Decomposition of risk based on the Martingale Representation Theorem for annuity and insurance portfolios

Edwin H. Ng

A Thesis
In
The Department
of
Mathematics and Statistics

Presented in Partial Fulfilment of the Requirements
for the degree of Master of Science (Mathematics) at
Concordia University
Montreal, Quebec, Canada

July, 2019

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Edwin H. Ng**

Entitled: **Decomposition of risk based on the Martingale Representation Theorem for annuity and insurance portfolios**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Examiner

Dr. Mélina Mailhot

_____ Examiner

Dr. Xiaowen Zhou

_____ Thesis Supervisor

Dr. Patrice Gaillardetz

Approved by _____

Chair of Department or Graduate Program Director

_____

Dean of Faculty

Date _____

ABSTRACT

Decomposition of risk based on the Martingale Representation Theorem for annuity and insurance portfolios

Edwin H. Ng

Numerous methods have been proposed throughout the literature for decomposing liabilities into risk factors. Such analysis is of great importance because it allows for explaining the impact of each source of risk in relation to the total risk, and thus it allows actuaries to have a certain degree of control over uncertainties. In an insurance context, such sources usually consist of the mortality risk, represented in this paper by the systematic and by the unsystematic mortality risk, and of the investment risk. The objective of this thesis is to consider the Martingale Representation Theorem (MRT) introduced by Schilling et al., (2015) for such risk decomposition, because this method allows for a detailed analysis of the influence of each source of risk. The proposed dynamic models used in this thesis are the Lee-Carter model for the mortality rates and, the arbitrage-free Nelson-Siegel (AFNS) models for the interest rates. These models are necessary in providing accuracy by improving the overall predictive performance. Once, the risk decomposition has been achieved, quantifying the relative importance of each risk factor under different risk measurements is then proceeded. The numerical results are based on annuities and insurances portfolios. It is found that for extended coverage periods, investment risk represents most of the risk while for shorter terms, the unsystematic mortality risk takes larger importance. It is also found that the systematic mortality risk is almost negligible.

Acknowledgements

First and foremost, I would like to express my profound gratitude to my family for their unceasing support and encouragement during my entire life.

My heartfelt gratitude goes to my very supportive supervisor, Professor Patrice Gaillardetz. His expertise, patience, and insightful guidance made this work possible and added considerably to my graduate experience. Besides my supervisor, I would like to thank the rest of my thesis committee; Professor Mailhot and Professor Zhou, for taking the time to read and provide the necessary valuable feedback. I would like to also take this opportunity to express a deep gratitude to all of the Mathematics and Statistics department faculty members at Concordia University for their help and advisement.

Last but not the least; my sincere thanks goes to all my friends who bared a listening ear and provided continuous support throughout the entire process.

Contents

List of Figures

List of Tables

Introduction

Risk is defined to be the financial consequences of future uncertain events about the deviation from expected outcomes. From an actuarial context such as insurances and annuities, risks originate mainly from uncertainties arising out of mortality and investment factors. The question is then the following: how significant is each factor? Answering this question allows actuaries to identify and to manage the various sources of risks in relation to the total risk and to quantify their relative contributions respectively. In doing so, actuaries are then able to reduce the overall uncertainties. This could be done for example in the following manners. One could proceed through classical pooling techniques, through risk transference techniques such as reinsurance, or through financial risk management techniques such as hedging. Thus, numerous authors in the literature have taken an interest in the problem of risk decomposition.

One widely used method of risk decomposing is the so-called *variance decomposition* as introduced by Bühlmann (1995), where he introduced the use of conditional expectation in identifying two risk components. This method is further developed by Parker (1997) where he interpreted the *variance decomposition* into a stochastic insurance environment in order to separate the riskiness of a portfolio of different insurance policies into the *mortality risk* and the *investment risk*. Within the same *variance decomposition* framework, Christiansen and Helwich (2008) later decompose the *mortality risk* itself into two other components: the *systematic mortality risk* and the *unsystematic mortality risk*. The risk of the portfolio loss is then made up of the sum of three addends that correspond exactly to the *unsystematic mortality risk*, the *systematic mortality risk*, and the *investment risk*.

The two latter risks are said to be systematic because they are non-diversifiable since they are associated with changes in the underlying mortality and interest intensity. On the other hand, the *unsystematic mortality risk* is diversifiable as it represents the risk associated with the randomness of death occurrences within a portfolio with known mortality intensity. Under the concept of risk pooling, one should then expect this risk to disappear as the number of policies in a given portfolio increases. Although the *variance decomposition* is successful in identifying the three main risk components in an insurance context, it nevertheless fails at satisfying the order invariance criteria from a meaningful risk decomposition. That is, different values attributed to the risk components are possible depending on the order choice for the conditional expectations. This can be problematic since one possible choice among others might be less interpretable.

The objective of this thesis is to further explore the solution to this problem through the *Martingale Representation Theorem Decomposition* as introduced by Schilling et al. (2015). The Martingale Representation Theorem states that if a random variable is generated by a state space of stochastic processes, then it could be represented as a stochastic integral of predictable processes with respect to those state processes. One such financial application is illustrated by Harrison and Pliska (1981) in a continuous trading context where a financial product's future price is replicated through self-trading strategies evolving over time and, measured in relation to the sources of fluctuation of the price. The self-trading strategies would correspond to the predictable processes from the Martingale Representation Theorem (MRT) and the sources of fluctuation to the state processes. In this thesis, the idea is then to reinterpret these strategies as the contributions of each source of risk to the total risk of the insurer. In fact, the *Martingale Representation Theorem Decomposition* is successful in identifying the sources of risk while still fully satisfying the criteria for a meaningful decomposition. Moreover, it is able to establish a diversifiable unsystematic risk that disappears as the portfolio size increases.

The thesis is structured as followed: Chapter 1 establishes the annuity and the insur-

ance portfolios as the modelling framework for this thesis, along with the basic assumptions and the working stochastic environment needed for any rigorous tasks. Chapter 2 gives an overview of the concept of risk decomposition as well as the criteria for a meaningful decomposition. The MRT decomposition is then introduced and derived in details.

Chapter 3 introduces the stochastic mortality model to be used to assess the mortality risks. The Lee-Carter model serves well as it shows many desired properties in producing a reasonable forecast. Those include exhibiting a high degree of intertemporal correlation across the ages and imposing an age-specific rate trend in all individuals. These properties are of particular interest as they explain the unanticipated mortality decline in relation to the population ageing. In a life insurance context, this is essential for the costs associated with greater longevity. The Lee-Carter model will then be fitted to Canadian mortality data ranging from 1921 to 2016.

Chapter 4 introduces the dynamic interest rate to be used to assess the investment risk. In order to have better modern interpretations of the uncertainties associated with the *investment risk*, a three-factor interest rate model is necessary in explaining the intertemporal variations of the term structure, where the three factors correspond to the classical yield factors of the general level of interest rates, the slope of the yield curve, and the curvature. The model used in this paper is the affine arbitrage-free class of Nelson-Siegel term structure model as derived by Christensen et al. (2011). This model proves to be rigorous because it is based on affine models where yields are theoretically easy to derive. But it is also based on the Nelson-Siegel model so that the model fits empirically well to time series of yields. The model will then be fitted to daily Canadian yield curves ranging from 1991-01-02 until 2017-10-31.

Chapter 5 applies MRT decomposition for the annuity and insurance portfolios introduced in Chapter 1 into the stochastic environment established by the mortality and investment models discussed in Chapters 3 and 4. Subsequent results will then be dis-

cussed in details.

Chapter 1

Modelling Framework

The financial and mortality frameworks are modelled using a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$, and T is a finite time horizon. T is interpreted as the time for which the insurance contract is covered. $\mathcal{F}_t$ represents the information available up to time t and $\mathcal{F} = \mathcal{F}_T$ is the complete information at final time T. It is assumed in this thesis that the risks in an insurance portfolio come from three sources: the *investment risk*, the *systematic mortality risk*, and the *unsystematic mortality risk*. Mathematically, this is expressed by setting the filtration $\mathbb{F}$ as being generated by both sub-filtration $\mathbb{G}$ and $\mathbb{I}$, that is $\mathbb{F} = \mathbb{G} \vee \mathbb{I}$. $\mathbb{G}$ represents the filtration associated with any rate-related processes that are said to be systematic. Here, $\mathbb{G}$ represents the filtration associated with the *investment risk* and the *systematic mortality risk*. The investment risk arises from the evolution of the short-rate $r(t)$, where the discount function of one unit payable at time t is

$$B(t) = e^{-\int_0^t r(s)ds} \, . \tag{1.0.1}$$

The financial market is also assumed to be arbitrage-free, with the existence of a risk-neutral probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$. The systematic mortality risk is associated with the mortality rate $m(t)$. Assuming a portfolio for P_0 insurance contracts, let the curtate lifetime of the insured i be denoted by the random variable K_i $(i = 1, 2, ..., P_0)$.

Then, the probability of survival past time t at age x is denoted by

$$ {}_tp_{x_i} = P(K_i \geq t \mid \mathcal{G}_t) = e^{-\int_0^t m(s)ds} , \tag{1.0.2} $$

where x_i is the age of the ith insured.
The sub-filtration $\mathbb{G}$, associated with changes in the systematic risks of the short-rate and of the mortality rate, is driven by state processes $X(t) = (X_1(t), ..., X_k(t))_{0 \leq t \leq T}$. As discussed later, they are assumed to be Itô processes.

Associated with the sub-filtration $\mathbb{I}$ is the *unsystematic mortality rate*, generated by the random occurrence of the number of deaths, and it is linked with the death indicator process of each insured. Let $\mathbb{I}^i = (\mathcal{I}^i{}_t)_{0 \leq t \leq T}$ denotes the filtration generated by the indicator process of the ith insured, $(\mathbb{1}_{\{K_i \leq t\}})_{0 \leq t \leq T}$, and $\mathbb{I} = \vee_{i=1}^{P_0} \mathbb{I}^i$, where P_0 denotes the initial population at time 0 for an insured cohort of age x. The death process $\mathbb{1}_{\{K_i \leq t\}}$ is the process taking the value one when the ith insured dies before time t. Let $N(t) = \sum_i^m \mathbb{1}_{\{K_i \leq t\}}$ denotes the counting process for the number of deaths. Since $N(t)$ is the sum of Bernoulli random variables, it is therefore a random binomial variable, i.e. $N(t) \sim$ *binomial*$(P_0 \ , \ {}_tq_x)$, where $N(0) = P_0$ and ${}_tq_x = 1 - {}_tp_x = P(K_i < t \mid \mathcal{G}_t)$ is the probability of death before year t for an insured at age x.

