Université catholique de Louvain

École polytechnique de Louvain (EPL)

Longevity risk

Modeling and pricing of the longevity risk

Dissertation presented by
Simon Boigelot

for obtaining the Master's degree in
Mathematical Engineering

Supervisor(s)
Pierre Devolder, Julien Hendrickx

Reader(s)
Donatien Hainaut

Academic year 2017-2018

Abstract

Longevity risk is undoubtedly one of the major risks for the next decades. On one hand, people living longer is synonym of a real evolution in the social environment and the healthcare industry. But, on the other hand, the longevity risk raises financial concerns within insurers, annuity providers and more importantly governments, as it puts their financial solvency at risk. Thereby threatening the global stability of the financial market. In this paper, we will present the most popular stochastic mortality models that try to capture this longevity phenomenon as precisely as possible. Both discrete models, such as the Lee-Carter and the Cairns Blake Dowd models, and more recent continuous models, such as the Feller and Ornstein-Uhlenbeck models, will be discussed. The longevity risk being systematic, it is not hedgeable via classical insurance pooling. Instead, all eyes are oriented toward capital markets which present an interesting securization opportunity.
The second part of this paper will therefore be devoted to understanding how the major longevity-linked securities, i.e. survivor-forwards, longevity swaps and longevity bonds, can be priced. Various possibilities will be explored with a particular stress on the equivalent utility pricing method.

Keywords : longevity risk, stochastic mortality models, affine continuous processes, pricing, longevity-linked securities, utility

Acknowledgements

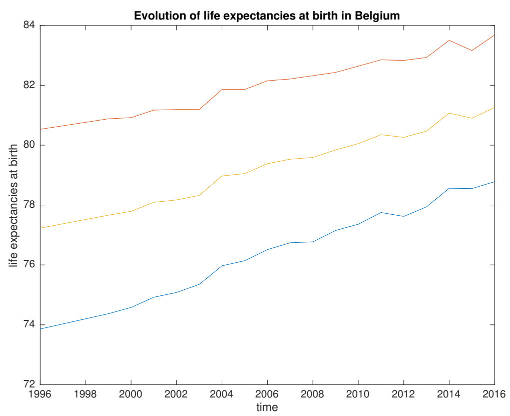
This thesis has been developed in the framework of my Master Degree in Mathematical Engineering at the Université catholique de Louvain. It is the achievement of five years of study and about ten months of considerable hours of work, thinking, research and writing. Needless to say that such process goes with ups and downs. I am therefore grateful to the people who supported me throughout this task and gave me the strength to rise both as a person and as a student. In priority, I would like to thank Professor dr. Pierre Devolder and Professor dr. Julien Hendrickx who, as promotors, offered me the possibility to pick a subject matter for this thesis that was dear to me. During the ten months that punctuated this thesis, Mr. Devolder always took the time to discuss my work at length as well as to give me precious advices and valuable comments. Then, I am particularly thankful to all the professors (from the Université catholique de Louvain and the University of Melbourne) I came across on my path to graduation. Together, they provided me with the necessary mathematical and scientific tools to draft this thesis and to face a challenging career. Finally, I would like to express all my gratitude to my friends and family for their continuous support throughout the years.

Contents

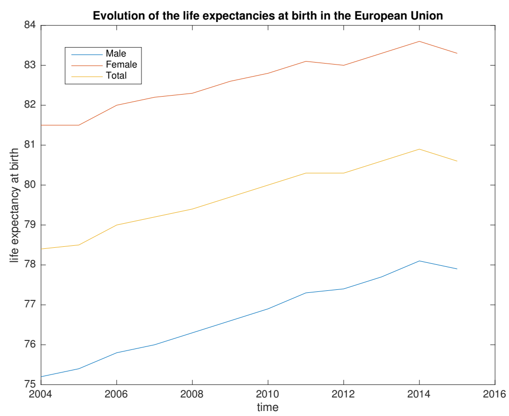
Chapter 1

Introduction

People live longer and longer. This is a fact. In Belgium, the life expectancy at birth raised by 5,22% over 20 years. This increase in lifetime is a global effect impacting most developed countries. In the European Union, for example, life expectancy at birth increased by 2,81% over the last 10 years. The increase in life expectancy might be a global phenomenon but the improvements in mortality are not similar all over the world and over all ages. On a historical point of view, the average lifespan nearly tripled, and much of this improvement happened during the last 150 years. In the first half of the 20^{th} century, birth mortality really improved, principally thanks to better birth methods. However, during the second half of the 20^{th} century, mortality improved mostly at higher ages, thanks to better medicine and treatments. Moreover, mortality improvements do not affect everybody in the same way. They highly depend on age groups, birth years and environments. Advanced models try to incorporate this cohort effect for the modeling of the longevity risk.



(A) Belgian life expectancies at birth in function of the calendar year



(B) Life expectancies at birth in function of the calendar year in the European Union

FIGURE 1.1

A higher life expectancy is admirable. It results from better food habits, development, healthcare and life style. Although the idea of living longer is synonym of progress, it also represents a big financial challenge for most life insurers, annuity providers and governments. Longevity risk is the risk to which a pension fund or life insurance company could be exposed as a result of higher-than-expected payout ratios. It exists due to the increasing life expectancy trends among policyholders and pensioners, and can result in payout levels that are higher than what a company or fund originally accounts for. This risk has to be considered when pricing annuities and forming provisions for future payments, as it affects the solvency of most life products. Indeed, taking the example of a life annuity and should the longevity improvement be ignored in its pricing, the issuer would then have to pay for a longer period than the one on which premiums (and thus reserves) are based. The amount of future payments are underestimated due to an underestimation of the survival probabilities. There is thus a risk that the net present value of the annuity payments will turn higher than expected, synonym of loss product (if no security margin is considered). Moreover, an imbalance between assets and liabilities (which are mainly reserves) management can cause the default risk of an insurer to increase, putting the life insurance industry and consequently the whole financial world into risk.

For governments and more importantly the individuals, the effect of the longevity risk on pension plans and age-related social insurance benefits is that pensioners might potentially outlive their resources resulting in lower living standards at older ages. Which is not desirable for a developed country. Furthermore, social security plans based on partition are also affected by the longevity risk as the ageing effect causes fewer and fewer active people to fund more and more pensioners. If nothing is done, the imbalance will be too important for these plans to be applicable.

The international actuarial association identifies four components in the longevity risk. Trend, level, volatility and catastrophe. The idea is to explain how mortality fluctuates (volatility) around a certain level knowing that fluctuations can be caused by two factors (trend and catastrophe). They could be a result of a simple trend over time. For example, longevity improvement thanks to better medication. Or, these fluctuations could be caused by catastrophic events such as nuclear accidents, tsunamis and earthquakes, to name a few.

The first two components regroup to form the so-called systematic risk. It is the risk of not concordant base assumptions for the estimation of the mortality rates. The last two form the specific risk representing fluctuations around the base case scenario. This paper will focus on the systematic risk. It is not hedgeable by pooling, which consists of distributing the risk over a large group of people. Indeed, all individuals are affected in almost the same way and classical reinsurance is therefore not consistent. Specifically, the most popular hedging technique used for the moment is the hedging through securization on capital markets. This can be done by purchasing or selling longevity derivatives such as longevity bonds or longevity swaps so as to have longevity cash inflows respectively increasing and decreasing when longevity cash outflows are increasing or decreasing

This paper is divided into two main parts. Firstly, we will discuss how longevity is currently modeled. We will go through a presentation of the most known models at the moment. Afterwards, we will focus on the impact of longevity risk on financial products. More precisely, the pricing of survivor forwards, longevity bonds and longevity swaps will be discussed in some more details.

Chapter 2

Notations

The notations used in this paper might be a little bit confusing and overwhelming. In this section, we will try to provide the signification and definition of certain expressions that we will encounter further in this paper.

Let $T_x(t)$ denote the future lifetime of an life aged x at time t. It is the number of years that the life still have to live. $T_x(t)$ is modeled as a positive continuous random variable on the interval $[0,\ \omega(t)-x]$ where $\omega(t)$ is the maximum attainable age at time t. The age of death can thus be expressed as $x+T_x(t)$ for someone born at time $t-x$. Let

- ${}_sp_x(t) = S_x(t,s) = P(T_x(t) > s)$ be the probability of a life aged x at time t to live for at least s more years. By definition, ${}_{\omega(t)-x}p_x(t) = 0$ and ${}_0p_x(t) = 1$. It is known as the survival function.
- ${}_sq_x(t) = F_x(t,s) = P(T_x(t) \leq s)$ be the probability of a life aged x at time t to die within s years, i.e., the distribution function of the random variable $T_x(t)$. It follows that ${}_{\omega(t)-x}q_x(t) = 1$ and ${}_0q_x(t) = 0$.

The most important expression to understand is the concept of force of mortality. It is defined as follows for a life aged x at time t.

$$
\begin{aligned}
\mu_x(t) &= \lim_{dx\to 0} \frac{P(T_0(t-x) \leq x+dx | T_0(t-x) > x)}{dx} \\
&= \lim_{dx\to 0} \frac{1}{dx} \frac{P(x \leq T_0(t-x) \leq x+dx)}{P(T_0(t) > x)} \\
&= \frac{1}{S_0(t-x,x)} \lim_{dx\to 0} \frac{S_0(t-x,x) - S_0(t-x,x+dx)}{dx} \\
&= -\frac{S_0'(t-x,x)}{S_0(t-x,x)}
\end{aligned}
$$

$\mu_x(t)$ thus represents the intensity of falling dead for someone aged x at time t. That is why it is sometimes called the intensity of mortality as well. Note that $\mu_x(t)$ is not a probability, but $\mu_x(t)dx$ can be seen as an approximation of the probability that a life aged x dies before reaching age $x + dx$ at time t.

$$\mu_x(t)dx \approx P(T_x(t) \leq dx) \text{ for very small dx}$$

The force of mortality is a function of two variables, x and t. In the past, we only considered x as being meaningful. However, this error is the purpose of this paper. Indeed, it seems quite intuitive that the probability of surviving from age 50 to age 55 in 2018 is not the same as what it was in 1960. we therefore need to add a temporal term in order to take the modifications of mortality rates over time into consideration. To determine the very important relation between the survival function and the force of mortality, we have to solve the differential equation obtained above. It is easy to find that

$$_xp_0(t-x) = \exp\left(-\int_0^x \mu_s(t-x+s)ds\right) \tag{2.1}$$

This is a really important result. It means that once one has the expression of $\mu_x(t)$ (which is not an easy matter though), one is able to compute all survival probabilities. Indeed, by independence, $_xp_0(t-x)_sp_x(t) = {}_{x+s}p_0(t-x)$. This is really intuitive. In order to live at least $x + s$ years at birth, one must first survive x years and then an additional s years when being aged x. Using (2.1), it follows that

$$\begin{aligned}
_sp_x(t) &= \exp\left(-\int_0^{x+s} \mu_r(t-x+r)dr + \int_0^x \mu_r(t-x+r)dr\right) \\
&= \exp\left(-\int_x^{x+s} \mu_r(t-x+r)dr\right) \\
&= \exp\left(-\int_0^s \mu_{x+r}(t+r)dr\right)
\end{aligned}$$

And

$$_sq_x(t) = 1 - {}_sp_x(t) = 1 - \exp\left(-\int_x^{x+s} \mu_r(t-x+r)dr\right)$$

Due to modeling restrictions, we almost always compute the force of mortality at integer ages. But as $T_x(t)$ is a continuous random variable, we also need to determine the force of mortality at non-integer ages. In this purpose, the widespread approximation used all over the world is the piecewise constant force of mortality (CFM) assumption. Mathematically, this approximation is translated into

$$\mu_{x+s}(t+u) = \mu_x(t) \qquad \text{for } 0 \leq s, u < 1$$

It then directly follows that ${}_sp_x(t) = \exp\left(-\int_0^s \mu_{x+u}(t+u)du\right) = \exp\left(-\mu_x(t)s\right)$. This result can then be used to find out that

$$ {}_sp_x(t) = \prod_{i=0}^{s-1} p_{x+i}(t+i) = \exp\left(-\sum_{i=0}^{s-1} \mu_{x+i}(t+i)\right) $$

The assumption of a constant force of mortality between integer ages will be used a wide number of times throughout this paper. It is only relevant for discrete time models of the force of mortality. We will come to that later on.

Chapter 3

Modeling

From the basics of actuarial sciences, the pricing of any insurance product on the duration of life depends on two main basis: demographical and financial assumptions. In the past, these assumptions were treated in a deterministic manner (for example, by the use of periodic life tables). However, nowadays, stochastic models are mostly used to project the unknown future lifetime of an individual and the unknown financial environment he will be living in during his life. Analyzing past data, there are two main behaviours to outline concerning the death curve and the survival function.

1. rectangularization of the survival function. The volatility of the duration of life around the mode of death decreases, leading to a lower dispersion of the ages of death around the most likely age of death. In rough words, people tend to die more and more at the same age.

2. expansion of the death curve. The death curve has a tendency to move to the right. Consequently, the age at which death is most likely to occur increases as time passes, thanks to improvements in economic and social conditions.

Bad estimations of the future life expectancy of policyholders can lead to bad pricing of life insurance products and alter the solvency of the insurance company.

Several approaches exist to model the death probabilities.

1. **Short rate models** in which we choose the dynamics of the force of mortality and derive the consequent survival probabilities. The force of mortality is thus a stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ containing all the "information" up to time t. Survival probabilities are then obtained as an expectation

$$ {}_sp_x(t) = \mathbb{E}\Bigg[\exp\left(-\int_0^s \mu_{x+u}(t+u)du\right)|\mathcal{F}_t\Bigg] $$

short rate models that we will encounter in this paper are the Lee-Carter model and its extensions, the P-spline model, the Orhnstein-Uhlenbeck model (with and without jumps), the Feller model (with and without jumps) and the general affine model. Of course, there exist plenty more!

2. **Forward mortality models** where the instantaneous force of mortality is modeled. Survival probabilities are then obtained via

$$_sp_x(t) = \exp\left(-\int_0^s h_x(t, t+u)du\right)$$

   Modeling is more complicated but the computation of the probabilities is easier as we do not need to compute a conditional expectation anymore. The basic example of such models is the Heath-Jarrow-Morton model.

3. **Market models** These are models in which the survival probabilities $_sp_x(t)$ are directly modeled. These models are the easiest to use but are unfortunately often less precise. In this paper, market models regroup the Cairns Blake Dowd model and its extensions as well as the reduction factor model.

In theory, there exists a lot of parametric stochastic models that can be used in order to determine future mortality rates. All of them are based on an extrapolating method. Unknown parameters are determined by fitting the model on historical data. The resulting model can then be seen as a relatively precise way of projecting future mortality rates thanks to time series. Models are decomposed into two main groups: the discrete time models and the continuous time models. In discrete time, research has mainly been focused on the statistical analysis and projection of annual mortality data using short-rate models. In continuous time, most research has also been focused on short-rate models including affine mortality models. Short-rate models model the force of mortality from the time to death and survival probabilities can be determined using the expressions derived above and the CFM. However, it is important to note that the output of the models are mortality intensities and not survival probabilities. This will have an impact on the final probabilities as we need further assumptions that will add to the model risk. The obtained probabilities are thus far from being exact. Yet, they represent a convenient approximation for insurance purposes.

As all models are forecasting methods, they are not immune of certain problems such as

1. model risk
2. parameter uncertainty
3. parameter stability
4. stochastic variation

The model risk consider the risk of the model being used to be false. It is just not a 100% exact representation of the reality. Parameter uncertainty always exists in forecasting methods because parameters are estimated based on past data and therefore subject to sample variations. Parameter stability stems for the evolution of the estimated parameters with time. Are they computed once and for all and then considered fixed? Must they be recomputed every year? What kind of time series should be applied? This is a sort of model incorporated into the initial model. Finally, stochastic variation is associated with confidence intervals. Only if the three problems previously discussed are correctly handled, will the confidence interval be of good accuracy. By definition of modeling, it will not be possible to obtain a very narrow confidence interval. Every time we try to fit a model on some data, errors arise just because the data is not exactly distributed as the model we want to use. Every time we make use of models, there are three risks we need to bear in mind

1. The process risk due to the stochastic nature of the model and the uncertainty it represents. Indeed, multiple models are based on Box-Jenkins processes in order to project time parameters.
2. The parameter risk. Parameters of the models are usually determined by maximizing the log-likelihood expression of an observed scenario. This expression is directly related to the assumed distribution of a certain random variable and is thus also only an approximation. Consequently, resulting parameters deviate from the true values.
3. The model risk. Risk linked to the error made when modeling. For example, the fact of adding a cohort effect sustaining that mortality rates decrease exponentially with the year of birth. This is of course not 100% true.

In a mortality study, process and parameter risks can be clearly illustrated in terms of the variance of residual life-time $T_x(t)$. Let us define a real valued function $\Gamma(x,t)$ that somehow describes the mortality of a life aged x at time t and thus born during year $\tau = t - x$. Let $\theta(\tau)$ define the vector of parameters of the projected mortality model. An appropriate choice of parameters $\theta(\tau)$ is then based on some mortality trend assumptions. Let us here focus on one year of birth. Suppose we can think of m different scenarios $\Theta = [\theta_1 \theta_2 ... \theta_m]$ for the evolution of the mortality of this cohort. Suppose that the scenarios are mutually exclusive and collectively exhaustive so that the probabilities p_j of occurrence of scenario θ_j respect $\sum_{j=1}^{m} p_j = 1$. In an age continuous context,

$$\mathbb{E}[T_x|\theta_j] = \int_0^{+\infty} t f_x(t|\theta_j)dt$$
$$Var[T_x|\theta_j] = \int_0^{+\infty} \left(t - \mathbb{E}[T_x|\theta_j]\right)^2 f_x(t|\theta_j)dt$$

where $f_x(t|\theta_j)$ is the conditional probability density function of the residual life time $T_x(t)$ with respect to the selected scenario. A statistical result states that

$$Var[X] = \mathbb{E}[Var[X|\theta]] + Var[\mathbb{E}[X|\theta]]$$

Consequently,

$$Var[T_x] = \sum_{j=1}^{m} p_j Var[T_x|\theta_j] + \sum_{j=1}^{m} p_j \left(E[T_x|\theta_j] - E[T_x]\right)^2 \tag{3.1}$$

where

$$E[T_x] = \sum_{j=1}^{m} p_j E[T_x|\theta_j]$$

The first term in (3.1) measures the risk from random fluctuations while the second term measures the risk of systematic deviations, i.e., the longevity risk.

As we shall see shortly, there are a lot of models that can be used for the modeling of the longevity risk. But then, how to determine if a model is appropriate or not? This can be done by checking if it satisfies a list of criteria that a good mortality model should meet. The idea here is to list the pros and cons of a certain model. This should afterwards help to determine which model to use. What we expect from a good mortality model is:

- the model should be consistent with historical data. In other words, the model has to stay within a predetermined accuracy from past results.
- The force of mortality should stay positive. Indeed, a negative force of mortality leads to an increasing survival function which make non sense.
- Long-term future dynamics of the model should be biologically reasonable. For example, a model doubling the life expectancy of human beings within five years is not biologically reasonable.
- Long-term deviations in mortality improvements should not be mean-reverting to a predetermined target, even if the target is time dependant. This is because the

force of mortality observed and/or extrapolated from the mortality tables does not seem to present a mean reverting behaviour, but rather an exponential one.

- The model should be comprehensive enough to deal appropriately with pricing valuation and hedging problem. In other words, the model should be easy to adapt and to use in order to price products depending highly on the life duration of the policyholder.
- It should be possible to value mortality linked derivatives using analytic methods or fast numerical methods.

What is sure is that there does not exist one perfect model that captures all the features listed above. The general way of decreasing the model risk, is to combine results obtained by different models $\mu_x(s)^i$ for $i = 1...m$ if we are using m models and then finally work with $\mu_x(s) = \sum_{i=1}^{m} w_i \mu_x(s)^i$ where w_i is a weight associated to model i that represents its importance or relevance.

It is clear that age is the main factor influencing death probabilities, along with time. There are two ways of computing these very important $\mu_x(t)$.

First, we could work with each cohort separately. For example, consider newborns in 2018, the idea would then be to use the model to determine $\mu_0(2018), \mu_1(2019), \mu_2(2020), ...$

Another method to model mortality intensities over two dimensions (age and time), that can add precision but asks for a larger work load, is to construct mortality tables based on the calibration of several cohorts. Instead of working with each cohort separately, we manipulate a group of multiple cohorts. This way, it is possible to find out how cohorts are related with each other. This is done by working with a vector of correlated Brownian motions. However, building mortality surfaces incorporating cohort effects is behind the scope of this paper. We will thus construct surfaces working with one cohort at a time, independently of all the others.

For what concerns the calibration of the models, there are mainly two approaches that can be considered. We could estimate the parameters via the least square error minimization or via the maximization of the log-likelihood of certain observations.

Minimization of the least square errors on the survival probabilities computed between an observed mortality table and one predicted by the stochastic model the gives next problem

$$\phi^* = \underset{\phi}{\operatorname{argmin}} \sum_{s=0}^{T} \sum_{T=0}^{\omega-x} \left({}_{T-s}p_{x+s}^{observed}(t+s) - {}_{T-s}p_{x+s}^{model}(t+s;\phi) \right)^2$$

We here introduce the method with survival probabilities but the truth is that it can be used with whatever quantity $(\mu_x(t), q_x(t), ...)$. This method has the advantage of being intuitive and sometimes easy to apply. We will use this approach for the calibration of the Lee-Carter and CBD models.

For the maximum likelihood estimation, we need to assume that the force of mortality remains constant within a same year. As a result, we have that $\mu_x(t) = -\log(1 - q_x(t))$ so that it is possible to obtain an expression for $q_x(t)$ in terms of $\mu_x(t)$ to calibrate on crude (and eventually smooth) mortality probabilities

$$q_x(t) = 1 - e^{-\mu_x(t)}$$

Once this is done, we need to build a likelihood expression to maximize using the density function of certain observations..

$$\phi^* = \underset{\phi}{\operatorname{argmax}} \sum_{t=0}^{T} \sum_{x=0}^{\omega(t)} \ln\left(f(\mu_x(t);\phi)\right)$$

In fact, this method often appears because it is applicable to all parametric mortality models. Suppose that the number of deaths, $D_x(t)$, follows a Poisson distribution with parameter $E_x(t)\mu_x(t,\phi)$ where ϕ represents the parameter vector of the model and $E_x(t)$ represents the exposure to risk. The idea behind that scheme is that $d_x(t) = E_x(t)\mu_x(t)$. Since

$$P\left(D_x(t) = d_x(t)\right) = \frac{[E_x(t)\mu_x(t,\phi)]^{d_x(t)}}{d_x(t)!} e^{-E_x(t)\mu_x(t,\phi)}$$

the likelihood expression of the observations is given by

$$L(\phi) = \prod_{x,t} P\left(D_x(t) = d_x(t)\right) = \prod_{x,t} \frac{[E_x(t)\mu_x(t,\phi)]^{d_x(t)}}{d_x(t)!} e^{-E_x(t)\mu_x(t,\phi)}$$

The parameter vector is then estimated by maximization of the log-likelihood expression

$$ll(\phi) = \sum_{x,t} \left(D_x(t)\log(E_x(t)\mu_x(t,\phi)) - E_x(t)\mu_x(t,\phi) - \log(D_x(t)!)\right)$$

As we will see in a few pages, for the CBD model, it suffices to substitute

$$\mu_x(t,\phi) = \log\left(1 + \exp(\phi_{1t} + \phi_{2t}(x - \bar{x}))\right)$$

in above expression in order to obtain the log-likelihood to maximize. For Lee-Carter, it becomes

$$\mu_x(t,\phi) = \exp\left(\phi_{1x} + \phi_{2x}\phi_{3t}\right)$$

This Poisson approximation is principally used for discrete time models as most continuous models, can be calibrated via likelihood without using this Poisson assumption. We will come to that later.

3.1 Discrete time models

3.1.1 Lee - Carter model (1992)

The Lee Carter model is the oldest but still the most popular of all the mortality models. It is formulated as :

$$\log \mu_x(t) = \beta_x^1 + \beta_x^2 \kappa_t \Leftrightarrow \mu_x(t) = \exp\Big(\beta_x^1 + \beta_x^2 \kappa_t\Big)$$

where $\mu_x(t)$ is the force of mortality of a person aged x at calendar year t (for example $\mu_{65}(2018)$). The parameters β_x^1 and β_x^2 reflect age-related effects while the parameter κ_t reflects random period related effects. Specifically, κ_t is an index of the level of mortality in calendar year t. β_x^1 is an age-specific component, which is independent of the time, representing the age group shift effect and β_x^2 is a parameter representing how quickly the mortality rate at age x changes when the mortality index changes, i.e., the age group's reaction to a change of period. A big β_x^2 might for example correspond to infantile years where the volatility in $\mu_x(t)$ is quite high while a smaller one usually appears at older ages.

The Lee Carter model is often chosen because of the separation between the time and age effects. Once β_x^1 and β_x^2 have been computed, these parameters can be used for any time t of interest.

The idea is to use historical data to estimate these parameters for a range of ages, x, and periods, t. There are two issues with parameter estimation

1. The estimates depend on the chosen period, e.g., data from 1900-2000 will almost certainly produce different estimates to data from 1950-2000.
2. Solutions are not unique. Indeed, the problem structure is invariant under the transformations:

$$(\beta_x^1, \beta_x^2, \kappa_t) \to (\beta_x^1, \frac{\beta_x^2}{c}, c\kappa_t)$$

$$(\beta_x^1, \beta_x^2, \kappa_t) \to (\beta_x^1 - c\beta_x^2, \beta_x^2, \kappa_t + c)$$

This is overcome by setting two additional constraints, forming an identifiability problem:

$$\sum_x \beta_x^2 = 1 \qquad \sum_t \kappa_t = 0$$

Some researchers prefer the model to fit almost perfectly the last years without actually caring about the earlier fitting errors. Others prefer a good global fit even if some interpolation errors occur in the last years. The latter view is the most usual way of calibrating the model. Indeed, as already mentioned, computing very precise forces of mortality is not very useful as one still needs to make use of assumptions to compute the final survival probabilities.

Summing the formula for $\log \mu_x(t)$ and applying these constraints results in β_x^1 being the average of the $\log \mu_x(t)$ values over time.

$$\beta_x^1 = \frac{\sum_{t=t_1}^{t_n} \log \mu_x(t)}{t_n - t_1 + 1} = \log \left(\prod_{t=t_1}^{t_n} \mu_x(t)^{\frac{1}{t_n - t_1 + 1}} \right)$$

Moreover,

$$\kappa_t = \sum_x (\log \mu_x(t) - \beta_x^1)$$

because of the first constraint. Consequently, β_x^2 can be estimated by regressing, without a constant term, $\log \mu_x(t) - \beta_x^1$ on κ_t separately for each age group x and applying the least squares method

$$\min_{\beta_x^2} \left[\sum_{x,t} \left(\log \mu_x(t) - \beta_x^1 - \beta_x^2 \kappa_t \right)^2 \right] \rightarrow \beta_x^2 = \frac{\sum_{t=t_1}^{t_n} \kappa_t (\log \mu_x(t) - \beta_x^1)}{\sum_{t=t_1}^{t_n} \kappa_t^2}$$

Another method also yields the two unknown parameters β_x^2 and κ_t by the Singular Value Decompostion (SVD) of the matrix $Z := \log \mu_x(t) - \beta_x^1$. Given the decomposition $Z = USV^T$, let us define

$$b_x = \frac{u_1}{\sum_i u_{1i}}$$
$$k_t = \sqrt{s_{11}} \left(\sum_i u_{1i} \right) v_1^T$$

where u_1 is the dominant eigenvector of matrix ZZ^T (or the first column of matrix U), v_1 is the dominant eigenvector of matrix $Z^T Z$ (or the first column of matrix V) and s_{11} is the biggest eigenvalue of either U or V (the first element of the diagonal matrix S). Usually, the time parameter is then recalibrated as described below so that the observed number of deaths equals the estimated number of deaths. Finally, the parameters are a little bit modified so as to satisfy the two constraints for uniqueness:

$$\beta_x^{1,\star} = \beta_x^1 + b_x \bar{k}$$
$$\beta_x^2 = b_x \left(\sum_i b_{1i} \right)$$
$$\kappa_t = (k_t - \bar{k}) \left(\sum_i b_{1i} \right)$$

where $\bar{k} = \frac{1}{t_n - t_1 + 1} \sum_{t=t_1}^{t_n} k_t$ is the arithmetic average of k_t with respect to time t, and $(\sum_i b_{1i})$ is the sum of all the estimated b. The two methods give very similar results.

Once the model parameters have been determined, a better model consists in reestimating the values of κ_t through a certain amount of iterations so that the expected number of deaths fits the observed number of deaths for each period. More precisely,

$$\sum_{x=x_1}^{x_k} d_x(t) = \sum_{x=x_1}^{x_k} E_x(t) \exp\Big(\beta_x^1 + \beta_x^2 \kappa_t\Big) \qquad \forall t = t_1, ..., t_n$$

- if $\sum_{x=x_1}^{x_k} d_x(t) = \sum_{x=x_1}^{x_k} E_x(t) \exp(\beta_x^1 + \beta_x^2 \kappa_t)$ STOP.
- else if $\sum_{x=x_1}^{x_k} d_x(t) < \sum_{x=x_1}^{x_k} E_x(t) \exp(\beta_x^1 + \beta_x^2 \kappa_t)$ the mortality rates have to be decreased by adjusting κ_t

$$\kappa_t^* = \begin{cases} \kappa_t(1+d) & \text{if } \kappa_t < 0 \\ \\ \kappa_t(1-d) & \text{if } \kappa_t > 0 \end{cases}$$

  where d is a small positive number.
- else the mortality rates have to be increased by adjusting κ_t

$$\kappa_t^* = \begin{cases} \kappa_t(1-d) & \text{if } \kappa_t < 0 \\ \\ \kappa_t(1+d) & \text{if } \kappa_t > 0 \end{cases}$$

  where d is a small positive number.

Note that in order to compute the interpolating parameters, the forces of mortality have to be known. These can be calculated based on passed data. The idea is to divide the mortality data into different groups. If we denote by $d_x(t)$ the number of deaths at age x observed during year t and $E_x(t)$ the total exposure to risk at age x and year t then

$$\mu_x(t) = \frac{d_x(t)}{E_x(t)}$$

The total exposure to risk $E_x(t)$ is the time of observation of people aged x during year t. If the number of survivors of age x at time t is $l_x(t)$ then we simply have

$$E_x(t) = 1 \cdot l_{x+1}(t+1) + \sum_{i=1}^{l_x(t)} \mathbb{I}_{T_x^i(t)<1} \cdot T_x^i(t) \qquad \mathbb{I}_{T_x^i(t)<1} = \begin{cases} 1 \text{ if life i dies between age x and x+1} \\ 0 \text{ otherwise} \end{cases}$$

The projections of the period dependant parameter are estimated using time series models

$$\kappa_t = \kappa_{t-1} + a + \epsilon_t \tag{3.2}$$

Where a is the drift and ϵ_t is white noise.

$$\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

with σ_ϵ^2 the historical variance.

$$\sigma_\epsilon^2 = \frac{\sum_{t=t_2}^{t_n} (\Delta\kappa_t - a)^2}{t_n - t_1}$$

It is also important to note that the process κ_t needs to be weakly stationary which means that its mean and covariance have to be constant over time. If this is not the case, the process is transformed into a stationary one by differencing the time serie, i.e., $\Delta\kappa_t = \kappa_t - \kappa_{t-1}$ is modeled (and so on, $\Delta^2\kappa_t = \Delta\kappa_t - \Delta\kappa_{t-1}$). Hence, if d differentiation are needed to obtain stationarity:

$$\Delta^d\kappa_t = \Delta^d\kappa_{t-1} + a + \epsilon_t$$

The parameters defining this stochastic process have to be determined so that they fit the reestimated κ_t as close as possible, usually using a least squares procedure (but maximum likelihood estimations are also possible).

$$\min_a \sum_{t=t_2}^{t_n} (\kappa_t - \kappa_{t-1} - a)^2 \rightarrow a = \frac{\sum_{t=t_2}^{t_n} \kappa_t - \kappa_{t-1}}{t_n - t_2 + 1} = \frac{\kappa_{t_n} - \kappa_{t_1}}{t_n - t_1}$$

The time serie (3.2) is an ARIMA(0,1,0) process. Note that different ARIMA(p,d,q) processes can be used. The general form is:

$$\Delta^d Z_t = a + \sum_{i=1}^{p} \alpha_i \Delta^d Z_{t-i} + \sum_{i=0}^{q} \beta_i \epsilon_{t-i}$$

However, ARIMA(0,1,0) is still the most encountered in the literature.

