0.021070	0.079333	0.086588
72	2	0.023550	0.022594	0.074700	0.089260
73	3	0.025680	0.024057	0.069843	0.091636
74	4	0.027990	0.025548	0.065130	0.094073
75	5	0.030680	0.027219	0.060933	0.097150
76	6	0.033820	0.029084	0.057173	0.100963
77	7	0.037340	0.031025	0.053554	0.105194
78	8	0.041280	0.033018	0.050047	0.109862
79	9	0.045710	0.035052	0.046655	0.115054
80	10	0.050840	0.037204	0.043483	0.121174
81	11	0.056690	0.039376	0.040412	0.128143
82	12	0.063200	0.041409	0.037318	0.135726
83	13	0.070390	0.043206	0.034191	0.143880
84	14	0.078380	0.044723	0.031078	0.152760
85	15	0.087330	0.045925	0.028023	0.162578
86	16	0.097360	0.046728	0.025037	0.173459
87	17	0.108570	0.047035	0.022130	0.185492
88	18	0.121010	0.046732	0.019308	0.198691
89	19	0.134700	0.045724	0.016589	0.213037
90	20	0.149610	0.043945	0.014000	0.228460
91	21	0.165730	0.041397	0.011580	0.244939
92	22	0.183030	0.038141	0.009369	0.262453
93	23	0.201470	0.034300	0.007398	0.280999
94	24	0.220990	0.030043	0.005690	0.300592
95	25	0.240780	0.025500	0.004241	0.320322

Table 6.1: Partial standard life table and adjusted life table that achieves a life expectancy equal to 8.5 years

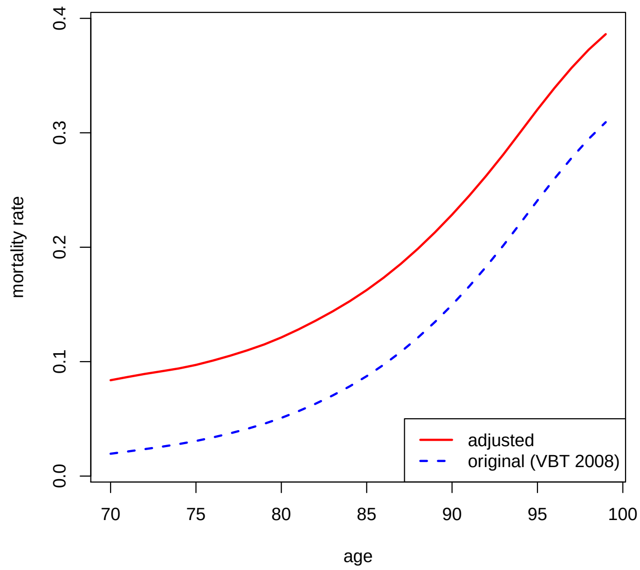


Figure 6.1: The comparison of mortality rate curves

settlement pricing that is adjusted by the private information (expected length of life is 8.5 years) is always higher that the price without the adjustment.

When we have our pricing methods adjust with the private information of the policyholder, the probabilistic life settlement pricing will produce the higher value than the deterministic life settlement pricing. This is because the probabilistic life settlement pricing considers the uncertainty of the year of death contribute a value to prices. Since no matter what mortality table we are using, the life settlement pricing with the adjustment of the policyholder's private information is always the highest, we only compare the pricing results from the 2008 VBT with the pricing from the modified Lee-Carter model with Lévy processes and the Esscher transform. As mentioned before, the 2008 VBT table is the most popular one for life settlement

required	Disabled Retiree		Healthy Annuitant		CSO 2001		2008 VBT	
rate of return	not adjusted	deterministic	not adjusted	deterministic	not adjusted	deterministic	not adjusted	deterministic
1%	37,122.82	41,260.97	34,083.33	41,252.18	33,934.03	41,259.48	34,895.90	41,249.14
2%	32,745.97	38,094.60	28,862.04	38,072.68	28,686.22	38,091.47	29,852.70	38,065.30
3%	29,091.04	35,330.87	24,611.75	35,293.86	24,426.17	35,326.56	25,704.35	35,281.79
4%	26,016.99	32,904.81	21,129.08	32,852.29	20,945.07	32,900.09	22,271.93	32,835.71
5%	23,413.66	30,763.66	18,257.13	30,696.11	18,082.06	30,759.37	19,415.50	30,675.55
6%	21,194.36	28,864.31	15,874.11	28,782.76	15,712.54	28,861.29	17,025.01	28,758.86
7%	19,290.44	27,171.32	13,884.84	27,077.07	13,739.34	27,170.31	15,013.47	27,050.55
8%	17,647.18	25,655.39	12,214.56	25,549.84	12,086.31	25,657.03	13,311.73	25,521.41
9%	16,220.67	24,292.14	10,804.17	24,176.71	10,693.41	24,296.95	11,864.55	24,147.04
10%	14,975.47	23,061.21	9,606.67	22,937.25	9,513.03	23,069.59	10,627.60	22,906.96
11%	13,882.80	21,945.48	8,584.53	21,814.26	8,507.23	21,957.73	9,565.11	21,783.88
12%	12,919.16	20,930.51	7,707.56	20,793.17	7,645.60	20,946.83	8,648.08	20,763.16
13%	12,065.25	20,004.01	6,951.40	19,861.60	6,903.66	20,024.53	7,852.90	19,832.36
14%	11,305.12	19,155.55	6,296.26	19,008.97	6,261.58	19,180.32	7,160.26	18,980.85
15%	10,625.54	18,376.15	5,725.99	18,226.23	5,703.22	18,405.17	6,554.28	18,199.49
16%	10,015.48	17,658.13	5,227.38	17,505.56	5,215.39	17,691.36	6,021.88	17,480.43
17%	9,465.68	16,994.84	4,789.52	16,840.24	4,787.27	17,032.19	5,552.19	16,816.89
18%	8,968.36	16,380.51	4,403.40	16,224.42	4,409.90	16,421.88	5,136.17	16,202.99
19%	8,516.91	15,810.13	4,061.54	15,653.00	4,075.88	15,855.39	4,766.29	15,633.59
20%	8,105.72	15,279.31	3,757.68	15,121.54	3,779.04	15,328.32	4,436.20	15,104.23

Table 6.2: The comparison of the life settlement prices calculated without adjustment of information and the price from deterministic life settlement pricing by using four mortality tables

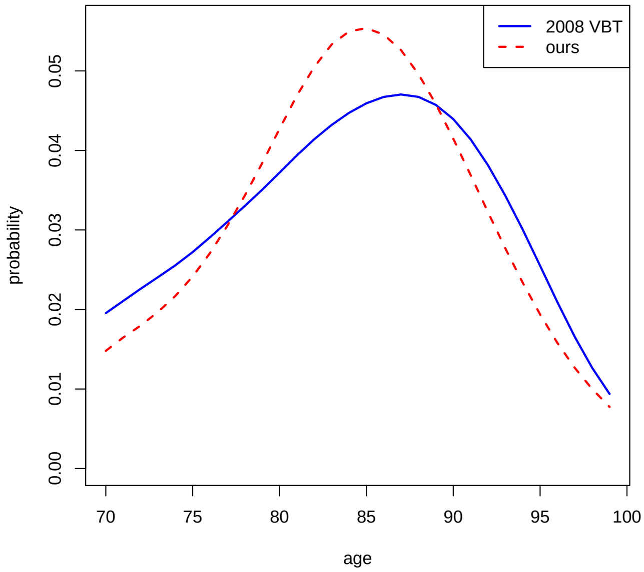


Figure 6.2: The comparison of the curves of density of death year

pricing and it is the latest mortality table.

Figure 6.2 compares the 2008 VBT distribution of the death year with our adjusted model when the expected number of years to live is 8.5. The 2008 VBT shows the expected number of years remaining to be approximately 16.5 years. The 2008 VBT shows the higher probabilities for the first seven year. Our model shows the higher probabilities from year seven to nineteen. Using information theory and adjusting for mortality improvement suggests, that this policyholder has a disposition to live longer than current mortality tables suggest. Figure 6.3 shows the adjusted probability distribution for the 2008 VBT and our model. The adjusted probability distribution is the primary cause of the difference in price . The adjusted probability distribution of the 2008 VBT is higher than the probability of our model in the first

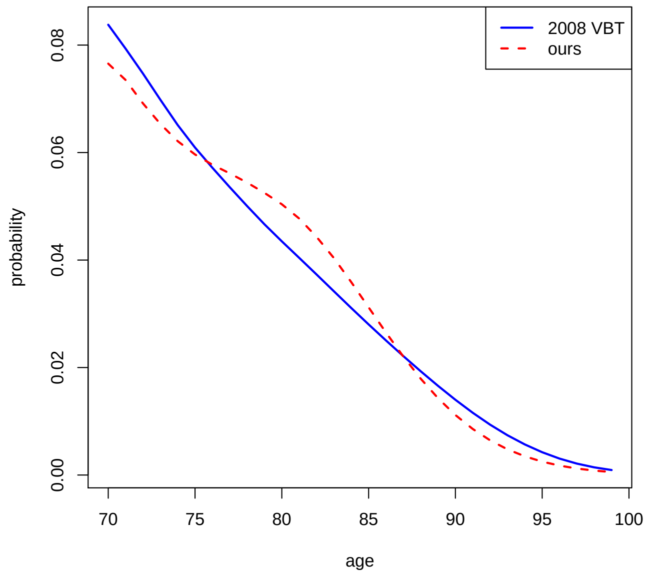


Figure 6.3: The comparison of the curves of adjusted density of death year

five years. Due to this difference, we expect the life settlement price calculated from the 2008 VBT will be higher than ours. The discounted value of a life policy is very high for the first five years. Although our model shows the adjusted probability will be higher in the later 10 years, our model produces a lower price for a life settlement.

Table 6.3 shows the present value of a life settlement calculated with life expectancy of 8.5 years . The deterministic life settlement pricing shows the lowest prices because it does not consider that the uncertainty of the year of death has a value. Compare the price of the 2008 VBT and our model when the required rate of return is higher than 10%; our model produces a price of \$22,459.07, between \$450.00 and \$600.00 less than the price calculated using the 2008 VBT. This difference is because the 2008 VBT has a higher probability of death in the first five

rate	deterministic	2008 VBT	ours
1%	41,052.48	41,259.48	41,224.31
2%	37,546.01	38,091.47	38,001.25
3%	34,356.27	35,326.56	35,170.19
4%	31,452.17	32,900.09	32,672.48
5%	28,805.93	30,759.37	30,459.51
6%	26,392.70	28,861.29	28,490.87
7%	24,190.20	27,170.31	26,732.74
8%	22,178.47	25,657.03	25,156.75
9%	20,339.60	24,296.95	23,738.95
10%	18,657.47	23,069.59	22,459.07
11%	17,117.61	21,957.73	21,299.90
12%	15,706.99	20,946.83	20,246.72
13%	14,413.87	20,024.53	19,286.94
14%	13,227.65	19,180.32	18,409.75
15%	12,138.79	18,405.17	17,605.97
16%	11,138.65	17,691.36	16,867.01
17%	10,219.42	17,032.19	16,186.36
18%	9,374.04	16,421.88	15,557.93
19%	8,596.10	15,855.39	14,976.02
20%	7,879.82	15,328.32	14,436.11

Table 6.3: Present value of the life settlement starting from VBT table 2008 and our model both adjusted to have life expectancy of 8.5 years

years when the discounted value of a life settlement is high.

6.6 Risk Considerations

In the previous section of the discussion of the pricing a life settlement, we consider only a pure premium or expected actuarial present value. We do not discuss a risk premium. In the life settlement market, pricing is based on the expected rate of return of a life settlement, and the risk premium means there will be an additional expenses and a cost for the investors who bears the mortality risk. Moreover, since the nature of the financial market of life settlements is actually close to an auction market for policies, the actual risk premium varies company by company. Besides, the risk premium can be correlated with other factors and the level of impacts may be different for different companies. Hsieh et al. (2012) analyze several hundred of successful life settlement transactions from a major life settlement company, Conventry[4]. They find that the risk premium can be correlated to the issuing insurance company rating, the age of the policyholder, and the length of time the policy has been in force.

There is no common approach to calculate the risk premium for a life settlement. The risk associated with a life settlement policy is related to the policyholder's individual risk. Moreover, each life settlement company has its own pricing methodology. Since purchasing a life settlement can be viewed as an investment project, the internal rate of return (IRR) is a good way to evaluate the transaction. The IRR is frequently used for understanding the profitability of investment opportunities in corporate finance. If the required rate of return (may be a risk-free rate) is known, then the difference between the IRR and the required rate of return is called the implied risk premium (Damodaran, 1999). Murphy (2006) indicates that the usual IRR of investing a life settlement ranges from 15% to 18%. The average T-bill rate in 2006 is 4.68%. If the average T-bill rate is assigned as the required rate of return, the implied risk premium ranges from 10.32% to 13.32%.