1.1 Description of the portfolios

In the following, two types of insurance portfolios are introduced. Let L_t be a random variable denoting the present value at time t of the prospective loss of the portfolios. Notice that L_0 represents the present value of the loss at time 0.

The T-year temporary **annuity policy** consists of periodical payments made to the policyholders in the condition of their survival for a maximum of T payments. That is,

$$ L_0 = \sum_{j=1}^{T} (P_0 - N(t_j)) e^{-\int_0^{t_j} r(s)ds} , \tag{1.1.1} $$

where t_j $(j = 1, ..., T)$ is the periodical time in which payments are made and it is assumed here that the t_j's are annual.

The prospective loss random variable underlying a T-year term life insurance policies with the death benefit payable at the end of the period of death is

$$L_0 = \sum_{j=1}^{T} (N(t_j) - N(t_{j-1}))e^{-\int_0^{t_j} r(s)ds} \ . \tag{1.1.2}$$

This could also be expressed as the difference of two annuities,

$$L_0 = \sum_{j=1}^{T} (P_0 - N(t_{j-1}))e^{-\int_0^{t_j} r(s)ds} - \sum_{j=1}^{T} (P_0 - N(t_j))e^{-\int_0^{t_j} r(s)ds} \ . \tag{1.1.3}$$

Chapter 2

Martingale Representation Theorem Decomposition

Quantifying the contributions of each source of risks to the total risk has been an important matter in the insurance context. This allows for a deeper understanding of the uncertainties at hand and allows actuaries to manage and to have a certain degree of control on them. It also allows for identifying and for pinpointing the most critical source of risk. As stated previously, the typical sources of risk in an insurance context are the *investment risk*, the *systematic mortality risk*, and the *unsystematic mortality risk*. Those represent the risks associated with changes in the interest, the risk associated with changes in the mortality intensity, and the risk associated with the actual observed number of deaths in a given portfolio, respectively. A risk decomposition method is a quantifying method that assigns each source of risk a risk factor that represents their respective contribution to the total risk. In order to assert a reasonable decomposition methodology, let first establish a number of properties defining a risk decomposition method as being *meaningful*.

2.1 Meaningful risk decompositions

Christiansen et al. (2015) proposed a number of properties that define a risk decomposition method as being *meaningful*. Assuming that the *total risk* is represented by

the loss random variable L with $E[L] = 0$, where $E[L]$ is the expected value of the loss variable, and assuming k different sources of risk for L, let $X = (X(t))_i$, $i = 1, 2...k$, be the corresponding k sources of risk such that L is $\sigma(X)$-measurable. More generally, the total risk is represented by $L - E[L]$. The followings are the proposed properties for a *meaningful risk decomposition*. Let R_i , $i = 1, 2...k$, denote the corresponding risk factor associated to each source of risk X_i.

1. **Randomness**: The loss random variable L can be decomposed into k random risk factors R_i, $i = 1, 2, ..., k$, corresponding exactly to the k sources of risk $X(t)$. Denote $(X_1, ..., X_k) \longleftrightarrow (R_1, ..., R_k)$ to be the relation between the k sources of risk $(X_1, ..., X_k)$ and the k risk factors $(R_1, ..., R_k)$.

2. **Attribution**: The risk factor R_i corresponds exactly to the risk X_i. Letting L to be $\sigma(X_i)$-measurable only and letting X_i to be independent of $(X_j)_{j\neq i}$, then all other risk factors should be zero, i.e. $R_j = 0$ for all $i \neq j$.

3. **Uniqueness**: The risk decomposition is unique in the sense that if $(X_1, ..., X_k) \longleftrightarrow (R_1, ..., R_k)$ and if $(X_1, ..., X_k) \longleftrightarrow (\tilde{R}_1, \ldots, \tilde{R}_k)$, then $R_i = \tilde{R}_i$ for $i = 1, ..., k$.

4. **Order invariance**: The decomposition is invariant to the order of the risks $1, 2, \ldots, k$. That is, let permutation $\pi : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, k\}$ and if $(X_1, ..., X_k) \longleftrightarrow (\tilde{R}_1, \ldots, \tilde{R}_k)$, then $(X_{\pi(1)}, \ldots, X_{\pi(k)}) \longleftrightarrow (\tilde{R}_{\pi(1)}, \ldots, \tilde{R}_{\pi(k)})$ is equivalent.

5. **Scale invariance**: The decomposition should be such that the risk factors are invariant to any change of scales from the sources of risk. Assuming $(X_1, ..., X_k) \longleftrightarrow (R_1, ..., R_k)$ and let $f_i : \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth, invertible function such that $\tilde{X}_i = f_i(X_i)$. If $(\tilde{X}_1, ..., \tilde{X}_k) \longleftrightarrow (\tilde{R}_1, ..., \tilde{R}_k)$, then $R_i = \tilde{R}_i$, for all $i = 1, ..., k$.

6. **Aggregation**: The risk decomposition should be such that the decomposed risk factors R_i aggregate to the total risk L. In other words, let L be $\sigma(X)$-measurable and $(X_1, ..., X_k) \longleftrightarrow (R_1, ..., R_k)$, then there exists a function $A : \mathbb{R}^k \longrightarrow \mathbb{R}$ such that $L = A(R_1, ..., R_k)$.

   A special case is letting A being the sum function such that $L = \sum_{i=1}^{k} R_i$.

2.2 Martingale Representation Theorem Decomposition

The purpose of this section is to formulate the Martingale Representation Theorem Decomposition which constitutes the basis of this thesis. The concept of the Martingale Representation Theorem (MRT) is first introduced. The total risk, as represented by $L - E[L]$, is assumed to be a martingale and this assumption is critical in imposing the insurance policies to be fair.

2.2.1 Martingale Representation Theorem

Let the state processes adapted to the sub-filtration $\mathbb{G}$ be $X(t) = (X_1(t), ..., X_k(t))$, $t \geq 0$, and let $X(t)$ be a k-dimensional Itô process such that

$$dX(t) = \theta(t)dt + \sigma(t)dW(t) \ , \tag{2.2.1}$$

where $X(0) = x_0$ is a deterministic initial value and $W(t)$ is a k-dimensional Brownian motion. $\theta(t)$ is a k-dimensional drift vector and $\sigma(t)$ is a $k \times k$ dimensional diffusion matrix. The X_i's are assumed to be independent with each other.

The corresponding compensated process would then be

$$dM(t) = \sigma(t)dW(t) \ , \tag{2.2.2}$$

the martingale part of $X(t)$. More precisely, from the Doob-Meyer Decomposition Theorem (See Protter, 2005, p. 130, Theorem 34), an adapted, integrable semi-martingale process X_t such as the above can be expressed as the sum of a unique martingale M and of a unique predictable process A, i.e. $X(t) = X(0) + A(t) + M(t)$. If $A(t)$ is increasing, then $X(t)$ is a submartingale and if $A(t)$ is decreasing, then $X(t)$ is a supermartingale. From (2.2.1), $dM(t) = (dM_1(t)...dM_k(t))$ is given by (2.2.2) and $dA(t) = \theta(t)dt$ is the unique predictable process. Moreover, $A(t)$ is said to be a *compensator* and $M(t)$, the *compensated process*.

As for the sub-filtration $\mathbb{I}$, it is assumed to be generated by the binomial process N. Again from Doob-Meyer Decomposition Theorem, since N is a submartingale, it can also be expressed as the sum of a predictable process and of a martingale.

Proposition 1. Let the counting process $N(t)$ with left-hand limits. Then, the unique compensator of $N(t)$ is given by $\int_0^t (P_0 - N(s-))m(s)ds$, and its compensated process follows as

$$M^N(t) = N(t) - \int_0^t (P_0 - N(s-))m(s)ds \ , \qquad 0 \leq s \leq t \leq T \ . \tag{2.2.3}$$

Proof. The *proof* can be found in Bielecki and Rutkowski, 2004, p.153, Proposition 5.1.3. □

The following introduces the Martingale Representation Theorem under a Lévy setting as proposed by Jeanblanc et al. (2009, p.621, Proposition 11.2.8.1).

Theorem 1. Let X and N be Lévy processes as defined in the above and let $\mathbb{F}$ be their natural filtration. Let $\hat{L} = L_0 - E[L_0]$ be a $\mathbb{F}$-local martingale. Then there exist unique $\mathbb{F}$-predictable processes $(\tilde{\psi}_1, ..., \tilde{\psi}_k, \psi^N)$ such that

$$\begin{aligned} \hat{L} = L_0 - E[L_0] &= \sum_{i=1}^{k} \int_0^T \tilde{\psi}_i(t) dW_i(t) + \int_0^T \psi^N(t) dM^N(t) \\ &= \sum_{i=1}^{k} \int_0^T \tilde{\psi}_i(t)\sigma_i^{-1}(t)\sigma_i(t) dW_i(t) + \int_0^T \psi^N(t) dM^N(t) \ , \end{aligned} \tag{2.2.4}$$

where L_0 is the discounted at time zero of the insurer's future net liability and $(L_0 - E[L_0])$ is the risk associated with the portfolio. $\left(\int_0^T \tilde{\psi}_i(t) dW_i(t)\right)_{i=1}^{k}$ are the *risk factors* associated with the sub-filtration $\mathbb{G}$ and adapted to the source of risks X. Finally, $\int_0^T \psi^N(t) dM^N(t)$ is the *risk factor* associated with the sub-filtration $\mathbb{I}$ and adapted to the risk arising from the random occurrence of death N.

Imposing $\psi_i(t) = \tilde{\psi}_i(t)\sigma_i^{-1}(t)$ and $dM_i(t) = \sigma_i(t)dW_i(t)$,

$$L = L_0 - E[L_0] = \sum_{i=1}^{k} \int_0^T \psi_i(t)dM_i(t) + \int_0^T \psi^N(t)dM^N(t) \ . \tag{2.2.5}$$

Proof. See Jeanblanc et al. (2009). □

2.2.2 The Predictable Processes ψ

Although the MRT states that such a decomposition is theoretically possible, the values and the form of the predictable processes ψ are not given and are yet to be found. In their paper, Schilling et al. (2015) derived the unique predictable processes ψ and their steps are followed in this thesis.