Once all the parameters are calibrated, forecasts of the force of mortality are easily produced. Noting $\mu_x(t')$ as being the last available number in the historical data, one obtains

$$\mu_x(t'+s) = \mu_x(t') \exp\Big(\beta_x^2(\kappa_{t'+s} - \kappa_{t'})\Big) \qquad \forall x$$

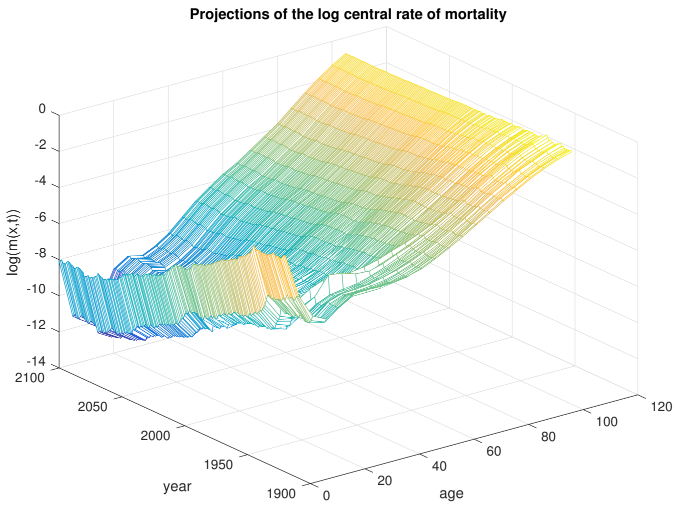
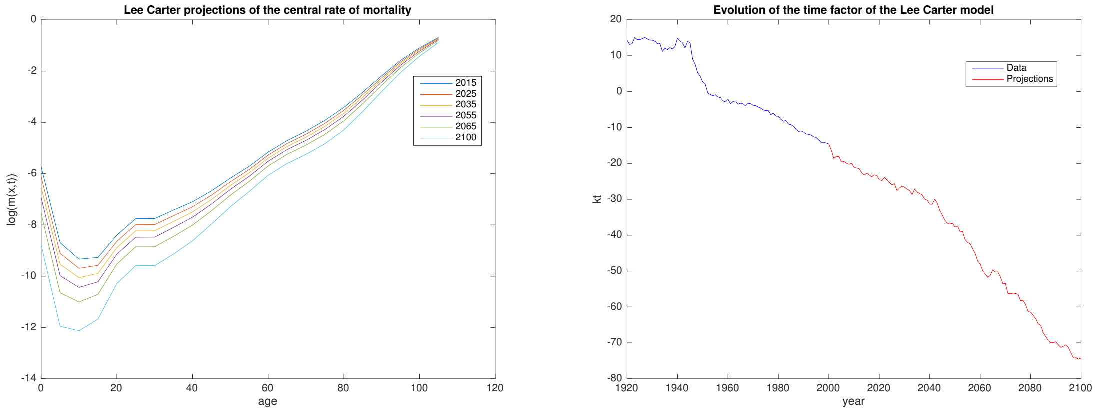


FIGURE 3.1: Evolution of $\log(\mu_x(t))$ as a function of time and age. Values until 2000 are determined by data while the others are computed using the Lee-Carter model.

Figure 3.2a captures a few projections made with this model. The historical data is provided by the "Human Mortality Database"[1]. The simulations are done using the Belgian population. The global improvement in mortality is clearly observable on the surface plot thanks to the decreasing trend. Furthermore, we also observe a significant discontinuity around 1940-1945. This is no surprise as it corresponds to World War II. The shape of a single curved is discussed below.

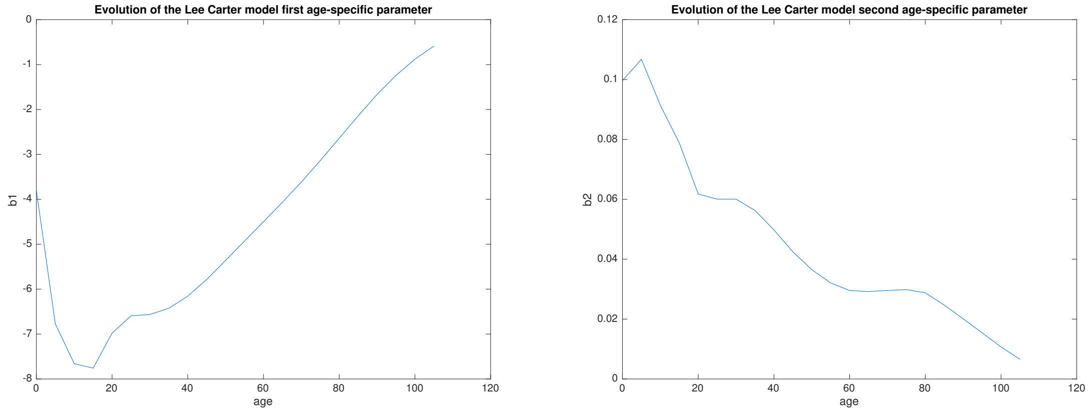
[1]https://www.mortality.org/cgi-bin/hmd/country.php?cntr=BELlevel=1



(A) Projections of the central rate of mortality over the age period 0 - 105 for several calendar years.

(B) Evolution of the index of the level of mortality during calendar year t associated with the Lee-Carter model.

FIGURE 3.2



(A) Evolution of β_x^1 associated with the Lee-Carter model for the age period 0 - 105.

(B) Evolution of β_x^2 associated with the Lee-Carter model for the age period 0 - 105.

FIGURE 3.3

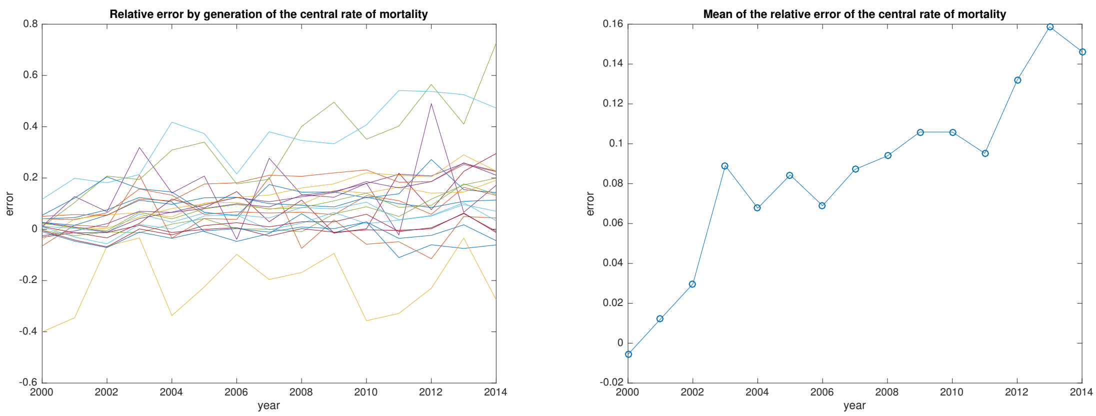
For each calendar year, the evolution of the mortality intensities is similar. It starts quite high due to the newborns' mortality then decreases until reaching a bump over the age period 18-25. The bump is mostly due to alcohol or car accidents. For the rest of the ages, the mortality rates then evolve exponentially (almost linear curve on a logarithmic scale). All of these observations are depicted on Figure 3.2a. Figure 3.2b gives an expected result. The time parameters are decreasing with time. Since the smaller κ_t, the smaller $\mu_x(t)$ and the smaller $q_x(t)$, this is translated in English by saying that people will live longer in the future. For what concerns β_x^1 and β_x^2, we observe on Figure 3.3a that β_x^1 is the leading term in the expression of $\mu_x(t)$ as both evolves in kind of the same way. This is of course expected. Even tough the calendar year has an influence, it is intuitive that age will still be the main factor to death. The decreasing trend of β_x^2 observed on Figure 3.3b shows that the time effect is the most important at younger ages.

Let now have a look at the projection error in order to validate the model. The graphs presented below are graphs representing the average of the relative errors based on a total of 15 000 simulations. Relative errors are computed over a period of 15 years (2000 - 2015) as follows

$$e(x,t) = \frac{\mu_x^{projected}(t) - \mu_x^{actual}(t)}{\mu_x^{actual}(t)} \tag{3.3}$$

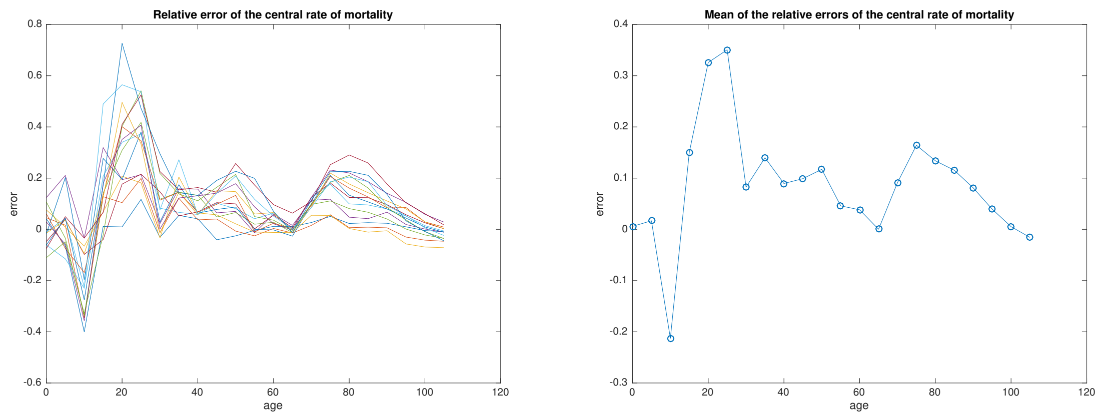
As we can see on Figures 3.4a and 3.4b, most of the generations do have positive error averages. As a result, we can conclude that the Lee-Carter model tends to overestimate mortality intensities. Using Lee-Carter's probabilities is thus a good idea for an insurer selling mortality products as he then almost surely takes a security margin. However, there still remains a small longevity risk even if this risk is minimal compared to what it was using deterministic mortality tables. Numerical comparisons will be made further in this paper when we will discuss pricing. Figure 3.4b also pinpoints the fact that the error tends to increase with time. This is a natural result since uncertainty and inaccuracies also increase with time. It is already complicated to predict something for next year. Imagine how difficult and inaccurate it is to predict something 50 years in advance.

Figures 3.5a and 3.5b outline a decreasing variance of the error with age. We therefore conclude that mortality intensities at older ages are usually better estimated than the ones at younger ages.



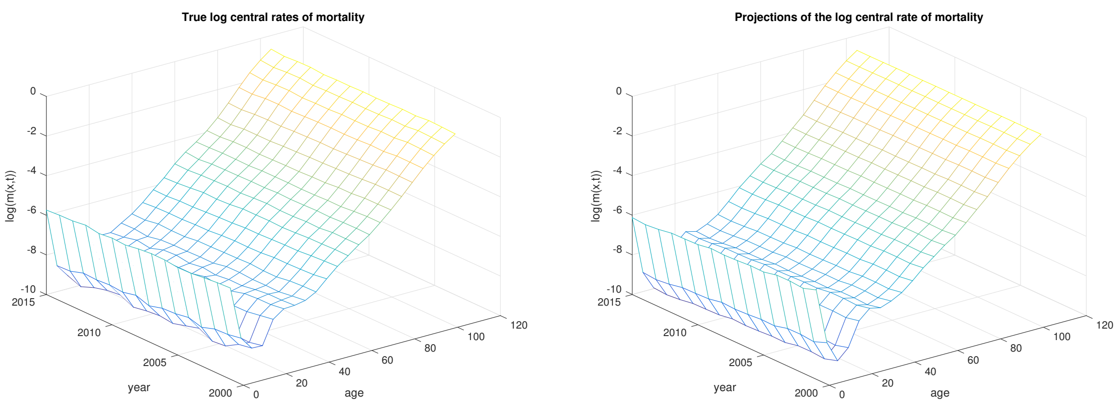
(A) (B)

FIGURE 3.4



(A) (B)

FIGURE 3.5



(A) 3D plot of the evolution of $\log(\mu_x(t))$ as a function of time and age. These are the exact values determined with the Human Mortality Database data.

(B) 3D plot of the evolution of $\log(\mu_x(t))$ as a function of time and age. These are the values as computed by the Lee-Carter model.

FIGURE 3.6

Note that this model is frequently used as basis in the development of new mortality models because it is easily adaptable and extendable to different cohorts and because it can be applied over the whole age range. Furthermore, the model is a function of age and time which is an improvement to the traditional hazard rate model which only depends on age. The projected values of the mortality rates increase with age and decrease with time. Nevertheless, this model also presents some serious drawbacks

1. it is a one factor model and therefore cannot differentiate mortality improvements between age periods.

2. β_x^2 sets the level of uncertainty in future death rates at age x. Hence, it is not possible to decouple the uncertainty level and improvement in future mortality rates.

3. Smoothness issues might arise in the estimated age effect, β_x^2. Let define the standardized residual $\epsilon(x,t) = (D_x(t) - \mu_x(t)E_x(t))/\sqrt{\mu_x(t)E_x(t)}$. If the model and estimation method are correct, these residuals should be independent and normally distributed. The literature states that this is usually not the case. However, this problem has been tackled by other models.

3.1.2 Renshaw and Haberman model with cohort effect (2003)

This model is an extension of the Lee Carter model discussed above. It just takes an additional cohort effect into consideration. The model is

$$\log \mu_x(t) = \beta_x^1 + \beta_x^2 \kappa_t + \beta_x^3 \gamma_{t-x}$$

With the normalization conditions that

$$\sum_x \beta_x^2 = 1 \qquad \sum_t \kappa_t = 0$$

$$\sum_x \beta_x^3 = 1 \qquad \sum_{t,x} \gamma_{t-x} = 0$$

β_x^1, β_x^2 and κ_t are the same parameters as those described for the Lee Carter model. β_x^3 and γ_{t-x} are respectively the cohort effect and the parameter corresponding to it. A cohort is a group of people born the same year, (t-x). They are therefore supposed to have similar environments and health care programs which affect the mortality and life expectancy of people.

Besides taking account of a cohort effect, this model has the advantage that the analysis of the standardized residuals reveals very little dependency on the year of birth. That's why the Lee-Carter model still is the most effective model on a result-versus-calibration difficulty basis. The biggest inconvenient related to this model is its lack of robustness. A change in the range of ages, used to fit the parameters, can give drastically different outcomes.

The particular case where $\beta_x^2 = 1$ and $\beta_x^3 = 1$ corresponds to the Curie model or Age-Period-Cohort model. Curie then uses P splines to fit remaining parameters to avoid smoothness issues.

3.1.3 Cairns Blake Dowd model (2006)

This model is focused on older ages, say 60 and above, with its main application being for financial products for retirement. The basic CBD model is

$$\log \frac{q_x(t)}{1-q_x(t)} = \kappa_t^1 + \kappa_t^2(x-\bar{x})$$

where $q_x(t)$ is the mortality rate at age x in year t and $\bar{x}$ is the average age in the sample used for fitting the model. The parameters κ_t^1 and κ_t^2 represent period effects and can be estimated using ordinary least squares. κ_t^1 and κ_t^2 are modeled by a bivariate random walk with drift of the form:

$$\kappa_t = a + \kappa_{t-1} + C\epsilon_t$$

where $\kappa_t = (\kappa_t^1, \kappa_t^2)^T, a = (a_1, a_2)^T$ is the drift and $\epsilon_t = (\epsilon_t^1, \epsilon_t^2)^T$ is a two dimensional white noise. $C \in \mathcal{R}^{2\times 2}$ is a constant upper triangular matrix derived by the unique Cholesky decomposition of the variance-covariance matrix $V = CC^T$ of the parameter vector κ_t. κ_t^1 presents the level of the logit[2] mortality curve. A reduction in κ_t^1 represents a vertical displacement of the curve and thus an improvement in mortality. κ_t^2 presents the steepness of the logit mortality curve. An increase in κ_t^2 means that the mortality improves more quickly at younger ages than at older ages. The least squares method consists in

$$\min_{\kappa_t^1,\kappa_t^2} \sum_{x,t} \Big(\text{logit}(q_x(t)) - \kappa_t^1 - \kappa_t^2(x-\bar{x})\Big)^2$$

Differentiation with respect to κ_t^1 and κ_t^2 leads to a system of equations for every t

$$\begin{cases} \sum_{x=x_l}^{x_u} \left(\text{logit}(q_x(t)) - \kappa_t^1 - \kappa_t^2(x-\bar{x})\right) & = 0 \\ \sum_{x=x_l}^{x_u} \left[(x-\bar{x})\left(\text{logit}(q_x(t)) - \kappa_t^1 - \kappa_t^2(x-\bar{x})\right)\right] & = 0 \end{cases}$$

Solving this system leads to

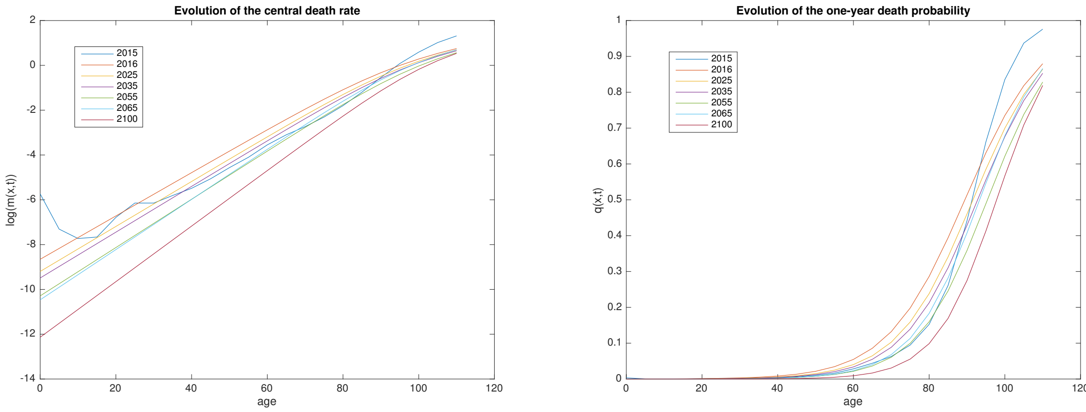
$$\kappa_t^1 = \frac{\sum_{x=x_l}^{x_u} \text{logit}(q_x(t)) - \left[\left(\sum_{x=x_l}^{x_u}\left((x-\bar{x})(x_u-x_l+1) - \sum_{x=x_l}^{x_u}(x-\bar{x})\right)\right)\text{logit}(q_x(t))\right]\sum_{x=x_l}^{x_u}(x-\bar{x})}{(x_u-x_l+1)\sum_{x=x_l}^{x_u}(x-\bar{x})^2 + \left(\sum_{x=x_l}^{x_u}(x-\bar{x})\right)^2}$$

$$\kappa_t^2 = \frac{\sum_{x=x_l}^{x_u}\left((x-\bar{x})(x_u-x_l+1) - \sum_{x=x_l}^{x_u}(x-\bar{x})\right)\text{logit}(q_x(t))}{(x_u-x_l+1)\sum_{x=x_l}^{x_u}(x-\bar{x})^2 + \left(\sum_{x=x_l}^{x_u}(x-\bar{x})\right)^2}$$

[2] $\text{logit}(x) = \ln\left(\frac{x}{1-x}\right)$

After fitting the values of κ_t^1 and κ_t^2, it is necessary to determine the drift and coefficient matrix of the bivariate random walk in order to compute forecasts.
Similar to the Lee-Carter model, the drift a of the randow walk and the variance σ_ϵ^2 of the white noise are respectively given by:

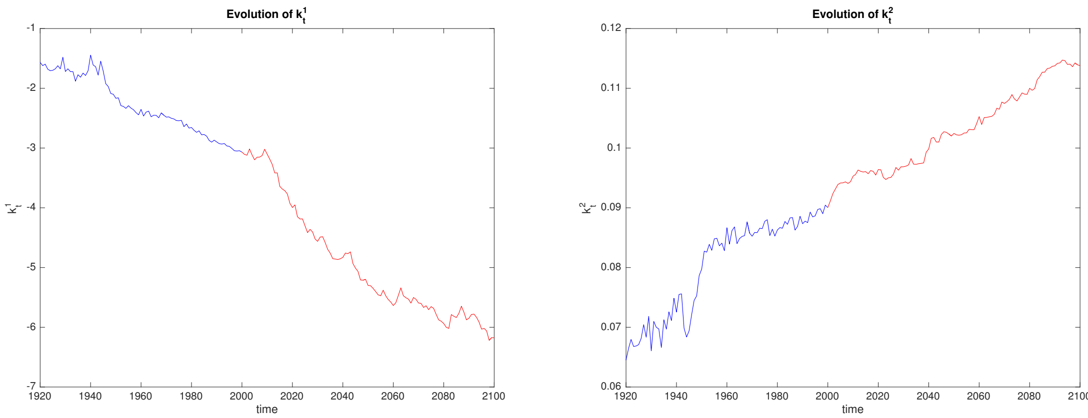
$$a = \begin{bmatrix} \frac{\kappa_{t_n}^1 - \kappa_{t_1}^1}{t_n - t_1} \\ \frac{\kappa_{t_n}^2 - \kappa_{t_1}^2}{t_n - t_1} \end{bmatrix} \qquad \sigma_\epsilon^2 = \begin{bmatrix} \frac{\sum_{t=t_2}^{t_n}\left(\Delta\kappa_t^1 - a_1\right)^2}{t_n - t_1} \\ \frac{\sum_{t=t_2}^{t_n}\left(\Delta\kappa_t^2 - a_2\right)^2}{t_n - t_1} \end{bmatrix}$$



(A) Evolution of $\log(\mu_x(t))$ for different calendar years over the age period 0-105.

(B) Evolution of $q_x(t)$ for different calendar years over the age period 0-105.

FIGURE 3.7

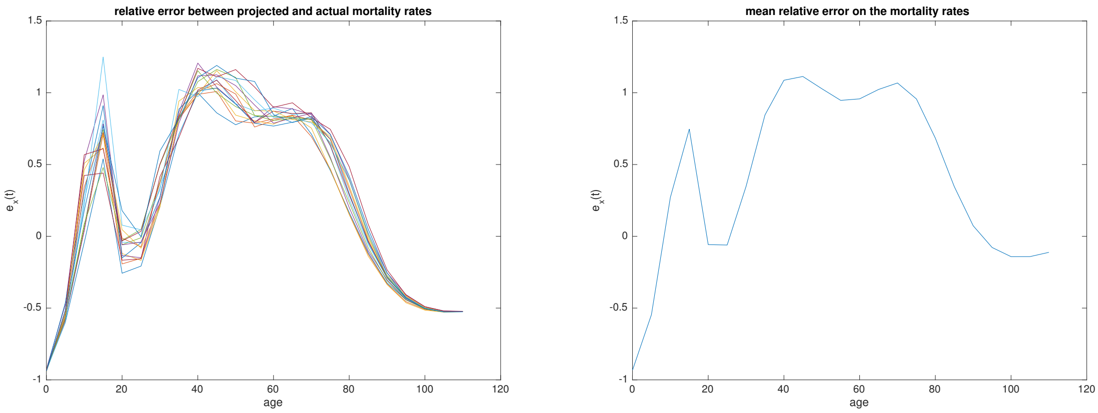


(A) Evolution of κ_t^1 over the time period 1920-2100. In blue, the estimations and in red the projections

(B) Evolution of κ_t^2 over the time period 1920-2100. In blue, the estimations and in red the projections

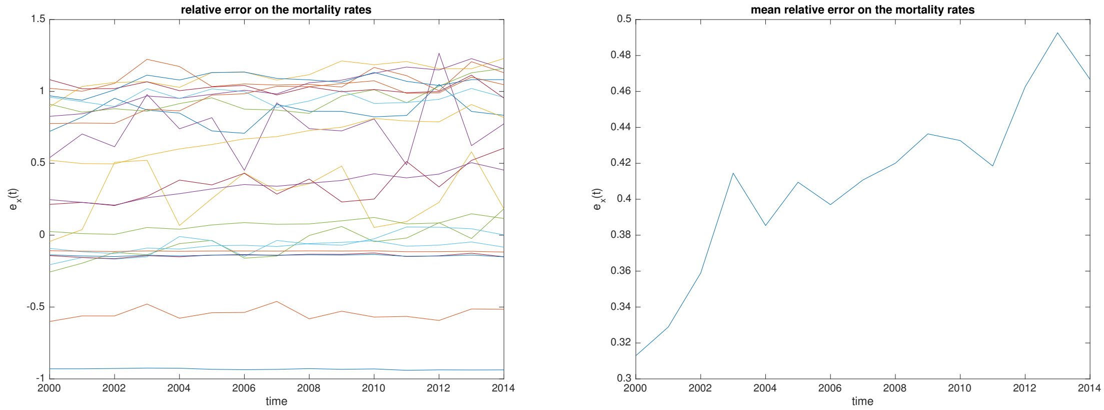
FIGURE 3.8

We notice on Figure 3.7a that the force of mortality for calendar year 2015 has a different shape than all the others as it is made of crude mortality intensities and not a result of the model. The infant mortality and accident bumps do not exist with the CBD model. However, mortality intensities at higher ages have quite the same pattern. As a result, we will mostly use the CBD model for old policyholders. Figure 3.7b outlines the fact that the mortality rates decrease with time while still evolving in an exponential way. This is in line with what would be expected except the important difference between the 2015 rates and the 2016 rates. This difference is due to the model being not perfectly precise. The risk linked to that difference is part of what we called earlier the model risk. κ_t^1 is the intercept of the model. It represents the level of mortality at time t for all ages. Its downward trend as observed on Figure 3.8a outlines the fact that the mortality rates have decreased over time for all ages. The time index κ_t^2 represents the slope of the model. Each age is affected differently. Its increasing trend observed on Figure 3.8b is translated in the fact that the mortality improvements have been bigger at lower ages than at higher ages.



(A) (B)

FIGURE 3.9: Distribution of the relative error between projected and actual mortality rates $q_x(t)$. The projections have been made using the CBD model. Errors are computed for the years 2000 to 2015 and displayed on the left. Their mean is plotted on the right.



(A) (B)

FIGURE 3.10: Distribution of the relative error between projected and actual mortality rates $q_x(t)$. The projections have been made using the CBD model. Errors are computed for the generations $0-5, 5-10, ..., 100-105$ and displayed on the left. Their mean is plotted on the right.

On Figure 3.9b, we observe the trend of the relative projection error. As we could notice from Figure 3.7b, the fit is not very good. The error starts quite big because of the infant mortality that is not represented by the model. There is then a big fluctuation corresponding mainly to the mortality bump that is not captured by the model either. Finally, as ages continue to grow, the error decreases before reaching on almost 0 value at very big ages. That good behaviour at old ages can increase the expected remaining life time of humans, without being very precise about the exact value of that remaining life expectancy, because of the big errors in the early ages. These errors are moreover positive, meaning that the projected mortality rates are often bigger than the actual ones. Once again, survival probabilities are underestimated. From this analysis, the Lee-Carter seems to be a better model for mortality projections.

Figure 3.10b outlines the fact that the error is increasing with time. This is exactly the same as what we obtained for the Lee-Carter model. Values far in the future are complicated to compute precisely.

The CBD model is new data invariant. This means that adding a new year of data does not affect the value of the parameters already computed. The CBD model is the only model with this characterization. Furthermore, just like the Lee Carter model, it is possible to project mortality rates by fitting time series models to the period effects and extend the model with cohort effects.

3.1.4 CBD model with cohort effect (2007)

This model is an extension of the CBD model discussed above. It expands the latter to the third power, adding an age-period-cohort effect. The general form is

$$\log \frac{q_x(t)}{1-q_x(t)} = \kappa_t^1 + \kappa_t^2(x-\bar{x}) + \kappa_t^3\left((x-\bar{x})^2 - \sigma_x^2\right) + \gamma_{t-x}$$

with

$$\bar{x} = \frac{\sum_{x=x_l}^{x_u} x}{x_u - x_l + 1} \qquad \sigma_x^2 = \frac{\sum_{x=x_l}^{x_u}(x-\bar{x})^2}{x_u - x_l + 1}$$

where σ_x^2 and $\bar{x}$ are respectively the variance and the mean of the range of ages used to fit the model parameters, γ_{t-x} is the cohort effect and κ_t^1, κ_t^2, κ_t^3 correspond respectively to the general mortality improvement over time (the level of mortality), the specific improvement for all ages (the slope coefficient, taking into account the fact that the mortality improves slower for older ages than for younger) and finally the age-period related coefficient (the curvature coefficient).

3.1.5 The P-spline model (2004)

This model uses penalized splines (P-splines) to fit the mortality rates in order to deduce future mortality pattern. Generally, the P-spline model takes the form

$$\log \mu_x(t) = \sum_{i,j} \theta_{ij} B_{ij}(x,t)$$

where θ_{ij} are parameters to be determined and $B_{i,j}(x,t)$ are the basic cubic functions used to fit the historical data. $\mu_x(t)$ is the central rate of mortality of an individual aged x at time t. We usually introduce roughness penalties on the parameters θ_{ij} in order to adjust the log-likelihood function. Excessive smoothness in the period dimension can lead to systematic over- or under- estimation of mortality rates, as the model fails to capture what may be genuine environmental fluctuations from one year to the next that affect all ages in the same direction.

3.1.6 Mortality forecasts via reduction factors

A popular way of forecasting mortality rates is by mean of certain reduction factors. Reduction factors are functions of two variables: age x and period since last observation

$t - t'$. For example a common reduction factor used in the UK is as follows:

$$RF(x, t - t') = \alpha_x + (1 - \alpha_x)(1 - f(x))^{\frac{t-t'}{20}}$$

where

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

and

$$f(x) = \begin{cases} 0.55 & \text{if x} < 60 \\ 1 + 0.87\frac{0.55(x-110)+0.29(x-60)}{50} & \text{if } 60 \leq \text{x} \leq 110 \\ 0.29 & \text{if x} > 110 \end{cases}$$

Reduction factors are constructed without a specific modeling structure. Suppose that the current calendar year is t', then for each $t \geq t'$, the death probabilities are given by

$$q_x(t) = q_x(t')RF(x, t - t')$$

A few projections are plotted on Figure 3.11.

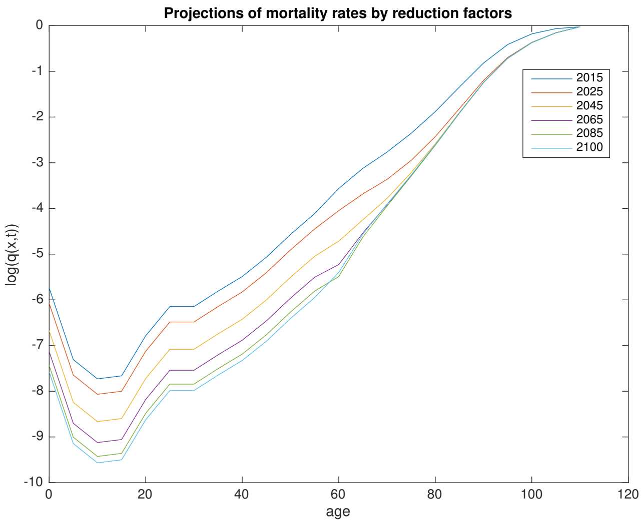


FIGURE 3.11: Mortality rates projection by use of the reduction factors method described above.

The shape of the mortality intensities is comparable with what we obtained with the Lee-Carter model. The advantage of this scheme is that once the reduction factors are known, death probabilities are very easy to compute. The drawback, obviously, consists in computing these reduction factors appropriately.