From Table 6.3, we can obtained the prices of a life settlement under different

[4]Conventry is a global financial services firm leading the development of the life settlement market. It is also one of the originators and the creators of the life settlement market in the United State.

required rate of return	purchase price (our model)	expected IRR	expected risk premium
1%	41,224.31	4.3356%	3.3356%
2%	38,001.25	6.9273%	4.9273%
3%	35,170.19	9.5165%	6.5165%
4%	32,672.48	12.0994%	8.0994%
5%	30,459.51	14.6725%	9.6725%
6%	28,490.87	17.2327%	11.2327%
7%	26,732.74	19.7778%	12.7778%

Table 6.4: Calculating the risk premium of a life settlement by using IRR

discount rates. Because the discount rate has been under 7% since 1991, we assume that the life settlement company purchases the policy with different required rate of return ranging from 1% to 7%. Under our model, the purchase price is calculated with the adjusted length of years to live at 8.5 (the last column of Table 6.3). We then can calculate the expected IRR for each price and thus generate the implied risk premium. Table 6.4 shows the results. When the discount rate is about 5%, the implied risk premium is 9.6725% that is similar to calculations arrived at by Murphy (2006).

6.7 Summary

A life settlement is a financial instrument that enables a third party to purchase an existing life insurance policy from a policyholder. Because the policy was issued for an individual, the mortality risk mostly depends on the individual's health conditions. When the individual negotiates the transaction with a life settlement company, some medical records may be disclosed. The traditional pricing methods either do not modify the pricing model for the availability of the personal medical information or do adjust the pricing method with the a multiplier method that actually cannot reflect all available information. We propose a method to adjust the probability distribution of death year by the minimum discrimination information estimation. Our method is capable of adopting different kinds of information for adjustment, such as the number of years that the policyholder is expected to live and a given probability for which year that the policyholder is going to die.

We first use a variety of published mortality tables, such as the PR-2000, the 2008 VBT, and the CSO 2001 to implement our method. We also include the deterministic life settlement pricing for comparison. The deterministic life settlement pricing produces the lowest price since it does not place a value on the randomness of the death year. Settlement prices from published mortality tables are similar, and we compare the prices from 2008 VBT with prices from our proposed model. Our model produces lower settlement prices because we take mortality improvement into consideration. This lowers the probability of death in the early years due to the adaption of the mortality improvement.

There is no common way to calculate risk premium for a life settlement. The pricing is generally calculated based on the expected actuarial present value. We propose the application of the IRR to calculate the implied risk premium. The IRR has been frequently used in corporate finance for measuring the profitability of an investment opportunity. Given a required rate of return, we can calculate the purchase price and the expected IRR. The difference between the IRR and the required rate of return is the implied risk premium. Our estimation results are close to the practical experience in Murphy (2006).

Chapter 7

Pricing Mortality/Longevity Linked Securities

7.1 Introduction

Underestimating life expectancy whether by individuals or governments causes longevity risks. Indeed, in recent decades, the individual cost of aging has risen, but our focus here is the increased liabilities pension funds and insurance companies have incurred because people in developed countries are living longer than expected.(Biffis and Blake, 2009). As mentioned previously, improved medical care and technological advances will continue to lower the death rate and simultaneously longevity risks will increase. Public spending on senior citizens is increasing and pension funds and insurance companies are increasing their payouts to pensioners and retirees. Table 7.1 shows age related pension and public health expenditure projections. We see that the projected increases are high in Europe, Japan and the United States.

	Country	Pension Expenditure		Public Health Expenditure		Total	
		2010	2030	2010	2030	2010	2030
EMG6	Brazil	8.5	9.8	3.6	5.1	12.1	14.9
	China	2.2	2.4	1.9	2.8	4.1	5.2
	India	1.7	2.1	1.1	1.5	2.8	3.6
	Mexica	2.4	4.5	2.7	3.8	5.1	8.3
	Russia	9.4	14.0	3.5	4.6	12.9	18.6
	Turkey	7.3	10.5	3.5	4.8	10.8	15.3
G6	France	13.5	14.2	9.0	10.5	22.5	24.7
	Germany	10.2	11.5	8.1	9.0	18.3	20.5
	Italy	14.0	14.8	6.9	7.5	4.1	22.3
	Japan	10.3	10.1	6.8	7.8	17.1	17.9
	United Kingdom	6.7	7.6	7.3	10.6	4.1	18.2
	United States	4.9	6.0	7.6	12.7	12.5	18.7

Source: IMF, Credit Suisse.

Table 7.1: Age related government expenditure projections (in percent of GDP)

Better means of mitigating longevity risks are needed to keep pension funds and insurance companies solvent so that they can adequately meet the needs of people living longer than ever expected.

Traditionally, pension funds and insurance companies have used insurance and reinsurance to mitigate risk. Today, these approaches lack the capacity and liquidity to support the gigantic global exposure currently estimated to be about $20 trillion (Loeys et al., 2007; Biffis and Blake, 2009). Capital markets do not have the same liquidity and the capacity issues associated with insurance and reinsurance and thus can provide more transparent and competitive pricing for longevity risks. In 2006 United Kingdom pension funds began transferring their longevity risks to capital markets (ARTEMIS, 2012). Market size from 2007 to 2010 was approximately £8 billion each year (Chang, 2011). In 2011, market size rose to £12.4 billion. Although the numbers are not currently available, in 2012 hedging transactions are expected to exceed £30 billion (Kiff, 2012).

Currently there are four ways to transfer longevity risks to capital markets: buy-outs, buy-ins, longevity bonds, and longevity swaps. Buy-out and buy-in financial instruments attempt to mitigate not only longevity risks but also demographic and market risks. A buy-out transaction means that all pension fund liabilities are ceded to an insurer[1] through a bulk annuity. The pension fund is fully discharged of liabilities and the uncertainties of asset returns. The pensioner retains no connection to the original pension fund. Therefore, a buy-out not only transfers longevity risks but also other types of risk, such as demographic risks and market risks. The first buy-out transaction was made between Paternoster and Cuthbert Health Family Plan in November 2006[2]. Various buy-out transactions followed increasing in value each time. In June 2012, a buy-out transaction between General Motors and Prudential was valued at $26 billion .

Increased life expectancy and high cost to transfer the full amount of risk makes buy-out transactions prohibitively expensive. The transaction of a buy-out keeps reaching a record high, many pensioners are now facing large funding deficits due to the underestimate of longevity risks and cannot always afford the cost to transfer the full amount of the risk. Mercer Global launched a pension buyout Index[3] in 2010.

[1]The insurer is referred as a regulated life insurer defined by the Financial Services Authority (FSA) in the United Kingdom.

[2]The Cuthbert Health Family Plan is the pension scheme of a former Lloyd's of London underwriter. There are approximately 50 funds in the scheme. The transaction amount is not disclosed but commonly understood to be around £10 million.

[3]The Pension Buyout index is based on the pricing data from Aviva, Legal General and Pensions Insurance Corporation to provide companies with a monthly snapshot of the affordability of a buy-

They anticipate it will cost a mature pension scheme with liabilities and assets of £100 million an additional £44 million to complete a buy-out transaction with an insurer (Investment and Pensions Europe, 2010; Mercer Global, 2010). Buy-out transactions are costly, therefore, most longevity risks hedgers prefer buy-in transactions that have lower premiums and are easier to implement (Chang, 2011).

A buy-in hedge is similar to a buy-out although the pensioner continues to be connected with the original pension fund. A buy-in can be viewed as a partial buy-out. In a buy-in transaction, the pensioner purchases an insurance contract that guarantees a portion of the benefits to transfer the liabilities. For example, the fund could elect to transfer only the pension of deferred members[4] that are a subgroup of the pension members. The transaction can also include the transfer of liabilities that come from the payable over a limited time-horizon, such as the liabilities with more than 10 years maturity. A buy-in transaction has a lower premium and is usually a part of the de-risking strategy for reducing the risk exposure of a pension plan. The first buy-in transaction was made between Lane Clark and Peacock & Hunting PLC in the United Kingdom in January 2007, and the value of the transaction was £100 million. Many buy-in transactions followed since then but they all took place in the United Kingdom. The first buy-in deal outside the United Kingdom took place in Canada in 2009 involving Sun Life Financial, and the value of transaction was $50 million. Hickory Springs Manufacturing Company and Prudential Retirement transacted the first buy-in in the United States worth $75 million dollars. The fact that the cost of a buy-in is lower makes that buy-ins are now more popular than buy-

out or buy-in.

[4]A deferred member is a member who has left service or opted out and is currently entitled to a deferred pension.

outs among pensioners for the de-risking purpose and sometimes for the investment strategy tilting toward liability hedging (LCP, 2011; Biffis and Blake, 2009).

Unlike the buy-ins and the buy-outs, which are instruments to mitigate all sources of risks, the transaction of a longevity bond and a longevity swap are used to transfer only the longevity risks. The first attempt of issuing a longevity bond was in November 2004 when BNP Paribas announced the issue of a 25-year bond linked to a cohort survivor index based on the realized mortality rates of English and Welsh males aged 65 in 2003 . The bond was issued by European Investment Bank (EIB) and is commonly known as the EIB bond. However, the bond was not actually launched due to insufficient investor demand.

The mortality catastrophe bond is another type of bond that is linked to mortality rates by the mortality index. The short-term structure and the catastrophe bond feature make the mortality catastrophe bonds successfully marketed by Swiss Re since 2003. Learning lessons from the failure of the EIB bond and the success of mortality catastrophe bonds, Swiss Re launched a longevity catastrophe bond in 2010, which has the design similar to mortality catastrophe bonds. Beginning in 2012 and continuing until 2020, the Swiss Re longevity bonds hedge relative changes in the mortality rate of United States males between the ages of 55 to 65 and United Kingdom males between the ages 75 to 85. A new sense of urgency to manage longevity risks combined with investor anticipated short-term wins makes longevity catastrophe bonds very popular financial instrument investments. There is a difference between the EIB bond and a longevity catastrophe bond. The EIB bond was designed to offer pensioners with a price which they could hedge for the

Index	Time Launched	Population	Index Group
Credit Suisse Longevity Index	2005	U.S.	Overall, gender and age-specific sub-indices
J.P. Morgan Life-Metrics Index	2007	US, England and Wales, the Netherlands and Germany	Overall, gender and age-specific sub-indices
Deutsche Börse Xpect Age and Cohort Indices	2008	Germany, the Netherlands and England and Wales	Overall, gender and age-specific sub-indices

Source: Credit Suisse and the Bloomberg Professional service.

Table 7.2: Longevity indices

longevity risks, but the longevity catastrophe bond was designed to offer a premium to investors who are willing to assume some longevity risks from Swiss Re.

Longevity swaps[5] are another derivatives that concern only longevity risks, thus they have lower transaction costs. A longevity swap is essentially tied to a longevity index, such as J.P. Morgan LifeMetrics, the Credit Suisse Longevity Index, and Deutsche Börse Xpect Age and Cohort Indices. Table 7.2 shows the detailed information of the indices[6]. Credit Suisse and Babcock completed the first longevity swap in May 2009, in the amount of £750 million. Several longevity swaps have followed including a successful 10-year q-forward lunched by J.P. Morgan in 2008. This J.P. Morgan contract is the world's first longevity derivative that involves market

[5]There are two types of the longevity swaps, indemnity-based and index based. This paper concerns index-based swaps because they are easier to be construct than indemnity-based swaps and popular among pension funds.

[6]The Goldman Sachs QxX index launched in 2008 is not included in the discussion. The target population of the QxX index does not reflect the general population. The data refers to 46,290 people aged 65 or older with an impairment other than HIV-AIDS. The QxX was discontinued in 2010 due to lack of commercial activity.

transactions. In February 2011, J.P. Morgan executed another 10-year q-forward contract with the Pall U.K. pension fund in the amount of £70 million (Mortimer, 2012; Cobley, 2012). It is the world's first longevity swap for non-pensioners [7]. After the introduction of longevity swaps the buy-out and buy-in markets shrank by more than 50 percent (Wang, 2011). In 2012 Deutsche Bank and the Dutch insurer Aegon transacted a 12 billion Euro dollars longevity swap, the largest to date. The demand of low-cost longevity swaps shows that investment banks are eager to write more of them. Thus, the investment banks recognize the necessity of a standard of a longevity index and the format of a contract, which leads to the launch of the Life and Longevity Markets Association (LLMA) in 2010.