Proposition 2. Let X be a k-dimensional diffusion process and let fixed time $U \in [0, T]$. Let $L_0 = e^{-\int_0^U r(s,X(s))ds}$ with measurable function $r : [0, U] \times \mathbb{R}^k \longrightarrow \mathbb{R}$, and define $f(t, \omega) = E[e^{-\int_t^U r(s,X(s))ds} \mid X(t) = \omega]$. Then the unique integrands of the MRT decomposition for $L_0 - E[L_0]$ are given by:

$$\psi_i(t) = \mathbb{1}_{[0,U]}(t)e^{-\int_0^t r(s,X(s))ds} \frac{\partial f}{\partial X_i}(t, X(t)) \ , i = 1, ..., k \ , \tag{2.2.6}$$

and

$$\psi^N(t) = 0 \ . \tag{2.2.7}$$

Note that the measurable function r is interpreted as the short-rate.

Proof. The proof follows Itô's lemma applied to a k-dimensional semimartingales. First, since $X(t)$ is a Markov process, then

$$f(t, X(t)) = E[e^{-\int_t^U r(s,X(s))ds}|X(t)] = E[e^{-\int_t^U r(s,X(s))ds}|\mathcal{G}_t] \ .$$

Moreover, define

$$\tilde{f}(t,X(t)) = E[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_t] = e^{-\int_0^t r(s,X(s))ds} E[e^{-\int_t^U r(s,X(s))ds} \mid \mathcal{G}_t]$$
$$= e^{-\int_0^t r(s,X(s))ds} f(t,X(t)) \ . \tag{2.2.8}$$

Then, L_0 and $E[L_0]$ can be expressed as $L_0 = e^{-\int_0^U r(s,X(s))ds} = \tilde{f}(U,X(U))$ and $E[L_0] = \tilde{f}(0,X(0))$.

Applying Itô's Lemma on $\tilde{f}$ leads to

$$\tilde{f}(t,X(t)) - \tilde{f}(0,X(t)) = \sum_{i=1}^{k} \int_0^t \frac{\partial \tilde{f}}{\partial X_i}(s,X(s))dM_i(s) + \int_0^t a(s)ds \ ,$$

where $a(t)$, $0 \leq t \leq T$, is short-hand for all ds-quantities. In fact, those quantities must be null. The above equation is equivalent to:

$$E[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_t] - E[e^{-\int_0^U r(s,X(s))ds}] = \sum_{i=1}^{k} \int_0^t \frac{\partial \tilde{f}}{\partial X_i}(s,X(s))dM_i(s) + \int_0^t a(s)ds \ , \tag{2.2.9}$$

and using conditional expectations, the left-hand side can be shown to be a martingale. From the right-hand side, from the martingale preservation property, the stochastic integrals with respect to the M_i are martingales as well. But by the Doob-Meyer decomposition, $a(s)$ must be null in order for the whole left-hand side to be indeed a martingale.

Finally, setting $t = U$ in (2.2.9) and recalling that $\tilde{f}(t,X(t)) = e^{-\int_0^t r(s,X(s))ds} f(t,X(t))$,

$$L_0 - E[L_0] = \sum_{i=1}^{k} \int_0^U e^{-\int_0^s r(s,X(s))ds} \frac{\partial f}{\partial X_i}(s,X(s))dM_i(s) \ .$$

□

The following Lemma gives the basis in deriving the MRT decomposition needed for the insurance portfolios.

Lemma 1. Let L_0 be a loss random variable of the form $L_0 = (P_0 - N(t_j))F$ for $0 \leq t_j \leq T$, where F is $\mathcal{G}_T$-measurable and integrable and $N(t_j)$ is $\mathcal{I}_{t_j}$-measurable. Then there

exist $\mathbb{G}$-predictable processes $\varphi_1, \ldots, \varphi_k$ such that

$$E\left(e^{-\int_0^{t_j} m(s,X(s))ds} F \mid \mathcal{G}_t\right) = E\left(e^{-\int_0^{t_j} m(s,X(s))ds} F\right) + \sum_{i=1}^{k} \int_0^t \varphi_i(s) dW_i(s) \ , \tag{2.2.10}$$

is satisfied, for $t \leq T$.

Then the martingale representation of L is

$$\begin{aligned} L_0 = E\left(L_0\right) &+ \sum_{i=1}^{k} \int_0^{t_j} (P_0 - N(s-)) e^{\int_0^s m(u,X(u))du} \varphi(s) dW_i(s) \\ &+ \sum_{i=1}^{k} \int_{t_j}^{T} (P_0 - N(t_j)) e^{\int_0^{t_j} m(u,X(u))du} \varphi(s) dW_i(s) \\ &- \int_0^{t_j} E\left(e^{\int_s^{t_j} m(u,X(u))du} F | \mathcal{G}_s\right) dM^N(s) \ . \end{aligned}$$

Proof. A proof can be found in Christensen et al. (2015). □

Now that the general form of the MRT decomposition has been defined, the following proposition gives the predictable processes ψ such that (2.2.10) is satisfied.

Proposition 3. Let X be the above k-dimensional diffusion process. Let $L_0 = (P_0 - N(t_j))e^{-\int_0^U r(s,X(s))ds}$ for measurable function $r : [0,U] \times \mathbb{R}^k \longrightarrow \mathbb{R}$, and $0 \leq t_j \leq T$, $0 \leq U \leq T$.

1. Suppose $U > t_j$. In this thesis, U will later be assumed to be $U = t_{j+1}$. Define the following:

   - $f^{A_1}(t,\omega) = E[e^{-\int_t^{t_j} m(s,X(s))ds} e^{-\int_t^U r(s,X(s))ds} \mid X(t) = \omega]$
   - $f^{B_1}(t,\omega) = E[e^{-\int_t^U r(s,X(s))ds} \mid X(t) = \omega]$.

   Then the unique integrands of the MRT decomposition for $L_0 - E[L_0]$ are given by:

$$\begin{aligned} \psi_i(t) = &\mathbb{1}_{[0,t_j]}(t) \ (P_0 - N(t-)) e^{-\int_0^t r(s,X(s))ds} \frac{\partial f^{A_1}}{\partial X_i}(t, X(t)) \\ &+ \mathbb{1}_{(t_j,U]}(t) \ (P_0 - N(t_j)) e^{-\int_0^t r(s,X(s))ds} \frac{\partial f^{B_1}}{\partial X_i}(t, X(t)) \ , \\ &i = 1, ..., k \ , \end{aligned} \tag{2.2.11}$$

$$\psi^N(t) = -\mathbb{1}_{[0,t_j]}(t)\ e^{-\int_0^t r(s,X(s))ds} f^{A_1}(t, X(t))\ . \tag{2.2.12}$$

2. Suppose $U \leq t_j$. In this thesis, U will later be assumed to be $U = t_j$. Define the following:

   - $f^{A_2}(t,\omega) = E[e^{-\int_t^{t_j} m(s,X(s))ds} e^{-\int_t^U r(s,X(s))ds} \mid X(t) = \omega]$
   - $f^{B_2}(t,\omega) = E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid X(t) = \omega]$.

   Then the unique integrands of the MRT decomposition for $L_0 - E[L_0]$ are given by:

$$\begin{aligned}\psi_i(t) = \mathbb{1}_{[0,U]}(t)\ (P_0 - N(t-))e^{-\int_0^t r(s,X(s))ds}\frac{\partial f^{A_2}}{\partial X_i}(t, X(t)) \\ + \mathbb{1}_{(U,t_j]}(t)\ e^{-\int_0^U r(s,X(s))ds}\frac{\partial f^{B_2}}{\partial X_i}(t, X(t))\ , \\ i = 1, ..., k\ ,\end{aligned} \tag{2.2.13}$$

$$\psi^N(t) = -\mathbb{1}_{[0,U]}(t)\ e^{-\int_0^t r(s,X(s))ds} f^{A_2}(t, X(t)) - \mathbb{1}_{(U,t_j]}(t)\ e^{-\int_0^U r(s,X(s))ds} f^{B_2}(t, X(t))\ . \tag{2.2.14}$$

Proof. 1. Consider $U > t_j$ and $L_0 = (P_0 - N(t_j))e^{-\int_0^U r(s,X(s))ds}$. The idea behind the proof is to apply Lemma 1 using $F = e^{-\int_0^U r(s,X(s))ds}$.

According to Lemma 1 and setting $t = U$ in (2.2.10), there exist processes φ_i such that

$$\begin{aligned}e^{-\int_0^{t_j} m(s,X(s))ds} e^{-\int_0^U r(s,X(s))ds} - E\left[e^{-\int_0^{t_j} m(s,X(s))ds} e^{-\int_0^U r(s,X(s))ds}\right] \\ = \sum_{i=1}^k \int_0^U \varphi_i(s) dW_i(s)\end{aligned}$$

is satisfied.

The left-hand side can be rewritten as

$$e^{-\int_0^{t_j} m(s,X(s))ds} e^{-\int_0^U r(s,X(s))ds} - E\left[e^{-\int_0^{t_j} m(s,X(s))ds} e^{-\int_0^U r(s,X(s))ds}\right]$$

$$= \left(e^{-\int_0^{t_j} m(s,X(s))ds} E\left[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_{t_j}\right] - E\left[e^{-\int_0^{t_j} m(s,X(s))ds} e^{-\int_0^U r(s,X(s))ds}\right]\right)$$

$$+ \left(e^{-\int_0^{t_j} m(s,X(s))ds}\left[e^{-\int_0^U r(s,X(s))ds} - E[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_{t_j}]\right]\right) \ , \tag{2.2.15}$$

such that it can be derived into two parts respectively in order to find the φ_i. For the first part, let apply Itô's formula to f^{A_1}. Define

$$\tilde{f}^{A_1}(t, X(t)) = e^{-\int_0^t m(s,X(s))ds} e^{-\int_0^t r(s,X(s))ds} f^{A_1}(t, X(t)) \ ,$$

with $0 \leq t \leq t_j < U$. By Itô's lemma with $\tilde{f}^{A_1}$,

$$\begin{aligned}
&\tilde{f}^{A_1}(t, X(t)) - \tilde{f}^{A_1}(0, X(0)) \\
&\qquad = E\left(e^{-\int_0^{t_j} m(s,X(s))ds} E\left[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_{t_j}\right] \mid \mathcal{G}_t\right) \\
&\qquad - E\left(e^{-\int_0^{t_j} m(s,X(s))ds} E\left[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_{t_j}\right]\right) \\
&\qquad = \sum_{i=1}^k \int_0^t e^{-\int_0^s m(u,X(u))du} e^{-\int_0^s r(u,X(u))du} \frac{\partial f^{A_1}}{\partial X_i}(s, X(s)) dM_i(s) + \int_0^t a(s)ds \ ,
\end{aligned}$$

where $a(s)$ is short-hand for all ds-quantities.