3.2 Continuous time models

In the classical approach, mortality models are specified through their hazard rate function, which depends upon the age x and the chronological time. Stochastic mortality models introduce uncertainty on the hazard rate function. As a result, $T_x(t)$ is modeled as a stochastic stopping time. More precisely, $T_x(t)$ will be the time when a point process $N_x^t(s)$ jumps from 0 to 1. A life aged x at time t will then be considered alive at time t+s if $N_x^t(s) = 0$ and will be considered dead at time t+s if $N_x^t(s) = 1$. This can be summarized as

$$N_x^t(s) = \begin{cases} 0 & \text{if s} < T_x(t) \\ 1 & \text{if s} \geq T_x(t) \end{cases} \Leftrightarrow N_x^t(s) = \mathbb{I}_{\{T_x(t) \leq s\}}$$

In this setup, the survival probabilities can be expressed as

$${}_sp_x(t) = P\left(T_x(t) > s\right) = \mathbb{E}[\mathbb{I}_{\{T_x(t)>s\}}]$$

A point process is a continuous time stochastic process $(N_t)_{t\geq 0}$ adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$ with $N_0 = 0$ and $N_t \leq \infty$ almost surely and such that its sample path is cadlag (from the French expression: "continu à droite, limite à gauche"), piecewise constant and with jumps of discontinuity equal to 1. A corrolary of the Doob-Meyer decomposition is that if $(N_t)_{t\geq 0}$ is a point process, $\mathcal{F}_t$-adapted, such that $\mathbb{E}[N_t] < \infty$ $\forall t > 0$, then there exists a unique predictable right continuous process Λ_t such that $\Lambda_0 = 0$, $\mathbb{E}[\Lambda_t] < \infty$ and such that $M_t = N_t - \Lambda_t$ is a martingale. If Λ_t can be written as an integral

$$\Lambda_t = \int_0^t \mu_s ds$$

the stochastic process μ_s is called the intensity of the point process. We will denote by $(\mathcal{G}_t)_{t\geq 0}$ the filtration associated to μ_t. N_x^t is a point process of intensity $\mu_x(t)$, if conditionally to the sample path of $\mu_x(t)$ (history in $\mathcal{G}_t$), $N_x(t)$ is a non homogeneous Poisson process. Therefore, the survival probability given $\mathcal{G}_t$ is

$$S_x(t,s) = P(N_x^t(s) = 0|\mathcal{G}_s \cup \mathcal{F}_0) = \mathbb{E}[\mathbb{I}_{\{T_x(t)>s\}}|\mathcal{G}_s \cup \mathcal{F}_0] = \exp\left(-\int_0^s \mu_{x+u}(t+u)ds\right)$$

In reality, $\mathcal{G}_t$ is not observed but $\mathcal{F}_t$ is and we then need to use the Tower property in order to obtain following expression for the survival probabilities

$$\begin{aligned}
S_x(t,s) &= P(N_x^t(s)=0|\mathcal{F}_0)\\
&= \mathbb{E}[\mathbb{I}_{\{T_x(t)>s\}}|\mathcal{F}_0]\\
&= \mathbb{E}\left[\mathbb{E}[\mathbb{I}_{\{T_x(t)>s\}}|\mathcal{G}_t \cup \mathcal{F}_0]\middle|\mathcal{F}_0\right]\\
&= \mathbb{E}\left[\exp\left(-\int_0^s \mu_{x+u}(t+u)du\right)|\mathcal{F}_0\right]
\end{aligned}$$

The probability of being alive at time t+T for a life of age x at time t, alive at time t+s (spot survival probability) is then given by (t $\leq$ s $\leq$ T)

$$S_x(t,s,T) = {}_{T-s}p_{x+s}(t+s) = \mathbb{E}\left[\exp\left(-\int_0^{T-t} \mu_{x+s+r}(t+s+r)dr\right)\middle|\mathcal{F}_s\right]$$

Note that $S_x(t,0,T) = S_x(t,T)$. The intensity of the non homogeneous Poisson process, which we will refer to as the force of mortality, is often modeled as an affine stochastic process. Affine models are models of the type

$$d\mu_x(t) = \alpha(x,t,\mu)dt + \sigma(x,t,\mu)dW(t)$$

$W(t)$ being a standard brownian motion with respect to the filtration $(\mathcal{G}_t)_{t\geq 0}$ and α, σ^2 being linear functions of the stochastic process μ:

$$\alpha(x,t,\mu) = \alpha_1(x,t) + \alpha_2(x,t)\mu_x(t)$$

$$\sigma^2(x,t,\mu) = \sigma_1(x,t) + \sigma_2(x,t)\mu_x(t)$$

α is known as the drift whereas σ is the volatility. For the model to make sense, we have to follow cohorts. As a result, $\mu_x(t)$ is now the force of mortality of a life aged x during the calendar year t. Taking t as the reference time, i.e., $t = t_0 = 0$, this becomes the force of mortality of a life aged at time 0. To follow that cohort, $\mu_x(t)$ will now be used to represent the force of mortality of a life aged $x+t$ at time t. The Ito's lemma states that any function $f(t, N_x(t), \mu_x(t)) : \mathbb{R}^+ \times [0,1] \times \mathbb{R}^+ \to \mathbb{R}$, continuously derivable with respect to time and two times continuously derivable with respect to $\mu_x(t)$, satisifies

$$\begin{aligned}
df = &\left(\frac{\partial f}{\partial t} + \alpha(x,t,\mu)\frac{\partial f}{\partial \mu} + \frac{1}{2}\sigma^2(x,t,\mu)\frac{\partial^2 f}{\partial \mu^2}\right)dt + \sigma^2(x,t,\mu)\frac{\partial^2 f}{\partial \mu}dW_t\\
&+(f(t,1,\mu_x(t)) - f(t,0,\mu_x(t)))dN_x(t)
\end{aligned} \tag{3.4}$$

It follows from (3.4) that $S_x(t,s,T)$ is solution of the stochastic differential equation

$$\frac{\partial S}{\partial s} + \alpha(x,t,\mu)\frac{\partial S}{\partial \mu} + \frac{1}{2}\sigma^2(x,t,\mu)\frac{\partial^2 S}{\partial \mu^2} - S(t,0,\mu_x(t))\mu_x(s)\mathbb{I}_{\{N_x(s)=0\}} = 0 \tag{3.5}$$

with the limit condition $S_x(t,T,T) = 1$ if $N_x^t(T) = 0$ and $S_x(t,T,T) = 0$ if $N_x^t(T) \neq 0$. Affine models for the force of mortality are usually used because they allow to find survival probabilities of the form

$$S_x(t,s,T) = \exp\left(A(s,T) - B(s,T)\mu_x(s)\right)$$

Replacing that expression into (3.5), the system in A and B to be solved is:

$$\begin{cases} \frac{\partial A}{\partial t} - \alpha_1(x,t)B + \frac{1}{2}\sigma_1(x,t)B^2 & = & 0 \\ \frac{\partial B}{\partial t} + \alpha_2(x,t)B - \frac{1}{2}\sigma_2(x,t)B^2 + 1 & = & 0 \end{cases}$$

with limit conditions: $A(T,T) = B(T,T) = 0$. These equations are called the generalized Riccati ODEs. Henceforth, we will focus on affine processes from now on. They can be classed into two groups: the mean reverting processes and the non mean reverting processes.

mean reverting processes Examples of mean reverting processes are given below

- CIR process: $d\mu_x(t) = k(\gamma - \mu_x(t))dt + \sigma\sqrt{\mu_x(t)}dW(t)$
- mean reverting with jumps: $d\mu_x(t) = k(\gamma - \mu_x(t))dt + dJ(t)$
- VAS process: $d\mu_x(t) = k(\gamma - \mu_x(t))dt + \sigma dW(t)$

γ is called the mean-reversion level. It is the level towards which $\mu_x(t)$ tends to converge if we wait sufficiently. The pace at which that convergence occurs is given by k. Mean reverting processes are not very famous in the literature for mortality modeling. Their resulting survival function fails to capture the rectangularization phenomenon and produces unrealistic survival probabilities at very old ages.

non mean reverting processes Non mean reverting affine processes are thus the model by default. $\mu_x(t)$ is usually chosen to be either a Feller or an Ohrnstein-Uhlenbeck process. The Ornstein Uhlenbeck process is a derivative of the more general VAS process expressed above in the case where $\gamma = 0$. The Feller process is a derivative of the CIR process with $\gamma = 0$.

3.2.1 The Ornstein Uhlenbeck process without jumps

The model for the intensity of mortality is given by

$$d\mu_x(t) = a\mu_x(t)dt + \sigma dW(t) \qquad \text{with a} > 0 \text{ and } \sigma \geq 0$$

where W(t) is a standard Brownian motion. Solving this stochastic differential equation leads to

$$\mu_x(t) = \mu_x(0)e^{at} + \sigma \int_0^t e^{a(t-u)} dW(u)$$

We can rewrite above expression in order to isolate $\mu_x(0)$

$$\mu_x(0) = \mu_x(s)e^{-as} - \sigma \int_0^s e^{-au} dW(u)$$

The final result for t>s is given by

$$\mu_x(t) = \mu_x(s)e^{a(t-s)} + \sigma \int_s^t e^{a(t-u)} dW(u)$$

The integral of $\mu_x(t)$ is then equal to

$$\int_s^T \mu_x(t)dt = \frac{1}{a}(e^{a(T-s)} - 1)\mu_x(s) + \sigma \int_s^T \int_s^t e^{a(t-u)} dW(u)dt$$

changing the order of integration leads to

$$\int_s^T \mu_x(t)dt = \frac{1}{a}(e^{-a(T-s)} - 1)\mu_x(s) + \sigma \int_s^T \int_u^T e^{a(t-u)} dW(u)dt$$

Which then gives

$$\int_s^T \mu_x(t)dt = \frac{1}{a}(e^{-a(T-s)} - 1)\mu_x(s) + \sigma \int_s^T \frac{1}{a}(e^{-a(T-u)} - 1)dW(u)$$

Let $B(s,T) = \frac{1}{a}(e^{-a(T-s)} - 1)$, we then easily prove that

$$\begin{aligned}
\mathbb{E}\left[\int_s^T \mu_x(t)dt \middle| \mathcal{F}_s\right] &= \mu_x(s)B(s,T) \\
\mathbb{V}\left[\int_s^T \mu_x(t)dt \middle| \mathcal{F}_s\right] &= \sigma^2 \mathbb{E}\left[\int_s^T B(u,T)^2 du\right] \\
&= \frac{\sigma^2}{a^2}\mathbb{E}\left[\int_s^T e^{2a(T-u)} - 2e^{a(T-u)} + 1du\right] \\
&= \frac{\sigma^2}{a^2}\left(\frac{a}{2}B(s,T)^2 - B(s,T) + T - s\right)
\end{aligned}$$

Consequently, we conclude that

$$\int_s^T \mu_x(t)dt \sim \mathcal{N}\left(\mu_x(s)B(s,T), \frac{\sigma^2}{a^2}\left(\frac{a}{2}B(s,T)^2 - B(s,T) + T - s\right)\right.$$

so that

$$S_x(t,s,T) = \mathbb{E}\left[\exp\left(-\int_s^T \mu_x(u)du\right)\middle|\mathcal{F}_s\right]$$

is actually the expectation of a log-normal random variable. It is the probability of survival from age x+s to age x+T given all the information gathered until time s. As a result,

$$\begin{aligned} S_x(t,s,T) &= \mathbb{E}\left[\exp\left(-\int_s^T \mu_x(u)du\right)\middle|\mathcal{F}_s\right] \\ &= \exp\left(-\mathbb{E}\left[\int_s^T \mu_x(u)du\middle|\mathcal{F}_s\right] + \frac{1}{2}\mathbb{V}\left[\int_s^T \mu_x(u)du\middle|\mathcal{F}_s\right]\right) \end{aligned}$$

Simply replacing the expression of the expectation and the variance then leads to

$$S_x(t,s,T) = \exp\left(A(s,T) - B(s,T)\mu_x(s)\right)$$

where B(s,T) and A(s,T) are given by

$$\begin{aligned} B(s,T) &= \frac{1}{a}(e^{a(T-s)} - 1) \\ A(s,T) &= \frac{\sigma^2}{2a^2}\left(\frac{a}{2}B(s,T)^2 - B(s,T) + T - s\right) \\ &= \frac{\sigma^2}{4a}B(s,T)^2 - \frac{\sigma^2}{2a^2}\left(B(s,T) - (T-s)\right) \\ &= \frac{\sigma^2}{4a^3}e^{2a(T-s)} - \frac{\sigma^2}{a^3}e^{a(T-s)} + \frac{3\sigma^2}{4a^3} + \frac{\sigma^2}{2a^2}(T-s) \end{aligned}$$

Since we used the least square error approach for the discrete models, we will now use the maximum likelihood estimation to compute parameters a and σ. Therefore, we need to remind two properties of the stochastic integral

1. $\mathbb{E}\left[\int_0^t X(s)dW(s)\right] = 0$
2. ITO's isometry: $\mathbb{E}\left[\left(\int_0^t X(s)dW(s)\right)^2\right] = \mathbb{E}\left[\int_0^t X^2(s)ds\right]$

Consequently, we determine that $\mu_x(t) \sim \mathcal{N}\left(\mu_x(0)e^{at}, \sigma^2\frac{e^{2at}-1}{2a}\right)$. Moreover, we have

$$\mu_x(t+\delta) \sim \mathcal{N}\left(\mu_x(t)e^{a\delta}, \sigma^2\frac{e^{2a\delta}-1}{2a}\right)$$

It is therefore possible to compute projections as follows

$$\mu_x(t+\delta) = \mu_x(t)e^{a\delta} + \sigma\sqrt{\frac{e^{2a\delta}-1}{2a}}z$$

with z being a realization of a random variable $Z \sim \mathcal{N}(0,1)$. As a result, if $\mu_x(t)$ and $\mu_x(t+\delta)$ are two observations separated by a time δ then the conditional density function of $\mu_x(t+\delta)$ given $\mu_x(t)$ is given by

$$f(\mu_x(t+\delta)|\mu_x(t);a,\hat{\sigma}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}}\exp\left(\frac{-1}{2\hat{\sigma}^2}(\mu_x(t+\delta)-\mu_x(t)e^{a\delta})^2\right)$$

with $\hat{\sigma}^2 = \sigma^2\frac{e^{2a\delta}-1}{2a}$.
This is just the distribution of a random variable normally distributed. The likelihood associated with a vector of n observations can be expressed as

$$L(a,\hat{\sigma}) = \prod_{t=0}^{n-1} f(\mu_x(t+\delta)|\mu_x(t);a,\hat{\sigma}) = (\frac{1}{\sqrt{2\pi\hat{\sigma}^2}})^n \exp\left(\frac{-1}{2\hat{\sigma}^2}\sum_{t=0}^{n-1}(\mu_x(t+\delta)-\mu_x(t)e^{a\delta})^2\right)$$

The objective is to find values of a and $\hat{\sigma}^2$ that maximize above expression. Since the logarithmic function is monotonic we can apply it on $L(a,\hat{\sigma})$ without changing the argument of the maximum. We now obtain

$$l(a,\hat{\sigma}^2) = \frac{-n}{2}\log\left(2\pi\hat{\sigma}^2\right) - \frac{1}{2\hat{\sigma}^2}\sum_{t=0}^{n-1}(\mu_x(t+\delta)-\mu_x(t)e^{a\delta})^2$$

Optimal conditions are given by

$$\begin{cases} \frac{\partial l}{\partial a} = 0 \\ \frac{\partial l}{\partial \hat{\sigma}^2} = 0 \end{cases} \Leftrightarrow \begin{cases} \frac{\delta e^{a\delta}}{\hat{\sigma}^2}\sum_{t=0}^{n-1}(\mu_x(t+\delta)-\mu_x(t)e^{a\delta})\mu_x(t) = 0 \\ \frac{-n}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\sum_{t=0}^{n-1}(\mu_x(t+\delta)-\mu_x(t)e^{a\delta})^2 = 0 \end{cases}$$

Solving the two-equation system in the two unknowns leads finally to

$$a = \frac{1}{\delta}\log\left(\frac{\sum_{t=0}^{n-1}\mu_x(t)\mu_x(t+\delta)}{\sum_{t=0}^{n-1}\mu_x(t)^2}\right)$$

$$\hat{\sigma}^2 = \frac{\sum_{t=0}^{n-1}(\mu_x(t+\delta)-\mu_x(t)e^{a\delta})^2}{n}$$

Estimation and projection of the force of mortality described by a OU process is easy because of its normal distribution. However, if the force of mortality is normally distributed, this also means that it is possible for it to be negative

$$P(\mu_x(t) < 0) = P\left(Z < \frac{-\mu_x(0)e^{at}}{\sigma\sqrt{\frac{e^{2at}-1}{2a}}}\right) = \Phi\left(\frac{-\mu_x(0)e^{at}}{\sigma\sqrt{\frac{e^{2at}-1}{2a}}}\right)$$

This model is inadequate for the modeling of the intensity on a theoretical point of view because the survival probability for t large enough becomes an increasing function of age and the probability of surviving forever tends to infinity. Moreover, the probability of $\mu_x(t)$ being negative is not zero. However, the model turns out to be quite adequate in the applications as the probability of a negative force of mortality is very low. This is because, in practice, $\hat{\sigma}^2$ is small and a is big. As shown on Figure 3.12, the decreasing property of the force of mortality with time is clear. Unfortunately, this model is not able to capture the impact of a high infant mortality nor to capture the accident bump.

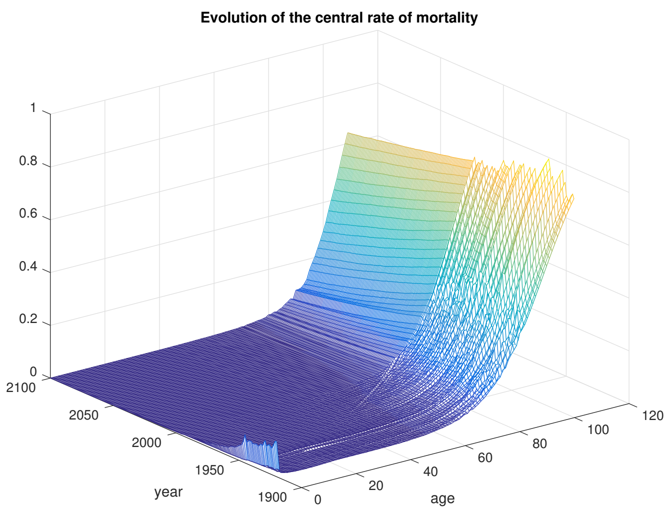
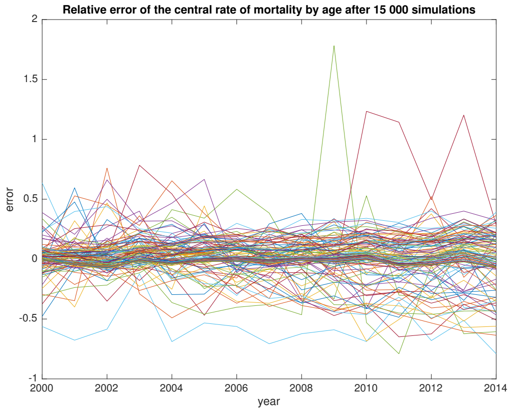
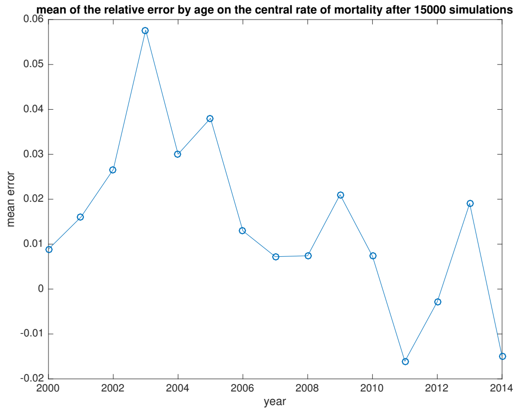


FIGURE 3.12: projections of $\mu_x(t)$ using the OU model on Belgian data. The projections cover the time period 2000-2100 with the first 15 years being crude data for validation.

Figures 3.13a, 3.13b, 3.14a and 3.14b are a representation of the relative errors of the central rate of mortality. These errors are computed as in (3.3). It is important to comment those results so as to analyze how the model behaves compared to known results.

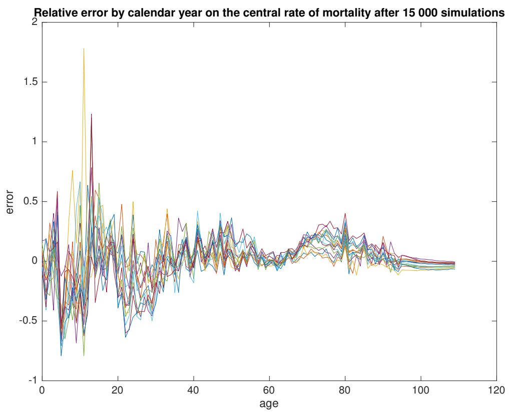


(A)

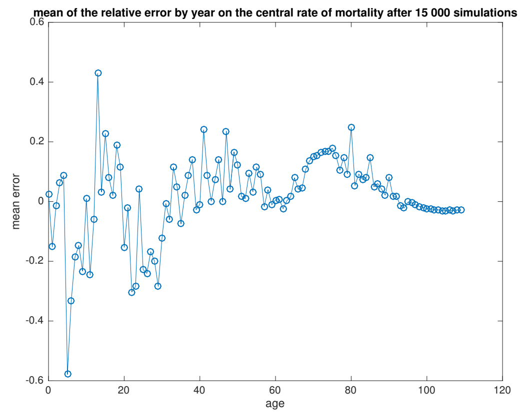


(B)

FIGURE 3.13



(A)



(B)

FIGURE 3.14

Figure 3.13a shows that most of the errors are small (approximately $\pm 10\%$) and mostly positive. Hence, the model tends to overestimate the force of mortality and thus underestimate survival probabilities. Just like the Lee-Carter model. However, the error is a little bit smaller. Indeed, from Figure 3.13b, we observe that the error presents small variations around zero. This is the most accurate model seen so far. We see on Figure 3.14b that the mean of the error remains positive for higher ages. Consequently, mortality intensities are overestimated at older ages. There is no such trend at younger ages where the error is kind of equally distributed between positive and negative values.

3.2.2 Ornstein Uhlenbeck process with jumps

We will now quickly introduce an extension of the OU model that adds a jump component J. Jump additions are relevant because the intensity of mortality might change suddenly due to unexpected external events. Uncertainty in the mortality rates projection is usually obtained be adding one or multiple stochastic brownian processes into the model. Jump processes might also be of interest to replace or extend them. The idea behind jump processes is intuitive. On one hand, suppose a country enters in war, the mortality rates of the population will highly increase during the fight period (a jump in mortality rates). On the other hand, suppose someone finds a cure to cancer, mortality rates will then decrease (again, a jump in mortality rates). These two examples outline the two different effects a jump process can have on the force of mortality. The first one has a direct effect on $\mu_x(t)$ if t is a year of war. The same effect would happen if a certain natural disaster happens during year t. It is a temporary change in mortality rates. The second one is less temporary. It needs time to have a global effect on the mortality rates. The effect will therefore not be observable directly on the force of mortality but on the improvement rate of the force of mortality (this is for example the case of the longevity improvement of humans). To summarize, jumps should correspond to discontinuity points of the intensity process. Let J be a pure compound Poisson jump process with Poisson arrival times of mean $l > 0$ and exponential jump sizes with mean $\lambda < 0$. Assume independence between the Brownian motion $W(t)$ and the Poisson process. The stochastic differential equation describing $\mu_x(t)$ is now given by:

$$d\mu_x(t) = a\mu_x(t)dt + \sigma dW(t) + dJ(t)$$

If $\beta(T) = B(0,T)$ and $\alpha(T) = A(0,T)$, the processes defining the survival function of a person aged x at time t for a period T are now solutions of:

$$\begin{cases} \beta'(T) &= & -1 - a\beta(T) & \text{with boundary condition } \beta(0) = 0 \\ \alpha'(T) &= & -\frac{\sigma\beta^2(T)}{2} + l\frac{\lambda\beta(T)}{1-\lambda\beta(T)} & \text{with boundary condition } \alpha(0) = 0 \end{cases}$$

Once these functions are computed, the survival probabilities are simply $S_x(t,0,T) = \exp(\alpha(T) - \beta(T)\mu_x(0))$. The differential equation on $\beta(T)$ is the same as previously but the jump component modify the equation in $\alpha(T)$. The solution of this system is given by

$$\begin{cases} \beta(T) &= & \frac{1}{a}(e^{aT} - 1) \\ \alpha(T) &= & -\left(\frac{\sigma}{2a^2} + \frac{l\lambda}{a-\lambda}\right)T + \frac{\sigma}{a^3}e^{aT} - \frac{\sigma}{4a^3}e^{2aT} - \frac{3\sigma}{4a^3} + \frac{l}{a-\lambda}\log\left(1 - \frac{\lambda}{a} + \frac{\lambda}{a}e^{aT}\right) \end{cases}$$

The condition that has to be fulfilled in order to obtain a solution that makes sense is $\beta(T) > \frac{1}{\lambda}$ if $\lambda > 0$ and $\beta(T) < \frac{1}{\lambda}$ if $\lambda < 0$. Indeed, the logarithm can only be taken on a positive number. Also, this problem does not solve the problem of the force of mortality becoming negative with positive probability. Hence, the survival function can either be increasing or decreasing with time. The condition to be fulfilled in order to have a decreasing survival function is:

$$\frac{d}{dT}S_x(t,T) < 0 \Rightarrow -\frac{\sigma\beta^2(T)}{2} + l\frac{\lambda\beta(T)}{1-\lambda\beta(T)} + \mu_x(t)\left(1 + a\beta(T)\right) < 0.$$

which can be translated in

$$\frac{1}{a}\log(1+\beta_2 a) < T < \frac{1}{a}\log(1+\beta_1 a)$$

for $\lambda > 0$ or

$$T > \frac{1}{a}\log(1+\beta_2 a) \vee T < \frac{1}{a}\log(1+\beta_1 a)$$

for $\lambda < 0$, where

$$\beta_1 = \frac{(-\sigma - 2\lambda a\mu_x(t)) - \sqrt{(-\sigma + 2\lambda a\mu_x(t))^2 - 8\lambda^2\sigma(-\mu_x(t)+l)}}{2\lambda\sigma^2}$$

and

$$\beta_2 = \frac{(-\sigma - 2\lambda a\mu_x(t)) + \sqrt{(-\sigma + 2\lambda a\mu_x(t))^2 - 8\lambda^2\sigma(-\mu_x(t)+l)}}{2\lambda\sigma^2}$$

3.2.3 The Feller process without jump

Another very famous short rate model for the force of mortality is the Feller (or Cox Ingressol Ross) model without mean reverting term.

$$d\mu_x(t) = a\mu_x(t)dt + \sigma\sqrt{\mu_x(t)}dW(t) \qquad \text{with a} > 0 \text{ and } \sigma > 0$$

It can be seen as a model forcing the force of mortality to remain close to 0 with a "force" of a. The big difference with the Ornstein Uhlenbeck model is that the volatility now depends on $\mu_x(t)$ as well. Hence when $\mu_x(t)$ tends to zero, the volatility vanishes and a positive a then forces the force of mortality back up. It thus insures that the force of mortality remains positive at all times!
Solving this stochastic differential equation gives

$$\mu_x(t) = \mu_x(0)e^{at} + \sigma\int_0^t e^{a(t-s)}\sqrt{\mu_x(s)}dW(s)$$

Consequently, the survival probability takes the form:

$$S_x(t,s,T) = e^{A(s,T)-B(s,T)\mu_x(s)}$$

with $A(s,T)$ and $B(s,T)$ solutions of

$$\begin{cases} \frac{\partial B(s,t)}{\partial t} & = & 1 + aB(s,t) - \frac{1}{2}\sigma^2 B^2(s,t) & \text{with boundary condition } B(t,t) = 0 \\ \frac{\partial A(s,t)}{\partial t} & = & 0 & \text{with boundary condition } A(t,t) = 0 \end{cases}$$

Solving this (which is far from being easy) leads to

$$\begin{cases} B(s,T) & = & \frac{2(e^{\lambda(T-s)}-1)}{(\lambda-a)e^{\lambda(T-s)}+a+\lambda} \\ A(s,T) & = & 0 \end{cases}$$

with $\lambda = \sqrt{a^2 + 2\sigma^2}$.

The survival probability is bounded between 0 and 1 which is desirable. The calibration of this model gives very good results and a negative value for the force of mortality, leading to increasing survival probabilities with time, is avoided if $\mu_x(0) > 0$.

In order to calibrate the parameters of the model, we will work on the more general CIR process

$$d\mu_x(t) = k(\gamma - \mu_x(t))dt + \sigma\sqrt{\mu_x(t)}dW(t)$$

We then recover the parameters of the Feller process by taking $\gamma = 0$ and $k = -a$ as constraints. The CIR model suggests that the force of mortality is pulled toward γ at a pace of k. As has been done for the Ornstein-Uhlenbeck process, we will maximize the likelihood of some observations in order to calibrate the unknown parameters k, γ and σ. Feller, in his paper of 1951, proves that the distribution of $\mu_x(t+\delta)$ given $\mu_x(t)$ is a non central chi square distribution with $2q+2$ degrees of freedom and a non centrality parameter equal to

$$\lambda = 2ce^{a\delta} = \frac{-4ae^{a\delta}\mu_x(t)}{\sigma^2(1-e^{a\delta})}$$

Suppose that two observations are separated by a time δ, the conditional probability density of an observation $\mu_x(t+\delta)$ given an observation $\mu_x(t)$ is therefore of the form

$$\begin{aligned} f(\mu_x(t+\delta)|\mu_x(t);k,\gamma,\sigma) &= ce^{-u_x(t)-v_x(t+\delta)}\left(\frac{v_x(t+\delta)}{u_x(t)}\right)^{\frac{q}{2}} B_q(2\sqrt{u_x(t)v_x(t+\delta)}) \\ &= \frac{1}{2}\exp\left(-\frac{1}{2}\left(\mu_x(t+\delta)+\lambda\right)\right)\left(\frac{\mu_x(t+\delta)}{\lambda}\right)^{\frac{q}{2}} B_q(2\sqrt{\mu_x(t+\delta)\lambda}) \end{aligned}$$

with $B_q(.)$ being the modified Bessel function of the first kind and of order q and with the parameters as follows

$$
\begin{aligned}
c &= \tfrac{2k}{\sigma^2(1-e^{-k\delta})} \\
u_x(t) &= c\mu_x(t)e^{-k\delta} \\
v_x(t+\delta) &= c\mu_x(t+\delta) \\
q &= \tfrac{2k\gamma}{\sigma^2} - 1
\end{aligned}
$$

The likelihood expression of N observations is

$$
\begin{aligned}
L(k,\gamma,\sigma) &= \prod_{t=0}^{N-1} f(\mu_x(t+\delta)|\mu_x(t);k,\gamma,\sigma) \\
&= c^N \exp\left(-\sum_{t=0}^{N-1} u_x(t) - \sum_{t=0}^{N-1} v_x(t+\delta)\right) \prod_{t=0}^{N-1} \left(\frac{v_x(t+\delta)}{u_x(t)}\right)^{\frac{q}{2}} B_q\left(2\sqrt{u_x(t)v_x(t+\delta)})\right)
\end{aligned}
$$

From which it is easy to derive the log-likelihood function:

$$
\begin{aligned}
ll(k,\gamma,\sigma) = N\log(c) - \sum_{t=0}^{N-1} u_x(t) - \sum_{t=0}^{N-1} v_x(t+\delta) + \sum_{t=0}^{N-1} \frac{q}{2}\log\left(\frac{v_x(t+\delta)}{u_x(t)}\right) + \\
\sum_{t=0}^{N-1} \log\left(B_q(2\sqrt{u_x(t)v_x(t+\delta)})\right)
\end{aligned}
$$

In practical, we can also work via discretization of the stochastic differential equation

$$\mu_x(t) - \mu_x(t-1) \approx a\mu_x(t-1)\Delta t + \sigma\sqrt{\mu_x(t-1)}\Delta W_t$$

and use the fact that $\mu_x(t) - \mu_x(t-1) \sim \mathcal{N}\left(a\mu_x(t-1), \sigma\sqrt{\mu_x(t-1)}\right)$. Numerical results are plotted below.