The LLMA is professional organization for those involved with life and longevity risks related products. The LLMA, establishes guidelines, standards, best practices, and pricing structures to promote liquidity and transparency in the trading of financial instruments that are linked to longevity or mortality related risks. Deutsche Bank, J.P. Morgan, Morgan Stanley, Munich Re, Prudential PLC, Swiss Re, and several other insurance companies and pension funds are active members of the LLMA.

In the following section we show the pricing for a q-forward. We give an example of a q-forward pricing under an LLMA pricing structure. We also show the pricing of a mortality catastrophe bond and the EIB bond. We apply our proposed model to arrive at the mortality prediction .

[7] A non-pensioner means either an active member or a deferred member of a pension scheme in an occupational pension scheme. An active member is a member of a pension scheme who is in the employment and accruing benefits under the scheme.

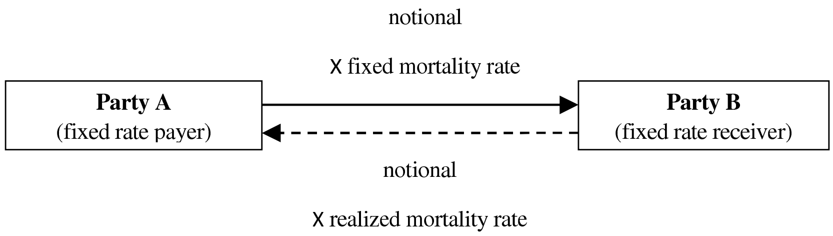


Figure 7.1: A q-forward transaction at maturity

7.2 q-forward

A q-forward is named after the letter q that is denoted as the mortality rate in actuarial science. A q-forward is linked to the mortality rate and is used to hedge the longevity exposure particularly for pension or annuity funds. Figure 7.1 illustrates the transaction of a q-forward at maturity for two parties. Party B is the hedger, typically a pension fund. A decreasing mortality rate means that more people will live longer than expected, meaning the pension fund will have to pay out more than originally anticipated. The pension fund hedges this longevity risks by exchanging the varying payment generated from a floating mortality rate for the fixed payment. In other words, the pension fund wants to ensure a positive cash flow to hedge the longevity risks when the mortality rate is lower than expected. Party A is the fixed mortality rate payer, usually an investment bank. The bank receives the payment linked to the varying mortality rate and takes over the longevity risks.

Denote the realized mortality rate at maturity as $m_{realized}$, and the fixed mortality rate in the contract as m_{fixed}. The net settlement amount for the bank at maturity is shown as follows

$$\text{notional amount} \times (m_{realized} - m_{fixed}), \tag{7.1}$$

and the net settlement amount at maturity for the pension fund, the counter party, is shown as follows

$$\text{notional amount} \times (m_{fixed} - m_{realized}).$$

Table 7.3 shows an illustration of a q-forward net settlement for the bank when outcomes of the realized reference rate vary. If the realized rate is lower than the fixed rate, the bank needs to pay the net settlement to the pension fund, which means the fund is protected from a decreasing mortality rate. Table 7.4 shows the illustrative term sheet for a single q-forward for hedging the longevity risks. We notice that the fixed rate is pre-determined in the contract. The present value of the net settlement of the bank is to discount (7.1) with the risk-free rate or the required rate of return of the bank. The discounted net settlement should be the premium for the bank, but the realized mortality rate is unknown. Therefore, replace the realized mortality rate to the expected mortality rate and then discount value of the net settlement is the premium that the bank charges the hedger.

Ideally, the net settlement is zero when the fixed rate and the realized rate are the same, meaning that the risk is hedged perfectly and the bank does not receive any payment from the transaction. However, the risk from the mortality improvement is inherent in the realized mortality rate varying over the contract duration, and

Reference Rate (Realized Rate)	Fixed Rate	Notional (GBP)	Settlement (GBP)
1.00%	1.2%	50,000,000	10,000,000
1.00%	1.2%	50,000,000	5,000,000
1.20%	1.2%	50,000,000	0
1.30%	1.2%	50,000,000	-5,000,000

Source: J.P. Morgan.

Table 7.3: An illustration of q-forward net settlement for a bank when outcomes of the realized reference rate vary.

the bank takes this risk from the transaction. The bank should be rewarded for taking the risk, so the bank will propose a fixed rate that is generally lower than the expected mortality rate in order to have a positive cash flow at maturity. The fixed rate is calculated according to not only the information of the expected mortality rate but also the risk tolerance of the bank, so it is also called a forward rate.

7.2.1 LLMA structure

The LLMA structure of a q-forward provides a simple way to quantify the mortality risk. The fixed rate, assigned by the bank, contains the information of the future mortality rate and the risk appetite of the bank. The LLMA proposes the structure which was originally used by J.P. Morgan when it proposed the q-forward in 2007 shown as follows

$$m_{x,t} = m_{x,0} \prod_{i=1}^{t} (1 - (\hat{m}_{x,i} + \xi)), \tag{7.2}$$

where the x can be an age, an age group, or a group of age groups, the $m_{x,t}$ is the forward rate at future time t, $\hat{m}_{x,i}$ is the best estimate of the mortality improvement

Notional Amount	50 million GBP
Trade Date	31 December 2006
Effective Date	31 December 2006
Maturity Date	31 December 2016
Reference Year	2015
Fixed Rate	1.20 %
Fixed Amount Payer	J.P. Morgan
Fixed Amount	Notional Amount × Fixed Rate × 100
Reference Rate	LifeMetrics graduated initial mortality rate for 65-year-old males in the reference year for England and Wales national population Bloomberg ticker: LMQMEW65 Index <GO>
Floating Amount Player	XYZ Pension
Floating Amount	Notional Amount × (Realized) Reference Rate × 100
Settlement	Net Settlement= Fixed Amount-Floating Amount

Source: J.P. Morgan.

Table 7.4: An illustrative term sheet for a single q-forward to hedge longevity risks

rate, and ξ is the adjustment term for the risk appetite. The forward rate is the fixed rate given in the standard q-forward contract. The $\hat{m}_{x,i}$ involves the information of how the future mortality rate evolves. We can re-write (7.2) by dropping ξ as shown below to understand the LLMA structure better

$$m_{x,t} = m_{x,0} \prod_{i=1}^{t} (1 - \hat{m}_{x,i}). \tag{7.3}$$

The (7.3) is similar to the estimation of the mortality rate with the mortality improvement rate shown as follows

$$m_{x,t} = m_{x,0} \prod_{i=1}^{t} (1 - r_{x,i}). \tag{7.4}$$

where $r_{x,k}$ is the mortality improvement rate for the age x at time i defined as follows

$$r_{x,i} := 1 - \frac{m_{x,i}}{m_{x,i-1}}.$$

Compare (7.4) and (7.3), we can find that the best estimate $\hat{m}_{x,i}$ is estimated as a mortality improvement rate.

The LLMA suggests an average mortality improvement rate as the best estimate of the mortality improvement rate. If the reference mortality rate is the rate of an age group or a group of age groups, such as males, 65 to 69 years old, the best estimate can be the average predicted mortality improvement rate of the males who belong to the range of age over the contract duration. If the reference rate is for a specific age, the best estimate of the mortality improvement rate can be estimated by the average predicted mortality improvement rate over the contract duration. Therefore, the (7.3) can further be simplified as follows

$$m_{x,t} = m_{x,0}(1 - \hat{m}_x^t)^t,$$

where $\hat{m}_x^t$ is the average predicted mortality improvement rate over the contract duration t.

The (7.3) shows the forward rate is the expected mortality rate at maturity and contains no information of the bank's risk appetite. If the bank uses (7.3) as the fixed rate, it means that the bank expects zero net payment because the forward rate is the expected mortality rate. In the q-forward transaction, the bank receives the floating payment and the decreasing mortality rates result in a lower payment.

The variation and the trending-downward mortality rate is the risk taken over by the bank. Therefore, the forward rate proposed by the bank will be lower than the forward rate in (7.3), i.e. the expected mortality rate, so that the bank can have a positive future cash flow at maturity. The mortality improvement rate is thus estimated as a best estimate plus an adjustment term that incorporates the information of the bank's risk appetite shown in (7.2). In the standard q-forward contract, the fixed rate is given and we can calculate the ξ which is used to measure the risk appetite of the transaction. If the bank perceives more improvement of the mortality rate, the bank increases the value of the ξ.

7.2.2 Graduation

The purpose of graduation is essential because the age-specific mortality rates for each calendar year are generally not available. Mortality data are mostly presented in an abridged form, such as the values of death rates are shown at age 0, age group 1 to 4, 5 to 14, 15 to 24, and so on up to 75 to 84, and the open age group 85 and over. Such a layout is not sufficient for the computation of some monetary functions involving life contingencies, such as annuities that require probabilities of death for every single year of age.

The published data for mortality rates are age-group-specific, and they should be graduated to be age-specific mortality rates for most actuarial applications such as the construction of a life table. The graduated mortality rates should incorporate the important property of the mortality rate: the mortality rates for adjacent ages are similar and highly correlated. Besides, the graduation process should not destroy the integrity of the data, and the process should magnify the information yet reduce the noises in the mortality rates.

The traditional graduation methods are parametric, such as Beers ordinary minimized fifth difference method and the Gompertz-Makeham method. Although they are commonly seen in demographical and actuarial science literature, these parametric models can be applied only when we understand the data very well. Both LLMA and J.P. Morgan suggest a non-parametric method, a smoothing cubic splines method, which provides a supurb fit and is much better than the traditional methods for graduation purpose. J.P. Morgan (2007b) runs several tests to compare the goodness of fit of these three graduation techniques, including signs test, runs test, Chi-square goodness of fit test, serial correlations, standardized deviations test, and cumulative deviations test. The report concludes that the smoothing cubic splines method provides the flexibility for fitting the data best.

Splines are piecewise polynomial functions with the assumption that each spline is differentiable (for some defined number of times) at knots, which are connecting points of those polynomial functions. Therefore, splines are determined by the locations of knots and the coefficients of the polynomials. The cubic spline is a smoothing function defined by a polynomial of degree three, which fits cubic curves among data points.

The smoothing cubic splines method is the splines method with a modification for better goodness of fit. The modification is made through a smoothing parameter used to penalize too much smoothing which deteriorates the goodness of fit (see Appendix C). A small value of the smoothing parameter could result in a too rough curve while a large value could result in non-monotonic mortality rates. J.P. Morgan (2007b) runs several tests to find an appropriate estimation interval for the value of the smoothing parameter. The report concludes that the appropriate interval of the value for the mortality rate of the United States is between 0.2 and 0.8, while the interval for the England and Wales data is between 0.2 and 0.4. The LLMA (2012) suggests the value 1/3 for the smoothing parameter.

The data we applied for construction of in-sample estimates and out-sample predictions are the death rates for males in England and Wales from 1960 to 2004[8]. The data are available for download from Human Mortality Database. We first fit data with the modified Lee-Carter model and obtain the estimates of the parameters for the Normal Inverse Gaussian (NIG) distribution. We then use these estimates and the modified Lee-Carter model with the corresponding NIG Lévy process and the Esscher transform to predict 10-year mortality rates by running 100,000 simulations. Since the reference mortality rate of the J.P. Morgan q-forward shown in Table 7.4 is the mortality rates for 65-year-old males, we apply smoothing cubic splines method on our estimation results which are age-group specific mortality rates to obtain the age-specific mortality rate for 65-year-old males. The selected value of the smoothing parameter is within 0.3 to 0.4, which is consistent to J.P. Morgan (2007b) and the reports of LLMA.

Figure 7.2 shows the fitted and prediction results compared to the LifeMetrics index. The LifeMetrics index is estimated by eight models[9] with some adjustments from the raw data of population and deaths. We first observe that our fitted mor-

[8]See previous chapters for details.

[9]J.P. Morgan (2007) applies Lee and Carter (1992), Renshaw and Haberman (2006), Currie (2006), Currie et al. (2004), Cairns et al. (2006), and three extensions to Cairns et al. (2007).