Similarly as before, the left-hand side of the equation is a martingale so that the ds-term has to vanish. Furthermore, letting $t = t_j$,

$$e^{-\int_0^{t_j} m(s,X(s))ds} E\left[e^{-\int_0^U r(s,X(s))ds} \mid \mathcal{G}_{t_j}\right] - E\left[e^{-\int_0^{t_j} m(s,X(s))ds} e^{-\int_0^U r(s,X(s))ds}\right]$$

$$= \sum_{i=1}^k \int_0^{t_j} e^{-\int_0^s m(u,X(u))du} e^{-\int_0^s r(u,X(u))du} \frac{\partial f^{A_1}}{\partial x_i}(s, X(s)) dM_i(s) \ . \tag{2.2.16}$$

For the second part of (2.2.15), applying Proposition 2 on $L^*_{t_j} = E\,[\, e^{-\int_{t_j}^{U} r(s,X(s))ds}\,]$ instead, with function f^B, and considering $t_j \le t \le U$,

$$\left(e^{-\int_0^{t_j} m(s,X(s))ds}\left[e^{-\int_0^{U} r(s,X(s))ds} - E[e^{-\int_0^{U} r(s,X(s))ds} \mid \mathcal{G}_{t_j}]\right]\right)$$

$$= \left(e^{-\int_0^{t_j} m(s,X(s))ds}e^{-\int_0^{t_j} r(s,X(s))ds}\left[e^{-\int_{t_j}^{U} r(s,X(s))ds} - E[e^{-\int_{t_j}^{U} r(s,X(s))ds} \mid \mathcal{G}_{t_j}]\right]\right)$$

$$= e^{-\int_0^{t_j} m(s,X(s))ds}e^{-\int_0^{t_j} r(s,X(s))ds}\sum_{i=1}^{k}\int_{t_j}^{U} e^{-\int_{t_j}^{s} r(u,X(u))du}\frac{\partial f^{B_1}}{\partial X_i}(s,X(s))dM_i(s)$$

$$= e^{-\int_0^{t_j} m(s,X(s))ds}\sum_{i=1}^{k}\int_{t_j}^{U} e^{-\int_0^{s} r(u,X(u))du}\frac{\partial f^{B_1}}{\partial X_i}(s,X(s))dM_i(s)$$

$$= \sum_{i=1}^{k}\int_{t_j}^{U} e^{-\int_0^{t_j} m(s,X(s))ds}e^{-\int_0^{s} r(u,X(u))du}\frac{\partial f^{B_1}}{\partial X_i}(s,X(s))dM_i(s)\ . \tag{2.2.17}$$

Thus, adding (2.2.16) and (2.2.17) leads to

$$\begin{aligned}
&e^{-\int_0^{t_j} m(s,X(s))ds}e^{-\int_0^{U} r(s,X(s))ds} - E\left[e^{-\int_0^{t_j} m(s,X(s))ds}e^{-\int_0^{U} r(s,X(s))ds}\right] \\
&\qquad = \sum_{0}^{k}\int_{0}^{t_j} e^{-\int_0^{s} m(u,X(u))du}e^{-\int_0^{s} r(u,X(u))du}\frac{\partial f^{A_1}}{\partial X_i}(s,X(s))dM_i(s) \\
&\qquad + \sum_{0}^{k}\int_{t_j}^{U} e^{-\int_0^{t_j} m(u,X(u))du}e^{-\int_0^{s} r(u,X(u))du}\frac{\partial f^{B_1}}{\partial X_i}(s,X(s))dM_i(s) \\
&\qquad = \sum_{0}^{k}\int_{0}^{U} [\ \mathbb{1}_{(0,t_j]}(t)\ e^{-\int_0^{s} m(u,X(u))du}e^{-\int_0^{s} r(u,X(u))du}\frac{\partial f^{A_1}}{\partial X_i}(s,X(s)) \\
&\qquad + \mathbb{1}_{(t_j,U]}(t)\ e^{-\int_0^{t_j} m(u,X(u))du}e^{-\int_0^{s} r(u,X(u))du}\frac{\partial f^{B_1}}{\partial X_i}(s,X(s))\ ]\ \sigma_i(s)dW_i(s) \\
&\qquad = \sum_{i=1}^{k}\int_{0}^{U} \varphi_i(s)dW_i(s)\ .
\end{aligned}$$

Finally, applying Lemma 1,

$$\begin{aligned} L_0 - E\left[L_0\right] &= (P_0 - N(t_j))e^{-\int_0^U r(s,X(s))ds} - E\left[(P_0 - N(t_j))e^{-\int_0^U r(s,X(s))ds}\right] \\ &= \sum_0^k \Big(\int_0^{t_j} (P_0 - N(s-))e^{-\int_0^s r(u,X(u))du} \frac{\partial f^{A_1}}{\partial X_i}(s, X(s))dM_i(s) \\ &+ \int_{t_j}^U (P_0 - (N(t_j)))e^{-\int_0^s r(u,X(u))du} \frac{\partial f^{B_1}}{\partial X_i}(s, X(s))dM_i(s) \Big) \\ &- \int_0^{t_j} e^{-\int_0^s r(u,X(u))du} f^{A_1}(s, X(s))dM^N(s) \ . \end{aligned}$$

2. For $U \leq t_j$, considering the same procedures as in the first part of the proof yields the desired statement.

□

2.2.3 MRT Risk Decomposition

Thus, the risk decomposition is obtained for the following portfolios:

1. An annuity portfolio of the form $L_0 = \sum_{j=1}^T (P_0 - N(t_j))e^{-\int_0^{t_j} r(s)ds}$, with the risk decomposition obtained following Proposition 3, part 2, with $U = t_j$ for each t_j, $j = 1, \ldots, T$. The MRT risk decomposition is given:

$$\begin{aligned} L_0 - E[L_0] &= \sum_{i=1}^k \sum_{j=1}^T \int_0^{t_j} (P_0 - N(t-))e^{-\int_0^t r(s,X(s))ds} \frac{\partial f^{A_2}}{\partial X_i}(t, X(t))dM_i(t) \\ &- \sum_{j=1}^T \int_0^{t_j} e^{-\int_0^t r(s,X(s))ds} f^{A_2}(t, X(t))dM^N(t) \qquad (2.2.18) \\ &= \sum_{i=1}^k R_i + R_N \ . \end{aligned}$$

The first sum represents the risk factors over T periods for the state variable X of dimension k. They represents the *systematic risks* associated with the underlying rates of the portfolio. The risk factor R_N represents the *unsystematic mortality risk* generated by the number of deaths over T periods. Moreover,

$$f^{A_2}(t, \omega) = E[e^{-\int_t^{t_j} m(s,X(s))ds} e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega] \ . \qquad (2.2.19)$$

2. For a life insurance portfolio of the form

$$L_0 = \sum_{j=1}^{T}(P_0 - N(t_{j-1}))e^{-\int_0^{t_j} r(s)ds} - \sum_{j=1}^{T}(P_0 - N(t_j))e^{-\int_0^{t_j} r(s)ds} \ ,$$

a risk decomposition can be obtained by applying the MRT decomposition for each sum. Applying Proposition 3, setting $U = t_{j+1}$ in the first part, and $U = t_j$, to the second part, yields

$$\begin{aligned}
L_0 - E[L_0] &= \sum_{i=1}^{k}\sum_{j=1}^{T}\left(\int_0^{t_{j-1}}(P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_1}}{\partial X_i}(t, X(t))dM_i(t)\right) \\
&+ \sum_{i=1}^{k}\sum_{j=1}^{T}\left(\int_{t_{j-1}}^{t_j}(P_0 - N(t_{j-1}))e^{-\int_0^t r(s)ds}\frac{\partial f^{B_1}}{\partial X_i}(t, X(t))dM_i(t)\right) \\
&- \sum_{i=1}^{k}\sum_{j=1}^{T}\left(\int_0^{t_j}(P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_2}}{\partial X_i}(t, X(t))dM_i(t)\right) \\
&+ \left(-\sum_{k=1}^{T}\int_{0^+}^{t_{j-1}} e^{-\int_0^t r(s)ds} f^{A_1} dM^N(t) + \sum_{k=1}^{T}\int_{0^+}^{t_j} e^{-\int_0^t r(s)ds} f^{A_2} dM^N(t)\right) \\
&= \sum_{i=1}^{k} R_i + R_N \ ,
\end{aligned} \tag{2.2.20}$$

with

$$\begin{aligned}
f^{A_1}(t,\omega) &= E[e^{-\int_t^{t_{j-1}} m(s,X(s))ds} e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega] \ , \\
f^{A_2}(t,\omega) &= E[e^{-\int_t^{t_j} m(s,X(s))ds} e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega] \ ,
\end{aligned}$$

and

$$f^{B_1}(t,\omega) = E[e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega] \ .$$

2.3 The MRT as a meaningful risk decomposition

Schilling et al. (2015) proved in their paper that the MRT decomposition does indeed satisfy the six properties defining the *meaningfulness* of a risk decomposition.

2.3.1 The Variance Decomposition Method

In the literature, a commonly used risk decomposition method for the risk $L = L_0 - E[L_0]$ would be the so-called *variance decomposition*, which is based on conditional expectation. For simplicity, suppose for example that only X_1 and X_2 are the sources of risk. Then the *variance decomposition method* proceeds in such a way that it first assumes and conditions X_1 as the cause of risk. Then, the remaining un-captured risk would then be attributed to X_2. That is, the decomposition of L is as such:

$$L_0 - E(L_0) = [E(L_0 \mid X_1) - E(L_0)] + [L_0 - E(L_0 \mid X_1)] = R_1 + R_2 \ .$$

Moreover, the number of sources of risk can be increased accordingly.