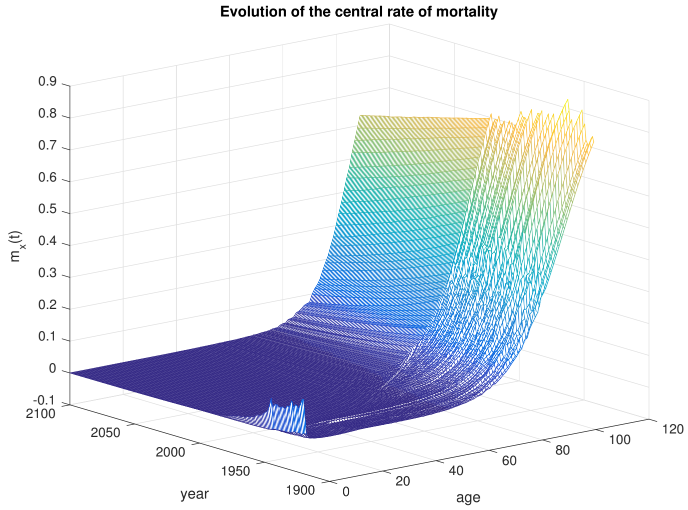


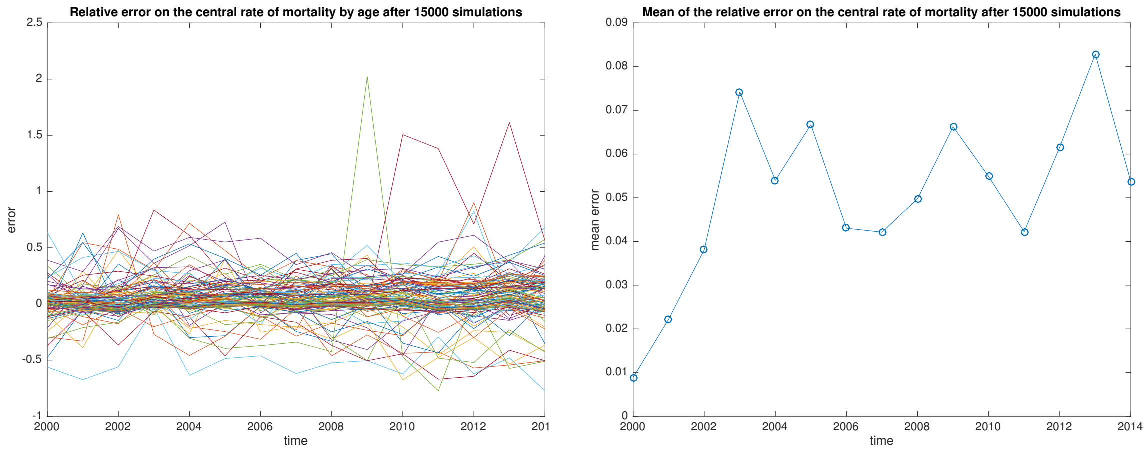
FIGURE 3.15

As for the Ornstein-Uhlenbeck model, the force of mortality does not capture the infant mortality phenomenon nor the accident bump. The slightly decreasing surface outlines decreasing mortality intensities with time as desired. Also notice the change in the death curve shape. It is becoming more and more steep with the calendar year. We can relate that to the rectangularization characteristic of the cumulative distribution function discussed earlier in this paper.

Figures 3.16a and 3.16b illustrate that for the most of the ages, the error is mainly positive. As for the OU model, the Feller model also underestimates survival probabilities. The error being larger when the calendar year is further in the future. This makes sense as there are then more uncertainties.

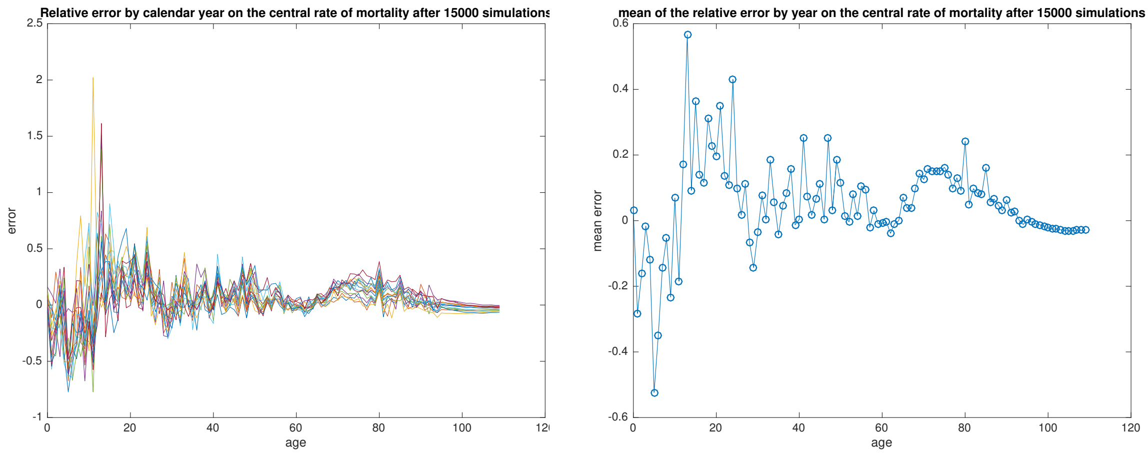
As expected, the relative error is more important at younger ages as the infant mortality is not represented by the model. However, the older the ages, the smaller the volatility of the error around 0.

It is also interesting to note that the CIR and OU models give similar results.



(A) (B)

FIGURE 3.16



(A) (B)

FIGURE 3.17

3.2.4 The Feller process with jump

$$d\mu_x(t) = a\mu_x(t)dt + \sigma\sqrt{\mu_x(t)}dW(t) + dJ(t) \qquad \text{with a} > 0 \text{ and } \sigma > 0$$

The functions $\alpha(T) = A(0,T)$ and $\beta(T) = B(0,T)$ that enter the survival probability solve the following system

$$\begin{cases} \beta'(T) &=& -1 + a\beta(T) + \frac{1}{2}\sigma^2\beta^2(T) & \text{with boundary condition } \beta(0) = 0 \\ \alpha'(T) &=& \frac{l\lambda\beta(T)}{1-\lambda\beta(T)} & \text{with boundary condition } \alpha(0) = 0 \end{cases}$$

Again, the introduction of the jump component leads to the requirement that $\beta(t) > \frac{1}{\lambda}$ ∀t. The literature outlines that the extra stochastic component to the model gives significantly better results in terms of goodness of the fit and adds richness and flexibility.

As we could have guessed, there exists an enormous variety of continuous models to describe the dynamics of $\mu_x(t)$. Each model may or may not present desirable features such as mean reversion, positive drift, affinity, proportional volatility,... The following model tries to reunite all these characteristics in one general expression.

3.2.5 A general affine model

A general affine model can be formulated as:

$$d\mu_x(t) = \left[\mu_x(t)\left(\frac{\bar{\mu}'_{x+t}}{\bar{\mu}_{x+t}} + \beta'_x(t) - k\right) + k\tilde{\mu}_{x+t}\right]dt + \gamma_x(t)\sqrt{\mu_x(t) - \mu^*_{x+t}}dW(t) \quad (3.6)$$

The different elements of this equation have following meanings

- An initial life table $\bar{\mu}_x(t)$ corresponding to the present force of mortality. It is, for example, the deterministic life table used by an insurer.
- A deterministic improvement function $\beta_x(t)$, depending on age and time (prospective aspect). This term allows to consider improvements in mortality as observed in practice on crude data.
- A mean reversion effect to a limit table $\tilde{\mu}_x(t)$. Mean reversions enable to center the force of mortality around a certain level.
- A stochastic noise with a volatility controlled by a lower bound table $\mu^*_x(t)$. As for the Feller process, this allows to make the volatility vary with respect to the lower bound table. A force of mortality way higher than the lower bound table, gives a high volatility whilst the extreme scenario $\mu_x(t) = \mu^*_{x+t}$ vanishes the volatility

The survival probabilities are therefore given by expression

$$S_x(t, s, T) = \exp\left(A(s, T) - B(s, T)\mu_x(s)\right)$$

with functions $A(t, T)$ and $B(t, T)$ solutions of

$$\left\{\begin{array}{rcl} \frac{\partial A}{\partial t} - k\tilde{\mu}_{x+t}B - \frac{1}{2}\gamma_x(t)^2\mu^*_{x+t}B^2 & = & 0 \\ \frac{\partial B}{\partial t} + \left(\frac{\bar{\mu}'_{x+t}}{\bar{\mu}_{x+t}} + \beta'_x(t) - k\right)B - \frac{1}{2}\gamma^2_x(t)B^2 + 1 & = & 0 \end{array}\right.$$

3.2.6 The forward mortality modeling framework

We here introduce forward mortality modeling frameworks where we model the instantaneous force of mortality $h_x(t,T)$. In what follows, $h_x(t,T)$ represents the force of mortality at time T of a life aged x at time t < T, determined at time t. We understand why the denomination forward is relevant. It is a little bit as if we calculated today a future value. Of course, the forward force of mortality satisfies

$$\lim_{T \to t} h_x(t,T) = \mu_x(t)$$

where $\mu_x(t)$ is the force of mortality of a life aged x at time t. By definition of forward mortality intensities

$$h_x(t,s) = -\frac{\partial S_x(t,s)}{\partial s}$$

Consequently, forward mortality intensities allow to compute survival probabilities as follows

$${}_sp_x(t) = \exp\left(-\int_0^s h_x(t,t+u)du\right)$$

These are models for the two-dimensional (in age x and maturity T) forward mortality surface, $h_x(t,T)$. The Heath Jarrow Morton framework encompasses a lot of different models. It is therefore sort of more general. The generic form of the model is

$$dh_x(t,T) = \alpha(t,T)dt + \beta(t,T)^T dW_t$$

where $\alpha(t,T)$ is a scalar and $\beta(t,T)$ is a $(1 \times n)$ vector and W_t is a n dimensional vector formed of independent brownian motions under a martingale probability measure $\mathcal{Q}$. However, we cannot just choose any $\alpha(t,T)$ and $\beta(t,T)$ for the model. In order to be well defined owe need

1. $\alpha(t,T)$ and $\beta(t,T)$ to be $\mathcal{F}_t$ adapted, where $\mathcal{F}_t$ is the filtration associated with the multidimensional brownian motion W_t.
2. $\int_0^T |\alpha(t,T)|dt$ is almost surely finite
3. $\int_0^T ||\beta(t,T)||^2 dt$ is almost surely finite
4. $\int_0^T \int_0^u ||\beta(t,u)||^2 dtdu$ is finite
5. $\mathbb{E}\left[\int_0^u |\beta(t,u)^T dW_u|du\right] \leq \infty$

Out of the record, the HJM model is usually used for interest rate modeling but as there is almost a one to one correspondence between interest rates and mortality intensities and between zero coupon prices and survival probabilities, interest rate models are very suitable for mortality models. We already encountered this when analyzing the OU and CIR model as both models where in fact originally developed for interest rate modeling.

Let us take a minute to express what we mean by a martingale probability measure $\mathcal{Q}$. This will be explained in more details when we will treat the pricing of mortality derivatives. Briefly, a well known stochastic finance result states that if a certain spot market is arbitrage free and complete, then there exists a unique probability measure $\mathcal{Q}$, called the risk neutral measure, under which all derivative prices (discounted cash flows) are martingales. Basically, a stochastic process M_t, $\mathcal{G}_t$ adapted is a martingale under $\mathcal{Q}$ if $\mathbb{E}^{\mathcal{Q}}[M_t|\mathcal{G}_s] = M_s$. It is thus sort of a zero gain game. Under $\mathcal{Q}$, the stochastic differential equation of a martingale must have a zero drift. It is then possible to show that there must exist a relation between $\alpha(t,T)$ and $\beta(t,T)$.

$$\alpha(t,T) = \beta(t,T)^T \int_t^T \beta(t,s)ds$$

The model therefore benefits of only one degree of freedom, $\beta(t,T)$. Once it is determined, $\alpha(t,T)$ follows. For example, choosing $\beta(t,T) = \sigma \exp(-a(T-t))$ will transform the HJM model in a OU model.

3.3 From discrete to continuous mortality models

Certain continuous mortality models can be obtained by extension of discrete models to a continuous time frame. Let us consider a certain cohort of age x during calendar year t, t being the reference time. Hence, $\mu_x(0)$ is the force of mortality of a life aged x at time 0 (during calendar year t). Let denote by $\mu_x(t)$ his force of mortality at time t (and thus age x+t). Pure deterministic models consider this force to be only a function of age, i.e.,

$$\mu_x(t) = \mu_{x+t}(0) = \bar{\mu}_{x+t}$$

Improved mortality models allow the force of mortality to change between integer ages, capturing a longevity or mortality factor into the model. Considering the Lee-Carter model as described earlier in this paper, we have in the discrete scenario

$$\mu_x(s) = \exp\Big(\beta_x^1 + \beta_x^2 \kappa_s\Big)$$

A possible extension of this model in continuous time is

$$\mu_x(s) = \exp(\alpha_{x+s} + \beta_x(s))$$

with $\beta_x(0) = 0$ and $\bar{\mu}_{x+s} = \exp(\alpha_{x+s})$. $\beta_x(s)$ is actually a correction to apply at time s on the initial life table. For example, if we want to match the Lee Carter model, we will have to take $\beta_x(s) = \theta_{x+s}s$. Although the force of mortality might now be function of age and calendar year, the model still does not take some uncertainties into consideration. As for now, it is totally deterministic. Uncertainty was added to the discrete Lee Carter model by considering a unitary degree autoregressive process in order to model κ_t. What we could do here, is add an exponential martingale to the continuous model. As a result, we end up with

$$\mu_x(s) = \bar{\mu}_{x+s} \exp(\beta_x(s))\Gamma(s)$$

where $\Gamma(s) = \exp\left(\int_0^s \gamma_x(u)dW(u) - \frac{1}{2}\int_0^s \gamma_x^2(u)du\right)$ is an exponential martingale. Indeed, denoting by $\mathcal{F}_t$ the filtration at time t associated to the standard brownian motion W(t), we easily show that

- $\Gamma(s)$ is adapted to $\mathcal{F}_s$.
- $\mathbb{E}[\Gamma(s)]$ exists for all s.
- $\mathbb{E}[\Gamma(t)|\mathcal{F}_s] = \Gamma(s)$ for t $\leq$ s.

The exponential martingale forms a multiplicative noise on the deterministic force of mortality. It is mainly chosen over all other multiplicative noises because of the necessity for $\mu_x(s)$ to remain positive. The continuous Lee Carter force of mortality could thereby be of the form

$$\mu_x(s) = \bar{\mu}_{x+s} \exp\Big(\beta_x(s) + \int_0^s \gamma_x(u)dW(u) - \frac{1}{2}\int_0^s \gamma_x^2(u)du\Big)$$

Differentiation leads to the stochastic differential model equation

$$d\mu_x(s) = \mu_x(s) \underbrace{\left(\frac{\bar{\mu}'_{x+s}}{\bar{\mu}_{x+s}} + \beta'_x(s)\right)}_{\text{drift}} ds + \underbrace{\gamma_x(s)}_{\text{volatility}} \mu_x(s)dW(s)$$

However, this model is not affine (since the stochastic term is in $\mathcal{O}(\mu_x(s)^2)$) and therefore not easy to work with. A remedy is to scale the volatility part so that

$$d\mu_x(s) = \mu_x(s) \underbrace{\left(\frac{\bar{\mu}'_{x+s}}{\bar{\mu}_{x+s}} + \beta'_x(s)\right)}_{\text{drift}} ds + \underbrace{\gamma_x(s)}_{\text{volatility}} \sqrt{\mu_x(s)}dW(s)$$

and we then obtain a variation of the Cox-Ingressol-Ross model.

What has just been done for the Lee Carter model can be executed for almost all discrete models. Consider for example the CBD model

$$\mu_x(s) = \log\left(1 + \exp\left(\kappa_s^1 + \kappa_s^2(x + s - \bar{x})\right)\right)$$

with κ_s^1 and κ_s^2 two autoregressive processes. Replacing κ_t^1 and κ_t^2 by two functions of time $\alpha(s)$ and $\beta(s)$ then leads to

$$\mu_x(s) = \log\left(1 + \exp\left(\alpha(s) + \beta(s)(x + s - \bar{x})\right)\right)$$

Considering linear functions for $\alpha(s)$ and $\beta(s)$ on then obtain

$$\mu_x(s) = \log\left(1 + \exp\left(\alpha s + \beta s^2 + \beta s(x - \bar{x})\right)\right)$$

Via Taylor expansion, we have for small x that $\log(1 + x) \approx x$, so that the force of mortality can be computed as

$$\mu_x(s) = \exp\left(\alpha s + \beta s^2 + \beta s(x - \bar{x})\right)$$

Again, this is purely deterministic and we can add an exponential martingale in order to add uncertainty

$$\mu_x(s) = \exp\left(\alpha s + \beta s^2 + \beta s(x - \bar{x})\right)\Gamma(s)$$

with $\Gamma(s)$ as before so that we obtain

$$\mu_x(s) = \exp\left(\alpha s + \beta s^2 + \beta s(x - \bar{x}) + \int_0^s \gamma_x(u)dW(u) - \frac{1}{2}\int_0^s \gamma_x^2(u)du\right)$$

Finally, the affine stochastic differential equation modeling the force of mortality is given by

$$d\mu_x(s) = \underbrace{(\alpha + 2\beta s + \beta(x - \bar{x}))}_{\text{drift}}\mu_x(s)ds + \underbrace{\gamma_x(s)}_{\text{volatility}}\sqrt{\mu_x(s)}dW(s)$$

Suppose $\gamma_x(s) = \gamma$. It follows that $\Gamma(s) = \exp\left(\int_0^s \gamma_x(u)dW(u) - \frac{1}{2}\int_0^s \gamma_x^2(u)du\right) = \exp\left(\gamma\int_0^s dW(u) - \frac{1}{2}\gamma^2 s\right)$ so that

$$\Gamma(s) \sim \mathcal{LN}\left(-\frac{1}{2}\gamma^2 s, \gamma\sqrt{s}\right)$$

. Consequently

$$\mathbb{E}\Big[\Gamma(s)\Big] = 1$$
$$\mathbb{V}\Big[\Gamma(s)\Big] = \exp\Big(\gamma^2 s\Big) - 1$$

Note that there is no golden rule in choosing the functions $\alpha(s)$ and $\beta(s)$. The typical way is to fit a polynomial function because polynomials benefit of nice mathematical properties in term of differentiability and continuity. The same remark is applicable for the function $\gamma_x(s)$. This is then a whole new problem that won't be tackled any further in this paper.

Chapter 4

Pricing of the longevity risk

Longevity risk is receiving more and more attention from analysts and experts. Longevity risk designate the uncertainty attached to aggregate mortality rates. Strong deviations around the assumptions used to price long-term life products can affect the solvency and profitability of the institutions that sell them. Indeed, insurers make most of their profit thanks to the difference between actual results obtained and expected results computed based on safe assumptions. Usually, assumptions are made on two independent aspect of a life product

1. The mortality of the policyholders.
2. The return on the investments of premiums.

The big issue here is the time frame of life insurance contracts. Life insurance contracts often cover a period of 30 to 60 years. It is already difficult to make projections of the mortality rates and financial markets for next year. Imagine what it is to make projections for 60 years. There are a lot of uncertainties, this is the least to say. For example, a life table that may seem to be on the safe side at the beginning of the contract, may turn out not to be so after 20 years. Moreover, in 1970, with interest rates around 4-5%, nobody could have possibly forecast the interest rates being around 0.01% today as a result of the financial crisis of 2008. The technical interest guaranteed by the insurer to his policyholders at that time is impossible to reach nowadays resulting in an imbalanced portfolio with liabilities becoming very threatening.

Although misestimation on the return on investment is likely to be the major risk, longevity risk is also significant for an insurer as he needs to completely understand and manage to risks he is exposed to in order to price his contracts appropriately.

Longevity risk is a systematic risk. It is therefore not possible to use classical reinsurance via pooling as a hedging technique. People exposed to this risk therefore need to find another way to manage it.

The historical approach to cover mortality risk was to keep a certain percentage of the amount exposed to risk or reserves in case of deviations between actual and expected mortality.

Nowadays, this has been replaced by stress tests. A typical stress test is to adjust the deterministic mortality downwards by 20-30% and see what happens financially. This still remain quite basic though. Today, stochastic mortality models become more and more popular. The trend component of stochastic models will allow companies to quantify their potential exposure to longevity. However, stochastic mortality models are not precise enough and still bear some risks (model and parameter risks).

People then came with the idea of natural hedging. Imagine that an annuity issuer also holds a life insurance business line. The decrease in mortality will affect adversely the annuity contracts (because more annuities have to be paid) but will have opposite effect on the insurance block (with less payments in case of death). Hence losses in one business line can be recovered in another business line. This technique would be very effective if the policyholders that purchase annuity contracts are the same as the ones that purchase life insurance contracts. In other words, a well diversified portfolio with significant exposure to both mortality and longevity risk allow companies to maximize the impact of natural hedging. Unfortunately, this is usually not the case. Indeed, it would not really make sense for someone entering an annuity contract (and thus expecting a long remaining lifetime) to enter a term insurance contract (usually for people with a small remaining lifetime) as well. This phenomenon is called the adverse selection.

Securization through capital markets is maybe the solution. It corresponds to a transfer of longevity/mortality risk by the purchase of some specific investment products. Popular products are for example survivor bonds and survivor swaps. The idea is to develop financial products whose payoffs are related to the survival index of a reference population. That way, high survival probabilities leads to high payoffs making it possible to match longevity cash inflows and outflows.

Capital markets are so big that they offer the opportunity to spread longevity risk among a huge number of market participants. For longevity products, participants would include

- investors and banks looking for diversification of their portfolio with a zero beta product (financial product uncorrelated to the market). Indeed, secured insurance products are often uncorrelated to other economic assets.
- Hedgers that are exposed to longevity risk and wish to lay it off. For example, annuity providers lose money if mortality improves over what was expected. On

the other hand, life insurers stand to gain in that case and vice versa. This suggests that life insurers and annuity providers could hedge each other risks through the purchase or selling of longevity-linked securities.

- Speculators seeking to make money on the death of others might be interested as well.
- One of the most exposed institution to longevity risk are the governments. There are thus good reasons to believe that governments would be interested in a mortality-linked market. It is of their duty to ensure that the number of companies declaring bankruptcy remains low so that the economy as a whole benefits from a greater stability. Moreover, the governments are significant holders of the longevity risk in its own right, via their pension systems and via their obligation to provide health care for the elderly.

The major issue with longevity securization through capital markets is the pricing of the products. They are usually very expensive and thus not yet interesting. It is therefore essential to develop adequate pricing strategies. Annuity providers, insurance companies and pension funds hence need to price the longevity risk and take that value into consideration when pricing a product related to it. Different ways exist in order to price this risk

- Change of measure
- Standard deviation principle
- CAPM and CCAPM ratios
- Sharpe ratio
- Utility principle
- Wang transform
- Risk measure such as VaR (Value at Risk) or Tail VaR.

In a first step, we will now briefly introduce the different pricing strategies that are currently being used. We will quickly cover the change of measure, the CAPM, CCAPM and Sharpe ratios, the Wang transform and the utility equivalence technique. VaR and Tail VaR methods will not be presented in this paper. The utility method will be analyzed in further details and used in order to price the three main longevity derivatives. Namely survivor forwards (s-forwards), longevity swaps and longevity bonds.

4.1 Change of Measure

The change of measure was already quickly introduced earlier in this paper. We now go into more details. The idea behind it is to find a certain probability measure $\mathcal{Q}$, equivalent to the conventional measure $\mathcal{P}$, so as to add a security margin called the risk premium. In certain cases, and that is when the change of measure is the most effective and intuitive, it is possible to find one and only one equivalent probability measure $\mathcal{Q}$ under which all derivatives have an average return equal to the risk free rate r. That special measure is called the neutral risk measure and we then speak of risk neutral pricing. The use of the risk neutral measure is the classical methodology used to price different securities. However, risk neutral pricing is not convenient to price longevity-linked securities. In order to explain why, we need to first explain the developments behind the risk neutral valuation. This needs a close understanding of the notions arbitrage free, completeness and martingale probability measure. Let us consider a spot market defined by a set of k securities represented by a vector $S_t \in \mathbb{R}^k$. A self financed spot trading strategy is a vector $\phi_t \in \mathbb{R}^k$ such that ϕ_i corresponds to the number of units of the asset S_i. Self financing suppose that no external funds are added to the portfolio between each time considered (the only changes in ϕ are done internally). All self financing trading strategies are regrouped in a vector space denoted Φ. S and Φ suffices to define a finite spot market $\mathcal{M} = (S, \Phi)$. A market $\mathcal{M}$ is called arbitrage free if there are no arbitrage opportunities. An arbitrage is a trading strategy that does not ask for a personal initial investment and is such that, for all possible scenarios, it is not possible to lose money and such that there exist at least one scenario where one is certain to earn money. Mathematically, if X(ω,t) denotes the value of a portfolio under scenario $\omega \in \Omega$ (set of all possible scenarios) at time t, then X(.,T) is called arbitrage at time T if

1. X(ω,0) $= 0 \; \forall \omega$
2. X(ω,T) $\geq 0 \; \forall \omega$
3. $\exists \omega^*$ such that X(ω^*,T) > 0

Moreover, given a contingent claim X that pays X at maturity T, the spot trading strategy $\phi_t^{'} \in \Phi$ is said replicating if $\phi_T^{'} S_T = X$. X is said attainable if there exists a self trading strategy that replicates it. We then say that a market is complete if there exist at least one replicating strategy in $\mathcal{M}$ for ANY contingent claim X $\in \mathcal{X}$. It can be proven that if the market is arbitrage free, then all attainable contingent claims X are uniquely replicated in $\mathcal{M}$. The process $\phi_t S_t$ associated to this claim X is called the arbitrage price. We now understand why completeness is a highly desirable feature. Indeed, it enables to price all securities via arbitrage. How to prove these characteristics? We therefore introduce the notion of martingale probability measure. A probability measure $\mathbb{P}^*$ equivalent to $\mathbb{P}$ (the two probability measures share the same set of events but with different probabilities) is called a martingale measure for $\mathcal{M}$ if for every trading strategy $\phi \in \Phi$, the relative process $V^*(\phi) = \phi_t S_t e^{-rt}$ follows a $\mathcal{P}^*$ martingale with respect to the filtration $\mathcal{F}_t$

$$\mathbb{E}_{\mathbb{P}^*}\Bigg[V^*(\phi_t)|\mathcal{F}_s\Bigg] = V^*(\phi_s)$$

We will use the notation $\mathcal{P}(\mathcal{M})$ to denote the set of martingale probability measures of the spot market. The end is near as it can then be proven that

1. A spot market is arbitrage free if and only if the class $\mathcal{P}(\mathcal{M})$ is non empty and the price process for any contingent claim X that settles at time T is given by $\pi_t(X) = e^{rt}\mathbb{E}\Bigg[Xe^{-rT}|\mathcal{F}_t\Bigg] \qquad t \leq T$
2. An arbitrage free spot market $\mathcal{M}$ is complete if and only if there exists a unique martingale measure $\mathcal{P}^*$ for $\mathcal{M}$. This measure is called the risk neutral measure.

What is now not of application when we desire to price longevity-linked securities? First, the longevity market lacks of liquidity which in turn induces incompleteness. This is usually the case when non-hedgeable claims exists. Consequently, the replication technique discussed above is not applicable. Incompleteness also means that there is no universal pricing probability measure. Instead, we will work with one of the possible martingale probability measures available. In the literature, the choice is often made to choose a probability measure equivalent to the historical probability measure. We will now present the mathematical developments behind a change of measure and illustrate how this technique can be used to price longevity bonds and longevity swaps.

Generalized Girsanov's theorem Let us fix a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ on which the standard Brownian motion is defined as $W : \mathbb{R}^+ \times \Omega \to \mathbb{R}^p$, and the filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{R}^+}$ generated by W. Let T $\in \mathbb{R}^+$ and consider a process $h : \mathbb{R}^+ \times \Omega \to \mathbb{R}^p$. Suppose that h is adapted to the filtration $\mathbb{F}$ and that it respects Novikov's condition:

$$\mathbb{E}\Bigg[\exp\left(\frac{1}{2}\int_0^T ||h_t||^2 \mathrm{dt}\right)\Bigg] < +\infty,$$

then the stochastic process W* : $[0, T]\times\Omega \to \mathbb{R}$ defined by

$$W_t^* = W_t + \int_0^t h_s \mathrm{ds}$$

for all t $\in$ [0,T] is a standard brownian motion with respect to the probability measure $\mathcal{Q} : \mathcal{F} \to [0, 1]$ which is defined by the Radon-Nikodym's density with respect to $\mathcal{P}$,

$$\frac{d\mathcal{Q}}{d\mathcal{P}} = \exp\left(\int_0^T h_s^T \mathrm{d}W_s - \frac{1}{2}\int_0^T ||h_s||^2 \mathrm{ds}\right)$$

To summarize, Girsanov's theorem says that there exist a probability measure $\mathcal{Q}$ associated to $\mathcal{P}$ such that

$$\mathcal{N}(\mu, \sigma^2) \text{ under } \mathcal{P} \to \mathcal{N}(0, \sigma^2) \text{ under } \mathcal{Q}$$

The association is done via a specific density of Radon-Nikodym defined by:

$$\rho(w(t)) = \exp\left(\frac{-\mu w(t) - \frac{1}{2}\mu^2}{\sigma^2}\right)$$

In the case of longevity-linked securities market, let us introduce a new probability measure $\mathcal{Q}$, such that under this measure, the stochastic process defined by:

$$W^*(t) = W(t) + \int_0^t h(s, \mu_x(s))ds$$

$$\Leftrightarrow dW(t) = dW^*(t) - h(t, \mu_x(t))dt$$

is a standard Brownian motion. The process $h(t, \mu_x(t))$ is called the "market price of the longevity risk". This term is added to the brownian motion and allows a safer pricing of the security. It can be seen as an additional margin to the ideal price. This function is chosen a priori. Different choices lead to different probability measures. The survival probability including a market price of risk is then given by

$$S_x^*(t, s, T) = \mathbb{E}^{\mathcal{Q}}\left[\exp\left(-\int_s^T \mu_x(u)du\right)\Big|\mathcal{F}_s\right]$$

where $\mathbb{E}^{\mathcal{Q}}$ denotes the expectation under the probability measure $\mathcal{Q}$. Affine short rate models can be rewritten as

$$\begin{aligned} d\mu_x(t) &= \alpha(x, t, \mu)dt + \sigma(x, t, \mu)(dW^*(t) - h(t, \mu)dt) \\ &= (\alpha(x, t, \mu) - \sigma(x, t, \mu)h(t, \mu))\, dt + \sigma(x, t, \mu)dW^*(t) \\ &= \alpha^*(x, t, \mu)dt + \sigma(x, t, \mu)dW^*(t) \end{aligned}$$

where

$$\alpha^*(x, t, \mu) = \alpha_1^*(x, t) + \alpha_2^*(x, t)\mu_x(t)$$

$$\sigma^2(x, t, \mu) = \sigma_1(x, t) + \sigma_2(x, t)\mu_x(t)$$

The idea is to choose a market price of risk in order to stay in affine models under $\mathcal{Q}$. That way, the good features concerning the solution of affine models that were discussed in the first part of this paper are not lost and the survival probabilities including market risk are given by:

$$S_x^*(t,T) = \exp\left(A^*(0,T) - B^*(0,T)\mu_x(0)\right) \tag{4.1}$$

where

$$\begin{cases} \frac{\partial A^*}{\partial t} - \alpha_1^*(x,t)B^* + \frac{1}{2}\sigma_1(x,t)B^{*2} &= 0 \\ \frac{\partial B^*}{\partial t} + \alpha_2^*(x,t)B^* - \frac{1}{2}\sigma_2(x,t)B^{*2} + 1 &= 0 \end{cases}$$

with limit conditions: $A^*(T,T) = B^*(T,T) = 0$. An example of market price of risk for models where $\sigma_1(x,t) = 0$ is

$$h(t,\mu) = \lambda_x(t)\sqrt{\mu} + \frac{\lambda_x^*(t)}{\sqrt{\mu}}$$

which gives

$$\begin{cases} \alpha_1^*(x,t) &= \alpha_1(x,t) - \lambda_x^*(t)\sqrt{\sigma_2(x,t)} \\ \alpha_2^*(x,t) &= \alpha_2(x,t) - \lambda_x(t)\sqrt{\sigma_2(x,t)} \end{cases}$$

For example, applying a change of measure on a model of type (3.6) with the particular form of market price of risk described above leads to survival probabilities given by (4.1) with $A^*(0,T)$ and $B^*(0,T)$ solutions of:

$$\begin{cases} \frac{\partial A^*}{\partial t} - \left(\lambda_x^*(t)\gamma_x(t) - k\tilde{\mu}_{x+t} - \gamma_x(t)\lambda_x(t)\mu_{x+t}^*\right)B^* - \frac{1}{2}\gamma_x^2(t)\mu_{x+t}^*B^{*2} &= 0 \\ \frac{\partial B^*}{\partial t} + \left(\frac{\bar{\mu}'_{x+t}}{\bar{\mu}_{x+t}} + \beta'_x(t) - k - \gamma_x(t)\lambda_x(t)\right)B^* - \frac{1}{2}\gamma_x^2(t)B^{*2} + 1 &= 0 \end{cases}$$

Let us present a more explicit example. Suppose that the affine model for the force of mortality under the probability measure $\mathcal{P}$ is given by

$$d\mu_x(t) = \alpha(t,\mu_x(t))dt + \sigma(t,\mu_x(t))dW(t)$$

where

$$\alpha(t,\mu_x(t)) = \alpha_1 - \alpha_2\mu_x(t)$$
$$\sigma(t,\mu_x(t)) = \sqrt{\sigma_1 + \sigma_2\mu_x(t)}$$

Now, let pick the function $h(t,\mu_x(t)) = -\left(\frac{c_1}{\sqrt{\sigma_1+\sigma_2\mu_x(t)}} + c_2\sqrt{\sigma_1+\sigma_2\mu_x(t)}\right)$.
We know that under $\mathcal{Q}$, the stochastic process defined by the equation

$$dW_t^* = dW_t + h(t,\mu_x(t))dt = dW_t - \left(\frac{c_1}{\sqrt{\sigma_1+\sigma_2\mu_x(t)}} + c_2\sqrt{\sigma_1+\sigma_2\mu_x(t)}\right)dt$$

is a brownian motion. Hence, the dynamics of $\mu_x(t)$ under $\mathcal{Q}$ is simply

$$\begin{aligned}
d\mu_x(t) &= \alpha(t, \mu_x(t)dt + \sigma(t, \mu_x(t))dW(t) \\
&= \alpha(t, \mu_x(t))dt + \sigma(t, \mu_x(t)) \left(dW^*(t) + \left(\frac{c_1}{\sqrt{\sigma_1 + \sigma_2 \mu_x(t)}} + c_2 \sqrt{\sigma_1 + \sigma_2 \mu_x(t)} \right) dt \right) \\
&= (\alpha_1 - \alpha_2 \mu_x(t))\, dt + \sqrt{\sigma_1 + \sigma_2 \mu_x(t)} \left(\frac{c_1}{\sqrt{\sigma_1 + \sigma_2 \mu_x(t)}} + c_2 \sqrt{\sigma_1 + \sigma_2 \mu_x(t)} \right) dt \\
&\quad + \sqrt{\sigma_1 + \sigma_2 \mu_x(t)} dW_t^* \\
&= (\alpha_1 + c_1 + c_2\sigma_1 + (c_2\sigma_2 - \alpha_2)\mu_x(t))\, dt + \sqrt{\sigma_1 + \sigma_2 \mu_x(t)} dW_t^* \\
&= -(c_2\sigma_2 - \alpha_2) \left(-\frac{\alpha_1 + c_1 + c_2\sigma_1}{c_2\sigma_2 - \alpha_2} - \mu_x(t) \right) dt + \sqrt{\sigma_1 + \sigma_2 \mu_x(t)} dW_t^* \\
&= \alpha^*(t, \mu_x(t))dt + \sigma^*(t, \mu_x(t))dW_t^*
\end{aligned}$$

where we recognize a CIR process for the force of mortality.