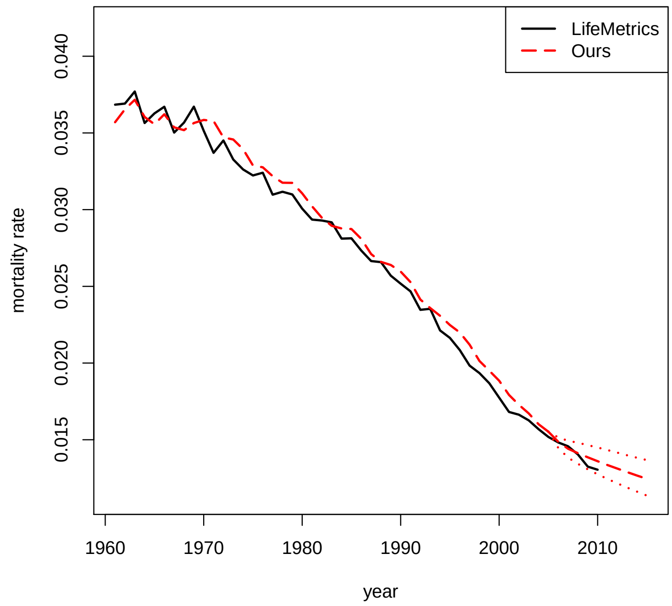


Figure 7.2: The mortality estimation and prediction for age 65 males in England and Wales

tality curve is close to the LifeMetrics index for the in-sample fitting (from 1961 to 2005). Besides, our model also preserves the jumps that are shown in the LifeMetrics index. We predict the mortality rate for year 2006 to 2015, which is the contract duration of the J.P. Morgan q-forward shown in Table 7.4. Since the LifeMetrics index is updated to year 2010, we can compare part of our prediction results with the LifeMetrics index. We can find that our prediction results of the mortality rate for year 2006 to 2010 are also close to the LifeMetrics index.

7.2.3 Estimation

Consider the fixed rate as 1.2% given in the J.P. Morgan q-forward shown in Table 7.4. We calculate the premium for the q-forward by applying formula (7.2). The best estimate of the mortality improvement rate for the year i during the contract duration is formulated as the mortality improvement rate, $\hat{m}_{i,65}/\hat{m}_{i-1,65}$, and the initial mortality rate (of year 2005) is $m_{0,65} = 1.5170\%$. We find the corresponding ξ by solving the equation shown as follows

$$1.2000\% = 1.5170\% \prod_{i=1}^{10} \left(1 - \left(\frac{\hat{m}_{i,65}}{\hat{m}_{i-1,65}} + \xi\right)\right).$$

The ξ is 0.3769% and the resulting premium is £1,534,992.

7.2.4 The fixed rate

The fixed rate in the standard q-forward is given, but how does the bank determine that it is the right number? Nothing is mentioned regarding this in the LLMA documents. However, Loeys (2007) proposes applying the concept of an interest rate forward on the construction of the fixed rate. An interest rate forward is the market's expectation of the interest rate at a specific future time plus term premium to compensate investors for taking the duration risk. Therefore, the forward rate is the expected rate plus a mark-up value for the reward of taking the risk. The mark-up value increases when the duration becomes longer, because the variation and uncertainty of the change of the interest rate increases over time. This is similar to a q-forward except that the interest rate is replaced by a mortality rate.

The reference mortality in the J.P. Morgan q-forward contract shown in Table 7.4 reflecting the mortality rate for 65 year old English and Welsh males in 2015, meaning that they were 55 years old in 2005. The longevity risks of this cohort of males is the risk that the pension fund wants to transfer. The risk that the bank carries over is actually from this cohort so the bank should calculate the forward rate according to this cohort. The risk premium during the contract duration is calculated by the Sharpe ratio suggested by Loeys (2007).

The Sharpe ratio is originally defined as excess return to cash dividends divided by the volatility of return. Since the mortality rate in a q-forward transaction functions similar to the interest rate, the Sharpe ratio is defined as the excess mortality rate to expected mortality rate divided by the volatility of mortality rate. The long-term returns of equities and bonds can generate a Sharpe ratio between 0.2 to 0.3. Moreover, Burton et al. (2008) mention that a simple buy-and-hold strategy for ten year swaps can return a Sharpe ratio of 0.25. Loeys (2007) suggests a Sharpe ratio of 0.25 for the calculation of the forward rate of a q-forward, since the rate of return of a q-forward has to be at least as good as the return from equities and bonds. Thus a q-forward can attract the interest of investors.

Loeys (2007) proposes the formula for calculating the forward of a q-forward as follows

$$q_{forward} = (1 - t \times \text{Sharpe Ratio} \times \sigma) \times q_{expected}, \tag{7.5}$$

where t is the contract duration, σ is the annualized historical standard deviation of the reference age or age group, and $q_{expected}$ is the expected mortality rate of the reference year. We will use (7.5) to calculate the forward rate for the J.P. Morgan

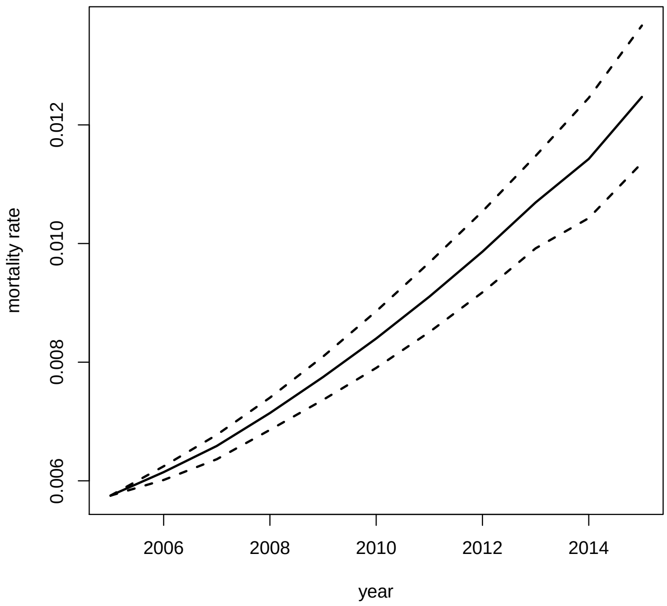


Figure 7.3: Mortality estimations and predictions for 55 year-old English and Welsh males in 2005

contract shown in Table 7.4 with the Sharpe ratio 0.25.

The first step in calculating the forward rate is to forecast the ten year mortality rates for the cohort of English and Welsh males who were 55 years old in 2005. We apply the modified Lee-Carter model with NIG Lévy process and the Esscher transform and simulate 100,000 times. Figure 7.3 shows the expected mortality rates and the 95% confidence interval of the expected mortality rate of the cohort. The confidence interval becomes larger over time which means the uncertainty of the mortality rate movement increases. Table 7.5 shows the estimation results for males aged from 56 to 65 in 2006 to 2015. The numbers in bold are the expected mortality rates for the English and Welsh cohort. The mortality rates of the cohort

		year				
		2006	2007	2008	2009	2010
	56	**0.5990**	0.5891	0.5891	0.5790	0.5691
	57	0.6788	**0.6588**	0.6475	0.6362	0.6251
	58	0.6765	0.7273	**0.7142**	0.7014	0.6889
	59	0.7476	0.8052	0.7900	**0.7749**	0.7603
age	60	0.8284	0.8909	0.8732	0.8562	**0.8399**
	61	1.0147	0.9846	0.9647	0.9458	0.9277
	62	1.1195	1.0862	1.0643	1.0435	1.0239
	63	1.2320	1.1958	1.1722	1.1500	1.1289
	64	1.0139	1.3063	1.2792	1.2539	1.2301
	65	1.1462	1.4406	1.4121	1.3856	1.3606
		year				
		2011	2012	2013	2014	2015
	56	0.5593	0.5497	0.5402	0.5310	0.5219
	57	0.6141	0.6034	0.5928	0.5827	0.5725
	58	0.6767	0.6647	0.6529	0.6417	0.6305
	59	0.7462	0.7320	0.7182	0.7051	0.6920
age	60	0.8241	0.8084	0.7932	0.7787	0.7643
	61	**0.9104**	0.8931	0.8765	0.8606	0.8449
	62	1.0051	**0.9863**	0.9682	0.9510	0.9340
	63	1.1087	1.0886	**1.0693**	1.0508	1.0326
	64	1.2072	1.1848	1.1633	**1.1426**	1.1225
	65	1.3365	1.3129	1.2903	1.2683	**1.2471**

Table 7.5: The mortality rate for a 55 year-old male in 2005

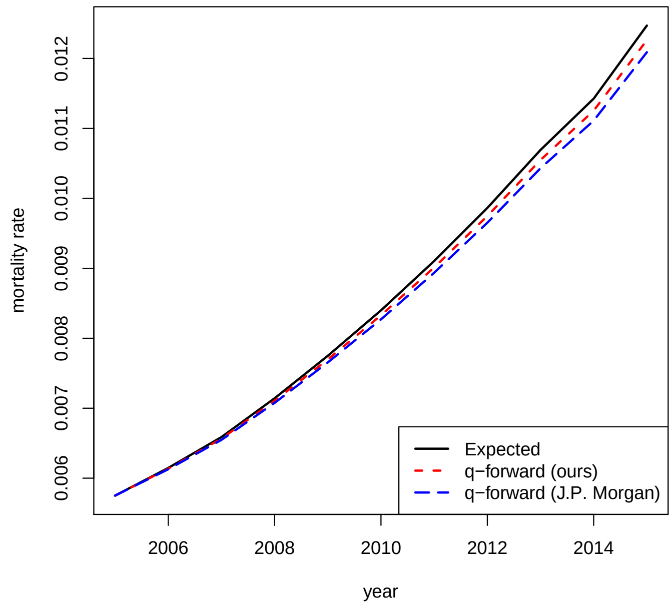


Figure 7.4: The expected mortality rates and forward rates for 65 year-old English and Welsh males in 2005

of males in England and Wales who was 55 years old in 2005 is increasing over time because the fact that the mortality rate increases with age.

The second step is to calculate the forward rate. The historical standard deviation of the mortality rate for the 65 year old English and Welsh males is 0.6692%. The expected mortality rate is 1.2471% as shown in Table 7.5. We can calculate the forward rate by applying formula (7.5) and the result is 1.2262%, which is higher than 1.2% given in the J.P. Morgan q-forward contract. Figure 7.4 shows the expected mortality rate, the forward curve from our model, and the forward curve from the J.P. Morgan q-forward contract. We can observe that the risk premium for the duration risk of the J.P. Morgan q-forward is very high. The corresponding

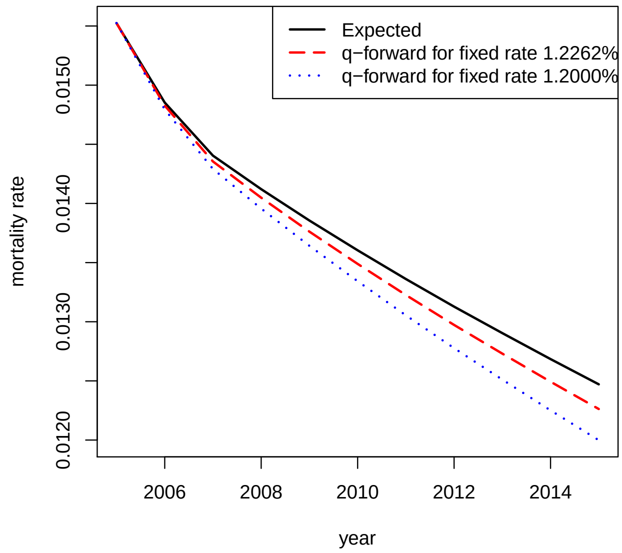


Figure 7.5: The expected mortality rate and the forward rate for 65 year-old English and Welsh males

Sharpe ratio is 0.4457, which is higher than the suggested value 0.25.

We can apply the forward rate 1.2262 % calculated from our model as a fixed rate to obtain ξ and risk premium for a q-forward. The resulting ξ is 0.1650% with the corresponding risk premium £680,138. Figure 7.5 shows the expected mortality rate curve, the forward rate curve generated from the 1.2262% fixed rate obtained from our model, the forward rate curve generated from the 1.2% assigned in the J.P. Morgan contract. Since J.P. Morgan proposes the 1.2% fixed rate much lower than ours, the risk adjustment term ξ is much higher than ours.