This is called the *variance decomposition* because the variance can easily be decomposed with $Var(L) = Var(R_1) + Var(R_2)$. It is easy to verify that the *variance decomposition* satisfy Properties 1, 2, 3, 5, and 6 of a meaningful risk decomposition. However, it does not satisfy Property 4, the order invariance property, and one could see intuitively that it is because the risk decomposition is different depending on which source of risk is assumed and conditioned first.

Chapter 3

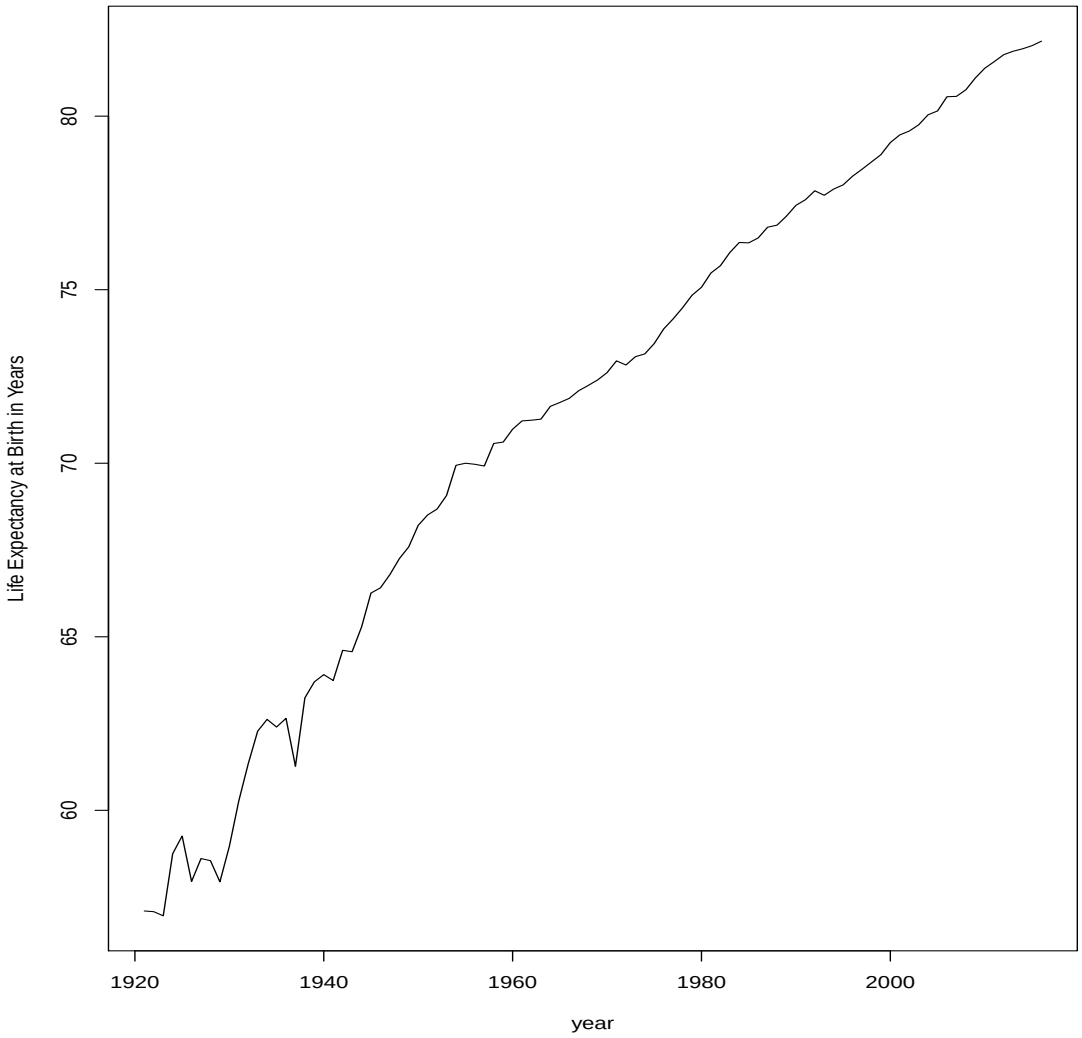
Stochastic Mortality Model

During the last couple of decades, there has been empirical evidence in the improvement of life expectancy in many countries, partly because of advancements in technology and in medical care, in better health system, in better lifestyles, etc. Such improvement in population longevity is widely studied in the actuarial field because population ageing also means that actuaries must re-evaluate the assumptions made to the underlying old-age mortality model and consequently, to re-evaluate the risk associated with the greater longevity. Figure 3.1 shows the Life Expectancy at Birth in Years in Canada from 1921-2016. There is a 25-year increase in life expectancy between 1921 and 2016. The dataset is available at the Human Mortality Database[1].

Because life expectancy has been improving over time, it would then be inappropriate to not consider time evolution in modelling mortality. Such flaws are observed for example in traditionally fixed deterministic models where they are not able to capture the observed declining mortality rate trend. Moreover, longevity has been increasing in quite an unpredictable way; there is a cluster of factors that are each responsible for population ageing but constitute a random process as a whole. Mortality should then be modelled under stochastic mortality models in order to better understand the uncertainty and the fluctuations in old age mortality evolution.

[1] http://www.mortality.org/

Figure 3.1: Life Expectancy at Birth in Years in Canada from 1921-2016



3.1 Data

As observed above, there has been an empirical improvement in life expectancy. In 2016, life expectancy was of 82 years compared to 57 years in 1921. This is because, over the years, the general trend seems to suggest that mortality intensity is in fact decreasing. Figure 3.2 contains empirical evidence of such, figuring death rates for fixed ages in Canada from 1921-2016, with the data extracted from the Human Mortality Database. The data is constructed in such a way that death rates are associated with age groups of length of 5 years, except at the extremities of age 0 and age 110+ where they consist for themselves. That is, age groups are grouped in the following manner: age 0, ages 1-4, ages 5-9,...,

ages 95-99, ages 100-104, ages 105-109, and finally ages 110+. This is more suitable since associating death rates to each individual age will later present over-fitting when fitting with mortality models, while associating death rates to age groups lend more flexibility when fitting for specific individual age.

3.2 Lee-Carter model

In order to capture the improvement in longevity and the decline in mortality, Lee and Carter (1992) proposed to model the death rates in such a way that their logistic function are bilinear functions of both age component and period effect. That is,

$$ln(m(x,t)) = a_x + b_x\kappa_t + \epsilon_{x,t} \ , \tag{3.2.1}$$

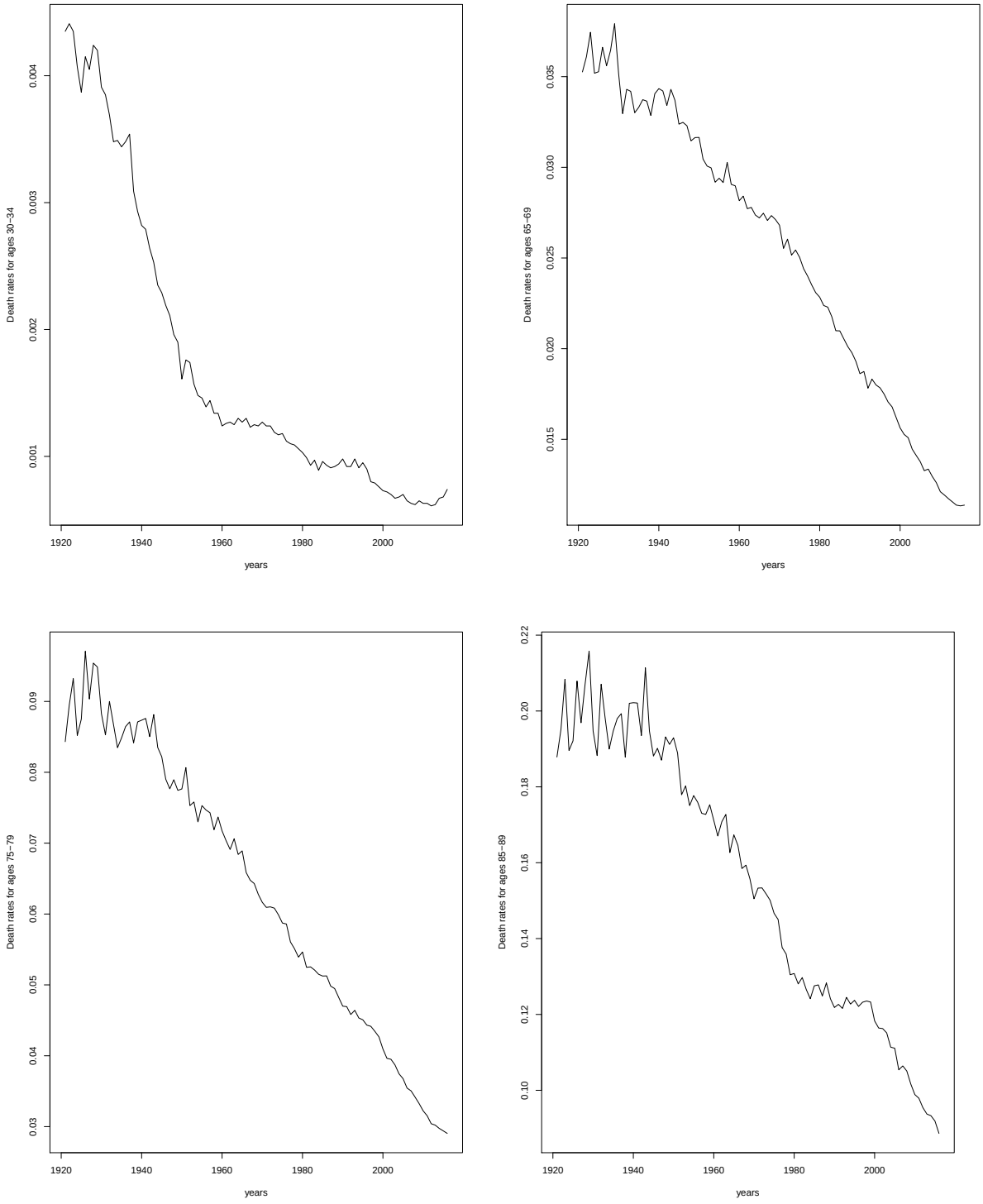
or equivalently,

$$m(x,t) = \ \exp\left[a_x + b_x\kappa_t + \epsilon_{x,t}\right] = \exp\left[a_x\right]\exp\left[b_x\kappa_t + \epsilon_{x,t}\right] \ , \tag{3.2.2}$$

where $x = x_1, x_2, ..., x_m$ are a set of consecutive ages and $t = t_1, t_2, ..., t_n$ are a set of consecutive calendar years. $m(x,t)$ is the death rate, or the mortality rate, and it is a function of both age component x and period effect t. The term $\exp\left[a_x\right]$ represents the general shape of the mortality rate across all ages x, κ_t represents the mortality index for a given period t, where it could also be interpreted as the trend in which mortality rates are evolving, and b_x represents the speed in which changes in the mortality rate evolve at a given age x in response to changes in the mortality index κ_t. In other words, it is the speed in which mortality rates at a given age are responding to the mortality improvement. Finally, $\epsilon_{x,t}$ represents the random error term. They are due to the noises represented by the age-specific historical influences that the model fails to capture. They are assumed to have mean 0 and variance σ_ϵ^2.