Intuitively, the change of measure allows to add a risk margin in the pricing of a certain product. This is done by changing the "world" in which we compute probabilities. Probabilities are not computed under $\mathcal{P}$ anymore but under $\mathcal{Q}$. Once this is done, a safe price of any security can be obtained via the actuarial equivalence between the expected present value of engagements of the seller of the product with the expected present value of the engagements of the buyer. Let illustrate that when pricing a longevity swap and a longevity bond.

4.1.1 Longevity bond pricing

Longevity bonds are by far the most developed security in order to hedge mortality thanks to capital markets. These are bonds where the payments are mortality dependent. More precisely, there exist three types of mortality bonds

1. principal-at-risk longevity bonds
2. coupon-based longevity bonds
3. hybrid longevity bonds

The first one is a bond where the investor loses part of the principal if the relevant mortality event arises. The second one is characterized by a constant principal but mortality dependent coupon payments. Finally, hybrid bonds are bonds where both the principal and the coupons are mortality dependent amounts.

Payment amounts are linked to the number of survivors in a certain reference population or cohort. The larger the number of survivors, the larger the cash flows. Classical longevity bonds are bonds with a stochastic maturity because the final payments finish, not after a certain maturity, but at the death of the last survivor of the reference population, i.e., the maturity of this bond can be modeled using a random variable T representing the death time of the last survivor. This is true in theory, in practice, most longevity bonds are issued with a predetermined deterministic maturity T (just like classical bonds).

The issuer of the bond could be any institution facing mortality risk while the purchaser will use this bond to hedge his longevity risk. Indeed, companies owning this product could balance their loss in life contracts with higher gains in bond payments. It has been suggested that governments could possibly provide longevity bonds as they have the duty of ensuring that the financial markets works correctly. Governments would be more suitable providers of such bonds because they do not require the risk premiums as private institutions do. As a result, the bond will be cheaper to buy. Although the financial interest, the issue of having governments as providers is that taxpayers might see their contribution increase in case of unfavorable experience.

There exists another security considered as a longevity bond, the geared longevity bond. It allows investors to hedge their risk for a much smaller initial investment. The pay-off at a certain time is a combination of a long forward contract, a long put option and a short call option.

$$p_t = S(t) - S_l(t) + \max\{S_l(t) - S(t), 0\} + \max\{S(t) - S_u(t), 0\} \qquad S(t) \in [S_l(t), S_u(t)]$$

S(t) being the survival function. The idea behind such bond is a little bit the same idea that pushed financials to create caps and floors. By restricting the payoffs with bounds, the financial product still proposes the desired effect but at a cheaper price. We will not discuss that kind of bond any further in this paper.

BNP Paribas and EIB were the first to announce the longevity bond in 2004. The coupons were indexed to a cohort survivor index in England and Wales. This cohort retires in 2004 at the age of 65. The maturity of this bond was 25 years. However, the bond was withdrawn in 2005 for redesign.

- the designed 25-year horizon was perhaps too short for an effective hedge since the longevity risk in the near future is short
- The facial value was large ($540m)

- The coupons were indexed based on the survival of a 65-year old cohort of males while hedging of longevity risk at younger ages and for females is also important
- High uncertainty about an adequate price
- Hedge failure or basis risk were large because the reference population was different of the policyholder population of an insurer
- Payments were nominal (they were independent of the inflation rate)

Swiss Re Vita 1 is a second example of longevity bond. On one hand, the issuer (Swiss Re) is a reinsurer tempting to hedge his mortality risks. He is searching protection against an important rise in mortality rates. On the other hand, the bondholder will have a good return on investment if the mortality improves. The issuer will face a loss (resp. gain) if the mortality rate, q_t, substantially exceeds (resp. undershoots) a predefined index, q_0. The q_t index is defined as a weighted basket of mortality rates from 5 different countries (USA 70%, UK 15%, France 7.5%, Switzerland 5% and Italy 2.5%). The gender distribution was 35% women and 65% men covering a broad range of ages. The instrument was designed with a maturity of three years paying quarterly variable coupons set at LIBOR + 135 basis points. The issue size was \$400m. The amount of principal lost followed the loss function

$$L(q_t) = \begin{cases} 0\% & \text{if } q_t < 1.3q_0 \\ \frac{q_t - 1.3q_0}{0.2q_0} & \text{if } 1.3q_0 \leq q_t < 1.5q_0 \\ 100\% & \text{if } q_t \geq 1.5q_0 \end{cases}$$

where

$$q_t = \sum_j C_j \sum_i \left(G^m A_i q^m_{i,j,t} + G^f A_i q^f_{i,j,t} \right)$$

and

$$\begin{aligned} q^m_{i,j,t} &= \text{mortality rate for males in the age group i for country j.} \\ q^f_{i,j,t} &= \text{mortality rate for females in the age group i for country j.} \\ C_j &= \text{weight attached to country j.} \\ A_i &= \text{weight attributed to age group i.} \\ G^m \text{ and } G^f &= \text{gender weights applied to males and females respectively.} \end{aligned}$$

The payment at maturity T then directly follows:

$$P_T = \begin{cases} 100\% - \sum_{i=1}^{T} L(q_i) & \text{if} \sum_{i=1}^{T} L(q_i) < 100\% \\ 0\% & \text{otherwise} \end{cases}$$

Suppose now that we observe a reference population of N_x individuals aged x during calendar year t. As usual, one takes this calendar year as the reference time $t = t_0 = 0$. Furthermore, suppose that the risk free interest rate r is constant (one could work with a stochastic risk free rate) leading to a constant discount rate $\delta = \ln(1+r)$. The price of the bond is the present value of all future expected cash flows, i.e. the coupons and the principal. Let C denote the initial price of the bond and suppose the bond is sold at par. c denotes the coupon rate and T denotes the maturity of the bond. We also assume that the capital C decreases with time proportionally to the number of survivors (hybrid longevity bond). Discounting at present gives

$$C = C\left(e^{-\delta T}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+T}}{N_x}\right] + c\sum_{s=1}^{T} e^{-\delta s}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+s}}{N_x}\right]\right)$$

from which c is determined.
Since N_x is known and $N_{x+i} \sim \mathcal{B}\left(N_x, {}_ip_x\right)$

$$\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+i}}{N_x}\right] = {}_ip_x^* = S_x^*(t,i) = \exp\left(A^*(0,i) + B^*(0,i)\mu_x(0)\right) \qquad \text{for i = 1,...,T}$$

At time t < T, the price of the bond is given by

$$\begin{aligned}
P(t,T) &= C\left(e^{-\delta(T-t)}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+T}}{N_x}\middle|\mathcal{F}_t\right] + c\sum_{s=t}^{T} e^{-\delta(s-t)}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+s}}{N_x}\middle|\mathcal{F}_t\right]\right) \\
&= C\left(e^{-\delta(T-t)}\frac{N_{x+t}}{N_x}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+T}}{N_{x+t}}\middle|\mathcal{F}_t\right] + c\sum_{s=0}^{T-t} e^{-\delta s}\frac{N_{x+t}}{N_x}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+t+s}}{N_{x+t}}\middle|\mathcal{F}_t\right]\right) \\
&= \frac{N_{x+t}}{N_x}C\left(e^{-\delta(T-t)}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+T}}{N_{x+t}}\middle|\mathcal{F}_t\right] + c\sum_{s=0}^{T-t} e^{-\delta s}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+t+s}}{N_{x+t}}\middle|\mathcal{F}_t\right]\right)
\end{aligned}$$

where

$$\mathbb{E}_{\mathcal{Q}}\left[\tfrac{N_{x+i}}{N_{x+t}}\middle|\mathcal{F}_t\right] = {}_{i-t}p_{x+t}^* = S_x^*(t_0,t,i) = \exp\left(A^*(t,i) + B^*(t,i)\mu_x(t)\right) \qquad \text{for i = t,...,T}$$

This is the functioning of a longevity bond in theory. If we are now interested in hedging the longevity risk, issuing bonds with mortality-based coupons is not enough. Indeed, longevity risk hedging goes through the comparison between mortality observed in a periodic mortality table and in a prospective mortality table. Let c and C denote the coupon rate and the facial value of the bond respectively. On one hand, suppose that the periodic survival probabilities of the cohort are given by ${}_tp_x^{per}$. On the other hand, we will denote the prospective survival probabilities obtained via one of the previous continuous models as ${}_tp_x = S_x^*(t_0,t) = \exp\left(A^*(0,t) + B^*(0,t)\mu_x(0)\right)$. In order to hedge the longevity risk efficiently, we consider a coupon of the form

$$C_t = cC\frac{{}_tp_x}{{}_tp_x^{per}}$$

This way, the holder of this bond will receive a coupon higher than cC if the true mortality increases and lower than cC if the true mortality decreases compared to the periodic mortality. Considering a discount rate of δ and a maturity of T, the price of the bond at time $t = t_0$ is obtained via discounting

$$\begin{aligned} P(T) &= C\left(c\sum_{s=1}^{T} e^{-\delta s}\frac{\mathbb{E}_{\mathcal{Q}}[{}_sp_x]}{{}_sp_x^{per}} + e^{-\delta T}\right) \\ &= C\left(c\sum_{s=1}^{T}\frac{e^{-\delta s}}{{}_sp_x^{per}}\mathbb{E}_{\mathcal{Q}}\left[\exp\left(-\int_0^s \mu_x(s)ds\right)\middle|\mathcal{F}_0\right] + e^{-\delta T}\right) \\ &= C\left(c\sum_{s=1}^{T}\frac{e^{-\delta s}}{{}_sp_x^{per}}\exp\left(A^*(0,s) + B^*(0,s)\mu_x(0)\right) + e^{-\delta T}\right) \end{aligned}$$

Now, at time $t < T$, the price of this bond is

$$\begin{aligned} P(t,T) &= C\left(c\sum_{s=t}^{T} e^{-\delta(s-t)}\mathbb{E}_{\mathcal{Q}}\left[\frac{{}_sp_x}{{}_sp_x^{per}}\middle|\mathcal{F}_t\right] + e^{-\delta(T-t)}\right) \\ &= C\left(c\sum_{s=0}^{T-t} e^{-\delta s}\mathbb{E}_{\mathcal{Q}}\left[\frac{{}_tp_x\,{}_sp_{x+t}}{{}_tp_x^{per}\,{}_sp_{x+t}^{per}}\middle|\mathcal{F}_t\right] + e^{-\delta(T-t)}\right) \\ &= C\left(c\frac{{}_tp_x}{{}_tp_x^{per}}\sum_{s=0}^{T-t} e^{-\delta s}\mathbb{E}_{\mathcal{Q}}\left[\frac{{}_sp_{x+t}}{{}_sp_{x+t}^{per}}\middle|\mathcal{F}_t\right] + e^{-\delta(T-t)}\right) \\ &= C\left(c\frac{{}_tp_x}{{}_tp_x^{per}}\sum_{s=0}^{T-t}\frac{e^{-\delta s}}{{}_sp_{x+t}^{per}}\exp\left(A^*(t,s) + B^*(t,s)\mu_x(t)\right) + e^{-\delta(T-t)}\right) \end{aligned}$$

4.1.2 Longevity swap pricing

Counterparties swap fixed series of payments (the fixed leg) for series of payments linked to the number of survivors in a cohort (the floating leg). Mortality swaps bear considerable similarity to reinsurance contracts. Moreover, they can be arranged at a lower transaction cost than a bond issue and are more easily cancelled. This is because the nominal C is usually not exchanged, only a certain proportion (interest) of it. They are more flexible and therefore adaptable to almost all circumstances. They do not need a lot of liquidity in the market because they mostly repose on the willingness of counterparties to exploit their comparative advantages or trade views on the development of mortality over time. As for the longevity bond, suppose that we observe a reference population of N_x individuals aged x during calendar year t. Furthermore, suppose that the risk free interest rate is constant (one could work with a stochastic risk free rate) leading to a constant discount rate $\delta = \ln(1+r)$. The price of the swap is the present value of all future expected cash flows, i.e. the coupons and the principal. Let C denote the principal of the swap. c denotes the leg rate and T denotes the maturity of the swap. Let the floating leg rate be $\left(1 - \frac{N_{x+t}}{N_x}\right)$. The nominal is assumed to be pro rated with respect to the death probabilities. Since there is no upfront payment, discounting at present gives

$$0 = C\left(c\sum_{s=1}^{T} e^{-\delta s} - \sum_{s=1}^{T} e^{-\delta s}\mathbb{E}_{\mathcal{Q}}\left[1 - \frac{N_{x+s}}{N_x}\right]\right)$$

from which one can deduce c.
Since N_x is known and $N_{x+i} \sim \mathcal{B}\left(N_x, {}_ip_x\right)$

$$\mathbb{E}_{\mathcal{Q}}\left[1 - \frac{N_{x+i}}{N_x}\right] = 1 - {}_ip_x^* = 1 - S_x^*(t,i) = 1 - \exp\left(A^*(0,i) + B^*(0,i)\mu_x(0)\right) \qquad \text{for i = 1,...,T}$$

At time t < T, the price of the swap is given by

$$\begin{aligned}
P(t,T) &= C\left(c\sum_{s=t}^{T} e^{-\delta(s-t)} - \sum_{s=t}^{T} e^{-\delta(s-t)}\mathbb{E}_{\mathcal{Q}}\left[1 - \frac{N_{x+s}}{N_x}\middle|\mathcal{F}_t\right]\right) \\
&= C\left(c\sum_{s=0}^{T-t} e^{-\delta s} - \sum_{s=0}^{T-t} e^{-\delta s}\mathbb{E}_{\mathcal{Q}}\left[1 - \frac{N_{x+t+s}}{N_x}\middle|\mathcal{F}_t\right]\right) \\
&= C\left(c\sum_{s=0}^{T-t} e^{-\delta s} - \sum_{s=0}^{T-t} e^{-\delta s}\left(1 - \mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+t+s}}{N_x}\middle|\mathcal{F}_t\right]\right)\right) \\
&= C\left(c\sum_{s=0}^{T-t} e^{-\delta s} - \sum_{s=0}^{T-t} e^{-\delta s}\left(1 - \frac{N_{x+t}}{N_x}\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+t+s}}{N_{x+t}}\middle|\mathcal{F}_t\right]\right)\right)
\end{aligned}$$

where

$$\mathbb{E}_{\mathcal{Q}}\left[\frac{N_{x+i}}{N_{x+t}}\middle|\mathcal{F}_t\right] = {}_{i-t}p_{x+t}^* = S_x^*(t_0,t,i) = \exp\left(A^*(t,i) + B^*(t,i)\mu_x(t)\right) \qquad \text{for i = t,...,T}$$

Mortality swaps are attractive because they are modulable to every requirement and therefore enable parties to manage their longevity in risk in a not too costly way. It can even be used in case of natural hedging. Moreover, bearing in mind that swap payments would be conditioned on particular time periods and reference populations, firms could use such swaps to manage their exposures across both reference population and across the mortality term structure.

4.2 Standard Deviation Principle

It is also possible to price a product without changing of probability measure, i.e., price a product under $\mathcal{P}$. However, simply using the actuarial equivalence under $\mathcal{P}$ is not enough as there is no security margin! Consequently, a pricing method called the standard deviation principle has been developed. It adds a certain risk loading to the survival probabilities computed under $\mathcal{P}$. The variance is used to express the risk as it is the first indicator of the volatility. Survival probabilities ${}_{T-t}p_{x+t}(t_0+t)$ are then computed as follows

$$S_x(t_0,t,T) = \mathbb{E}_{\mathcal{P}}\left[\exp\left(-\int_t^T \mu_x(s)ds\right)\middle|\mathcal{F}_t\right] + \rho\left(Var\left[\exp\left(-\int_t^T \mu_x(s)ds\right)\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}} \tag{4.2}$$

where ρ is the risk loading factor and $\mu_x(s)$ is the force of mortality at time s of a life aged x at time t_0. By the definition of the variance

$$Var\left[\exp\left(-\int_t^T \mu_x(s)ds\right)\middle|\mathcal{F}_t\right] = \mathbb{E}_{\mathcal{P}}\left[\exp\left(-\int_t^T 2\mu_x(s)ds\right)\middle|\mathcal{F}_t\right] - \left(\mathbb{E}_{\mathcal{P}}\left[\exp\left(-\int_t^T \mu_x(s)ds\right)\middle|\mathcal{F}_t\right]\right)^2$$

The first term on the right hand side of (4.2) is known to be such that:

$$\mathbb{E}_{\mathcal{P}}\left[\exp\left(-\int_t^T \mu_x(s)ds\right)\middle|\mathcal{F}_t\right] = \exp\left(A(t,T) - B(t,T)\mu_x(t)\right)$$

Letting $\widehat{\mu}_x(t) = 2\mu_x(t)$, the process $\widehat{\mu}_x(t)$ is solution of almost the same stochastic differential equation as $\mu_x(t)$:

$$d\widehat{\mu}_x(t) = \widehat{\alpha}(x,t,\mu)dt + \widehat{\sigma}(x,t,\mu)dW(t)$$

with

$$\begin{cases} \widehat{\alpha}(x,t,\mu) &=& \widehat{\alpha}_1(x,t) + \widehat{\alpha}_2(x,t)\mu_x(t) \\ \widehat{\sigma}^2(x,t,\mu) &=& \widehat{\sigma}_1(x,t) + \widehat{\sigma}_2(x,t)\mu_x(t) \end{cases}$$

Which has the same affine form as above. The following equation then holds

$$\mathbb{E}_{\mathcal{P}}\left[\exp\left(-\int_t^T 2\mu_x(s)ds\right)\middle|\mathcal{F}_t\right] = \exp\left(\widehat{A}(t,T) - 2\widehat{B}(t,T)\mu_x(t)\right)$$

with the functions $\widehat{A}(t,T)$ and $\widehat{B}(t,T)$ solutions of

$$\begin{cases} \frac{\partial\widehat{A}}{\partial t} - \widehat{\alpha}_1(x,t)\widehat{B} + \frac{1}{2}\widehat{\sigma}_1(x,t)\widehat{B}^2 &=& 0 \\ \frac{\partial\widehat{B}}{\partial t} + \widehat{\alpha}_2(x,t)\widehat{B} - \frac{1}{2}\widehat{\sigma}_2(x,t)\widehat{B}^2 + 1 &=& 0 \end{cases}$$

As a result, the survival probabilities with risk loading are finally given by

$$S_x(t_0,t,T) = \exp\left(A(t,T) - B(t,T)\mu_x(t)\right) \\ + \rho\left(\exp\left(\widehat{A}(t,T) - 2\widehat{B}(t,T)\mu_x(t)\right) + \exp\left(2\left(A(t,T) - B(t,T)\mu_x(t)\right)\right)\right)^{\frac{1}{2}} \tag{4.3}$$

Applying the standard deviation procedure on the general affine model (3.6), $\widehat{\mu}_x(t)$ is solution of following stochastic differential equation

$$d\widehat{\mu}_x(t) = \left[\widehat{\mu}_x(t)\left(\frac{\bar{\mu}'_{x+t}}{\bar{\mu}_{x+t}} + \beta'_x(t) - k\right) + 2k\tilde{\mu}_{x+t}\right]dt + \sqrt{2}\gamma_x(t)\sqrt{\widehat{\mu}_x(t) - 2\mu^*_{x+t}}dW(t)$$

which has the same affine form as previously so the survival probabilities in this case are given by (4.3) with the functions $\widehat{A}(t,T)$ and $\widehat{B}(t,T)$ solutions of:

$$\begin{cases} \frac{\partial \widehat{A}}{\partial t} - 2k\tilde{\mu}_{x+t}\widehat{B} - 2\gamma_x(t)^2\mu^*_{x+t}\widehat{B}^2 &= 0 \\ \frac{\partial \widehat{B}}{\partial t} + \left(\frac{\bar{\mu}'_{x+t}}{\bar{\mu}_{x+t}} + \beta'_x(t) - k\right)\widehat{B} - \gamma_x^2(t)\widehat{B}^2 + 1 &= 0 \end{cases}$$

Once these loaded survival probabilities have been computed, it is possible to price longevity-linked securities as best estimates using the actuarial equivalence. Using the same mathematics as developed for the pricing under $\mathcal{Q}$, longevity bonds and longevity swaps of maturity T are priced under $\mathcal{P}$ as

$$P^{\text{bond}}(t,T) = \frac{N_{x+t}}{N_x}C\left(e^{-\delta(T-t)}\mathbb{E}_{\mathcal{P}}\left[\frac{N_{x+T}}{N_{x+t}}\middle|\mathcal{F}_t\right] + c\sum_{s=0}^{T-t}e^{-\delta s}\mathbb{E}_{\mathcal{P}}\left[\frac{N_{x+t+s}}{N_{x+t}}\middle|\mathcal{F}_t\right]\right)$$
$$P^{\text{swap}}(t,T) = C\left(c\sum_{s=0}^{T-t}e^{-\delta s} - \sum_{s=0}^{T-t}e^{-\delta s}\left(1 - \frac{N_{x+t}}{N_x}\mathbb{E}_{\mathcal{P}}\left[\frac{N_{x+t+s}}{N_{x+t}}\middle|\mathcal{F}_t\right]\right)\right)$$

where

$$\mathbb{E}_{\mathcal{P}}\left[\frac{N_{x+i}}{N_{x+t}}\middle|\mathcal{F}_t\right] = {}_{i-t}p_{x+t} = S_x(t_0,t,i) \\ = \exp\left(A(t,i) + B(t,i)\mu_x(t)\right) \\ + \rho\left(\exp\left(\widehat{A}(t,i) - 2\widehat{B}(t,i)\mu_x(t)\right) + \exp\left(2\left(A(t,i) - B(t,i)\mu_x(t)\right)\right)\right)^{\frac{1}{2}}$$

This alternative approach is not very popular. It raises the question on how to select ρ and whether the standard deviation enough of a risk measure is.

4.3 CAPM & CCAPM approach

Friedberg and Webb (2005) developed the capital asset pricing model to determine the risk premium to be applied on certain securities. That same model was used later by Wang as a basis for the development of a model that now bears his name, the so-called Wang transform model. That model extends the CAPM model to risks that are not normally distributed. The CAPM and CCAPM formulas are widespread in finance in order to determine the expected return of a financial asset. It is based on the premise that investors have assumptions of systematic risk (also known as market risk or non-diversifiable risk) and need to be compensated for it in the form of a risk premium – an amount of market return greater than the risk-free rate. By investing in a security, investors want a higher return for taking on additional risk. Let R_s and R_m respectively denote the return of the security and the return of the market. The risk free rate is denoted by R_f, it is typically equal to the yield of a 10-year German (or US) government bond. However, healthier results are obtained using government bonds from the country in which the investment is being made and by matching the maturity of the bond with the one of the security. The CAPM formula is then given by

$$\mathbb{E}[R_s] - R_f = \beta(R_s, R_m)[\mathbb{E}[R_m] - R_f]$$

where

$$\beta(R_s, R_m) = \frac{Cov(R_s, R_m)}{\sigma_m^2}$$

The β is a measure of the correlation between the market as a whole and the security of interest. As mentioned earlier in this paper, one of the interest of longevity-linked securities is that they present very low β and henceforth enables investors to diversify their portfolio reasonably effectively. For example, the β of a US longevity bond was determined to be around 0.15 with a 95% confidence interval [-0.15, 0.46]. Therefore, if the market risk premium is 5%, the longevity risk premium is 75 bpm with confidence interval of [-75, 230] bpm. Assuming that asset returns have a normal distribution, the market price of risk is given by

$$\lambda_s = \frac{\mathbb{E}[R_s] - R_f}{\sigma_s}$$

which is known as the Sharpe ratio (it will be discussed in a moment). Because of this quite wide confidence interval, the CCAPM (consumed-based capital asset pricing model) approach was considered instead. The longevity risk premium is now determined using a consumption beta instead of a market beta.

$$\mathbb{E}[R_s] - R_f = -\frac{Cov(U'(c), R_b)}{\mathbb{E}[U'(c)]}$$

The consumption beta is not easy to compute though. It is measured by the movements of risk premium with consumption growth. A higher consumption beta implies a higher expected return on risk assets. For instance, a consumption beta of 2.0 would imply an increase of asset return of 2% if the market increased by 1%. The CCAPM incorporates many forms of wealth beyond security market wealth and provides a framework for understanding variation in financial asset returns over many time periods. This provides an extension of the CAPM, which only takes into account one-period asset returns. The CCAPM also provides a fundamental understanding of the relation between wealth, consumption and investor's risk aversion.

4.4 Sharpe ratio approach

The Sharpe ratio has become the most widely used method for calculating risk-adjusted returns for "classical" assets. The Achille's heel is that it can be inaccurate when considering securities that do not expected returns with a normal distribution. The Sharpe ratio on financial markets is the ratio between the excess return of a portfolio and the volatility associated with that portfolio.

$$SR = \frac{\mathbb{E}[R] - R_f}{\sigma(R)}$$

Sharpe ratios are usually computed before and after adding a certain asset to a portfolio. An increasing Sharpe ratio then outlines the fact that the diversification of the portfolio is better when that asset is added (and vice versa). In the insurance industry, such a ratio could be generalized as the ratio between the excess payoffs above the expected payments and the volatility of these risky expected payments.

$$SR = \frac{\Pi - \mathbb{E}[W_N]}{\sigma(W_N)}$$

where $\Pi = \pi(1 + L)$ is the total premium, i.e., considering a longevity risk loading. The idea behind this method is to determine the risk loading L such that the Sharpe ratio of the insurance portfolio is consistent with other asset classes in the economy.

For example, suppose that the return of a diversified portfolio of S&P 500 stocks over a certain period was somewhere between 11% with a standard deviation of 20%.

Suppose moreover that over that same period, the risk free rate averaged at around 6%. The Sharpe ratio is thus given by

$$SR = \frac{11-6}{20} = 0.25$$

Now, imagine that over that same period, an insurer issues a policy with a payoff of the form

$$w = \begin{cases} 2 & p^* \\ 0 & 1-p^* \end{cases}$$

for a prepayment of π where p^* is the probability of survival. If the insurer takes care of the longevity risk, he knows that p^* is stochastic. It is a random variable. For simplicity, let suppose that it is discrete and only takes two values

$$p^* = \begin{cases} p-\epsilon & \text{with probability } 0.5 \\ p+\epsilon & \text{with probability } 0.5 \end{cases}$$

ϵ being a small constant positive number. The amount at risk (the amount to pay) if the insurer sells N policies is then

$$W_N = \sum_{i=1}^{N} w_i$$

Assuming independence between the policies, it is clear that

$$\mathbb{E}[W_N] = NE[w_1] = N\mathbb{E}[\mathbb{E}[w_1|p^*]] = 2N\mathbb{E}[p^*] = 2Np$$

$$\begin{aligned} Var[W_N] &= \mathbb{E}\Big[Var[W_N|p^*]\Big] + Var\Big[\mathbb{E}[[W_N|p^*]\Big] \\ &= 4Np(1-p) - 4N\epsilon^2 + 4N^2\epsilon^2 \\ &= 4Np(1-p) + 4N\epsilon^2(N-1) \end{aligned}$$

The Sharpe ratio is then given by

$$SR^{ins} = \frac{N\pi(1+L) - 2Np}{\sqrt{4Np(1-p) + 4N\epsilon^2(N-1)}}$$

The insurer finds his risk loading L by solving

$$\frac{N\pi(1+L) - 2Np}{\sqrt{4Np(1-p) + 4N\epsilon^2(N-1)}} = 0.25 \Leftrightarrow L = \frac{0.25\sqrt{4Np(1-p) + 4N\epsilon^2(N-1)} + 2Np}{N\pi} - 1$$

It is important to outline that the last three approaches (CAPM, CCAPM and Sharpe) are of course not crude pricing techniques. However, they enable to compute the expected return of a certain asset or derivative such as a bond or a swap from which the risk premium is easily determined. This risk premium can then be used as a loading factor for the pricing of the underlying asset or derivative.