Our proposed model has an in-sample estimation similar to the LifeMetrics index that is computed from several other models with some complicated adjustments. Our model also provides a higher forward rate with a lower premium that reduces

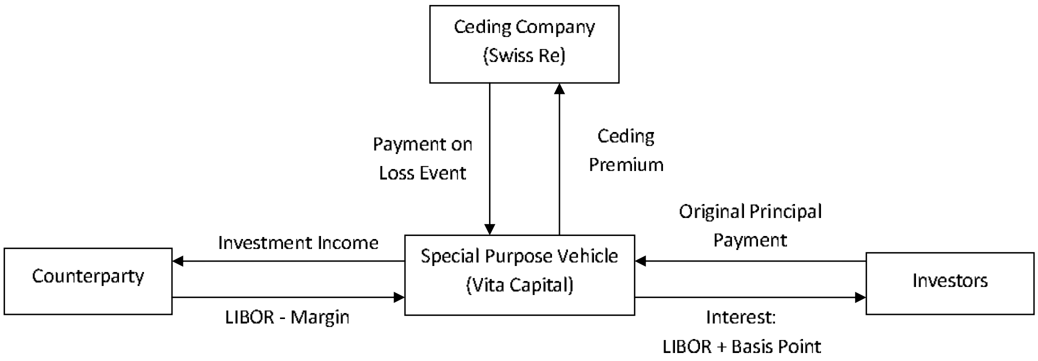


Figure 7.6: The basic structure of a mortality catastrophe bond lunched by Swiss Re

the costs of hedging the longevity risks. This could be the incentive to stimulate interest in hedging longevity risks.

7.3 Mortality Catastrophe Bond

The Swiss Re successfully lunched and marketed a mortality catastrophe bond in 2003. The basic structure of a mortality catastrophe bond is shown in Figure 7.6. The purpose of a mortality catastrophe bond is to mitigate the risk of a catastrophic event that would cause a life insurance company to compensate all the beneficiaries at once.

The bond is triggered by a dramatically rise in the mortality rate for a certain population, which is similar idea to the catastrophe bond. The first mortality catastrophe bond was issued by Vita Capital, as a special purpose vehicle enabling Swiss Re to remove extreme catastrophe risk from its balance sheet. The bond had a three year maturity, a principal of $400 million and a 135 basis point coupon rate plus

LIBOR. The reference mortality index is calculated from a certain age group and is a weighted average value from five countries[10]. The principal is paid in full if the reference mortality index does not exceed 1.3 times its 2002 base level before the maturity date. The scheduled payment is shown as follows

$$x_t = \begin{cases} \text{LIBOR+spread} & , t = 1, ..., T-1 \\ \text{LIBOR+spread} + \max(0, 1 - \sum_t L_t) & , t = T \end{cases}$$

where L_t is shown as follows

$$L_t = \begin{cases} 0 & , \text{ if } m_t < 1.3m_0 \\ (m_t - 1.3m_0)/0.2m_0 & , \text{ if } 1.3m_0 \leq m_t \leq 1.5m_0 \text{ for all } t \\ 1 & , \text{ if } 1.5m_0 < m_t \end{cases}$$

The goal of pricing is to know what the risk premium in term of interest rate.

Demographic information is not available for all countries; therefore, we only use United States mortality rate data from 1900 to 2002. The weighted average mortality index is calculated from the weight of each age group that is obtained from year 2000 mortality estimation for the standard population. Table 7.6 shows the weight for each age group of the United States based on the year 2000 standard population.

We first use the modified Lee-Carter model with the NIG distribution and Lévy process but without the Esscher transform to forecast the mortality rate. We run 100,000 simulations. For each simulation, we calculate the internal rate of return

[10] The weights for five countries are U.S.A. 70%, U.K. 15%, France 7.5%, Italy 5%, and Switzerland 2.5%.

age group	weight
<1	0.013818
1-4	0.055317
5-14	0.145565
15-24	0.138646
25-34	0.135573
35-44	0.162613
45-54	0.134834
55-64	0.087247
65-74	0.066037
75-84	0.044842
>84	0.155080

Table 7.6: The weights for age groups of the United States (year 2000)

(IRR) of the bond from an investor's perspective. The average IRR is 0.66%. We then forecast the mortality rate with the Esscher transform, and the average IRR is -0.22%. The mortality rate prediction with the Esscher transform considers the downward trend of the mortality rate, so it shows less inclination to have a large positive mortality rate jump. Therefore, it is less likely to have a mortality rate much higher than the base level of 2002. Moreover, if we consider the mortality catastrophe bond as a tool to mitigate the extreme risk of a catastrophic rise in the mortality rate, then the premium, in terms of interest rate, can be viewed as the difference between the risk-free rate and an IRR. In our example, the IRR is -0.22% and the risk-free rate is the LIBOR, so the premium that a pension fund pays to be protected from a mortality jump is (LIBOR + 0.22%). Although there are other ways to quantify the risk premium, such as Blake et al. (2006), applying an IRR calculation may be the most intuitive way to understand the cost of hedging

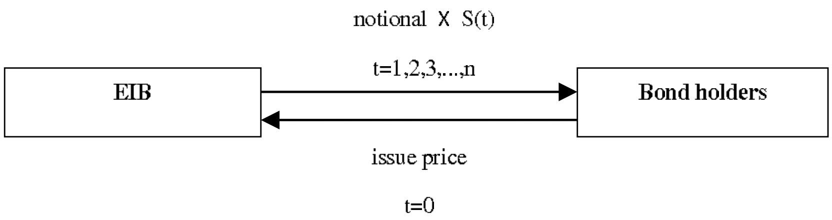


Figure 7.7: The cash flows from an EIB bond as viewed by investors

a catastrophic rise in the mortality rate.

7.4 EIB Longevity Bond

Although the 2004 EIB longevity bond never launched due to a combination of issues discussed previously, it did receive a great deal of public attention. The bond was issued by the European Investment Bank, and Partner Re was the reinsurer for the longevity risks. The face value of the bond is 540 million GBP with a 25-year maturity. The bond had floating coupon payments that were linked to a cohort survivor index based on the realized mortality rates of English and Welsh males age 65 in 2003. Figure 7.7 shows the cash flows from an EIB bond as viewed by investors.

Blake (2005) mentions that the risk-adjusted value for the bond should be shown as follows

$$V(0) = \sum_{i=1}^{25} P(0,i)\mathbb{E}_Q(S(i)|\mathcal{M}_0) \tag{7.6}$$

where $P(0,i)$ is the discount factor and $\mathbb{E}_Q(S(i)|\mathcal{M}_0)$ is the expected risk-adjusted survival index (or survival rate) at time i under conditional on the current information set.

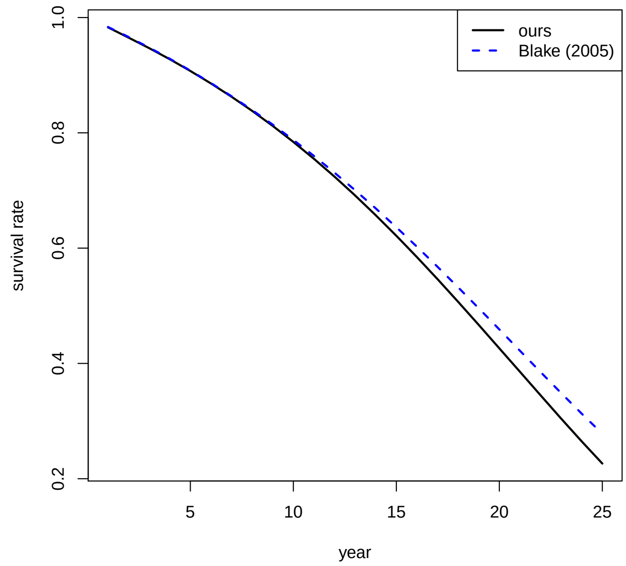


Figure 7.8: The comparison of the survival curves

To apply (7.6), we need to decide what discount rate and survival rate to use. For the expected risk-adjusted survival rate, we can apply our model to generate the mortality rate under a martingale measure. The resulting mortality rate has adjusts with the long-term mortality change and hence adjusts the mortality risk. We simply convert it to the survival rate and then apply the resulting survival rate to (7.6). Figure 7.8 shows the comparison of our results with Blake (2005). Our estimated survival curve is slightly lower than Blake (2005) but they are close to each other.

Now we need to decide what discount rate should be used. Blake (2005) suggests applying a 4% discount rate. We apply a 4% interest rate along with our mortality rate model and obtain 11.2533 for the price of the longevity bond that is close to

11.442 shown in Blake (2005). The choice of interest rate can be more sophisticated, because the duration of the bond lasts for 25 years. In financial modeling, a Cox-Ingersoll-Ross (CIR) model is usually applied for interest rate modeling. Introduced by John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross in 1985, it is widely applied to arrive at the valuation of interest rate derivatives. A CIR model for the interest rate dynamics is shown as follows

$$dr_i = \bar{k}(\bar{r} - r_i)dt + \bar{\sigma}\sqrt{r_i}dB_i$$

where r_i is the interest rate, $\bar{k}$, $\bar{r}$ and $\bar{\sigma}$ are the constants, and B_i is a standard Brownian motion. Deelstra and Delbaen (1995) furthermore show that the CIR model will can have the approximation property shown as follows

$$\mathbb{E}\left(\exp\left(-\int_0^t r_u du\right)\right) \sim \exp\left(\left(\frac{\bar{\sigma}^2}{2\bar{k}^2} - 1\right)\bar{r}i\right). \tag{7.7}$$

The $P(0, i)$ in (7.6) can be replaced by (7.7). By applying the historical annual LIBOR data from 1990 to 2003, (7.7) is estimated as $\exp(-0.039052)i$. The resulting price for the longevity bond is 11.2702 which is also close to 11.442 in Blake (2005).

7.5 Summary

Decrease in mortality rates has increased longevity risks creating an urgent need for pension funds and annuity providers to find ways to mitigate these risks. Insurance and reinsurance lack the capacity and liquidity to mitigate these risks so pension funds and annuity providers are now passing their longevity liability risks onto the capital markets.

Mortality/longevity linked derivatives such as q-forward, mortality catastrophe bonds, and the never launched but interesting EIB longevity bond are financial instruments pension funds and annuity providers may use to mitigate risk. Using our model we have shown how to price a q-forward under the LLMA structure. Our results show that we can obtain lower q-forward premiums reducing the cost of hedging longevity risks. The results also provide an incentive to stimulate more interest in using a q-forward to hedge longevity risks.

Mortality catastrophe bonds, similar to catastrophe bonds transfer the risk of a catastrophic mortality event to capital markets. Using our proposed model, we use the IRR to evaluate possible premium values. Since our model has adjusted to adopt the mortality improvement, the predicted mortality rates tend to be low. For a mortality catastrophe bond, the pensioner wants to hedge the risk when there is a sudden rise in mortality rate. The possible value of the premium is the difference between the LIBOR and the IRR .

We apply our model for pricing an EIB longevity bond proposed by Blake (2005). An EIB longevity bond is priced on the risk-adjusted survival rate, so we need to transfer the predicted mortality rate to the survival rate. We obtain the results close to Blake (2005).

Chapter 8

Summary

8.1 Findings and Contributions

Thanks to modern medicine, technology and increased standards of living in the developed world, mortality rates are in decline. However, this decline poses a longevity risk for pension and annuity providers. It is well known that actuarial tables do not accurately reflect mortality rates; therefore improved mortality models such as the Lee-Carter and its extensions have increased our understanding of the dynamics of the mortality rate. Working with these models reveals some shortcomings that we take into account with our proposed model. After comparing our mortality modeling results with previous models and seeing an improvement, we proceed to price mortality/longevity linked derivatives. Based on our proposed and improved mortality model, our results indicate lower premiums and reduced costs to hedge longevity risks. Since insurance and reinsurance are no longer capable of satisfying the liability needs of pension and annuity providers there is an urgent need for new financial instruments. Therefore, our proposed strategies should be an incentive for pension and annuity providers to consider financial instruments such as the q-forward, mortality catastrophe bonds and future EIB bonds to mitigate longevity risk.

Careful examination of the Lee-Carter model and its extensions reveal the following. The Lee-Carter model assumes that two factors, age effect and period effect, have impacts on the mortality rate. However, the model does not consider the jumps of mortality rate due to rare events or the current acceleration in the mortality improvement. The first extension of the Lee-Carter model adds non-Gaussian error terms representing the occurrences of abnormal jumps in mortality rates, but this extension does not solve the issue that the mortality dynamics are non-Gaussian. The second extension of the Lee-Carter model adds a cohort effect, which is called the age-period-cohort (APC) model. The cohort effect captures the phenomenon of accelerated mortality improvement for some birth cohorts. Although the APC model has shown success in mortality modeling in many countries, the cohort effect sometimes may not be significant. Additionally, it is almost impossible to convert the APC model to a stochastic mortality model making the pricing of the mortality/longevity linked derivatives impossible. Another extension is the stochastic mortality rate models. However, current stochastic mortality rate models are very complicated and require complex computational resources.