The above model possesses desired properties as it is able to impose a time-evolving trend upon the fitted data through its parameter κ_t, and upon all ages.

Figure 3.2: Death Rates, from 1921-2016, corresponding to age groups 30-34, 65-69, 75-79, 85-89



Notice that there is an arbitrary number of parametrizations for a given rate $m(x,t)$. To see this, consider

$$ln(m(x,t)) = \tilde{a}_x + \tilde{b}_x\tilde{\kappa}_t + \epsilon_{x,t} \ , \tag{3.2.3}$$

where $\tilde{a}_x = a_x + c_1 b_x$, $\tilde{b}_x = \frac{b_x}{c_2}$, and $\tilde{\kappa}_t = c_2(\kappa_t - c_1)$. Hence, (3.2.3) is equivalent to (3.2.1) for any arbitrary choice of scalars c_1 and c_2. Thus, let restrict the choice of parametrizations by normalizing the parameters as in Lee and Carter(1992):

$$\sum_x b_x = 1 \ , \tag{3.2.4}$$

and

$$\sum_t \kappa_t = 0 \ . \tag{3.2.5}$$

This therefore implies that

$$a_x = \frac{\sum_{t=t_1}^{t_n} ln(m(x,t))}{t_n - t_1 + 1} \ , \tag{3.2.6}$$

which is the average over time of the log of the age-specific death rates, $ln(m(x,t))$. Notice that the average over time of κ_t is zero.

3.3 Estimation

This section gives an overview of the estimation method proposed by Lee and Carter (1992). From (3.2.6), a_x could be estimated and (3.2.1) could be rearranged in the following way:

$$\hat{y}_{x,t} = b_x \kappa_t \ , \tag{3.3.1}$$

where $\hat{y}_{x,t} = ln(\hat{m}(x,t)) - a_x$.

On the left side of this new equation is the new estimated response variable $\hat{Y}_{x,t}$, while on the right side are the parameters b_x to estimate. However, since the predictors k_t are unknown indexes, ordinary regression methods are irrelevant in estimating the parameters b_x. Fortunately, the singular value decomposition (SVD) method offers a solution in finding a least squares solution without response variables.

3.3.1 Singular Value Decomposition

Theorem 2. The singular value decomposition (SVD) is the factorization of an $m \times n$ real matrix A. Let A be a $m \times n$ matrix such that $m \geq n$. Then, A can be decomposed such that

$$A_{m\times n} = U_{m\times r}\Sigma_{r\times r}V^T_{r\times n} \ . \tag{3.3.2}$$

Letting r be the rank of matrix A, U is an $m\times r$ matrix consisting of m orthonormalized eigenvectors associated with the m largest eigenvalues of AA^T. That is, the rows of U (called the *left singular vectors*) are the orthonormal eigenvectors of AA^T and $UU^T = I$. Σ is a diagonal $r \times r$ matrix with the diagonal elements being the *singular values*. They are the non-negative square roots of the eigenvalues of AA^T. Finally, V^T is a $r \times n$ matrix consisting of the orthonormalized eigenvectors of A^TA. That is, the rows of V (called the *right singular vectors*) are the orthonormal eigenvectors of A^TA and $VV^T = I$.

Proof. See Friedberg et al (2003, pp. 410) for more details. □

In the least-square estimation context, the SVD is useful since the data matrix A can be well approximated by a low rank matrix $\tilde{A}$ of rank k with $k < r$. In other words, $\tilde{A}$ is a truncated version of A.

$$\tilde{A}_{m\times n} = U_{m\times k}\Sigma_{k\times k}V^T_{k\times n} \ , \tag{3.3.3}$$

where the singular vectors of $\tilde{A}$ correspond to the first k singular vectors of A and it can be shown that $\tilde{A}$ is the best rank k approximation to A.

Briefly, SVD allows for reduction of the dimensionality of matrix data A while summarizing optimally the information (e.g. variance) provided by the original data A. For the purpose of this thesis least-squares approximation, choose $k = 1$. Thus, the following is obtained:

$$\tilde{A}_{m\times n} = U_{m\times 1}\Sigma_{1\times 1}V^T_{1\times n} = \sqrt{\lambda_1}u_1v_1^T \ , \tag{3.3.4}$$

where u_1 is the first row of U, v_1 is the first row of V, and $\sqrt{\lambda_1}$ is the first element of the diagonal matrix Σ.

3.3.2 Data Estimation

From (3.2.6), parameters a_x are calculated based on the data described in Section 3.1 and they are shown in Table 3.1.

Table 3.1: Estimated values for parameters a_x and b_x of the Lee-Carter model

ages	a_x	b_x
x_1 =0	−3.9088390	0.097464743
x_2 =1-4	−6.8959509	0.111307165
x_3 =5-9	−7.6849274	0.095827626
x_4 =10-14	−7.7429748	0.079704617
x_5 =15-19	−6.9532983	0.053748120
x_6 =20-24	−6.7024359	0.054023294
x_7 =25-29	−6.6951986	0.055701967
x_8 =30-34	−6.5629977	0.053224832
x_9 =35-39	−6.2814909	0.050913214
x_{10} =40-44	−5.9261071	0.044724695
x_{11} =45-49	−5.5058524	0.039022710
x_{12} =50-54	−5.0664319	0.034434783
x_{13} =55-59	−4.6234011	0.031774011
x_{14} =60-64	−4.1769200	0.030528534
x_{15} =65-69	−3.7361709	0.030116339
x_{16} =70-74	−3.2889979	0.029575536
x_{17} =75-79	−2.8160378	0.029446570
x_{18} =80-84	−2.3671886	0.023821005
x_{19} =85-89	−1.8937396	0.020930308
x_{20} =90-94	−1.4827932	0.013997600
x_{21} =95-99	−1.0974764	0.010124401
x_{22} =100-104	−0.7781022	0.005999671
x_{23} =105-109	−0.5242241	0.002742987
$x_{24} = 110^+$	−0.3609327	0.000845273

Following the procedure described in Subsection 3.3.1, the following equation is obtained:

$$\hat{y}_{x,t} = \begin{bmatrix} ln(m(x_1, 1921)) - a_{x_1} & \dots & ln(m(x_1, 2016)) - a_{x_1} \\ \vdots & \ddots & \vdots \\ ln(m(x_{24}, 1921)) - a_{x_{24}} & \dots & ln(m(x_{24}, 2016)) - a_{x_{24}} \end{bmatrix} = b_x \kappa_t \ . \qquad (3.3.5)$$

Following the SVD decomposition on $\hat{y}_{x,t}$, a rank 1 approximation of $\hat{y}_{x,t}$ is obtained such

that u_1 is interpreted as b_x. Since b_x is normalized to sum up to 1, the following is obtained

$$b_{x_i} = \frac{u_{1,x_i}}{\sum_{j=1}^{24} u_{1,x_j}} \ , \tag{3.3.6}$$

and κ_t results as

$$\kappa_t = \sqrt{\lambda_1} v_1 \sum_{j=1}^{24} u_{1,x_j} \ . \tag{3.3.7}$$

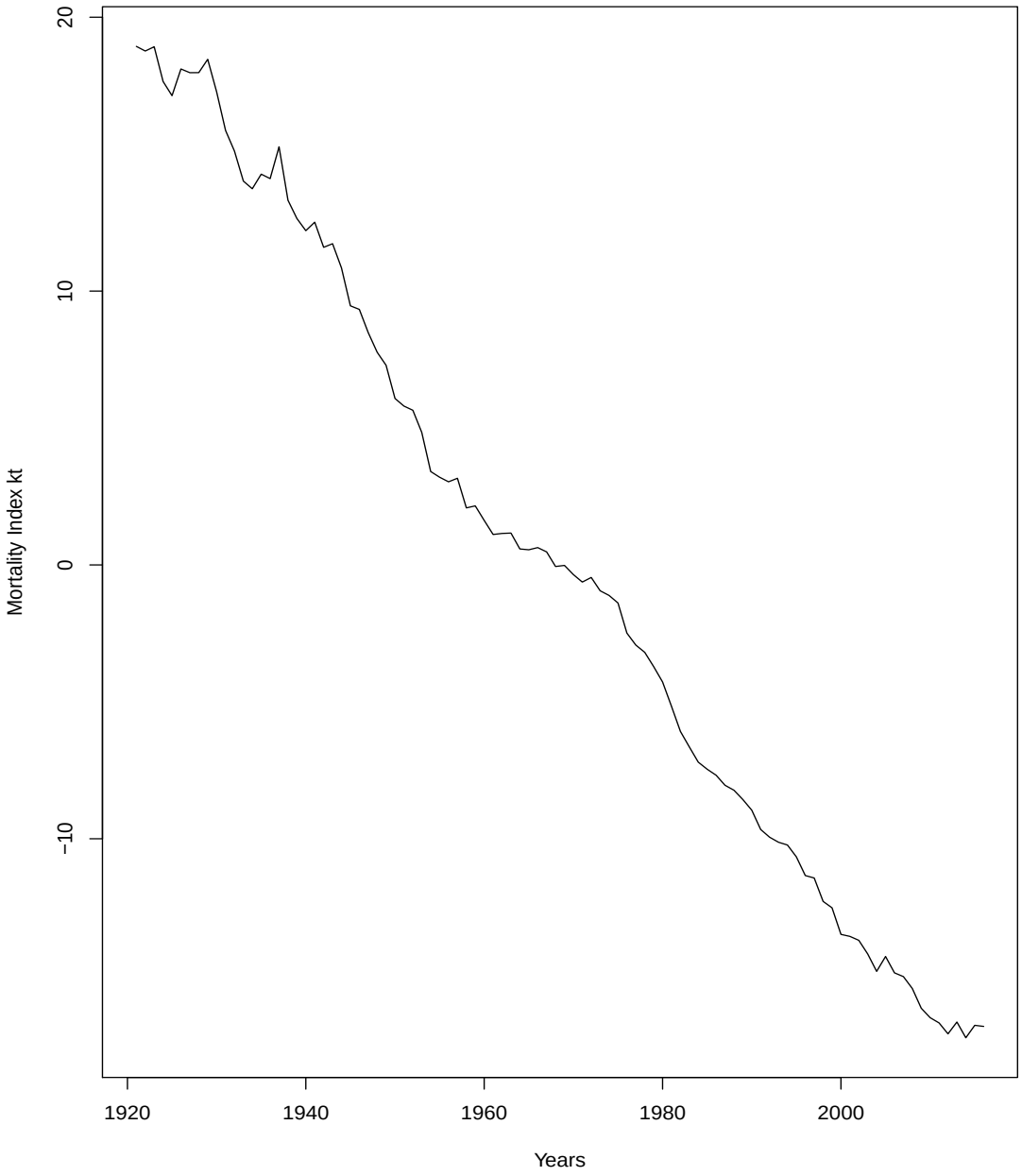
The extra term $\sum_{j=1}^{24} u_{1,x_j}$ in (3.3.7) compensates for the normalization of b_x. Parameters b_x obtained using (3.3.6) are shown in Table 3.1. Spline interpolation is used in order to obtain parameters a_x and b_x in between ages.