4.5 Wang transform

In 2000, Wang introduced a new transform and correlation measure that extends CAPM and Sharpe to pricing all kinds of assets and liabilities, having any type of probability distribution, whether traded or underwritten, in finance or insurance. This transform is just as easily applied to contingent payoffs that are co-monotone with their underlying assets or liabilities. In its simplest form, the new transform relies on a parameter called the "market price of risk", extending a familiar concept in CAPM to risks with non-normal distributions. The "market price of risk" can either be applied to or implied from a distribution in order to arrive at a "risk-adjusted price" for the underlying risk in question. The "market price of risk" increases continuously with duration, and is consistent at each horizon date between an underlying and its co-monotone contingent payoff. In order to understand how the pricing method works, it is useful to remind some basic concepts about risk measures. A risk measure is a functional R mapping a risk X to a non-negative real number $R[X]$, possibly infinite. There exist many possible risk measures such as the VaR, Tail-VaR, CTE,... Spectral risk measures are risk measures distorted by a distortion function. A distortion function is a function $g : [0, 1] \to [0, 1]$ so that

- g is non-decreasing
- g(0) = 0 and g(1) = 1

For example, we have for $R[X] = VaR[X; 1 - p]$

$$R_g[X] = \int_0^1 VaR[X; 1-p]dg(p) = \int_0^{+\infty} g\left(Pr[X > x]\right) dx = \int_0^{+\infty} g\left(\bar{F}_x(x)\right) dx$$

The distortion function actually add weights (dg(p)) to the risk measure. The increasing feature of g is thus important so as to consider positive weights. Another feature of a distortion function is concavity. As a result, distortion functions are continuous and $t \to g\left(\bar{F}_x(t)\right)$ is an excess function for a random variable X_g, i.e.

$$P[X_g > t] = g\left(\bar{F}_x(t)\right) \qquad t \in \mathbb{R}$$

The measuring process then proceeds in two steps

1. the original risk X is transformed into a new risk X_g with excess function

$$P[X_g > t] = g\left(\bar{F}_x(t)\right) \qquad t \in \mathbb{R}$$

2. the expectation $\mathbb{E}[X_g]$ is computed and

$$R_g[X] = \mathbb{E}[X_g]$$

Moreover, concavity leads to coherence. A risk measure is coherent if it respects following properties:

- translativity: $R[X + c] = R[X] + c$
- positive homogeneity: $R[cX] = cR[X]$ for $c > 0$
- subadditivity: $R[X + Y] \leq R[X] + R[Y]$
- monotonicity: $P[X \leq Y] = 1 \Rightarrow R[X] \leq R[Y]$

This will be useful when valuating longevity-linked securities.
Suppose we want to price a financial asset whose value is X at time T. Wang proposed to use the distortion transformation $x \to g(x) = 1 - \Phi\left(\Phi^{-1}(1-x) + \lambda\right)$ $x \in \mathbb{R}$ where Φ is the distribution function corresponding to the standard Normal distribution and λ is a parameter called the market price of risk, reflecting the level of systematic risk. Given a distribution function F for X, its Wang transform F^*_λ is defined as

$$F^*_\lambda(x) = \Phi\left(\Phi^{-1}(F(x)) + \lambda\right)$$

F^*_λ can actually be seen as a risk adjusted cumulative distribution function. As mentioned above, the Wang risk measure is the expected value of X associated with the distribution function F^*_λ

$$\mathbb{E}^*_\lambda[X] = \int_0^{+\infty} \bar{F}^*_\lambda(t)dt$$

where $\bar{F}^*_\lambda = 1 - F^*_\lambda$. $\mathbb{E}^*_\lambda[X]$ is the risk-adjusted value of the financial asset at time T, which can then be discounted to the reference time using the risk free rate to obtain a fair price.

Pricing using the Wang transform follows a little bit the idea of pricing via a change in probability measure. In both cases, probabilities (and thus expectations) are modified so as to price the derivative accordingly.

Let illustrate this pricing method on longevity bonds. In this context, let use following notations

1. $p_x(t)$ is the probability for a life aged x to survive to age x+1 in calendar year t.
2. $q_x(t) = 1 - p_x(t)$ is the probability for a life aged x to die before reaching age x+1 in calendar year t.
3. $\mu_x(t)$ is the force of mortality during calendar year t of a life aged x.

Moreover, let assume a constant force of mortality at integer ages. In this case, one have the well-known result

$$p_x(t) = \exp\left(-\int_0^1 \mu_{x+u}(t+u)du\right) = \exp\left(-\int_0^1 \mu_x(t)du\right) = \exp\left(-\mu_x(t)\right)$$

In what follows, we will use the stochastic Lee-Carter model, as described in a previous section, in order to project mortality intensities. Let L_{x_0} be the number of annuity contracts issued to a cohort of individuals aged x_0 at time t_0. Each contract is supposed to pay an annuity of 1 monetary unit if the policyholder is alive at the end of the year after having paid a single premium at time t_0. As a result, the amount the annuity provider will have to pay after t years (at time $t_0 + t$) is computed as:

$$F_t = L_{x_0} {}_t p_{x_0}^{prosp}$$

where ${}_t p_{x_0}^{prosp}$ denotes the probability of survival of an individual to at least age $x_0 + t$, computed using the prospective life table obtained via Lee Carter model. It follows that:

$$\begin{aligned} {}_t p_{x_0}^{prosp} &= \prod_{i=0}^{t-1} p_{x_0+i}(t_0+i) \\ &= \prod_{i=0}^{t-1} \exp\left(-\mu_{x_0+i}(t_0+i)\right) \\ &= \exp\left(-\sum_{i=0}^{t-1} \mu_{x_0+i}(t_0+i)\right) \\ &= \exp\left(-\sum_{i=0}^{t-1} \alpha_{x_0+i} + \beta_{x_0+i}\kappa_{t_0+i}\right) \end{aligned} \tag{4.4}$$

These survival probabilities are stochastic because of the ARIMA nature of the process κ_t. If the annuity provider did not take care of the longevity risk when pricing his products, he computed the survival probabilities of the cohort based on a periodic life table as if he was not aware that the mortality at a certain age improves with the calendar year t (i.e. $\mu_x(t)$ is usually bigger than $\mu_x(t+i)$ i = 1, 2,...). The annuity provider therefore expected to pay at time $t_0 + t$ an amount of

$$F_t^* = L_{x_0}{}_t p_{x_0}^{per}$$

The payable surplus at time $t_0 + t$ is therefore of

$$S_t = F_t - F_t^* = L_{x_0}({}_t p_{x_0}^{prosp} - {}_t p_{x_0}^{per})$$

This surplus represents a gain if $F_t < F_t^*$ and a loss if $F_t > F_t^*$. The annuity provider, in order to cover his longevity risk, buys at time t_0, N_{x_0} longevity bonds of maturity n. At time t ($t_0 < t \leq n$), the net revenue of his securities is thus of

$$R_t = N_{x_0} C_t$$

where C_t represents the coupon at time t. The longevity risk is then perfectly hedged if and only if

$$R_t = S_t \Leftrightarrow N_{x_0} C_t = L_{x_0}({}_t p_{x_0}^{prosp} - {}_t p_{x_0}^{per}) \Leftrightarrow C_t = \frac{L_{x_0}}{N_{x_0}}({}_t p_{x_0}^{prosp} - {}_t p_{x_0}^{per})$$

If one then choose N_{x_0} so that

$$\frac{L_{x_0}}{N_{x_0}} = k$$

Adding a security margin, the optimal floating coupon is $C_t = k({}_t p_{x_0}^{prosp} - {}_t p_{x_0}^{per}) + k^* = kI_t + k^*$. The coupon is thus composed of two aspects:

1. a constant coupon k pro rated with a mortality index I_t
2. a security margin k^* in order to obtain the final coupon C_t

Now that we obtained the expression of the optimal floating coupon, we can price the longevity bond by decomposing it into several zero coupon bonds of different maturities. Let P(s,t) be the price of a zero-coupon bond issued at time s and paying 1 monetary unit at time t. Noting $P^*(t_0, t_0 + n)$ the price of the coupon based longevity bond of maturity n at time t_0, the indifference principle states that

$$P^*(t_0, t_0 + n) = \sum_{t=1}^{n} \mathbb{E}_\lambda^*[C_t] P(t_0, t_0 + t) + P(t_0, t_0 + n)$$

where $\mathbb{E}_\lambda^*[C_t]$ is the certainty equivalent of the random cash flow C_t computed with the Wang transform. In order to write this equation, we considered that the longevity risk is independent of the financial risk, which is questionable. Coherence of the risk measure allows to rewrite the indifference equation as

$$\begin{aligned} P^*(t_0, t_0+n) &= \sum_{t=1}^{n} \mathbb{E}^*_\lambda[C_t] P(t_0, t_0+t) + P(t_0, t_0+n) \\ &= \sum_{t=1}^{n} \mathbb{E}^*_\lambda[k\,({}_tp^{prosp}_{x_0} - {}_tp^{per}_{x_0}) + k^*] P(t_0, t_0+t) + P(t_0, t_0+n) \\ &= \sum_{t=1}^{n} (k\,(\mathbb{E}^*_\lambda[{}_tp^{prosp}_{x_0}] - {}_tp^{per}_{x_0}) + k^*) P(t_0, t_0+t) + P(t_0, t_0+n) \end{aligned}$$

We now just need to determine $\mathbb{E}^*_\lambda[{}_tp^{prosp}_x]$ in order to obtain the final price of the longevity bond. Therefore, we first need to find the distribution function F_t of ${}_tp^{prosp}_{x_0}$

$$F_t(p) = P[{}_tp^{prosp}_{x_0} \leq p] = P\left[\sum_{i=0}^{t-1} \exp\left(\alpha_{x_0+i} + \beta_{x_0+i}\kappa_{t_0+i}\right) \geq -\ln(p)\right] \qquad 0 \leq p \leq 1$$

Finally, by definition of the Wang transform

$$\mathbb{E}^*_\lambda[{}_tp^{prosp}_x] = \int_0^1 \left(1 - \Phi\left(\Phi^{-1}(F_t(p)) + \lambda\right)\right) dp$$

Let us go through an illustrative numerical example to show how we can use the Wang transform to price longevity-linked products. Suppose we desire to price a zero coupon bond of maturity $T = \{5, 10, 15, 20, 25, 30, 35\}$. Suppose that the continuous compound rate is $\delta = 1\%$ so that the risk free rate is given by $r = \exp(\delta) - 1 \approx 1\%$. Let denote by ${}_Tp^*_{x_0}(t_0)$ the probability for a life aged x_0 in t_0 to live to at least age $x_0 + T$ computed using the Wang transform and by ${}_Tp^{ref}_{x_0}$ the same probability but computed according to the reference life table as given by (4.4). Consider $x_0 = 50$ and $t_0 = 2015$. The life table is closed at age $\omega(2015) = 110$. We first need to determine the market price of risk, $\lambda_{50}(2015)$. To this end, we need to compute the market value of an annuity for this cohort, $a^{market}_{50}(2015)$. This is simply the sum of pure endowments until death. Mathematically

$$a^{market}_{50}(2015) = \sum_{n=1}^{\omega(2015)-50} e^{-\delta n}\, {}_np^{market}_{50}(2015)$$

with ${}_np^{market}_{50} = \prod_{k=0}^{n-1} p^{market}_{50+k}$ are survival probabilities computed according to the rules recommended by the Belgian regulatory authorities (CBFA), that is

$$p^{market}_{50+k} = p^{XR}_{50+k-5}$$

The individual is made 5 years younger to take into account the adverse selection. The adverse selection is the fact that people subscribing life (resp. death) products are respectively in better (resp. worse) shape than the global population and henceforth have bigger (resp. smaller) survival probabilities. XR stands for mixt life table as it is nowadays prohibited to differentiate on the gender of a policyholder. Assuming an equal number of men and women (this is a discutable assumption).

$$p_x^{XR} = \frac{1}{2}\left(p_x^{MR} + p_x^{FR}\right) = \frac{1}{2}\left(s_m g_m^{c_m^x(c-1)} + s_f g_f^{c_f^x(c-1)}\right)$$

$$\begin{aligned} s_m &= 0.999441703848 & s_f &= 0.999669730966 \\ g_m &= 0.999733441115 & g_f &= 0.999951440172 \\ c_m &= 1.101077536030 & c_f &= 1.116792453830 \end{aligned}$$

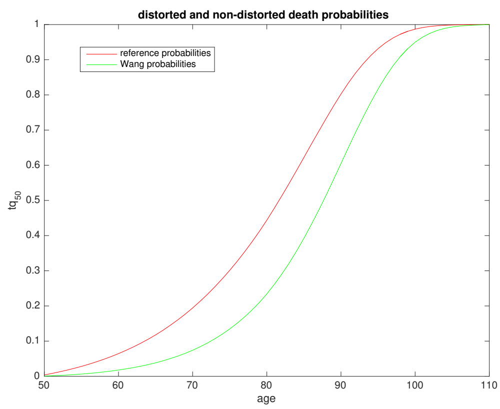
so that $a_{50}^{market}(2015) = 30.19773$. Thinking of the projected life table as the physical distribution, $\lambda_{50}(2015)$ is solution of

$$a_{50}^{market}(2015) = \sum_{t\geq 1} e^{-\delta t}\left(1 - \Phi\left(\Phi^{-1}({}_tq_{50}^{ref}) + \lambda_{50}(2015)\right)\right)$$

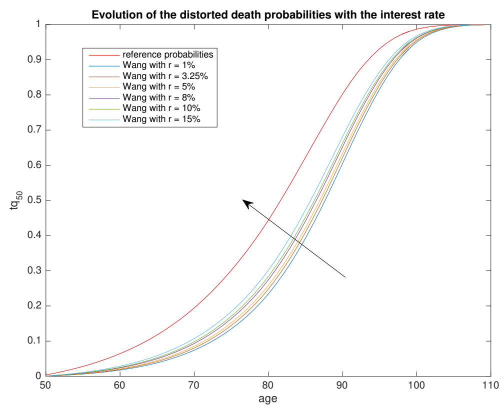
Solving this numerically gives $\lambda_{50}(2015) = -0.5839234$. Now that the market price is known, we can compute distorted probabilities as

$${}_tq_{50}^{*}(2015) = \Phi\left(\Phi^{-1}({}_tq_{50}^{ref}) + \lambda_{50}(2015)\right)$$

The impact of the distortion is clear on Figure 4.1a. The death probabilities under the Wang transform are smaller and the "S" curve is steeper. Death probabilities are therefore less concentrated around the mode that lies somewhere around 87 years.



(A) ${}_tq_{50}^{*}(2015)$ in green and ${}_tq_{50}^{ref}(2015)$ in red



(B) ${}_tq_{50}^{*}(r)$ for various r and ${}_tq_{50}^{ref}$ in red

FIGURE 4.1

The price of a zero coupon paying one monetary unit in $t_0 + T$ proportionally to the number of survivors of that cohort is now the simple expression

$$P(2015, 2015 + T) = e^{-\delta T} {}_{T}p^{*}_{50}(2015)$$

Results for different maturities are summarized in the table below.

T	5	10	15	20	25	30	35
P(2015,2015+T)	0.9466	0.8914	0.8321	0.7653	0.6853	0.5852	0.4539

Figure 4.1b is actually very interesting. It represents a few distorted death probabilities for various values of the risk free rate r. What we observe is that if the interest rate increases, the distorted probabilities tend to get closer to the reference death probabilities. Consequently, the larger the interest rate, the smaller the longevity risk. However, note that despite an interest rate of 15% (which is already very large), there still remain a considerable longevity risk. Interest rates thus have their influence, but it is not a critical one.

4.6 Equivalent utility principle

Another way of pricing longevity-linked securities is to work with the equivalent utility principle. Buyers (resp. sellers) are willing to buy (resp. sell) a product at a price in line with their utility. To outline the importance of utility, consider following example. Consider a billionaire and a homeless. Now imagine that a third party wants to give away 1000 dollars. Although 1000 dollars are 1000 dollars, it is clear that they will have a bigger value for the homeless than for the billionaire. This is the concept behind utility. In order to exist, the utility theory needs three axioms to be accepted

1. Comparability. Two situations A and B are always comparable. That is, one can make a choice between them.
2. Transitivity. If one prefers situation A to situation B which is preferred over situation C, then A is preferred to C.
3. Certain equivalent. For every situation A, there exists a certain equivalent c such that a situation B with a certain gain of c is equivalent to the situation A.

These are axioms, not proprieties. There is thus no demonstration but these rules are accepted anyway without formal proof. If this is the case, then there exists a utility function U such that selection criteria between two situations is based on the expected utility of the gain and not on the expected gain. Let now discuss the shape of such a utility function. It mostly relies on the fact that people make decisions rationally. That is, people tend to prefer more money to less money and people are usually averse to the risk, although this aversion varies from one individual to the other. The first hypothesis can be translated into mathematical language by stating that the utility function needs to be increasing as people prefer more wealth over less wealth. The second hypothesis suggests that we need to consider concave utility functions. This is a consequence from the fact that the risk aversion coefficient $r(x)$ is given by

$$r(x) = \frac{-U''(x)}{U'(x)}$$

Based on the equivalent utility principle, we should then be able to obtain the minimum risk premium asked by the longevity insurance sellers and the maximum acceptable risk premium considered by the longevity buyers. The size of the risk premium depends on the payoff structure of the product, the financial strength of the seller and buyer and on availability of the natural hedge (the impact of natural hedging is potentially significant).

The expected utility approach supposes that the decision maker tries to maximize his expected utility. In other words, a decision maker will prefer the random wealth Y over a random wealth X if and only if

$$\mathbb{E}[U(X)] \leq \mathbb{E}[U(Y)]$$

A decision maker is said to be risk averse if $\mathbb{E}[X]$ is always preferred to X ($\mathbb{E}[U(X)] \leq U\left(\mathbb{E}[X]\right)$). The risk premium is then defined as being the amount $\pi(X)$ such that

$$\mathbb{E}[U(X)] = U\left(\mathbb{E}[X] - \pi(x)\right)$$

On one hand, the utility of the buyers gives an upper bound on the price of a product (one is willing to buy something up to a certain point). On the other hand, the utility of the sellers gives a lower bound on the price of a product (one is willing to develop and sell a product up to profitability). We thus obtain a pricing interval to work on. The final price will stay within this interval but where in it usually depends on negotiation skills, marketing and market conditions. The lower bound is usually the minimum premium the insurer is willing to ask to cover the risk. It is solution of

$$\mathbb{E}[U_s(X + \pi_l)] = U_s(\mathbb{E}[X])$$

while the upper bound is usually the maximum premium the policyholder is willing to pay to have his risk covered. Henceforth, it is solution of

$$U_b(\mathbb{E}[X] - \pi_u) = \mathbb{E}[U_b(X)]$$

Note that U_s and U_b have been used to respectively represent the utility function of the seller and the one of the buyer. There is indeed no reason to consider these functions equal. Insurer-particular, government-insurer,... are intuitive example of couples where we would imagine the utility functions to be different.

Further in the utility theory development, we can look after the link between the risk aversion coefficient and the risk premium. Intuitively, we expect that the smaller the risk aversion coefficient, the smaller the risk premium should be. This can be proven analytically. Therefore, suppose that $U(x)$ is twice differentiable and let denote $x_u = \mathbb{E}[X]$ for a certain random variable X. Taylor's development of order 2 around x_u gives

$$U(x) \approx U(x_u) + U'(x_u)(x - x_u) + \frac{1}{2}U''(x_u)(x - x_u)^2$$

Taking the mathematical expectation of this expression leads to

$$\mathbb{E}[U(x)] = U(x_u) + \frac{1}{2}U''(x_u)\mathbb{V}(x)$$

Another Taylor expansion around x_u (of order 1 this time) gives

$$U(x_u - \pi(x)) = U(x_u) - U'(x_u)\pi(x)$$

The latter being equal to $\mathbb{E}[U(X)]$ by definition of the risk premium. As a result, we can write

$$U(x_u) + \frac{1}{2}U''(x_u)\mathbb{V}(x) = U(x_u) - U'(x_u)\pi(x)$$

Which finally provides desired result

$$\begin{aligned}\pi(X) &= -\frac{1}{2}\frac{U''(x_u)}{U'(x_u)}\mathbb{V}(x)\\ &= \frac{1}{2}r(x_u)\mathbb{V}(x)\end{aligned}$$

So the risk premium is proportional to both the risk aversion coefficient and the variance of the risk. This is no surprise to have an expression depending on the variance since the variance is the first and most intuitive indicator of the importance of the risk.

We will now price different securities based on the utility principle. This method is still widely used because of its intuitive basis. It is the method with the less mathematical developments and considerations, but the method that is the closest to the reality trying to understand human behaviour. As discussed in the first part of this paper, we can model mortality rates using a continuous or discrete time model. In what follows, mortality probabilities have been computed using the Lee Carter framework because of its nice representation of the force of mortality dynamics.

4.6.1 Longevity bond

pricing of a zero coupon longevity bond

Let work on an example to explain the utility pricing method of such a security. Suppose we want to price a zero coupon longevity bond paying K (normalized to one) in t years. The longevity factor is based on the surviving trend of a cohort of N males aged x_0 at time t_0. Hence, at time $t_0 + t$, the number of survivors is expressed by $S_t = N {}_tp_{x_0}$. The annuity provider must therefore pay an amount of $NK {}_tp_{x_0}$. Suppose also that the only risk considered is the longevity risk. The longevity risk is described by the deviation between the actual number of survivors and the expected one.

$$L_t = S_t - \mathbb{E}[S_t] = N\left({}_tp_{x_0}^{real} - \mathbb{E}[{}_tp_{x_0}^{real}]\right) \approx N\left({}_tp_{x_0}^{real} - {}_tp_{x_0}^{per}\right)$$

Now, let consider a concave utility function $U : w \to U(w)\ w \in \mathbb{R}^+$. Consider that the wealth of the seller is given by W_0 at time t_0 and $W_t = W_0 e^{\delta t}$ at time $t_0 + t$. It is known that

$$\mathbb{E}\Big[U\left(e^{-\delta t}(W_t - E[L_t])\right)\Big] = U\left(e^{-\delta t}(W_t - E[L_t])\right) \geq \mathbb{E}\Big[U\left(e^{-\delta t}(W_t - L_t)\right)\Big]$$

The smallest premium the seller is willing to ask to cover the longevity risk of the bond is the minimum premium π. The latter is solution of

$$U\left(e^{-\delta t}(W_t - \mathbb{E}[L_t])\right) = \mathbb{E}\Big[U\left(e^{-\delta t}(W_t - L_t) + \pi\right)\Big]$$

- Constant absolute risk averse (CARA) utility
  $U(w) = -\frac{1}{\alpha}\exp(-\alpha w)$ where α is the absolute risk aversion coefficient. As a result, the minimum risk premium is given by:

$$\begin{aligned} -\frac{1}{\alpha}\exp\left(-\alpha(e^{-\delta t}(W_t - \mathbb{E}[L_t])\right) &= \mathbb{E}\Big[-\frac{1}{\alpha}\exp\left(-\alpha(e^{-\delta t}(W_t - L_t) + \pi\right)\Big] \\ &= -\frac{1}{\alpha}\exp\left(-\alpha(e^{-\delta t}W_t + \pi)\right)\mathbb{E}\Big[\exp\left(\alpha L_t\right)\Big] \end{aligned}$$

so that

$$\pi = \frac{1}{\alpha}\log\left(\mathbb{E}\left[\exp\left(\alpha e^{-\delta t}(L_t - \mathbb{E}[L_t])\right)\right]\right)$$

- Constant relative risk averse (CRRA) utility
  $U(w) = \frac{w^{1-\gamma}}{1-\gamma}$. Consequently the minimum risk premium is solution of:

$$\begin{aligned} \frac{(e^{-\delta t}W_t)^{1-\gamma}}{1-\gamma} &= \mathbb{E}\left[\frac{\left(e^{-\delta t}(W_t - L_t) + \pi\right)^{1-\gamma}}{1-\gamma}\right] \\ \Leftrightarrow \qquad 1 &= \mathbb{E}\left[\left(1 + \frac{\pi e^{\delta t} - L_t}{W_t}\right)^{1-\gamma}\right] \end{aligned}$$

  Note that in this case the risk premium depends on the wealth level of the seller. The more wealthy the seller, the smaller the required risk premium.

The total price of the longevity bond is then given by

$$P = e^{-\delta t}\mathbb{E}[S_t] + \pi = Ne^{-\delta t}{}_tp_{x_0}^{per} + \pi$$

So that the total premium paid per person is equal to

$$P_i = P/N = e^{-\delta t}{}_tp_{x_0}^{per} + \pi/N$$

A small paragraph is needed concerning the selection and determination of the utility function of a third party. It makes sense to say that usually the wealthier a third party, the more risk he is willing to take and thus the smaller the risk premium. We can then propose a utility model based on the initial wealth level, W_0, of a third party

$$U(w) = -\frac{1}{\alpha(W_0)}\exp(-\alpha(W_0)w) \tag{4.5}$$

where $\alpha(W_0) = \alpha W_0^{-b}$ with $b \in [0,1]$. This utility function is often used because it keeps the convenient features of the negative exponential utility function (closed form for π) and because the expected utility is then separable between independent risks X and Y. Indeed

$$\mathbb{E}[U(X+Y)] = \mathbb{E}[U(X)]\mathbb{E}[U(Y)]$$

As we will use this utility function for numerical simulations, it is important to take the time to understand how each parameter influences it.

sensitivity to w direct differentiation leads to

$$\frac{\partial u}{\partial w} = \exp\left(-\frac{\alpha}{W_0^b} w\right) > 0$$

Everything equal, the bigger w, the bigger the utility as we discussed earlier. The increasing and concave feature of the utility function are well observable on Figure 4.2a.

sensitivity to b direct differentiation leads to

$$\frac{\partial u}{\partial b} = \underbrace{\frac{-W_0^b \ln(W_0)}{\alpha}}_{<0} \underbrace{\exp\left(-\frac{\alpha}{W_0^b} w\right)}_{>0} \underbrace{\left(1 + \frac{\alpha}{W_0^b} w\right)}_{>0} < 0$$

Everything equal, the bigger b, the smaller the utility. This is observable on Figure 4.2b. This is because the bigger b, the bigger the importance given to the initial wealth and henceforth the smaller the value of additional wealth.

sensitivity to W_0 direct differentiation leads to

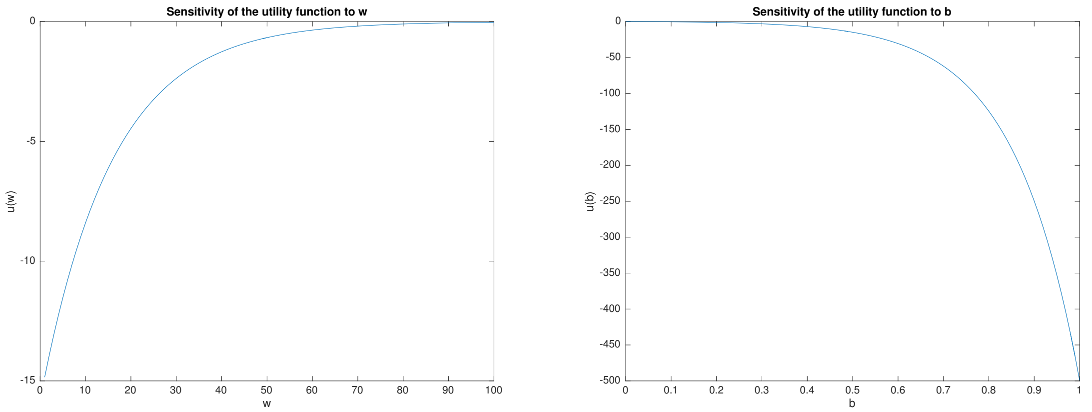
$$\frac{\partial u}{\partial W_0} = \underbrace{\frac{-bW_0^{b-1}}{\alpha}}_{<0} \underbrace{\exp\left(-\frac{\alpha}{W_0^b} w\right)}_{>0} \underbrace{\left(1 + \frac{\alpha w}{W_0^b} w\right)}_{>0} < 0$$

Everything equal, the bigger W_0, the smaller the utility. This is observable on Figure 4.3a. It is intuitive, the wealthier someone is, the less importance he brings to additional value.

sensitivity to α direct differentiation leads to

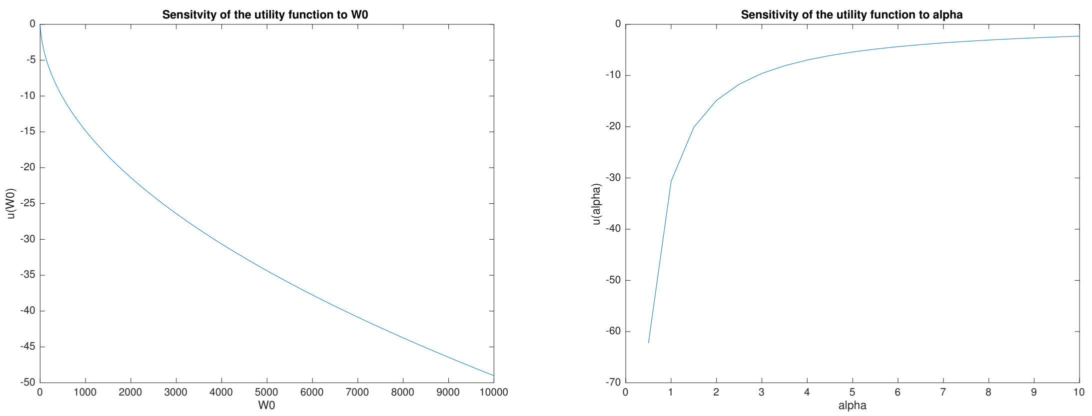
$$\frac{\partial u}{\partial \alpha} = \underbrace{\frac{W_0^b}{\alpha}}_{<0} \underbrace{\exp\left(-\frac{\alpha}{W_0^b} w\right)}_{>0} \underbrace{\left(\frac{1}{\alpha} + \frac{w}{W_0^b}\right)}_{>0} > 0$$

Everything equal, the bigger α, the bigger the utility. This is observable on Figure 4.3b



(A) $\alpha = 2$, $W_0 = 1000$, b $= 0.5$ (B) $\alpha = 2$, w $= 1$, $W_0 = 1000$

FIGURE 4.2



(A) $\alpha = 2$, w $= 100\ 000$, b $= 0.5$ (B) $W_0 = 1000$, x $= 1$, b $= 0.5$

FIGURE 4.3

Let now consider above utility equation in order to price the longevity risk. This can be done by taking ${}_tp_{x_0}^{real} = {}_tp_{x_0}^{prosp}$. Indeed, ${}_tp_{x_0}^{prosp}$ is a survival probability that takes the longevity phenomenon into account and is thus the closest guess to the reality. On the contrary, ${}_tp_{x_0}^{per}$ is a static survival probability. By static, we mean that the probability of living from age x to x+1 is constant whatever the calendar year t. Taking the difference of these two quantities leave us with a quantity completely dependant on the longevity of the cohort. Let us consider a cohort of 1000 people aged 50 in 2015. Suppose that the parameter b in the utility definition is taken to be equal to 0.5. We then obtain for different values of α, T and W_0 the following premiums to cover the longevity risk. A discount factor of 1% has been considered.

	$W_0 = 100$			$W_0 = 1000$			$W_0 = 10000$		
maturity	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
5	0.3473	1.2051	1.6826	0.2173	0.4169	0.7617	0.0694	0.1926	0.2976
10	2.4703	5.9804	8.0825	1.1463	2.3289	3.3416	0.2579	0.6195	1.2033
15	7.2689	14.7625	21.4131	2.7129	7.1135	10.4779	0.7786	1.6631	3.8012
20	14.9894	33.5137	38.6419	6.2226	14.7407	22.3777	1.1228	5.2627	8.8450
25	32.9249	53.7186	63.9309	11.8854	30.9311	44.6588	3.1110	10.6106	17.0449
30	60.4569	80.1978	103.7740	20.5816	46.5826	68.8493	7.0361	18.9850	30.5477
35	68.4798	124.1988	90.6713	27.3523	71.2820	95.9085	9.9453	24.9686	41.2652

Below is a table showing the total premium per person to be paid if one wants to obtain a payment of 1 after T years if alive and if one wants not to bear any longevity risk.

P_i	$W_0 = 100$			$W_0 = 1000$			$W_0 = 10000$		
maturity	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
5	0.9359	0.9368	0.9373	0.9358	0.9360	0.9364	0.9357	0.9358	0.9359
10	0.8673	0.8708	0.8729	0.8659	0.8671	0.8681	0.8651	0.8654	0.8660
15	0.7937	0.8012	0.8078	0.7891	0.7935	0.7969	0.7872	0.7881	0.7902
20	0.7160	0.7345	0.7396	0.7072	0.7157	0.7234	0.7021	0.7063	0.7098
25	0.6365	0.6573	0.6675	0.6155	0.6345	0.6483	0.6067	0.6142	0.6206
30	0.5471	0.5668	0.5904	0.5072	0.5332	0.5554	0.4936	0.5056	0.5171
35	0.4106	0.4663	0.4328	0.3695	0.4134	0.4380	0.3520	0.3671	0.3834

The premiums are decreasing with the maturity which makes sense. The longer the policyholder must wait for his payments, the higher the probability of him dying before the due date. As for the sensitivity to different utility functions, we can make the same observations as previously. Finally, since the survival probabilities are decreasing and the longevity risk premiums increasing but that the total premiums are decreasing, we conclude that the expected payoff has a bigger impact on the price than the longevity risk. That was expected.