We propose the modified Lee-Carter model with Normal Inverse Gaussian (NIG) Lévy process and Esscher transform. Borrowing from a similar model used in financial modeling, our proposed mortality model improves mortality modeling results and enables accurate pricing of mortality/longevity linked derivatives. Replacing current mortality modeling methods that are constructed on the level of mortality rate with adaptations of Mitchell et al. (2013) idea to build the mortality model on the growth rate of the mortality we then can apply Lévy processes, found in financial modeling, to price the derivatives. An accurate predicted mortality rate and a narrower confidence interval for the stable forecasting result when using our proposed model. The application of the Esscher transform provides a short cut for converting the regular mortality rate to the mortality rate under a martingale measure which adapts the information of the mortality improvement.

Previously, lack of urgency, inefficient pricing and poor design all factored into the limited interest in mortality/longevity linked derivatives, such as q-forward, EIB longevity bond, and mortality catastrophe bonds. However, pricing these instruments based on our proposed model generally shows that reasonably priced premiums will be sufficient to create the hedge incentive, because that the cost to pension and annuity providers will decrease. This is because our model tends to predict lower mortality rates that accurately reflect real life.

The decrease in mortality rates is also a cause for increased interest in life settlement instruments. In many cases, individuals are either opting or in need of the money accrued in annuity accounts. Using our proposed model we also price life settlements. Using an information theory approach, we propose a pricing procedure that adjusts the price of a life settlement according to a policyholder's medical information. Since the pricing of a life settlement is based on the expected rate of return, there is still no common way to calculate the risk premium for a life settlement transaction. Again borrowing from corporate finance, we use the internal rate of return (IRR) to calculate the implied risk premium. The results are close to the practical experience mentioned in Murphy (2006).

Pension and annuity marketplaces are in urgent need of financial instruments to hedge their unprecedented $20 trillion in liabilities. Without such business, many companies will go bankrupt leaving the insured with nothing or having to be bailed out by governments who really do not have the means to do so. To keep pension and annuity providers liquid, they must enter the capital markets and be able to invest in financial instruments that cover their liabilities and give investors incentive to take on the risk. Carefully examining the Lee-Carter mortality model and its extensions have led us to create a new model that is easy to understand and compute. Additionally, by borrowing from established financial models we are able to price the q-forward, bonds and life settlement policies in a compelling manner so that pension and annuity providers as well as investors will be interested in partaking in these markets.

8.2 Future Studies

8.2.1 The Improvement of the Model

Mortality indices are comprised mostly from the mortality rates of elderly people. It is said that when a long-time spouse dies, the survivor may soon follow (Redfern, 2012). Such phenomenon is called the widowhood effect. A study from the University of Glasgow asked more than 4,000 married couples between the ages of 45 to 65 to determine their risk of death after becoming a widow or widower. The survey results show that death rates increase about 30 percent in six months following a spouse's death. Similar research results were reached from a University of St. Andrew longitudinal study of a Scottish population (about 58,000 men and 58,000 women). The results show that the widowhood effect could contribute about a 40 percent higher risk of death for the survivor (Pelley, 2011). In statistics or financial modeling, the correlation of the survival rates is called coupled survival rate or joint survival probability. The joint survival probability has been applied in biostatistics, financial modeling, and risk management for years. There is still no related research that applies the ideal of the joint survival probability for the pricing mortality/longevity linked derivatives. There is still no related research that applies the ideal of the joint survival probability for the pricing mortality/longevity derivatives. The application of the joint survival rate can be included in the model for more accurate forecasting.

We apply our mortality modeling on mortality rates from only one country. Could it be possible that the morality rates of two countries are correlated? Since globalization has substantially increased international travel, deadly infectious diseases could break out and quickly cross the borders (Daulaire, 1999; Lin et al., 2013). Also, recent financial innovation provides insurers and reinsurers a viable option to transfer the risk of catastrophic mortality caused by either epidemics or disastrous terrorist attacks. Almost all the transactions of mortality catastrophe securities are bound with three or more population mortality indices[1]. It is natural to question if these mortality indices are correlated. Lin et al. (2013) apply the technique of exponential tilting on the Lee-Carter model with Brownian motion specification

[1]The only one exception is the Tartan mortality bond sold in 2006.

to estimate the correlation of mortality rate among the United States, the United Kingdom, Japan, and German. They conclude that the price of risk can be very different depending on the correlation among the countries.

Recently, financial literature has widely applied the exponential Lévy process with the NIG distributional specification. One major reason is that a Lévy process can technically be thought as the combination of a Brownian motion and a jump process[2]. Therefore, the alternative choice of applying Lévy processes is to model asset returns with a Brownian motion and a compound Poisson process as a jump process, which is common to see in financial modeling. Mortality dynamics are similar to the time evolution of the rate of return in finance. It may be that other ideas from financial modeling can be borrowed applied for stochastic mortality modeling. For example, our proposed model contains no diffusion term. However, if the correlations are concerned, such as the couple survival rates and multinational mortality rates, the diffusion term can be used to generate such relationships. Huang and Wu (2004) suggest that a diffusion return component is useful in their time-changed Lévy process setting for generating correlations with the diffusive activity rate process (Wu, 2005). The multidimensional Lévy copula can also be incorporated to include the correlations among the mortality indices of different countries.

The NIG distributional assumption has been emphasized in the previous chapters because it is capable to capture the frequent normal jumps and rare large jumps. However, some extreme events, such as sudden breakout of a worldwide epidemic disease or a disastrous flood occur with extremely low frequency, and the NIG dis-

[2]By Lévy Ito decomposition, a Lévy process can be separate into four parts, a drift term, a Brownian motion, a jump process, and a martingale.

tributional assumption may not be sufficient for prediction. To take the extreme events into account, the extreme theory should be included for better performance of forecasting. The extreme theory is helpful for pricing derivatives that are related to the extreme event, such as a mortality catastrophe bond.

8.2.2 The Public Policy

In the developed countries, both private companies and governments provide the social security benefits and pension benefits that serve as an insurance for individuals against the situation when they outlive their resources (Brown and Orszag, 2006; Antolin and Blommestein, 2007). However, companies and governments are facing increasing liabilities due to longevity. Longevity risks can be decomposed into two parts, the systematic longevity risks and the specific longevity risks (Blake et al., 2013). The systematic risk is a trend risk that was resulted from the change of life style or the advances of medical technology and affects entire populations. The specific risk is the risk that an individual's set of mortality rates differs from the expected rates. The specific risk theoretically can be eliminated in the aggregate (e.g., at the pension fund or social security) level by pooling over individuals using the law of large numbers to reduce the variability of the risk. The systematic risk is generally caused from mortality improvement and is difficult to be diversified by its nature.

While the public sector still has a large proportion of their promised pension and security benefits unfunded[3], the private pension sector faces a not-so-well-developed

[3]In the United Kingdom, the most recent estimates for state pension liabilities were 3,843 billion in respect of social pensions, and 1,156 billion were unfunded among them (Blake et al., 2013). In

annuity markets (Brown and Orszag, 2006). The private sector especially has to address the issue of adverse selection. Mitchell et al. (1999) mention that those who purchase annuities tend, on average, to live longer than those who do not purchase annuities. The correlation among individual mortality rates is thus positive (Bohn, 2005) making the traditional way for diversification, say, pooling, not work as efficiently. Annuity providers could use other tools to diversify their longevity risks, such as diversifying across the age distribution and diversifying internationally, although such efforts still cannot eliminate the risk entirely (Brown and Orszag, 2006).

Yarri (1965) mentions that an individual could obtain substantial welfare by having an annuity. While the private sector does not have the power to create a guideline and improve the management of longevity risks (after pooling), governments can use fiscal policy (including taxes, social insurance, transfers and public debts) and mandate a standard procedure for managing longevity risks if there is any. The government can create or promote longevity indices that can be used as benchmarks in the markets for mortality/longevity linked derivatives. The government has access to detailed information and legal documents for estimating mortality rates. The government can provide the indices for different population subgroups stratified by gender or other socio-economic factors. Our proposed model can be modified for calculating the longevity index easily.

Another perspective is that the government itself can issue longevity bonds to fund social security programs and be used by pensions to hedge aggregate longevity

the United States, the amount of unfunded social security benefits from Medicare for baby boomers are now estimated about 25 trillion (Cauchon, 2011).

risks. Blake et al. (2013) mention important reasons why the government should engage in sharing the longevity risks through the issuances of longevity bonds. The first reason is to ensure there is an efficient annuity market. While the private sector could sell annuities or increase the annuity prices due to the failure of hedging systematic longevity risks, the government could provide additional benefits to supplement pensioners' income, such as receiving lower income tax rate and expenditure taxes. The other reason is that the government can engage in intergenerational risk sharing by providing risk protection against systematic trend risk. The government would issue the longevity bond and receive the risk premium. Thus the current retired population pays the future generation a risk premium to hedge the current systematic risk. The design and pricing of the longevity bond is still developing, but it is a great financial instrument, particularly for the government, to manage the longevity risks.

Appendix A

The Modified Bessel Functions

The modified Bessel equation of order ν is shown as follows

$$x^2 \frac{dy^2}{d^2x} + x\frac{dy}{dx} - (x^2 + \nu^2)y = 0.$$

The solution to the modified Bessel equation of order ν shown as follows

$$y = CI_\nu(x) + DK_\nu(x), \;\; x > 0$$

where C and D are arbitrary constants,

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{(2k+\nu)},$$
$$K_\nu = \frac{\pi}{2} \frac{I_{-\nu}(x) - I_\nu(x)}{\sin \nu\pi},$$

and the gamma function $\Gamma(.)$ is defined as

$$\Gamma(n) = (n-1)!.$$

The modified Bessel function of the third kind with order ν is K_ν.

Appendix B

The Bessel Functions

The Bessel equation of order ν is shown as follows

$$x^2 \frac{dy^2}{d^2x} + x\frac{dy}{dx} + (x^2 - \nu^2)y = 0.$$

The solution to the Bessel equation of order ν shown as follows

$$y = AJ_\nu(x) + BY_\nu(x), \;\; x > 0$$

where A and B are arbitrary constants,

$$J_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{2m+\nu}$$
$$Y_\nu(x) = \frac{J_\nu(x)\cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)}.$$

The $J_\nu(x)$ is the Bessel function of the first kind with order ν. $Y_\nu(x)$ is the Bessel function of the second kind with order ν.

Appendix C

Smoothing Cubic Splines Method

Given n data points $\{x_i\}_{i=1}^{n}$, we can observe data and find some turning points if we wish to fit a curve. Let $\{t_k\}_{k=1}^{q}$ be knots such that $x_1 < t_1 < t_2 < ... < t_q < x_n$. For all x_i within $[t_k, t_{k+1})$, we define a smoothing function f as a cubic spline that is formulated as following

$$f(x) = \beta_{0,k} + \beta_{1,k}(x - t_k) + \beta_{2,k}(x - t_k)^2 + \beta_{3,k}(x - t_k)^3$$

$s.t.$

$$\beta_{0,k-1} + \beta_{1,k-1}(x - t_{k-1}) + \beta_{2,k-1}(x - t_{k-1})^2 + \beta_{3,k-1}(x - t_{k-1})^3$$

$$= \beta_{0,k} + \beta_{1,k}(x - t_k) + \beta_{2,k}(x - t_k)^2 + \beta_{3,k}(x - t_k)^3$$

$$\beta_{1,k-1} + 2\beta_{2,k-1}(x - t_{k-1}) + 3\beta_{3,k-1}(x - t_{k-1})^2$$

$$= \beta_{1,k}(x - t_k) + 2\beta_{2,k}(x - t_k) + 3\beta_{3,k}(x - t_k)^2$$

$$2\beta_{2,k-1}(x - t_{k-1}) + 6\beta_{3,k-1}(x - t_{k-1})$$

$$= 2\beta_{2,k}(x - t_k) + 6\beta_{3,k}(x - t_k)$$

$$\beta_{2,0} = \beta_{3,0} = \beta_{2,q} = \beta_{3,q}$$

The first three conditions are to ensure that the smoothing function f itself and its first two derivatives are continuous at knots. The last condition is the boundary condition, and it is to make f is linear on the two extreme intervals $[a, t_1]$ and $[t_q, b]$. The way that cubic splines estimate piecewise smoothing curves is to minimize integrated squared second derivatives, $\int(f'')^2$. The design is to have a smooth curve with small data variation yet does not consider the error between observed datum and fitted value.

The smoothing cubic splines integrate a function to measure the goodness of fit. Goodness of fit and smoothness are different fitting directions; the best fit is just data set itself and the fit is not smooth, yet the smoothest fit is a straight line and it is not a good fit. Smoothing spline was proposed to find a balance point between how well the function can fit and how smooth it can be.