From Figure (3.3.2), it is clear that the mortality index k_t is in decline over time and in a quite linear fashion. Hence, the Lee-Carter model is successful in capturing the general declining trend of the death rates. However, in fitting the model to the actual death rates, it is observed that age groups 95^+ are fitted very poorly. In fact, fitting mortality at advanced ages is known to pose serious challenges to researchers. As Gompertz (1825) pointed out with his own law of mortality, mortality rate modelling may not be applicable at extreme old ages. First, standard assumptions for mortality rate estimates may be irrelevant at older ages when the risk of death is extremely high. As seen in Table 3.2, death rates at higher ages may be linked more to noises than to the fitted models. Table 3.2 presents the Proportion of Unexplained Variances (PUV) over the total variances, for all age groups, i.e.,

$$PUV_{x_i} = \frac{Var(ln(m(t,x_i) - ln(\hat{m}(t,x_i))}{Var(ln(m(t,x_i))} \ . \tag{3.3.8}$$

It is clear that the Lee-Carter model does not cover all informations at ages 95^+. Thus, mortality at older ages are caused rather by unsystematic mortality than by systematic mortality and it is without surprise that the Lee-Carter model cannot fit higher ages mortality.

Figure 3.3: Mortality Index k_t over the years 1921-2016



3.4 Mortality Projection

This section proceeds in describing the methodology used to forecast the mortality under the Lee-Carter model so that it can later be applied to the risk decomposition context. First, notice that the only time-evolving uncertainty parameter in the Lee-Carter model is the factor κ_t; all other parameters are fixed parameters depending on the ages. So, it would make sense to model κ_t as a time-series model. As in Lee and Carter

Table 3.2: Proportion of Unexplained Variances for each age group

ages	Proportion of Unexplained Variances
x_1 =0	0.018739037
x_2 =1-4	0.004887704
x_3 =5-9	0.003585420
x_4 =10-14	0.005140763
x_5 =15-19	0.045366759
x_6 =20-24	0.028713555
x_7 =25-29	0.039388856
x_8 =30-34	0.034252903
x_9 =35-39	0.016419688
x_{10} =40-44	0.004366462
x_{11} =45-49	0.027211582
x_{12} =50-54	0.068440676
x_{13} =55-59	0.094954988
x_{14} =60-64	0.110649578
x_{15} =65-69	0.092880560
x_{16} =70-74	0.080520968
x_{17} =75-79	0.069519861
x_{18} =80-84	0.086311154
x_{19} =85-89	0.079725858
x_{20} =90-94	0.127103903
x_{21} =95-99	0.209314467
x_{22} =100-104	0.381587541
x_{23} =105-109	0.689810578
$x_{24} = 110^{+}$	0.947087928

(1992), κ_t is assumed to follow an ARIMA (0,1,0) time-series model, i.e. a random walk with a drift. κ_t is assumed as follows:

$$\kappa_t = \kappa_{t-1} + \mu + \epsilon_t \ , \tag{3.4.1}$$

where ϵ_t the noise term and it is assumed to be normal distributed with mean 0 and variance σ^2. More generally,

$$\kappa_t - \kappa_{t-\Delta t} \sim \mathcal{N}\left(\mu\Delta t, \Delta t\sigma^2\right) \ ,$$

and letting Δt approach to zero, the following instantaneous rate of change for κ_t is obtained:

$$\begin{aligned}\lim_{\Delta t\to 0}(\kappa_t-\kappa_{t-\Delta t}) &= \lim_{\Delta t\to 0}\mathcal{N}\left(\mu\Delta t,\Delta t\sigma^2\right)\\ &= \lim_{\Delta t\to 0}\left(\mu\Delta t+\mathcal{N}\left(0,\Delta t\sigma^2\right)\right)\\ &= \mu\lim_{\Delta t\to 0}\Delta t+\lim_{\Delta t\to 0}\mathcal{N}\left(0,\Delta t\sigma^2\right)\ .\end{aligned}$$

Thus,

$$d\kappa_t=\mu dt+\sigma dW_t\ , \tag{3.4.2}$$

where dW_t is a standard Wiener process. The long-term drift μ is interpreted as the average period-to-period change. That is,

$$\hat{\mu}=\frac{\sum_{t=t_2}^{t_n}(\kappa_t-\kappa_{t-1})}{t_n-t_1} \tag{3.4.3}$$

and σ^2 is simply estimated as the sample variance of the period-to-period changes, i.e.,

$$\hat{\sigma}=\sqrt{\frac{\sum_{t=t_2}^{t_n}((\kappa_t-\kappa_{t-1})-\hat{\mu})^2}{t_n-t_1}}\ . \tag{3.4.4}$$

Based on the fitted values for κ_t, $\hat{\mu}$ and $\hat{\sigma}$ are obtained such that

$$k_t=k_{t-1}-0.3767221+\epsilon_t\ , \tag{3.4.5}$$

where $\epsilon_t\sim N(0,\hat{\sigma}=0.5093948)$

3.4.1 Death Rates Dynamics

Let now proceed into using Itô's formula in order to forecast the death rates themselves by finding the dynamic of the force of mortality underlying the Lee-Carter model for an initial age x. Since the dynamics of $ln(m(t,x))$ is known and $ln(m(t,x))=a_{x+t}+b_{x+t}\kappa_t$, let apply Itô's lemma for $f(t,x)=e^x$ on $ln(m(t,x))$ in order to obtain the dynamics of

$m(t,x)$. Notice that

$$\begin{aligned} d(ln(m(t,x))) = a'_{x+t}dt + d(b_{x+t}\kappa_t) &= a'_{x+t}dt + \kappa_t b'_{x+t}dt + b_{x+t}d\kappa_t \\ &= \left[a'_{x+t} + \kappa_t b'_{x+t}\right]dt + b_{x+t}\left(\mu dt + \sigma dW_t\right) \\ &= \left[a'_{x+t} + \kappa_t b'_{x+t} + b_{x+t}\mu\right]dt + b_{x+t}\sigma dW_t \ . \end{aligned}$$

Moreover,

$$m(t,x) = e^{ln(m(t,x))} = f(t, ln(m(t,x))) \ ,$$

for all t. Using Itô's lemma with $f(t,z) = e^z$,

$$\begin{aligned} d(m(t,x)) &= f_z(t, ln(m(t,x)))d(ln(m(t,x))) + \frac{1}{2}f_{zz}(t, ln(m(t,x)))(d(ln(m(t,x))))^2 \\ &= e^{ln(m(t,x))}d(ln(m(t,x))) + \frac{1}{2}e^{ln(m(t,x))}(d(ln(m(t,x))))^2 \ . \end{aligned}$$

Drawing on the multiplication rules for differentials, $dW_t dW_t = dt$ and $dtdW_t = dtdt = 0$, then

$$\begin{aligned} d(m(t,x)) &= m(t,x)\left(\left[a'_{x+t} + \kappa_t b'_{x+t} + b_{x+t}\mu\right]dt + b_{x+t}\sigma dW_t\right) + \frac{1}{2}m(t,x)b^2_{x+t}\sigma^2 dt \\ &= m(t,x)\left[a'_{x+t} + \kappa_t b'_{x+t} + b_{x+t}\mu + \frac{1}{2}b^2_{x+t}\sigma^2\right]dt + b_{x+t}\sigma m(t,x)dW_t \ . \end{aligned}$$

Briefly,

$$\frac{d(m(t,x))}{m(t,x)} = \left[a'_{x+t} + \kappa_t b'_{x+t} + b_{x+t}\mu + \frac{1}{2}b^2_{x+t}\sigma^2\right]dt + b_{x+t}\sigma dW_t \ . \tag{3.4.6}$$

It is interesting to note that by letting the age factor increase as time progresses, the drift of the death rate process, $\left[a'_{x+t} + \kappa_t b'_{x+t} + b_{x+t}\mu + \frac{1}{2}b^2_{x+t}\sigma^2\right]$, is always positive so that $m(t,x)$ is a submartingale. This is appropriate as the mortality rate is expected to increase along the years so that older ages have higher mortality.

Chapter 4

Interest Rates Model

This chapter focuses on choosing a relevant interest rates model in order to better understand the dynamic of the bond market and the impact of investment risk on the insurance portfolios. Since contracts often span over long periods of time, it is important to choose an interest dynamic that better reflects the fluctuation of the rate over the years. In doing so, the focus consists in choosing an appropriate model that can provide good forecasting in relation to better risk management. Interest rates will be modelled through yield curve estimation because yield curves give a direct relationship between short-term, medium-term, and long-term bond rates. In fact, literature suggests that yield curves are mainly motivated by three factors: the level, slope, and curvature, representing the long-term factor and general shape of the yield curve, the short-term factor, and lastly the shape of the medium-term interest rates, respectively. Thus, in relation to the goal of this thesis, the MRT risk decomposition will be able to explain the contribution of each of these factors to the overall risk in the insurance setting. This chapter presents a model that is appropriate for these objectives.