It is also interesting to compare the premiums to the ones obtained with the Wang transform. For each maturity, despite being of the same magnitude, Wang premiums are bigger than the premiums computed with the equivalent utility principle. A few remarks concerning this result. First, this can be the result of the chosen utility function. Different values for b, α and W_0 may give premiums higher than the ones of Wang. Then, we assumed in the utility method that the seller asks the smallest risk premium. In other words, we here work with $\pi = \pi_l$. In practice, the commercial premiums may be higher than this and henceforth get closer to the Wang premiums. In conclusion, these two methods give similar results.

Let now analyze how important it is to consider the longevity risk when computing the needed reserves for solvency. Therefore, as the Human Mortality Database only provides data up to 2015, we will now consider a cohort of 1000 people aged 30 at time 1980 and various maturities T. The zero-bond provider will have to pay $N_T p_{x_0}^{real}$ at the end of the period. He therefore puts money aside based on a certain life table. If he considers (resp. does not consider) the longevity risk of the cohort, he will use a stochastic (resp. deterministic) life table for his reserve calculations which will be given by $N_T p_{x_0}^{prosp}$ (resp. $N_T p_{x_0}^{per}$). The surplus the issuer will have to pay from its equity funds is then given by the payoff less the reserves. The results are given below

Prospective	
T	$P_T = N({}_Tp_{x_0}^{real} - {}_Tp_{x_0}^{prosp})$
5	0.0067
10	0.0118
15	0.0104
20	0.0054
25	0.9374
30	6.8038
35	18.4235

Periodic	
T	$P_T = N({}_Tp_{x_0}^{real} - {}_Tp_{x_0}^{per})$
5	0.1592
10	1.4034
15	3.3137
20	6.4516
25	14.4783
30	31.2421
35	60.4148

A quick look through the numbers outlines the fact that working with prospective life tables helps to decrease the surplus needed quite drastically. As expected, the longevity effect is more observable and important for a bond of large maturity.

pricing of a coupon-based longevity bond

Assume a financial market with constant risk free rate r. Let δ be the compound discount rate and T the maturity of the bond. Suppose that the coupon rate is fixed and equal to c (%). The coupon basis, however, is variable and depends on the number of survivors of a certain cohort. A coupon-based bond can actually be seen as a sum of different zero coupon bonds of maturities 1,2,..., T. In the same context as the one described for zero coupon longevity bond, at time t, the exposure to longevity risk is given by

$$L_t = \begin{cases} cN({}_tp_{x_0}^{real} - {}_tp_{x_0}^{per}) & \text{if } 0 \leq \text{t} < \text{T} \\ (1+c)N({}_Tp_{x_0}^{real} - {}_Tp_{x_0}^{per}) & \text{if t = T} \end{cases}$$

Considering the CARA utility function, π is then solution of the equation

$$U\left(\sum_{t=1}^{T} e^{-\delta t}(W_t - \mathbb{E}[L_t])\right) = \mathbb{E}\left[U\left(\pi + \sum_{t=1}^{T} e^{-\delta t}(W_t - L_t)\right)\right]$$

And with developments totally identical to the ones above, we obtain

$$\pi = \frac{W_0^b}{\alpha} \log\left(\mathbb{E}\left[\exp\left(\frac{\alpha}{W_0^b}\sum_{t=1}^{T} e^{-\delta t}(L_t - \mathbb{E}[L_t])\right)\right]\right)$$

Numerical simulations are summarized in the table below. Results are obtained for a group of N = 1000 independent lives aged 50 in 2015, with a discount factor of 1% and a coupon rate of 5%. The numbers represent the cumulative risk premium asked by the insurer to hedge the longevity risk on the whole group considered. As expected, since there are now coupons to be paid after each year, resulting in more cash flows, the risk premium is more important. Beside that, the same comments as the ones produced on zero-coupon bonds are applicable.

P_i	$W_0 = 100$			$W_0 = 1000$			$W_0 = 10000$		
maturity	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
5	0.8855	2.7264	4.3606	0.2007	0.9550	1.4370	0.1757	0.4800	0.5939
10	5.5400	12.9736	18.8468	2.1036	4.9120	8.2914	1.0511	1.7385	3.2371
15	19.7534	38.7653	49.7218	6.6843	18.4442	25.5793	3.3435	6.3556	11.3206
20	47.4736	75.0642	89.1061	17.3799	42.9121	68.8446	6.3931	16.7750	26.3482
25	92.9093	146.5095	175.3395	39.9792	99.9189	167.4704	13.8065	36.2906	51.0347
30	140.7379	202.8173	211.0272	76.5818	148.1670	163.3892	26.8349	76.9870	131.7416
35	165.8956	260.2152	260.8645	96.7251	158.5939	235.1553	32.4906	85.7105	144.0556

It is important to note that this result of additivity is proper to the exponential utility function. Indeed, if we had to work with the CRRA utility function, then π implicitly solves the equation

$$\mathbb{E}\left[\left(\frac{\pi + \sum_{t=1}^{T} e^{-\delta t}(W_t - L_t)}{\sum_{t=1}^{T} e^{-\delta t}(W_t - \mathbb{E}[L_t])}\right)^{1-\gamma}\right] = 1$$

The total price of the longevity bond is then given by

$$P = \sum_{t=1}^{T} e^{-\delta t}\mathbb{E}[S_t] + \pi = cN\sum_{t=1}^{T} e^{-\delta t}\,{}_tp_{x_0}^{per} + Ne^{-\delta T}\,{}_Tp_{x_0}^{per} + \pi$$

So that the total premium paid per person is equal to

$$P_i = P/N = c\sum_{t=1}^{T} e^{-\delta t}\,{}_tp_{x_0}^{per} + e^{-\delta T}\,{}_Tp_{x_0}^{per} + \pi/N$$

P_i	$W_0 = 100$			$W_0 = 1000$			$W_0 = 10000$		
maturity	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
5	1.1770	1.1788	1.1805	1.1763	1.1771	1.1775	1.1763	1.1766	1.1767
10	1.3342	1.3417	1.3475	1.3308	1.3336	1.3370	1.3298	1.3304	1.3319
15	1.4746	1.4936	1.5045	1.4615	1.4732	1.4804	1.4581	1.4612	1.4661
20	1.6009	1.6285	1.6425	1.5708	1.5963	1.6222	1.5598	1.5702	1.5797
25	1.7099	1.7635	1.7923	1.6570	1.7169	1.7845	1.6308	1.6533	1.6680
30	1.7746	1.8367	1.8449	1.7105	1.7821	1.7973	1.6607	1.7109	1.7656
35	1.7558	1.8501	1.8508	1.6866	1.7485	1.8251	1.6224	1.6756	1.7340

pricing of a deferred longevity bond

A drawback of the EIB longevity bond is that the early coupon payments have very low longevity risk attached to them. Yet, these cash flows are also the most expensive part of the bond. Hence such bonds use a lot of capital to cover a short period of low-risk payments. A natural way to deal with this problem is for users to buy longevity bonds with deferred payments. A deferred longevity bond is a bond whose coupons are perceivable only at a certain time in the future. For a bond of maturity T, coupons will be paid starting at time s (1 < s < T). The pricing follows exactly the same calculations and assumptions as above. We simply get

$$P(s,T) = c\sum_{t=s}^{T} e^{-\delta(t-s)}\mathbb{E}[S_t] + e^{-\delta(T-s)}\mathbb{E}[S_T] + \pi$$

where π is solution of

$$U\left(\sum_{t=s}^{T} e^{-\delta(t-s)}(W_t - \mathbb{E}[L_t])\right) = \mathbb{E}\left[U\left(\pi + \sum_{t=s}^{T} e^{-\delta(t-s)}(W_t - L_t)\right)\right]$$

The general CARA utility then gives

$$\pi = \frac{W_0^b}{\alpha}\log\left(\mathbb{E}\left[\exp\left(\frac{\alpha}{W_0^b}\sum_{t=s}^{T} e^{-\delta(t-s)}(L_t - \mathbb{E}[L_t])\right)\right]\right)$$

while working with the CRRA utility yields a more complicated expression

$$\mathbb{E}\left[\left(\frac{\pi + \sum_{t=s}^{T} e^{-\delta t}(W_{t_0} - L_t)}{\sum_{t=s}^{T} e^{-\delta t}(W_{t_0} - \mathbb{E}[L_t])}\right)^{1-\gamma}\right] = 1$$

P(s,T) is the price of the longevity bond at time $t_0 + s$. To obtain the price of that same bond at time t_0, we just need to discount P over s years. That way, we find

$$P(0,T) = e^{-\delta s} P(s,T)$$

Longevity floors / Longevity caps

The payoff structure of a longevity floor is $\max\left(S_t^f - \mathbb{E}[S_t], 0\right) = (\mathbb{E}[S_t] - S_t)_-$. When the number of survivors is larger than the expected number of survivors, the financial institution faces a longevity loss. Longevity caps are quite the opposite. Their payoff structure is given by $\max\left(\mathbb{E}[S_t] - S_t^c, 0\right) = (\mathbb{E}[S_t] - S_t)_+$. Note that the survival probability is not the only possible underlying of the floors and caps. We could also imagine using longevity derivatives such as survivor futures or even zero coupon longevity bonds. The institution makes a longevity gain when the number of survivor is less than the expected number of survivors. Capping or flooring a bond has the consequence of reducing the payoff and thus in a second phase, the price of the bond. It therefore enables a cheap hedging against the systematic risk that is the longevity risk. Proceeding the same way as for the longevity bonds, the longevity risk premiums for the longevity floors and longevity caps under the global CARA utility are respectively given by

$$\pi^- = \frac{W_0^b}{\alpha} \log\left(\mathbb{E}\left[\exp\left(\frac{\alpha}{W_0^b} \sum_{t=1}^{T} e^{-\delta t}(L_t - E[L_t])\right)\right]\right)$$

$$\pi^+ = \frac{W_0^b}{\alpha} \log\left(\mathbb{E}\left[\exp\left(-\frac{\alpha}{W_0^b} \sum_{t=1}^{T} e^{-\delta t}(L_t - E[L_t])\right)\right]\right)$$

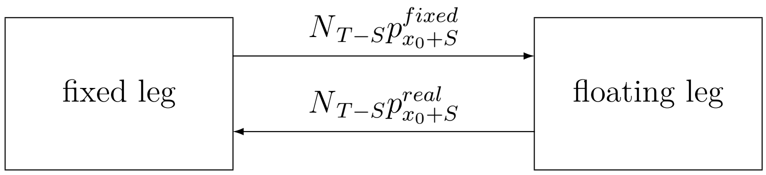
where

$$L_t = \begin{cases} (\mathbb{E}[S_t] - S_t)_- - \mathbb{E}\left[(\mathbb{E}[S_t] - S_t)_-\right] & \text{for a longevity floor} \\ (\mathbb{E}[S_t] - S_t)_+ - \mathbb{E}\left[(\mathbb{E}[S_t] - S_t)_+\right] & \text{for a longevity cap} \end{cases}$$

4.6.2 s-forwards

The denomination s-forward stands for survivor forward. It is a forward contract linked to a future survival rate. These products were the first used as hedging instruments against the longevity risk. It is important to carefully explore their conception and function because they usually are at the basis of other, more complex, longevity securities. For example, we can see a longevity swap as a sum of s-forward contracts with increasing maturities. An s-forward is defined as an arrangement between two parties in which they agree to exchange an amount proportional to the realized survival rate of a predetermined population in return to an amount proportional to a fixed survival rate that has been mutually agreed at the inception of the contract and payable at maturity. To simplify, it involves the exchange of a notional amount multiplied by the realized survival rate of an underlying population in return for the same notional amount multiplied by a fixed survival index agreed beforehand. A s-forward is thus a swap that exchanges a fixed survival rate for the realized survival rate at maturity. If the s-forward is fairly priced (both parties agree on a neutral survival rate), no cash is exchanged at inception. The only payment appears at maturity if the fixed and realized survival rates are different, and is proportional to their difference. Let N, S and T respectively denote the nominal amount, the starting time and the maturity of the s-forward contract. T-S is called the tenor. Its price at time t < S is F(t,S,T). Suppose that the survival rates are computed on a cohort aged x_0 at inception. The payoff at maturity, $F(T,S,T)$, is then given by

$$F(T,S,T) = N \left({}_{T-S}p^{real}_{x_0+S} - {}_{T-S}p^{fixed}_{x_0+S} \right)$$



$F(T,S,T)$ is a random variable because ${}_{T-S}p^{real}_{x_0+S}$ is unknown in advance. We could then consider paying a risk premium in order to replace $F(T,S,T)$ by $\mathbb{E}[F(T,S,T)]$ and so hedge the longevity risk. That way, the payoff to be received or paid is deterministic. Let now determine the premium of such a product. We will take the point of view of the buyer of such a product. The premium is the amount that the buyer is willing to pay in order to replace the real mortality exposure to the one determined in advance. The buyer receives the fixed leg and pays the floating leg. Suppose that the buyer here corresponds to an annuity provider that needs to pay an amount of 1 to each survivor of a certain cohort of N people after T years. Hence, following Jensen's inequality, we can write

$$U\left(e^{-\delta T}(W_T - \mathbb{E}[F(T,S,T)]) - \pi\right) = \mathbb{E}\left[U\left(e^{-\delta T}(W_T - F(T,S,T))\right)\right]$$

Using the same utility function as the one proposed in (4.5), we obtain

$$\begin{aligned} &-\frac{1}{\alpha(W_0)}\exp\left(-\alpha(W_0)\left(e^{-\delta T}(W_T - \mathbb{E}[F(T,S,T)]) - \pi\right)\right) \\ &= \mathbb{E}\left[-\frac{1}{\alpha(W_0)}\exp\left(-\alpha(W_0)\left(e^{-\delta T}(W_T - F(T,S,T))\right)\right)\right] \end{aligned}$$

Solving this equation then leads to the following premium

$$\begin{aligned} \pi &= \frac{1}{\alpha(W_0)}\log\left(\mathbb{E}\left[\exp\left(\alpha(W_0)e^{-\delta T}\left(F(T,S,T) - \mathbb{E}[F(T,S,T)]\right)\right)\right]\right) \\ &= \frac{1}{\alpha(W_0)}\log\left(\mathbb{E}\left[\exp\left(\alpha(W_0)Ne^{-\delta T}\left({}_{T-S}p^{real}_{x_0+S} - \mathbb{E}[{}_{T-S}p^{real}_{x_0+S}]\right)\right)\right]\right) \end{aligned}$$

For means of simulation, we will here assume that ${}_{T-S}p^{real}_{x_0+S}$ and ${}_{T-S}p^{fixed}_{x_0+S}$ are respectively given by ${}_{T-S}p^{prosp}_{x_0+S}$ (prospective probability) and ${}_{T-S}p^{per}_{x_0+S}$ (periodic probability). Let us fix the simulation parameters as follows:

- $N = 1000$
- $x_0 = 50$
- $t_0 = 2015$
- $b = 0.5$
- $\delta = 1\%$
- $S = 15$

	$W_0 = 100$			$W_0 = 1000$			$W_0 = 10000$		
maturity	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
5	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0
20	5.9851	13.9861	19.0510	2.0811	5.2868	9.1386	0.8568	2.3025	2.8300
25	25.9448	55.4185	56.0156	9.3183	24.2893	38.0431	4.1706	9.8151	14.6797
30	59.9566	86.2075	109.9617	22.8061	74.2488	76.9573	8.1949	23.0376	34.8667
35	83.0996	142.7323	148.9723	34.8702	79.0717	107.5966	12.7763	35.9794	53.4869

The table above shows some global risk premiums computed for certain parameters. By global, we mean this is the premium a reinsurer would ask to cover the longevity risk of the portfolio of 1000 people. As expected, hedging the longevity risk is more expensive on a long period (high maturity). We also clearly identify the influence of the parameters α and W_0 and the risk aversion coefficient. The larger α (resp. W_0), the larger (resp. smaller) the risk aversion coefficient, resulting in a larger (resp. smaller) risk premium. The three first lines are filled with zeros. This is because we work with $S = 15$ years. By definition, it is normal that all forward contracts with $T \leq S$ have a zero value. Indeed, it is then not possible to exchange something in T when the underlying exchange period starts in S.

Now, apart from the risk premium, we can compare how the reserving of s-forward contracts will differ when working with a periodic and a prospective life table. The latter trying to capture the longevity phenomenon of the cohort we consider. For example, suppose we buy a s-forward contract of maturity T to hedge the longevity exposure of a certain portfolio. We therefore work on a cohort of 1000 people aged 30 in 1980 and consider payoff structures of the form $F(T, S, T)$ with this time $S = 5$ years. The payoffs to be paid from the equity funds in the case of a prospective life table are given in the table on the left. The payoffs to be paid from the equity funds in the case of a periodic life table (the one of 1980) are given in the table on the right

$F(T,S,T) = N({}_{T-S}p^{real}_{x_0+S} - {}_{T-S}p^{prosp}_{x_0+S})$	
T	$F(T,S,T)$
5	0
10	-1.0386
15	-2.1991
20	-2.9054
25	-2.4885
30	0.5416
35	9.2801

$F(T,S,T) = N({}_{T-S}p^{real}_{x_0+S} - {}_{T-S}p^{per}_{x_0+S})$	
T	$F(T,S,T)$
5	0
10	0.2086
15	0.9727
20	3.4299
25	11.0451
30	24.9639
35	51.4344

The advantage of using prospective life tables when deciding on the fix life table is obvious. From the table on the right, the receiver (who receives the fixed leg), will earn money on short term s-forwards but will have to pay on very long term contracts but the amounts remain admissible. In contrast, from the table on the right, we notice that the receiver tends to receive less than what he actually gives. And this does not change with the maturity of the contract, on the contrary, he is losing more and more money as time flies.

4.6.3 Longevity swaps

Finally, as mentioned earlier, a longevity swap is a sum of s-forwards. If we denote by $F(t_0, T_{i-1}, T_i)$ a s-forward contract at time t_0 of tenor $T_i - T_{i-1}$, then the value at time

t_0 of a swap contract of maturity T can expressed as

$$S(t_0, T) = \sum_{i=1, t_0 \leq T_i} F(t_0, T_{i-1}, T_i)$$

Henceforth, the longevity risk premium associated with such a swap is given by

$$\begin{aligned}\pi &= \sum_{i=1, t_0 \leq T_i}^{T} \pi_i \\ &= \sum_{i=1, t_0 \leq T_i}^{T} \frac{1}{\alpha(W_0)} \log\left(\mathbb{E}\left[\exp\left(\alpha(W_0) N e^{-\delta T_i}\left({}_{T_i - T_{i-1}}p^{real}_{x_0 + T_{i-1}} - \mathbb{E}[{}_{T_i - T_{i-1}}p^{real}_{x_0 + T_{i-1}}]\right)\right)\right]\right)\end{aligned}$$

For simplicity in the numerical simulations, we assume that the time axis is decomposed such that $t_0 = T_0 < T_1 < ... < T_n = T$. We also consider $T_i = T_{i-1} + 5$. Again, fixing the numerical values of the parameters as follows

- $N = 1000$
- $x_0 = 50$
- $t_0 = 2015$
- $b = 0.5$
- $\delta = 1\%$

it is possible to find the value of the swap contract in $t = t_0$ for different maturities.

	$W_0 = 100$			$W_0 = 1000$			$W_0 = 10000$		
maturity	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 1$	$\alpha = 3$	$\alpha = 5$
5	0.2974	1.3029	1.8260	0.0742	0.4191	0.6208	0.0015	0.1199	0.1297
10	1.3650	4.0478	6.1033	0.2058	1.2125	1.9656	0.0211	0.4424	0.7148
15	3.7193	9.4302	13.0728	0.8958	3.3522	5.0397	0.1529	1.1314	2.0919
20	8.5141	19.7524	28.1245	2.0908	7.8761	11.8124	0.5826	2.3135	4.4846
25	17.0672	42.5753	58.0269	5.8229	16.9964	26.8744	1.5186	5.8234	9.8454
30	36.8219	78.9718	108.7250	12.8572	35.6511	54.5397	4.2092	13.6299	20.0131
35	68.7513	135.8805	171.1390	27.5769	65.9080	100.5901	7.5099	25.6922	39.6830

Finally, we now focus on the present value of the expected payments if prospective or periodic mortality tables are used to predict the real mortality. We therefore just adapt $t_0 = 1980$ and $x_0 = 30$.

$\sum_i e^{-\delta T_i} N({}_{T_i-T_{i-1}}p^{real}_{x_0+T_{i-1}} - {}_{T_i-T_{i-1}}p^{prosp}_{x_0+T_{i-1}})$	
T	$S(t_0, T)$
5	0.0063
10	0.0111
15	0.0099
20	0.0059
25	0.2467
30	7.7328
35	19.7199

$\sum_i e^{-\delta T_i} N({}_{T_i-T_{i-1}}p^{real}_{x_0+T_{i-1}} - {}_{T_i-T_{i-1}}p^{per}_{x_0+T_{i-1}})$	
T	$S(t_0, T)$
5	0.1514
10	1.2848
15	2.9649
20	5.6462
25	12.3106
30	26.0957
35	50.5761

Once again, expected payments for receiver swaps are way smaller if we consider prospective life table instead of periodic life table. This shows how important and how expensive the longevity risk is. It is certainly not viable on the long term to forget about it.

Chapter 5

Conclusion

In the first part of this paper, we explained how human mortality is modeled and outlined the fact that it is considerably evolving with time, heavily challenging financial institutions such as governments, insurers and annuity providers. The longevity risk forms a real threat for them on a solvency point of view and thereby weakens the financial sector in the world. This forces the major players to model and manage that risk appropriately.

Fortunately, people have already started developing multiple stochastic mortality models in order to capture the ageing trend among the base case scenario. The Lee-Carter and the CBD model are two very popular discrete models. The former is widely used among insurers even though it has been proven to underestimate the survival probabilities. Later, continuous time models, closer to the reality, were developed. The easy closed form for survival probabilities obtained with affine processes resulted in people focus mainly on the Ohrnstein-Uhlenbeck and Feller models. These models reduce the estimation error and are very flexible. They are thus very promising, explaining the global interest, even if the matching at younger ages is not totally adequate.

Longevity risk has been proven to be a systematic risk. It affects everybody in almost the same way, making classical pooling useless as a hedging instrument. Instead, longevity is securized through capital markets on which longevity-linked securities are exchanged. The payoffs of these financial instruments depend on the survival rate of a reference population. This allows insurers and banks to have cash inflows proportional to the longevity, and thus represents an opportunity to hedge the risk. The idea is to counterbalance longevity-linked cash outflows with longevity-linked cash inflows. Capital markets are very promising because of the enormous amount of players that might be interested; investors, banks, insurers, annuity providers and even governments.

The challenge, however, lies in the pricing of survivor products. Numerical pricing and reserving simulations put a stress on the fact that using prospective life tables considerably reduces the risk exposure of financial institutions such as banks and insurers. Thereby, their default risk decreases and the financial stability increases. Moreover, the importance of the risk premium outlines that the longevity risk is very expensive and certainly not to be omitted.

Longevity is a long-term phenomenon that took a real importance a few decades ago. It is becoming more and more problematic year after year. That is why we need further improvements on our current strategy in modeling, development of securities and pricing. From this will result an effective management method and, who knows, the opportunity to develop and speculate on a brand new capital market.

Chapter 6

Bibliography

[1] Bauer, D., & Russ, J. *Pricing Longevity Bonds Using Implied Survival Probabilities.* Working Paper, 2006.

[2] Biffis, E. *Affine processes for dynamic mortality and actuarial valuations.* Insurance: Mathematics & Economics, **37**(3), 2005 ,443 - 468

[3] Biffis, E., Denuit, M., & Devolder, P. *Stochastic mortality under measure change.* Working Paper, Institut des Sciences Actuarielles, Université Catholique de Louvain, Louvain-La-Neuve, Belgium, 2005.

[4] Brouhns, N., Denuit, M., & Vermunt, J.K. *Measuring the longevity risk in mortality projections.* Bulletin of the Swiss Association of Actuaries 2, 2002, 105-130.

[5] Cairns, A., Blake D. & Dowd K. *Modelling and Management of Mortality Risk: A Review.* Working Paper 22/08, 2008.

[6] Cairns, A., Blake D. & Dowd K. *Survivor Swaps.* Journal of Risk and Insurance, **73**, 2006,1-17.

[7] Cairns, A., Blake D. & Dowd K. *A two-factor model for stochastic mortality with parameter uncertainty.* Journal of Risk and Insurance, 2005.

[8] Cairns, A., Blake D. & Dowd K. *Pricing death: frameworks for the valuation and securization of mortality risk.* Astin Bulletin, **36**, 2006, 79-120.

[9] Cheridito, P., Filipovic, D. & Kimmel R. L. *Market price of risk specifications for affine models: Theory and Evidence.* Working Paper **36**, Princeton University, 2003.

[10] Dahl, M. *Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts.* Insurance: Mathematics & Economics,**35**(1), 2004, 113-136.

[11] Dahl, M.,& Moller,T. *Valuation and hedging of life insurance liabilities with systematic mortality risk*, Proceedings, 15th AFIR Colloquium, Zurich, 2005

[12] Denuit, M., Devolder, P. *Continuous time Lee Carter models and securization of longevity risk*, Working Paper 01/07, Institut des Sciences Actuarielles, Université catholique de Louvain, Louvain-la-Neuve, 2007.

[13] Denuit, M. (supplied Trufin J.), Hainaut, D. *LACTU2060: Life insurance II.* Lecture notes, Institut des Sciences Actuarielles, Université Catholique de Louvain, Louvain-La-Neuve, Belgium, 2017-2018.

[14] Denuit, M., Devolder, P., & Goderniaux, A.-C. *Securization of longevity risk: Pricing survival bonds with Wang transform in the Lee-Carter framework*, Journal of Risk and Insurance, 2007.

[15] Haberman, S., Renshaw A.E., *On the forecasting of mortality reduction factors.* Insurance: Mathematics & Economics, **32**, 2003, 379-401.

[16] Hainaut, D. *LACTU2060: Stochastic Finance.* Lecture notes, Institut des Sciences Actuarielles, Université Catholique de Louvain, Louvain-La-Neuve, Belgium, 2017-2018.

[17] Hull, J., & White, A., *Pricing interest-rate derivatives securities.* Review of Financial studies, **3**, 1990, 573-592.

[18] Jiajia, C., *Longevity Risk Pricing.* Living to 100 and Beyond Symposium, Orlando, 2008.

[19] Luciano, E., & Vigna, E.*Mortality risk via affine stochastic intensities: calibration and empirircal relevance.* Belgian Actuarial Bulletin, Vol. 8, No. 1, 2008.

[20] Luciano, E., & Vigna, E.*Non mean reverting affine processes for stochastic mortality.* Working Paper 04/05, ICER, 2005.

[21] Maghsoodi, Y. *Solution of the extended CIR term structure and bond option valuation.* Mathematical Finance, **6**, 1996, 89-109.

[22] Milevsky, M.A., Promislow, S.D, & Young, V.R. *Financial valuation of mortality risk via the instanteneous Sharpe Ratio.* Working Paper, 2005.

[23] Musiela, M., & Young, V.R. *An example of indifference prices under exponential preferences.* Finance and Stochastics, **8**, 2005, 229-239.

[24] Pitacco, E.*Survival models in a dynamic context: a survey*, Insurance: Mathematics & Economics, **35**(2), 2004, pp. 279-298.

[25] Schrager, D., *Affine stochastic mortality.* Insurance: Mathematics & Economics, **38**(1), 2006, 81-97.

[26] Wilmoth, J.R, Horiuchi, S., *Rectangularization revisited: variability of age at death within human populations.* Demography, **36**(4), 1999, 475-495.

[27] Young, V.R., & Zariphopoulou, T. *Pricing dynamics insurance risks using the principle of equivalent utility.* Scandinavian Actuarial Journal, 2002, 246-279.