We already assume the smoothing function as cubic splines, so we would like to find a smoothing function $\hat{f}(x)$ with corresponding $\beta's$ that also continuous itself and its first two derivatives by minimizing the penalized sum of squares (PSE)

$$\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda \int (f''(x))^2 dx$$

where λ is the smoothing parameter. The first part is the sum of squared error, and it controls the goodness of fit. The second part is to control the degree of smoothing. If f is very smooth, then $\int (f''(x))^2 dx$ will be large. The smoothing λ is used to control how much smoothing impact that f can have. The larger λ will lead a smoother function. The choice of λ is generally based on the cross validation method. The basic idea of the cross validation method is to leave one datum out every time to obtain a estimator and then to do so for all data. Repeat the process with different λ to obtain the optimal choice for λ that can minimize the PSE. Most statistical software have programs to implement the process. For detailed of the cross validation method, see Wang (2011) and Wood (2006).

Bibliography

[1] Abramowitz, M., (1974). *Handbook of Mathematical Function, with Formula, Graphs, and Mathematical Tables.* Dover Publications Incorporated.

[2] Antolin, P. and Blommestein, H., (2007). Governments and the market for longevity-indexed bonds. OECD Working Papers on Insurance and Private Pensions, No. 4, OECD Publishing. Available from http://www.oecd.org/daf/fin/insurance/37977290.pdf

[3] Applebaum, D., (2009). *Lévy Processes and Stochastic Calculus.* Cambridge University Press.

[4] ARTEMIS, (2012). £40 billion of UK pensioner longevity risk already transferred. 20 March 2012. Available from `http://www.artemis.bm/blog/2012/03/20/40-billion-of-uk-pensioner-longevity-risk-already-transferred/`

[5] Bacaër, N. (2011). *A short History of Mathematical Population dynamics.* Springer.

[6] Ballota, L., and Haberman, S., (2006). The fair valuation problem of guaranteed annuity option: The stochastic mortality environment case. *Insurance: Mathematics and Economics* **38**, 195-214.

[7] Barndorff-Nielsen, O. E. and Blaesid, P., (1983). Hyperbolic distributions. In s. Kotz and N.L. Johnson (Eds.), *Encyclopedia of Statistics*, Volume 3. Wiley, New York.

[8] Barndorff-Nielsen, O. E. and Shephard, N., (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B* **63**, 167-241.

[9] Barrieu, P., and Albertini, L., (2009). *The Handbook of Insurance-Linked Securities.* The Wiley Finance Series, Wiley.

[10] Biffis, E., and Blake, D., (2009). Mortality-linked securities and derivatives. Discussion Paper PI-0901. The Pensions Institute, Case Business School, City University, U.K. Available from
http://www.pensions-institute.org/workingpapers/wp0901.pdf

[11] Blaesild, P. and Jensen, J. L., (1981). Multivariate distributions of hyperbolic type. In C. Taillie, G. Patil, and B. Baldessari (Eds.), *Statistical Distributions in Scientific Work*, Volume . Dordrecht Reidel.

[12] Blake, D., (2012). Longevity 8: Eighth International Longevity Risk and Capital Market Solutions Conference. Eighth International Longevity Risk and Capital Market Solutions Conference. Available from
http://www.cass.city.ac.uk/_data/assets/pdf_file/0008/141569/David-Blake_Longevity-8-Welcome.pdf

[13] Blake, D., T. Boardman, and Cairns, A. (2013). Sharing longevity risk: Why governments should issue longevity bonds. Discussion Paper PI-1002, The Pensions Institute, Cass Business School, City University London. Available from
http://pensions-institute.org/workingpapers/wp1002.pdf

[14] Bohn, H., (2005). Who bears what risk? An intergenerational perspective. Pension Research Council Working Paper PRC WP 2005-7.

[15] Bølviken, E., and Benth, F. E., (2000). Quantification of Risk in Norwegian stocks via the normal inverse Gaussian distribution. *Proceedings of the AFIA 2000 Colloquium.* Tromsø, Norway, International Actuarial Association, 87-98.

[16] Booth, H. and Tickle, L., (2008). Mortality modeling and forecasting: A review of methods. ADSRI Working Paper No.3, Australia National University.

[17] Booth, H., Maindonald, J., and Smith, L. (2001). Age-time interactions in mortality projection: Applying Lee-Carter to Australia. Working papers in Demography No. 85, Australian National University, Canberra.

[18] Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies* **56**, 325-336.

[19] Bravo, J. M., (2010). Lee-Carter mortality projection with "limit life table". Proceeding of the EUROSTAT/UNECE joint session on demographic projection, Lisbon, Portugal, 28-30 April 2010.

[20] Bravo, J. M., E. Coelho, and Magalhães, M. G., (2010). Mortality projection in Portugal. Proceeding of the EUROSTAT/UNECE joint session on demographic projection, Lisbon, Portugal, 28-30 April 2010.

[21] Brockett, P. L., A. Charnes, and Cooper, W. W., (1980). MDI estimation via unification and extension of relevant theory and application. *Transactions of the Society of Actuaries* **43**, 73-135.

[22] Brockett, P. L., S.-L. Chuang, Y. Deng, and MacMinn, R. D., (2013). Incorporating longevity risk and medical information into life settlement pricing. *Journal of Risk and Insurance*, forthcoming.

[23] Brockett, P. L., and Cox, S. H., (1984). Statistical Adjustment of mortality tables to reflect known information. *Transactions of the Society of Actuaries* **36**, 63-75.

[24] Brown, J. R., and Orsazg, P. R., (2006). The political economy of government issued longevity bonds. Journal of Risk and Insurance 73, 611-631. Available from http://www.investmentnews.com/assets/docs/CI15770330.PDF

[25] Bühlmann, H., and Delnaen, F., (1997). No-arbitrate, change of measure and conditional Esscher transforms. Working paper.

[26] Burton, R., S. Dektor, and Wheeler, D., (2008). Time series prediction of US swap rates. Available from
`http://cs229.stanford.edu/proj2008/BurtonDektorWheeler-TimeSeriesPredictionofUSSwapRates.pdf`

[27] Cairns, A., D. Blake, and Dowd, K., (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance* **73**, 787-718.

[28] Cairns, A., D. Blake, K. Dowd, G.D. Coughlan, D. Epstein, A. Ong, and Balevich, I., (2007). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal* **13**, 1-35.

[29] Carr, P., Geman, H., Madan, D., Yor, M., (2002). The fine structure of asset returns: An empirical investigation. *Journal of Business* **75**, 305-332.

[30] Carr, P., and Wu, L., (2003). Finite moment log stable process and option pricing. *Journal of Finance* **58**, 753-777.

[31] Cauchon, D., (2011). Governments mountain of debt. USA today. June 7 2011. Available from http://usatoday30.usatoday.com/news/washington/2011-06-06-us-debt-chart-medicare-social-security_n.htm

[32] Cawston, S., and Vostrikova, L., (2009). On continuity properties for option prices in exponential Lèvy process. *Theor. Veroyarnost. i Primenen*, **27**, 645-670.

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

[34] Chang, J.-M., (2011). Hedging longevity risk through capital markets. Capital Market Opinion, KCMI. Available from `http://www.kcmi.re.kr/Eng/cmweekly/down1.asp?num=25&seq=1&filename=(2011-0419)Weekly_eng.pdf`

[35] Cobley, M., (2011). It's a longevity hedge ... but not as we know it. *Financial Times*, 1 February 2011, Asset Management section. Available from `http://www.efinancialnews.com/story/2011-02-01/longevity-hedge-pall-corporation`

[36] Cont, R., and Tankov, P., (2004). *Financial Modeling with Jump Process.* Chapman & Hall/CRC.

[37] Continuous Mortality Investigation Bureau, (2004). Considerations for the graduation of the CMI 1999-2002 mortality experience. CMI Working Paper No. 8. London: Institute of Actuaries and Faculty of Actuaries.

[38] Continuous Mortality Investigation Bureau, (2007). Stochastic projection methodologies: Lee-Carter model features, example results and implications. CMI Working Paper No. 25. London: Institute of Actuaries and Faculty of Actuaries.

[39] Credit Suisse Demographics Research (2010). A Demographic Perspective of Fiscal Sustainability: Not Just the Immediate Term Matters. Available from `https://doc.research-and-analytics.csfb.com/docView?language=ENG&source=emfromsendlink&format=PDF&document_ id=803008340&serialid=lS%2BeXWRxMfOgeXJimQO3yvsMYR37KakEObxISOjFQKo%3D`

[40] Credit Suisse Demographics Research (2012). How increasing longevity affects us all? Market, economic & social implications. Available from `https://doc.research-and-analytics.csfb.com/docView?language=ENG&form`

at=PDF&document_id=804917550&source_id=em&serialid=hq62DgWSWwwUFH7wTb
vN1BHbKip%2BG0lPWFEl1Qcq2c0%3D

[41] Currie, I. D. (2006). Smoothing and forecasting mortality rates with p-splines. Available from
http://www.ma.hw.ac.uk/ iain/research/talks.html

[42] Currie, I. D., M. Durban, and Eilers, P. H. C., (2004). Smoothing and forecasting mortality rates. *Statistical Modeling* **4**, 279-298.

[43] Dahl, M., and Møller, T., (2006). Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: Mathematics and Economics* **3**, 193-217.

[44] Damodaran, A., (1999). Equity risk premiums: Looking backwards and forwards. Available from
http://people.stern.nyu.edu/adamodar/pdfiles/country/ERP.pdf

[45] De Moivre, A., (1725). *Annuities upon Lives*, London: Printed by W.P. and sold by Francis Fayram, 1725.

[46] Deng, Y., Brockett, P., and MacMinn, R., (2012). Longevity/Mortality risk modeling and securities pricing. *Journal of Risk and Insurance* **79**, 697-721.

[47] Eberlein, E., and Keller, U., (1995). Hyberbolic distribution in finance. *Bernoulli* **1**, 281-299.

[48] Eriksson, A., Forsberg, L., and Ghysels, E. (2004). Approximating the probability distribution of functions of random variables: A new approach. Discussion paper CIRANO.

[49] Eriksson, A., Ghysels, E., and Wang, F., (2009). The Normal Inverse Gaussian distribution and the pricing of derivatives. Working paper.

[50] Esche, F., and Schweizer, M., (2005). Minimal entropy preserves the Lévy property: how and why. *Stochastic Processes and their Applications* **115**, 299-327.

[51] Esscher, F., (1932). On the probability function in the collective theory of risk. *Scandinavian Actuarial Journal* **15**, 175-195.

[52] Eubank, R. L., (1999). *Nonparametric Regression and Spline Smoothing.* Second Edition, Mercel Dekker, Inc.

[53] Filo, F., (2004). Option pricing under the variance gamma process, MPRA Paper No. 15395. Available from
http://mpra.ub.uni-muenchen.de/15395/1/MPRA_paper_15395.pdf

[54] Forman, B. (2010). Life settlements, presentation to the Actuaries Club of the Southwest, June 10 2010. Available from
http://www.docstoc.com/docs/74887957/Life-Settlements-Life-Settlement

[55] Gavrilov, L. A., and Gavrilova, N. S., (2011). Reliability Theory of Aging and Longevity. Long-Run Economic Perspectives of an Aging Society Conference Papers. Published online as working paper on June 2011. Available from
http://www.lepas-fp7.de/lepas-proc-2011-10.pdf

[56] Gerber, H. U., and Shiu, E. S. W., (1994). Option pricing by Esscher tramsforms. *Transactions of Society of Actuaries* **46**, 98-140.

[57] Geman, H., Madan, D., and Yor, M., (2001). Time changes for Lévy processes. *Mathematical Finance*, **11**, 79-96.

[58] Giacometti, R., Ortobelli, S., and Bertocchi, M. I. (2009). Impact of different distributional assumptions in forecasting Italian mortality rates. *Investment Management and Financial Innovations*, **6**, 186-193.

[59] Gompertz, B., (1824). On the nature of the function expression of the law of human mortality and on a new mode of determining the value of life contingencies. *Phil. Tran. Royal Society* **115**, 513-585

[60] Graunt, John, (1662). *Natural and political observations made upon the bills of mortality.* For details and discussion of the results see Hald (2005).

[61] Gustafsson, M., (2011). Cohort effects in Swedish mortality and their effects on technical provisions for longevity risk.Department of Mathematics, Stockholm University, Sweden.

[62] Haberman, S., and Renshaw, A., (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics* **48**, 35-55.

[63] Hainuat, D., and Devolder, P., (2008). Mortality modeling with Lévy process. *Insurance: Mathematics and Economics*, **42**, 409-418.

[64] Hald, A. (2005). *John Graunt and the Observations Made upon the Bills of Mortality*, 1662, in *History of Probability and Statistics and Their Applications before 1750*, John Wiley & Sons, Inc., Hoboken, NJ, USA.