Appendix A

Codes

A.1 Matlab

```
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Appendix: Matlab Codes              %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% we go until age 105 and start with data available in 1920
d_xt = xlsread('/Users/sboigelot/Desktop/TFE/Human Mortality
   Database','dxt1','CC5:FE114');
L_xt = xlsread('/Users/sboigelot/Desktop/TFE/Human Mortality
   Database','Lxt1','CC5:FE114');
m_xt = xlsread('/Users/sboigelot/Desktop/TFE/Human Mortality
   Database','mxt1','CC5:FE114');

% data for valuation of model
val_d_xt = xlsread('/Users/sboigelot/Desktop/TFE/Human
   Mortality Database','dxt1','FF5:FT114');
val_L_xt = xlsread('/Users/sboigelot/Desktop/TFE/Human
   Mortality Database','Lxt1','FF5:FT114');
val_m_xt = xlsread('/Users/sboigelot/Desktop/TFE/Human
   Mortality Database','mxt1','FF5:FT114');

% m = number of x
% n = number of t
[m,n] = size(d_xt);

% ages
for i = 2:m
```

```
    x(i) = x(i-1) + 1;
end

% 1. Lee-Carter
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% temporal mean of the log central mortality rates
% beta_1 is normally a mx1 vector
beta_x1 = mean(log(m_xt),2);
Z = log(m_xt)-beta_x1*ones(1,n);
% kt is a 1xn vector
kt = sum(Z);

% SVD decomposition
[U S V] = svd(Z);
u1 = U(:,1);
v1 = V(:,1);
sumu1 = sum(u1);
s1 = S(1,1);
b_x1 = u1/sumu1;
kt1 = s1*sumu1*v1';
sumx1 = sum(b_x1);
sumkt1 = sum(kt1);
meant1 = mean(kt1);
% model parameters SVD decomposition
ax = beta_x1+b_x1*meant1;
bx = b_x1/sumx1;
kt2 = (kt1-meant1*ones(1,n))*sum(b_x1);

%sum_bx = sum(bx);
%sum_kt = sum(kt2);
% kT = kt2(1,length(kt2));
% k1 = kt2(1,1);
% d = (kT-k1)/length(kt2);

% num is a mx1 vector and den is a scalar
num = Z*kt';
den = sum(kt.^2);
beta_x2 = num./den; % LSE method

% reestimation of kt to fit observed values
```

```
l = 10^-4; % d is a small number
% D is a 1xn vector (total observed number of deaths per year)
D = sum(d_xt);
% normally n*m times mxn
D2 = L_xt'*exp(beta_x1*ones(1,n)+beta_x2*kt);
dk = D2(1,:); % first line of D2
for j = 1:n
   i = 1;
   while dk(j) ~= D(j) && i <= 1000
       if dk(j) < D(j)
           kt(j) = kt(j)*(1+sign(kt(j))*l);
       else
           kt(j) = kt(j)*(1-sign(kt(j))*l);
       end
       i = i+1;
   end
end

t1 = 1920;
tn = 2000; % last time for which we have info
T = 2100; % projection limit
f = zeros(m,T-tn);

% fit ARIMA process for kt ... need data for this because first
    we need to
% determine which ARIMA to use ;)
% Let us work with ARIMA(0,1,0): kt+1 = kt + a + et+1
kt_proj = [kt zeros(1,T-tn)];
a = (kt(n) - kt(1))/(tn - t1 + 1);
ktt1 = kt(1:(end-1));
ktt2 = kt(2:end);
Dkt = ktt2-ktt1-a;
sigma = sum(Dkt.^2)/(tn - t1 + 1);

% Monte-Carlo simulations of future mortality

iter = 15000;
mean_error = zeros(m,15);
mean_f = zeros(m,T-tn);

for j = 1:iter
```

```
    for i = 1:T-tn
        kt_proj(n+i) = kt_proj(n+i-1) + a + sqrt(sigma)*randn
           (1);
        f(:,i) = m_xt(:,n).*exp(beta_x2*(kt_proj(n+i)-kt_proj(n
           )));
    end
    error = (f(:,1:15)-val_m_xt)./val_m_xt;
    mean_error = (mean_error*(j-1)+error)/j;
    mean_f = (mean_f*(j-1)+f)/j;
end

% Simulations SVD
% rand = randn(1,T-tn);
% for i = 2:(T-tn)
%     rand(i) = rand(i)+rand(i-1);
% end
% deltaT = 1:(T-tn);
% k_projected = kT*ones(1,T-tn) + deltaT*d + rand;
% lnm_xt_proj = ax*ones(1,T-tn)+bx*k_projected;
% m_xt_proj = exp(lnm_xt_proj);

figure(1)
plot(tn:(tn+15-1),mean_error);

figure(2)
plot(tn:(tn+15-1),mean(mean_error),'-o');

figure(3)
plot(x,mean_error');

figure(4)
plot(x,mean(mean_error'),'-o');

figure(5)
plot(x,log(m_xt(:,n))); hold on;
plot(x,log(f(:,10))); hold on;
plot(x,log(f(:,20))); hold on;
plot(x,log(f(:,40))); hold on;
plot(x,log(f(:,50))); hold on;
plot(x,log(f(:,85)));
```

```
figure(6)
mesh(x'*ones(1,T-t1+1),ones(m,1)*(t1:T),log([m_xt f]));


% 2. CBD MODEL
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% mxn matrix
logitq = log(q_xt./(ones(m,n)-q_xt));

% 1xn vector
sumlq = sum(logitq);

meanx = mean(x);
xf = x - meanx;
xf2 = xf.^2;
sumxf = sum(xf);
sumxf2 = sum(xf2);

num2 = length(x)*xf*logitq-sumxf*sumlq; % 1xn vector
den2 = length(x)*sumxf2 + sumxf^2; % scalar
kt2 = num2./den2; % 1xn vector
kt1 = (sumlq - sumxf*kt2)./length(x); % 1xn vector

t1 = 1920;
tn = 2000; % last time for which we have info
T = 2100; % projection limit
f = zeros(m,T-tn);
f2 = zeros(m,T-t1);

% fit ARIMA process for kt ... need data for this because first we need to
% determine which ARIMA to use ;)
% Let us work with ARIMA(0,0,1): kt+1 = kt + a + et+1
% Bivariate random walk with drift.
% We may want to make it stationary
kt = [kt1;kt2];
kt_proj = [kt zeros(2,T-tn)];
a = (kt(:,n) - kt(:,1))/(tn - t1 + 1);
ktt1 = kt(:,1:(end-1));
ktt2 = kt(:,2:end);
```

```
Dkt = ktt2-ktt1-a*ones(1,tn-t1);
sigma = sum(Dkt.^2,2)/(tn - t1 + 1);

% Cholesky decomposition of covariance matrix
Corr = corrcoef(kt1,kt2);
C = chol(Corr);

% Monte Carlo simulations for future mortality rates

iter = 1500;
error = zeros(m,15);
mean_error = zeros(m,15);
mean_f = zeros(m,T-tn);

 for j = 1:iter
    for i = 1:T-tn
        kt_proj(:,n+i) = kt_proj(:,n+i-1) + a + C*sqrt(sigma).*
           randn(2,1);
        f(:,i) = exp(ones(m,1)*kt_proj(1,n+i)+kt_proj(2,n+i)*xf
           ');
        f(:,i) = f(:,i)./(ones(m,1)+f(:,i));
    end
    error = (f(:,1:15)-val_q_xt)./val_q_xt;
    mean_error = (mean_error*(j-1)+error)/j;
    mean_f = (mean_f*(j-1)+f)/j;
 end
m_xt_proj = -log(ones(m,T-tn)-mean_f); % CFM assumption

figure(1)
plot(x,log(-log(ones(m,1)-q_xt(:,n)))); hold on;
plot(x,log(m_xt_proj(:,1))); hold on;
plot(x,log(m_xt_proj(:,10))); hold on;
plot(x,log(m_xt_proj(:,20))); hold on;
plot(x,log(m_xt_proj(:,40))); hold on;
plot(x,log(m_xt_proj(:,50))); hold on;
plot(x,log(m_xt_proj(:,85)));

figure(2)
plot(x,q_xt(:,n)); hold on;
plot(x,mean_f(:,1)); hold on;
plot(x,mean_f(:,10)); hold on;
```

```
plot(x,mean_f(:,20)); hold on;
plot(x,mean_f(:,40)); hold on;
plot(x,mean_f(:,50)); hold on;
plot(x,mean_f(:,85));

figure(3)
plot(1920:2000,kt_proj(1,1:n),'-b'); hold on;
plot(2000:2100,kt_proj(1,n:end),'-r');

figure(4)
plot(1920:2000,kt_proj(2,1:n),'-b'); hold on;
plot(2000:2100,kt_proj(2,n:end),'-r');

figure(5)
plot(x,-log(ones(m,15)-val_q_xt)'+log(ones(m,15)-mean_f(:,1:15)
    )');

figure(6)
plot(x,mean_error');

figure(7)
plot(x,mean(mean_error'));

figure(8)
plot(tn:tn+15-1,mean_error);

figure(9)
plot(tn:tn+15-1,mean(mean_error));

% 3. Reduction Factors
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% we only consider the last year of available data
q_xt = xlsread('Human Mortality Database','qxt','FT5:FT27');
m = length(q_xt);
x = zeros(1,m);
for i = 2:m
    x(i) = x(i-1) + 5;
end
T = 2100;
t = 2015;
```

```
q_proj = [x' q_xt*ones(1,T-t)];

alpha = ones(m,1);
f = ones(m,1);

% definition of the reduction function
for i = 1:m
    if q_proj(i,1) < 60
        alpha(i) = 0.13;
        f(i) = 0.55;
    elseif 0.13 <= q_proj(i,1) && q_proj(i,1) <= 110
        alpha(i) = 1 + 0.87*(q_proj(i,1)-110)/50;
        f(i) = 1 + 0.87*(-0.55*(q_proj(i,1)-110)+0.29*(q_proj(i,1)-60))/50;
    elseif q_proj(i,1) > 110
        f(i) = 0.29;
    end
end

% projections
qf = q_xt*ones(1,T-t);
time = zeros(T-t,1);
for i = 1:T-t
    if i == 1
        time(i) = 1;
    else
        time(i) = time(i-1) + 1;
    end
    qf(:,i) = qf(:,i).*(alpha + (ones(m,1) - alpha).*(ones(m,1)-f).^(time(i)/20));
end

figure(1)
plot(x,log(q_xt)); hold on;
plot(x,log(qf(:,10))); hold on;
plot(x,log(qf(:,30))); hold on;
plot(x,log(qf(:,50))); hold on;
plot(x,log(qf(:,70))); hold on;
plot(x,log(qf(:,85)));

% 4. Ohrnstein-Uhlenbeck model
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% du = au dt + sigma dWt

t1 = 1920;
tn = 2000;
T = 2100;
t = t1:T;

% Calibration of the OU process
den = sum(m_xt(:,1:end-1).^2,2); % u_i^2
s1 = m_xt(:,1:end-1);
s2 = m_xt(:,2:end);
num = sum(s1.*s2,2);
% mx1 vector
a1 = log(num./den);
a = a1*ones(1,tn-t1);
num2 = sum((s2-s1.*exp(a)).^2,2);
% mx1 vector
sigma = sqrt(num2/(n-1));

% Monte Carlo simulations

iter = 15000;
error = zeros(m,15);
mean_error = zeros(m,15);
mean_u = zeros(m,T-tn);

for j = 1:iter
    u_proj = zeros(m,T-tn);
    for i = 1:T-tn
        if i == 1
            u_proj(:,i) = m_xt(:,end);
        else
            u_proj(:,i) = u_proj(:,i-1).*exp(a1)+sigma.*sqrt((
               exp(2*a1)-ones(m,1))./(2*a1)).*randn(m,1);
        end
    end
    error = (u_proj(:,1:15)-val_m_xt)./val_m_xt;
    mean_error = (mean_error*(j-1)+error)/j;
    mean_u = (mean_u*(j-1)+u_proj)/j;
```

```
end

figure(1)
plot(tn:(tn+15-1),mean_error);
figure(2)
plot(tn:(tn+15-1),mean(mean_error),'-o');

figure(3)
plot(x,mean_error');

figure(4)
plot(x,mean(mean_error'),'-o');

total = [m_xt mean_u];
years = ones(m,1)*t;
ages = x'*ones(1,T-t1+1);

figure(5)
mesh(ages,years,total);

% 5. Feller model
%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Loading data computed using R
load('paramFeller.mat');
a = paramFellerdata(:,1);
sigma = paramFellerdata(:,2);
m = length(a);

% Monte Carlo simulations

iter = 15000;
error = zeros(m,15);
mean_error = zeros(m,15);
mean_u = zeros(m,T-tn);

for j = 1:iter
    u_proj = zeros(m,T-tn);
    for i = 1:T-tn
        if i == 1
            u_proj(:,i) = m_xt(:,end);
```

```
        else
            u_proj(:,i) = u_proj(:,i-1).*(1+a)+sigma.*sqrt(
               u_proj(:,i-1)).*randn(m,1);
        end
    end
    error = (u_proj(:,1:15)-val_m_xt)./val_m_xt;
    mean_error = (mean_error*(j-1)+error)/j;
    mean_u = (mean_u*(j-1)+u_proj)/j;
end

figure(1)
plot(tn:(tn+15-1),mean_error);

figure(2)
plot(tn:(tn+15-1),mean(mean_error),'-o');

figure(3)
plot(x,mean_error');

figure(4)
plot(x,mean(mean_error'),'-o');

total = [m_xt mean_u];
years = ones(m,1)*t;
ages = x'*ones(1,T-t1+1);

figure(5)
mesh(ages,years,real(total));

% 6. Pricing using equivalent utility principle
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


q_xt_tot = [q_xt val_q_xt];
p_xt_tot = 1-q_xt_tot;

%%%%%%%%%%%%%%%%%%%%%%%%%
% ESTIMATION
%%%%%%%%%%%%%%%%%%%%%%%%%

% m = number of x
```

```
% n : number of t
[m,n] = size(d_xt);
% central mortality rates

%m_xt = d_xt./L_xt;
%val_m_xt = val_d_xt./val_L_xt;

% temporal mean of the log central mortality rates
% beta_1 is normally a mx1 vector
beta_x1 = mean(log(m_xt),2);
Z = log(m_xt)-beta_x1*ones(1,n);

% x1 = 0, x2 = 1-4, x3 = 5-9,....
x = zeros(1,m);
for i = 2:m
    x(i) = x(i-1) + 1;
end

[U S V] = svd(Z);
u1 = U(:,1);
v1 = V(:,1);
sumu1 = sum(u1);
s1 = S(1,1);
beta_x2 = -u1;
kt1 = -s1*v1';

%%% START LC_model estimation %%%%

% kt1 = sum(Z);
%
% [U S V] = svd(Z);
% u1 = U(:,1);
% v1 = V(:,1);
% sumu1 = sum(u1);
% s1 = S(1,1);
% b_x1 = u1/sumu1;
% kt11 = s1*sumu1*v1';
%
% meant1 = mean(kt11);
% ax = beta_x1+b_x1*meant1;
% bx = b_x1/sumx1;
```

```
% kt2 = (kt11-meant1*ones(1,n))*sum(b_x1);
%
% kT = kt2(1,length(kt2));
% k1 = kt2(1,1);
% d = (kT-k1)/length(kt2);
%
% % num is a mx1 vector and den is a scalar
% num = Z*kt1';
% den = sum(kt1.^2);
% beta_x2 = num./den;
%
% % reestimation of kt to fit observed values
% l = 10^-4; % d is a small number
% % D is a 1xn vector (total observed number of deaths per year
  )
% D = sum(d_xt);
% % normally n*m times mxn
% D2 = L_xt'*exp(beta_x1*ones(1,n)+beta_x2*kt1);
% dk = D2(1,:); % first line of D2
% for j = 1:n
%    i = 1;
%    while dk(j) ~= D(j) && i <= 1000
%        if dk(j) < D(j)
%            kt1(j) = kt1(j)*(1+sign(kt1(j))*l);
%        else
%            kt1(j) = kt(j)*(1-sign(kt1(j))*l);
%        end
%        i = i+1;
%    end
% end
%
% a = (kt1(n) - kt1(1))/(tn - t1 + 1);
% ktt1 = kt1(1:(end-1));
% ktt2 = kt1(2:end);
% Dkt = ktt2-ktt1-a;
% sq_sigma_e = sum(Dkt.^2)/(tn - t1 + 1);

%%% END LC MODEL ESTIMATION %%%


%added for utility simulations
```

```
delta_kt1 = kt1(2:end)-kt1(1:(end-1));
c = (kt1(end)-kt1(1))/length(delta_kt1);
sq_sigma_e = (1/length(delta_kt1))*sum((delta_kt1-c).^2);
sigma_c = sqrt(sq_sigma_e)/sqrt(length(delta_kt1));

%%%%%%%%%%%%%%%%%%%
% SIMULATION
%%%%%%%%%%%%%%%%%%%

% fit ARIMA process for kt ... need data for this because first
    we need to
% determine which ARIMA to use ;)
% Let us work with ARIMA(0,0,1): kt+1 = kt + a + et+1
t1 = 1920;
tn = 2000; % last time for which we have info
tval1 = 2001;
tvaln = 2015;
T = 2100; % projection limit
iter = 1500;

%%%%%%%%%%%%%%%%%
% ZERO-COUPON BOND
%%%%%%%%%%%%%%%%%

 % reminder: tn = 2000 and tvaln = 2015

t0 = 1980; % work with cohort aged 50 in 2018
x0 = 30;
N = 1000;
b = 0.5; % importance of initial wealth [0,1]
delta = 0.01;
premium = zeros(7,9);
p1 = zeros(7,1);
p2 = zeros(7,1);
price_pi = zeros(7,9);
for W0 = [100 1000 10000]
    if W0 == 100
        l = 0;
    end
    for alpha = [1 3 5]
        l = l+1;
```

```
        anew = alpha*W0^(-b);
        k = 0;
        for mat = [5 10 15 20 25 30 35]
            k = k+1;
            p_real = 1;
            p_per = 1;
            mean_prem = 0;
            mean_p_long = 0;
            mean_mu = zeros(m,T-tn);

            for i = 1:mat
                p_per = p_per*p_xt_tot(x0+i,t0-t1+1);
                p_real = p_real*p_xt_tot(x0+i,t0-t1+i);
            end

            for j = 1:iter
                kt_proj = [kt1 zeros(1,T-tn)];
                for i = 1:(T-tn)
                    kt_proj(n+i) = kt_proj(n+i-1) + c + sqrt(sq_sigma_e)*randn(1);
                    ;
                end
                mu = exp(beta_x1*ones(1,T-tn)+beta_x2*kt_proj((n+1):end));
                mean_mu = (mean_mu*(j-1)+mu)/j;
                mu2 = [m_xt mu]; % mu(x,t) from time 1920 to 2100

                % consider cohort of people born in 1950 (are 65 yo in 2015, 50 yo in 2000)
                % add a for loop for each cohort

                p_long = 1;

                for i = 1:mat
                    p_long = p_long*exp(-mu2(x0+i,t0-t1+i));
                end
                mean_p_long = (mean_p_long*(j-1)+p_long)/j;
                prem = exp(anew*N*exp(-delta*mat)*(p_long-mean_p_long));
                mean_prem = (mean_prem*(j-1)+prem)/j;
```

```
                end
                p1(k) = N*(p_real-mean_p_long);
                p2(k) = N*(p_real-p_per);
                premium_long = 1/anew*log(mean_prem);
                premium(k,l) = premium_long;
                price_pi(k,l) = p_per*exp(-delta*mat)+round(
                   premium_long/N,4);
            end
        end
end
out1 = round(premium,4);
out2 = round(p1,4);
out3 = round(p2,4);


%%%%%%%%%%%%%%%%%%%%
% COUPON BONDS
%%%%%%%%%%%%%%%%%%%%

for W0 = [100 1000 10000]
    if W0 == 100
        l = 0;
    end
    for alpha = [1 3 5]
        l = l+1;
        anew = alpha*W0^(-b);
        k = 0;
        for mat = [5 10 15 20 25 30 35]
            k = k+1;
            p_real = ones(mat,1);
            p_per = ones(mat,1);
            mean_prem = 0;
            mean_p_long = zeros(mat,1);
            mean_mu = zeros(m,T-tn);
            payoff = c*ones(mat,1);
            payoff(end) = payoff(end)+1;

            for i = 1:mat
                if i == 1
                    p_per(i) = p_xt_tot(x0+i,t0-t1+1);
                    %p_real(i) = p_xt_tot(x0+i,t0-t1+i);
```

```
            else
                p_per(i) = p_per(i-1)*p_xt_tot(x0+i,t0-t1+1);
                %p_real(i) = p_real(i-1)*p_xt_tot(x0+i,t0-t1+i);
            end
        end

        for j = 1:iter
            kt_proj = [kt1 zeros(1,T-tn)]; % kt from time 1920 to 2100
            for i = 1:(T-tn)
                kt_proj(n+i) = kt_proj(n+i-1) + c + sqrt(sq_sigma_e)*randn(1);
            end
            mu = exp(beta_x1*ones(1,T-tn)+beta_x2*kt_proj((n+1):end));
            mean_mu = (mean_mu*(j-1)+mu)/j;
            mu2 = [m_xt mu]; % mu(x,t) from time 1920 to 2100

            p_long = ones(mat,1);

            for i = 1:mat
                if i==1
                    p_long(i) = exp(-mu2(x0+i,t0-t1+i));
                else
                    p_long(i) = p_long(i-1)*exp(-mu2(x0+i,t0-t1+i));
                end
            end
            mean_p_long = (mean_p_long*(j-1)+p_long)/j;
            prem = exp(anew*N*sum(exp(-delta*(1:mat)').*payoff.*(p_long-mean_p_long)));
            mean_prem = (mean_prem*(j-1)+prem)/j;
        end
        %p1(k) = N*(p_real-mean_p_long);
        %p2(k) = N*(p_real-p_per);
        premium_long = 1/anew*log(mean_prem);
        premium(k,l) = premium_long;
```

```
                price_pi(k,l) = sum(payoff.*p_per.*exp(-delta*(1:
                   mat)'))+round(premium_long/N,4);
            end
        end
end
out1 = round(premium,4);
out2 = round(p1,4);
out3 = round(p2,4);


% %%%%%%%%%%%%%%%%%%%%%%%%%
%  Deferred longevity bond
% %%%%%%%%%%%%%%%%%%%%%%%%%

mat = 35;
r = 0.01;
premium3 = 0;
be3 = 0;
s = 4;
for k = 1:(mat-s)
    for j = 1:iter
        kt_proj = [kt1 zeros(1,T-tn)];
        prosp = zeros(1,x(end)-x0);
        per = zeros(1,x(end)-x0);
        for i = 1:(T-tn)
            kt_proj(n+i) = kt_proj(n+i-1) + c + sqrt(sq_sigma_e
               )*randn(1);
        end
        mu = exp(beta_x1*ones(1,T-tn)+beta_x2*kt_proj((n+1):end
           ));
        mu2 = [m_xt mu];
        % consider cohort of people born in 1980 (are 65 yo in
           2045, 20 yo in 2000)
        % add a for loop for each cohort
        for i = 1:(x(end)-x0)
            if i == 1
                prosp(i) = exp(-mu2(x0+i,(t0-t1)+i));
                per(i) = p_xt(x0+i);
            else
                prosp(i) = prosp(i-1)*exp(-mu2(x0+i,t0-t1+i));
                per(i) = per(i-1)*p_xt(x0+i);
```

```
            end
        end
        % St
        S = N*(prosp-per);
        % E[St]
        mean_s = (mean_s*(j-1)+S)/j;
        p = exp(anew*(mean_s(k+s)-S(k+s)));
        mean_p = (mean_p*(j-1)+p)/j;
    end
    premium3 = premium3+exp(-r*k)*mean_p;
    be3 = be3 + exp(-r*k)*mean_s(k+s);
end
riskp3 = premium3/be3

%%%%%%%%%%%%%%%%%%%%%%%
% S-FORWARD
%%%%%%%%%%%%%%%%%%%%%%%


S = 5;
t0 = 1980; % work with cohort aged 50 in 2015
age = 30+S;
N = 1000;
b = 0.5; % importance of initial wealth [0,1]
delta = 0.01;
premium = zeros(7,9);
p1 = zeros(7,1);
p2 = zeros(7,1);
mean_mu = zeros(m,T-tn);
for W0 = [100 1000 10000]
    if W0 == 100
        l = 0;
    end
    for alpha = [1 3 5]
        l = l+1;
        anew = alpha*W0^(-b);
        k = 0;
        time = ([5 10 15 20 25 30 35]-S+abs([5 10 15 20 25 30
           35]-S))/2;
        tenor = time(time > 0);
        p_real = ones(length(tenor),1);
```

```
        p_per = ones(length(tenor),1);
        p_long = ones(length(tenor),1);
        prem = zeros(length(tenor),1);
        mean_prem = zeros(length(tenor),1);
        mean_p_long = zeros(length(tenor),1);
        for mat = time(time > 0)
            k = k+1;
            x0 = age;
            for i = 1:mat
                p_per(k) = p_per(k)*p_xt_tot(x0+i,t0-t1+1); %
                   crude survival proba by column
                p_real(k) = p_real(k)*p_xt_tot(x0+i,t0-t1+i); %
                    crude survival by diagonal
            end
            for j = 1:iter
                kt_proj = [kt1 zeros(1,T-tn)]; % kt from time
                   1920 to 2100
                for i = 1:(T-tn)
                    kt_proj(n+i) = kt_proj(n+i-1) + c + sqrt(
                       sq_sigma_e)*randn(1);
                end
                mu = exp(beta_x1*ones(1,T-tn)+beta_x2*kt_proj((
                   n+1):end));
                mean_mu = (mean_mu*(j-1)+mu)/j;
                mu2 = [m_xt mu]; % mu(x,t) from time 1920 to
                   2100

                % consider cohort of people born in 1950 (are
                   65 yo in 2015, 50 yo in 2000)
                % add a for loop for each cohort

                p_long(k) = 1;

                for i = 1:mat
                    p_long(k) = p_long(k)*exp(-mu2(x0+i,t0+S-t1
                       +i));
                end
                mean_p_long(k) = (mean_p_long(k)*(j-1)+p_long(k
                   ))/j;
                prem(k) = exp(anew*N*exp(-delta*mat)*(p_long(k)
                   -mean_p_long(k)));
```

```
                    mean_prem(k) = (mean_prem(k)*(j-1)+prem(k))/j;
                end
                p1(S/5+k) = N*(p_real(k)-mean_p_long(k));
                p2(S/5+k) = N*(p_real(k)-p_per(k));
                premium_long = 1/anew*log(mean_prem(k));
                premium(S/5+k,l) = premium_long;
                out1 = round(premium,4);
                out2 = round(p1,4);
                out3 = round(p2,4);
            end
        end
    end


%%%%%%%%%%%%%%%%%%
% Longevity swap %
%%%%%%%%%%%%%%%%%%

tenor = 5;
t0 = 1980; % work with cohort aged 50 in 2015
age = 30;
N = 1000;
b = 0.5; % importance of initial wealth [0,1]
delta = 0.01;
premium = zeros(7,9);
p1 = zeros(7,1);
p2 = zeros(7,1);
p11 = zeros(7,1);
p21 = zeros(7,1);
mean_mu = zeros(m,T-tn);
for W0 = [100 1000 10000]
    if W0 == 100
        l = 0;
    end
    for alpha = [1 3 5]
        l = l+1;
        anew = alpha*W0^(-b);
        k = 0;
        proba_per = ones(length(0:5:35)-1,1);
        proba_prosp = ones(length(0:5:35)-1,1);
        proba_real = ones(length(0:5:35)-1,1);
```

```
        prem = zeros(length(0:5:35)-1,1);
        mean_prem = zeros(length(0:5:35)-1,1);
        mean_p_long = zeros(length(0:5:35)-1,1);
        swap = zeros(length(0:5:35)-1,1);

        for mat = [5 10 15 20 25 30 35]
            k = k+1;
            x0 = age + mat - 5;

            for i = 1:tenor
                proba_per(mat/5) = proba_per(mat/5)*p_xt_tot(x0
                   +i,t0-t1+1); % crude survival proba by column
                proba_real(mat/5) = proba_real(mat/5)*p_xt_tot(
                   x0+i,t0+mat-5-t1+i); % crude survival by
                   diagonal
            end

            for j = 1:iter
                kt_proj = [kt1 zeros(1,T-tn)]; % kt from time
                   1920 to 2100
                for i = 1:(T-tn)
                    kt_proj(n+i) = kt_proj(n+i-1) + c + sqrt(
                       sq_sigma_e)*randn(1);
                end
                mu = exp(beta_x1*ones(1,T-tn)+beta_x2*kt_proj((
                   n+1):end));
                mean_mu = (mean_mu*(j-1)+mu)/j;
                mu2 = [m_xt mu]; % mu(x,t) from time 1920 to
                   2100

                % consider cohort of people born in 1950 (are
                   65 yo in 2015, 50 yo in 2000)
                % add a for loop for each cohort
                proba_prosp(mat/5)=1;
                for i = 1:tenor
                    proba_prosp(mat/5) = proba_prosp(mat/5)*exp
                       (-mu2(x0+i,t0+mat-5-t1+i)); % for
                       deferred we would need to change the time
                end
                mean_p_long(mat/5) = (mean_p_long(mat/5)*(j-1)+
                   proba_prosp(mat/5))/j;
```

```
                prem(mat/5) = exp(anew*N*exp(-delta*mat)*(
                   proba_prosp(mat/5)-mean_p_long(mat/5)));
                mean_prem(mat/5) = (mean_prem(mat/5)*(j-1)+prem
                   (mat/5))/j;
                swap(mat/5) = sum(1/anew*log(mean_prem(1:(mat
                   /5))));
            end
            premium(k,l)=swap(k);
            p1(k) = sum(N*(proba_real(1:k)-proba_prosp(1:k)).*
               exp(-delta*(5:5:5*k)'));
            p2(k) = sum(N*(proba_real(1:k)-proba_per(1:k)).*exp
               (-delta*(5:5:5*k)'));
        end
    end
end
```

A.2 R

```
# Appendix: R codes
##############################

require(datasets);
require(stats4);
require(R.matlab);

t0 <- 1841
t1 <- 1980;
tval <- 2015;
tn <- 2015;
tproj <- 2100;

qxt <- read.csv("qxt.csv",header=TRUE,sep=";");
qxt <- data.matrix(qxt);
qxt <- matrix(qxt,nrow = nrow(qxt),ncol=ncol(qxt),dimnames=NULL
   );
qxt <- qxt[,(t1-t0+1):(ncol(qxt)-(tn-tval+1))];

m <- nrow(qxt); # = 111 car ages de [U+FFFD]10
```

```
n <- ncol(qxt); # = 21 car data de 1980[U+FFFD]2000 pour la
    calibration


# 1. OU model via MLE
############################
# Ornstein-Uhlenbeck model via ML estimation: du = b(a-u)dt +
    sigma dw

x <- 30;  #cohorte a[U+FFFD]de 30
uyt <- -log(1-qxt[(x+1):(x+1+tval-t1),]);

# log likelihood function to optimise in theta. Here theta = [
    beta1, beta2, kt]. It is a 1d vector of length 2m+n
OU <- function(theta){
    a <- theta[1];
    b <- theta[2];
    sigma <- theta[3];
    u1 <- uyt[2:nrow(uyt),];
    u2 <- uyt[1:(nrow(uyt)-1),];
    # log likelihood expression
    L <- -log(sigma*sqrt(2*pi))-((u1 - a*b + u2*(a-1))**2)/(2*
        sigma**2);
    ll <- sum(L);
    return(-ll)
}
out <- optim(theta <- 10**(-3)*runif(3,0,1),OU);

a <- out$par[1];
b <- out$par[2];
sigma <- out$par[3];

# On projette de 2001[U+FFFD]2100 (validation sur les annees 2001
  [U+FFFD]2015)
uytproj <- uyt
uytfin <- matrix(0,nrow(uytproj),ncol(uytproj));

itermax <- 1500;
for (i in 1:itermax){
    for (j in 2:(nrow(uytproj))){
```

```
        uytproj[j,] <- uytproj[j-1,]*exp(-a)+b*(1-exp(-a))+
           sigma*sqrt((1-exp(-2*a))/(2*a))*rnorm(ncol(uytproj)
           ,0,1)
    }
    uytfin <- (uytfin*(i-1)+uytproj)/i;
}
qyt <- -log(1-uytfin);
library(R.matlab)
save(file="uytfin.Rdata",uytfin)
writeMat("uytfin.mat",uytfindata = uytfin)
save(file="qyt.Rdata",qyt)
writeMat("qyt.mat",qytdata = qyt)

xplot <- 0:(nrow(qyt)-1);
plot(xplot,qyt[,1],col = "blue",type = "n",xlab ="t",ylab = "q_
   (x+t)");
for(i in 1:ncol(qyt)){
    lines(xplot,qyt[,i],col=i,type="p")
}

# 2. Feller model
###########################
# CIR model calibrated by MLE: du = audt + sigma sqrt(u) dw

uxt <- -log(1-qxt[1:(nrow(qxt)-1),]);

# log likelihood function to optimise in theta
CIR <- function(theta,x){
    a <- theta[1];
    sigma <- theta[2];
    u1 <- uxt[x,2:ncol(uxt)];
    u2 <- uxt[x,1:(ncol(uxt)-1)];
    # log likelihood expression
    L <- -log(sigma*sqrt(2*pi*u2))-((u1 - u2*(a+1))**2)/(2*
       sigma**2*u2);
    ll <- sum(L);
    return(-ll)

}

out <- matrix(0,nrow(uxt),2);
```

```
for (i in 1:nrow(uxt)){
    out[i,] <- optim(par = c(1,1),fn = CIR,x = i)$par;
}

library(R.matlab)
save(file="paramFeller.Rdata",out);
writeMat("paramFeller.mat",paramFellerdata = out);

# 3. Wang transform
################################

i = c(0.01, 0.0325, 0.05, 0.08, 0.1, 0.15);
v = 1/(1+i);
x0 = 50;
xmax = 110;
t0 = 2015;
tmax = 110;
kmax = xmax-x0;
k = c(0:(kmax-1));
sm = 0.999441703848;
gm = 0.999733441115;
cm = 1.101077536030;
sf = 0.999669730966;
gf = 0.999951440172;
cf = 1.116792453830;
pmarket_m = sm*gm**(cm**(x0+k-5)*(cm-1));
pmarket_f = sf*gf**(cf**(x0+k-5)*(cf-1));

b1 = readMat("Lee Carter/betax1_com.mat")$beta.x1;
b2 = readMat("Lee Carter/betax2_com.mat")$beta.x2;
kt = readMat("Lee Carter/kt_com.mat")$kt;
kt0 = kappa[length(kt)];
alpha = b1[x0:length(b1)];
beta = b2[x0:length(b2)];
n <- length(kt);
ktp <- kt[2:n];
ktm <- kt[1:(n-1)]
# log likelihood function to optimise in theta.
CoM <- function(theta){
    delta <- theta[1];
    sigma <- theta[2];
```

```
    L <- dnorm(ktp,ktm+delta,sigma,log=TRUE);
    ll <- sum(L);
    return(-ll)
}
out <- optim(theta <- c(1,1),CoM);
delta = out$par[1];
p_m = pmarket_m;
p_f = pmarket_f;
pref = pmarket_m;
pref[1] = exp(-exp(alpha[1]+beta[1]*kt0));
for (i in 2:kmax){
    p_m[i] = prod(pmarket_m[1:i]);
    p_f[i] = prod(pmarket_f[1:i]);
    pref[i] = exp(-sum(exp(alpha[1:i]+beta[1:i]*(kt0+c(0:(i-1))
       *delta))));
}
ptot = (p_m+p_f)/2;
qref = 1-pref;

a65_m = rep(0,length(v));
a65_f = rep(0,length(v));
a65_tot = rep(0,length(v));
a65_ref = rep(0,length(v));
lambda = rep(0,length(v));

fun <- function(x,v,a65_tot){
    return(sum(v**(k+1)*(1-pnorm(qnorm(qref,0,1)+x)))-a65_tot);
}

qwang = matrix(1,length(qref),length(lambda));

for(i in 1:length(v)){
    a65_m[i] = sum(v[i]**(k+1)*p_m);
    a65_f[i] = sum(v[i]**(k+1)*p_f);
    a65_tot[i] = sum(v[i]**(k+1)*ptot);
    a65_ref[i] = sum(v[i]**(k+1)*pref);
    lambda[i] = uniroot(function(x) fun(x,v[i],a65_tot[i]),c
       (-1,1))$root;
    qwang[,i] = pnorm(qnorm(qref)+lambda[i]);
}

```

```
finalref = c(qref,1);
finalwang = rbind(qwang,rep(1,length(v)));

plot(x,finalref,type="l",col="red")
for (i in 1:length(v)){
lines(x,finalwang[,i],type='l',col=1:length(v));
}

library(R.matlab)
save(file="qwang.Rdata",finalwang);
writeMat("qwang.mat",qwangdata = finalwang);

library(R.matlab)
save(file="qref.Rdata",finalref);
writeMat("qref.mat",qrefdata=finalref);
```

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve **www.uclouvain.be/epl**