[65] Hsieh, M.-H., J.-L. Peng, C. Tsai, and Wang, J., (2012). Explaining the risk premium of life settlements. Presentation at the 8th Longevity Risk and Capital Markets Solutions Conference, Waterloo, Ontario, Canada, September 7, 2012.

[66] Huang, J., and Wu, L., (2004). Specification analysis of option pricing models based on time-changed Lévy processes. *Journal of Finance* **59**, 1405-1440.

[67] Hubalek, F., and Sgarra, C., (2006). Esscher transforms and the minimal entropy martingale measure for exponential Lévy models. *Quantitative Finance*, **6**, 125-145.

[68] *Investment & Pensions Europe*, (2010). BMW UK pension complete £3bn longevity swap. 22 Febuary 2010. Available from
`http://www.ipe.com/news/bmw-uk-pension-completes-3bn-longevity-deal_34134.php`

[69] Jarque, C. M., and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economic Letters* **6**, 255-259.

[70] J.P. Morgan, (2007a). q-forwards: Derivatives for transferring longevity and mortality risk. Pension Advisory Group, J.P. Morgan.

[71] J.P. Morgan, (2007b). LifeMetrics, a toolkit for measuring and managing longevity and mortality risks, technical document.

[72] Keller, U., (1997). Realistic modelling of financial derivatives. Dissertation. Mathematische Fakultat der Albert-Ludwigs-Universitat Freiburg im Breisgau.

[73] Kitchen, C., (2009). Normal Inverse Gaussian (NIG) process in mathematical finance. Available from
`http://finance.math.ucalgary.ca/papers/CliffTalk26March09.pdf`

[74] Kliff, J., (2012). Financial impact of longevity risk. Eighth International Longevity Risk and Capital Market Solutions Conference. Available from
`http://www.cass.city.ac.uk/_data/assets/pdf_file/0007/141586/John-Kiff_presentation.pdf`

[75] Koissi, M.-C., and Shapiro, A. F., (2012). The Lee-Carter model under the condition of variables age-specific parameters. Working papers, Department of Mathematics, Western Illinois University.

[76] Konstantopoulos, T., Z. Zerakidze, and Sokhadze, G., (2011). The Radon-Nikodym theorem. In *International Encyclopedia of Statistical Science*, 2011. Edited by L. Miodrag. Springer.

[77] Krol, K., and Küchler, U., (2009). Minimal entropy martingale measure for Lévy processes. Working paper.

[78] LCP, (2011). LCP pension buy-out 2011. Available from `http://www.lcp.uk.com/news--publications/publications-and-research/2011/lcp-pension-buy-outs-2011/`

[79] Lane Clark & Peacock, (2012). LCP pension buy-ins, buy-outs and longevity swaps 2012. Available from `http://www.lcp.uk.com/media/458733/lcp%20buy-out%20report%202012.pdf`

[80] Lawson, C., and Hanson, R., (1974). *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, N.J..

[81] Lee, R. and Carter, L., (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, **87**, 659-671.

[82] Lee, R. and Miller, T., (2000). Evaluating the performance of Lee-Carter mortality forecasts. Working paper, University of California at Berkeley.

[83] Li, S.-H., and Chan, W.-S., (2007). The Lee-Carter model for forecasting mortality, revisited. *North American Acturial Journal*, **11**, 68-89.

[84] Life & Longevity Markets Association, (2010). Longevity pricing framework. Available from `http://www.llma.org/files/documents/Longevity_Pricing_Framework_Final.pdf`

[85] Life & Longevity Markets Association, (2012). Longevity index technical document, version 1.0. Available from `http://www.llma.org/files/documents/LLMA_Longevity_Index_Technical_Document_19.3.12_v1_final.pdf`

[86] Lillestøl, J., (2000). Risk analysis and the NIG distribution. *Journal of Risk* **2**, 41-56.

[87] Lin, Y., S. Liu, and Yu, J., (2013). Pricing mortality securities with correlated mortality indexes. *Journal of Risk and Insurance*, forthcoming.

[88] Loeys, J., (2007). Longevity: A market in the making. Global Market Strategy, J.P. Morgan.

[89] Loeys, J., N. Panigirtzoglou, and Riberio, R., (2007). Longevity: A market in the making. J. P. Morgan's Global Market Strategy. Available from `https://www.jpmorgan.com/cm/BlobServer/LM_longevity.pdf?blobkey=id&blobwhere=1158464081204&blobheader=application%2Fpdf&blobcol=urldata&blobtable=MungoBlobs`

[90] Mencia, F. J., and Sentana, E., (2004). Estimation and testing of dynamic models with generalized hyperbolic innovations. CMFI Working Paper 0411, Madrid, Spain.

[91] Mercado, D., (2009). Goldman's longevity index has short life. 18 December 2009, *Investment News*, News and Opinion section. Available from `http://www.investmentnews.com/article/20091218/FREE/912189978`

[92] Mercer Global, (2010). Mercer pension buyout index launched to assist companies and the trustees in pension scheme de-risking. 22 February 2010. Available from `http://www.mercer.com/press-releases/1373420`

[93] Merton, R., (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* **3**, 125-144.

[94] Milevsky, M. A., and Promislow, S. D., (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics,* **29**, 299-318.

[95] Mitchell, D., P. Brockett , R. Mendoza-Arriage, and Muthuraman, K., (2013). Modeling and forecasting mortality rate. *Insurance: Mathematics and Economics*, forthcoming.

[96] Mitchell, O. S., J. M. Poterba, M. J. Warshawsky, and Brown, J. R., (1999). New evidence on the moneys worth of individual annuities. American Economic Review 89, 1299-1318.

[97] Miyahara, Y., (2012). *Option pricing in incomplete markets: Modeling based on geometric Lévy process and minimal entropy martingale measure.* Imperial College Press.

[98] Mortality laws in Encyclopedia of Actuarial Science, available from `http://maths.swan.ac.uk/staff/zs/mortality_laws.pdf`

[99] Mortimer, S. (2011), JP Morgan takes on risk of UK pension plan, *The Independent*, 2 February 2011, Business section. Available from
`http://www.independent.co.uk/news/business/news/jp-morgan-takes-on-risk-of-uk-pension-plan-2201406.html`

[100] Murphy, M., (2006). Life settlements: A high-return emerging investment category. Available from
`http://EzineArticles.com/?expert=Michael_Murphy`

[101] Muspratt, C., (2006). Paternoster in pensions deal. *The Telegraph*, 10 November 2006, Finance section.

[102] Olivieri, A. and Pitacco, E., (2010). *Introduction to Insurance Mathematics: Technical and Financial Features of Risk Transfers.* Springer-Verlag Berlin Heidelberg.

[103] Önalan, O., (2009). Martingale measures for NIG Lévy processes with applications to mathematical finance. *International Research Journal of Finance and Economics* **34**, 56-68.

[104] Pascucci, A., (2011). *PDE and Martingale Methods in Option Pricing*, Bocconi & Springer Series, Springer-Verlag Italia. *The American Journal of Sociology* **114**, 738-780.

[105] Pelley, L. (2011). Widowhood can shorten a partner's lifespan. CBC News, 17 May 2011. Available from
`http://www.cbc.ca/news/health/story/2011/05/06/f-milestogo-broken-heart-syndrome.html`

[106] Quinns, S., (2008). The transformation of morals in markets: Death, benefits, and the exchange of life insurance policies.

[107] Pension Advisory Group, J.P. Morgan, (2007). LifeMetrics, a toolkit for measuring and managing longevity and mortality risks, technical document. Available from
`http://www.jpmorgan.com/cm/BlobServer/lifemetrics_technical.pdf`

[108] Prause, K., (1997). Modelling financial data using generalized hyperbolic distributions. *FDM Preprint* **48**, Albert-Ludwigs-Universität Feiburg.

[109] Prause, K., (1999). The generalized hyperbolic model: Estimation, financial derivatives, and risk measurement. Dissertation, Institute für Mathematiche Stochastik, Albert-Ludwigs-Universität Feiburg.

[110] Promislow, S.D., (2006). *Fundamentals of Actuarial Mathematics.* John Wiley and Sons, Ltd.

[111] Protter, P. E., (2002). *Stochastic Integration and Differential Equation*, Applications of Mathematics (New York), Springer-Verlag Berlin.

[112] Raible, S., (2000). Lévy processes in finance: Theory, numerics, and empirical facts. Dissertation, Institute für Mathematiche Stochastik, Albert-Ludwigs-Universität Feiburg.

[113] Redfern, R., (2012). When long-time spouse dies, survivor may soon follow. Available from
`http://www.toledoblade.com/local/2012/02/12/When-long-time-spouse-dies-survivor-may-soon-follow.html`

[114] Renshaw, A., and Haberman, S., (2003). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics* **33**, 255-272.

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

[116] Renshaw, A., Haberman, S., and, Hatzoupoulos, P., (1996). The modelling of recent mortality trends in United Kingdom male assured lives. *British Actuarial Journal* **2**, 449-477.

[117] Rydberg, T. H. (1997). The normal inverse Gaussian Lévy process: Simulation and approximation. *Communications in Statistics: Stochastic Models* **13**, 887-910.

[118] Sato, K.-I., (2000). *Lévy processes and Indefinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

[119] Schoutens, W., (2003). *Lêvy processes in Finance: Pricing Financial Derivatives.* John Wiley & Sons Ltd.

[120] Shemyakin, A., and Youn, H., (2006). Copula models of joint survival analysis. *Applied Stochastic Models in Business and Industry* **22**, 211-224.

[121] Siegert, P., (2010). Overview of the life insurance secondary market from an investment perspective. Available from
`http://insurancestudies.org/wp-content/uploads/2010/10/ISI_LifeINsuranceSettlementSeries_IX.pdf`

[122] Sisk, R., (2011). Census: A half million centenarians by 2050, New York Daily News, 25 April 2011. Available from
`http://nydailynews.com/blogs/dc/2011/04/census-a-half-million-centenarians-by-2050`

[123] Sithole, T. Z., Haberman, S., and Verrall, R. J., (2000). An investigation into parametric models for mortality projects with applications to immediate annuitants and life office pensioners' data. *Insurance: Mathematics and Economics* **27**, 285-312.

[124] Stone, C. A. and Zissu, A., (2006). Securitization of senior life settlements: Managing extension risk. *The Journal of Derivatives* **13**, 66-72.

[125] Sullivan, J., "University of Oregon sociologist devised key method of population forecasting," *The Seattle Times*, 5 February 2001.

[126] Sweeting, P., (2010). The Swiss Re longevity bond. Available from
`http://blog.paulsweeting.com/?q=node/48`

[127] Tankov, P., (2007). Lévy processes in finance and risk management. World Congress on Computational Finance.

[128] *Transamerica Reinsurance Risk Management Newsletter*, (2010). Longevity-the nest growth market. Available from
`http://www.scor.com/images/stories/pdf/library/archives-sgla-sgrc/longevity-thenextgrowthmarket.pdf`

[129] *The Four Pillars*, (2008). Longevity Four: Fourth International Longevity Risk and Capital Market solutions. Conference Amsterdam, 25-26 September 2008.

[130] Wang, C., Huang, H., and Liu, I., (2011). A quantitative comparison of the Lee-Carter model under different types of non-Gaussian innovations. *The Geneva Papers* **36**, 675-696.

[131] Wang, J. Z., (2007). Fitting and forecasting mortality for Sweden: Applying the Lee-Carter model. Report of Department of Mathematical Statistics, Stockholm University.

[132] Wang, Y., (2011). *Smoothing Splines: Methods and Applications.* Monographs on Statistics and Applied Probability 121, Chapman & Hall/CRC.

[133] Willets, R., (2003). The cohort effect: Insights and explanations. Willets Consulting Limited.

[134] Wood, S. N., (2006). *Generalized Additive Models: An Introduction with R.* Texts in Statistical Science, Chapman & Hall/CRC.

[135] World Population Ageing (2009), Department of Economics and Social Affairs, United Nations. Available from
http://www.un.org/en/development/desa/publications/world-population-ageing-2009.html

[136] Wu, L., (2005). Modeling Financial Security Returns Using Lévy Processes. Zicklin School of Business, Baruch College, City University of New York. Available from
http://faculty.baruch.cuny.edu/lwu/papers/Lévyhandbook.pdf

[137] Wu, L., (2012). Option pricing with time-changed Lévy processes. Available from
http://faculty.baruch.cuny.edu/lwu/papers/Lévy_Jane.pdf

[138] Yari, M., (1965). Uncertain life time, life insurance, and the theory of consumer. Review of Economic Studies, Vol. 32, No. 2.