4.1 The affine arbitrage-free class of Nelson-Siegel term structure models

Christensen et al. (2011) derive a term structure model where they combine elements of both models of the Diebold-Li extension of the Nelson-Siegel (DNS) term structure and of the arbitrage-free affine term structure models. The former model is easy to estimate and it provides good fits to empirical data because it provides an easy interpretation of the factors of level, slope, and curvature. However, the DNS model does not impose arbitrage-free restrictions.

The affine term (AF) structure models provide an arbitrage-free framework. They are also very convenient as they impose that the log bond prices be linear functions of the spot rate. However, AF models often exhibit poor empirical fitting and their estimations are often of a computational burden. The literature on the AF models is mainly concerned about fitting the short-rate dynamics, after which yield curves are derived. Thus, fitting yield curves under AF models often results in poor forecasting.

As such, Christensen et al. (2011) derived a new class of models called the *affine arbitrage-free class of Nelson-Siegel* (AFNS) term structure models where they combined advantages of both the DNS and AF models in terms of empirical tractability and of imposing an arbitrage-free condition.

4.1.1 The Nelson-Siegel model

The Nelson-Siegel model (1987) is given below as a simple function:

$$y(\tau) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) , \tag{4.1.1}$$

where $y(\tau)$ is the yield curve with τ years to maturity, and β_1, β_2, β_3 and λ are model parameters.

Parameter λ is responsible for the overall exponential rate:

- Small values of λ produce slow decay and are thus more appropriate in fitting the

curve at long maturities.

- On the contrary, large values of λ produce fast decay and are better fit for short maturities.

4.1.2 The dynamic Nelson-Siegel model

Diebold and Li (2006) extended the Nelson-Siegel model in order to allow for dynamic understanding of the evolution of the yield curve over time. They impose that parameters β_1, β_2, β_3 from (4.1.1) be dynamic factors such that

$$y(\tau) = Y^{(1)} + Y^{(2)} \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + Y^{(3)} \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) , \tag{4.1.2}$$

- $Y^{(1)}$ is interpreted as the *level of the curve*, that is, the long-term factor where it governs the yield curve level. As τ increases in maturity, the yield curve approaches to $Y^{(1)}$. In other words, $y(\infty) = Y^{(1)}$.

- $Y^{(2)}$ is interpreted as the *slope of the curve* because it is responsible for the short-term maturities. Starting at 1, its loading $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} \right)$ quickly decays monotonically to 0 as maturity increases and thus, factor $Y^{(2)}$ only have an impact in short-term maturities.

- $Y^{(3)}$ is then interpreted as the *curvature of the curve*, being responsible for medium-term maturities. Its loading $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$ starts at 0, increases and then decays back to 0. Thus, it cannot be responsible for short-term nor for long-term maturities.

Thus, the dynamic Nelson-Siegel model provides a good empirical interpretation of the yield curve in that it considers the modern factors of level, slope, and curvature of the yield curve. However, it does not impose any restrictions in term of arbitrage-freeness in bond pricing.

4.1.3 The Arbitrage-Free class of Nelson-Siegel model

In order to overcome this weakness, Christensen et al. (2011) further impose that the factors $Y_t = (Y_t^{(1)}, Y_t^{(2)}, Y_t^{(3)})$ presented in (4.1.2) be Vasicek processes. The following gives a detailed explanation of the Arbitrage-Free class of Nelson-Siegel model (AFNS).

Under the real world probability $\mathbb{P}$, the AFNS assumes that the instantaneous risk-free rate r is explained by only two risk factors, the level and the slope factors of the yield curve such that,

$$r_t = Y_t^{(1)} + Y_t^{(2)} \ . \tag{4.1.3}$$

Thus, the instantaneous risk-free rate is indeed the sum of only the level factor $Y_t^{(1)}$ and of the slope factor $Y_t^{(2)}$ under the real world probability $\mathbb{P}$, while the curvature factor $Y_t^{(3)}$ is latent under the real world probability but manifests itself only under the risk-neutral $\mathbb{Q}$-measure. This fact is supported by the empirical literature, where the curvature lacks sensible interpretation in the real world $\mathbb{P}$-measure (see Diebold, Rudebush, and Aruoba, 2006). On the one hand, Diebold et al. (2006) show that shocks on the curvature factor produced negligible responses to the macroeconomy and to the macroeconomic factors. On the other hand, they show that there is a strong connection between the inflation and the level factor, and the slope factor has also a close connection to the funds rate.

Under the risk-neutral $\mathbb{Q}$-measure, let assume that the factors of Y_t are described by the following stochastic differential equations (SDEs):

$$\begin{bmatrix} dY_t^{(1)} \\ dY_t^{(2)} \\ dY_t^{(3)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix} \left[\begin{bmatrix} \theta_1^{\mathbb{Q}} \\ \theta_2^{\mathbb{Q}} \\ \theta_3^{\mathbb{Q}} \end{bmatrix} - \begin{bmatrix} Y_t^{(1)} \\ Y_t^{(2)} \\ Y_t^{(3)} \end{bmatrix} \right] dt + \begin{bmatrix} \sigma_{Y^{(1)}} & 0 & 0 \\ 0 & \sigma_{Y^{(2)}} & 0 \\ 0 & 0 & \sigma_{Y^{(3)}} \end{bmatrix} \begin{bmatrix} dW_{t,1}^{\mathbb{Q}} \\ dW_{t,2}^{\mathbb{Q}} \\ dW_{t,3}^{\mathbb{Q}} \end{bmatrix}, \lambda > 0 \ . \tag{4.1.4}$$

Then, the zero-coupon bond prices are given under the no-arbitrage condition by the

following Riccati equation:

$$P(t,T) = E^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_u du\right) \mid \mathcal{G}_t\right]$$
$$= \exp\left(B^{(1)}(t,T)Y_t^{(1)} + B^{(2)}(t,T)Y_t^{(2)} + B^{(3)}(t,T)Y_t^{(3)} + A(t,T)\right) \ , \quad (4.1.5)$$

or equivalently, the zero-coupon bond yields are given by

$$y(t,T) = -\frac{\left(B^{(1)}(t,T)Y_t^{(1)} + B^{(2)}(t,T)Y_t^{(2)} + B^{(3)}(t,T)Y_t^{(3)} - A(t,T)\right)}{T-t} \ , \quad (4.1.6)$$

where $B^{(1)}(t,T)$,$B^{(2)}(t,T)$, and $B^{(3)}(t,T)$ satisfy the following system of ordinary differential equations (ODEs):

$$\begin{bmatrix} \frac{dB^{(1)}(t,T)}{dt} \\ \frac{dB^{(2)}(t,T)}{dt} \\ \frac{dB^{(3)}(t,T)}{dt} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{bmatrix} \begin{bmatrix} B^{(1)}(t,T) \\ B^{(2)}(t,T) \\ B^{(3)}(t,T) \end{bmatrix} \ , \quad (4.1.7)$$

and $A(t,T)$ is given by

$$\frac{dA(t,T)}{dt} = -B(t,T)'K^{\mathbb{Q}}\theta^{\mathbb{Q}} - \frac{1}{2}\sum_{j=1}^{3}(\Sigma' B(t,T)B(t,T)'\Sigma)_{j,j} \ , \quad (4.1.8)$$

with boundary conditions $B^{(1)}(T,T) = B^{(2)}(T,T) = B^{(3)}(T,T) = A(T,T) = 0$, and

$$\Sigma = \begin{bmatrix} \sigma_{Y^{(1)}} & 0 & 0 \\ 0 & \sigma_{Y^{(2)}} & 0 \\ 0 & 0 & \sigma_{Y^{(3)}} \end{bmatrix} .$$

The solutions to the ODEs are given by

$$B^{(1)}(t,T) = -(T-t) \ ,$$
$$B^{(2)}(t,T) = -\frac{1-e^{-\lambda(T-t)}}{\lambda} \ ,$$
$$B^{(3)}(t,T) = (T-t)e^{-\lambda(T-t)} - \frac{1-e^{-\lambda(T-t)}}{\lambda} \ ,$$

and

$$A(t,T) = (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_{2,1}\int_t^T B^{(2)}(s,T)ds + (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_{3,1}\int_t^T B^{(3)}(s,T)ds \\ + \frac{1}{2}\sum_{j=1}^{3}\int_t^T (\Sigma' B(s,T)B(s,T)'\Sigma)_{j,j}\, ds \ .$$

Thus,

$$y(t,T) = Y_t^{(1)} + \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}Y_t^{(2)} + [\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}]Y_t^{(3)} - \frac{A(t,T)}{T-t} \ . \qquad (4.1.9)$$

Please refer to Christensen et al. (2011) for further details.

Next, let define the term $A(t,T)$. Following Christensen (2011) et al., the mean reversion under the risk-neutral world is restricted to be zero such that $\theta^{\mathbb{Q}} = 0$, while $\theta^{\mathbb{P}}$ remains to be estimated. Notice again that the drift terms $(K^{\mathbb{P}}, \theta^{\mathbb{P}})$ can be chosen independently from the $\mathbb{Q}$-measure drift terms $(K^{\mathbb{Q}}, \theta^{\mathbb{Q}})$. Under affine properties, the solution for $A(t,T)$ is given under a closed form expression.

Since the state variables are also assumed to be independent, Christensen et al. (2011) showed that the term $A(t,T)$ is summarized as follows:

$$\frac{A(t,T)}{T-t} = \sigma^2_{Y^{(1)}}\frac{(T-t)^2}{6} + \sigma^2_{Y^{(2)}}\left[\frac{1}{2\lambda^2} - \frac{1}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t} + \frac{1}{4\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right]$$

$$+ \sigma^2_{Y^{(3)}}\left[\frac{1}{2\lambda^2} + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{4\lambda}(T-t)e^{-2\lambda(T-t)} - \frac{3}{4\lambda^2}e^{-2\lambda(T-t)} - \frac{2}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t} + \frac{5}{8\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right] .$$

The relationship between the real-world dynamics under the $\mathbb{P}$-measure and the risk neutral dynamics under the $\mathbb{Q}$-measure can be expressed as

$$dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}} + \Gamma_t\, dt \ ,$$

where Γ_t refers to the risk premium. Let Γ_t be the affine risk premium specifications such

that

$$\Gamma_t = \begin{bmatrix} \gamma_1^{(0)} \\ \gamma_2^{(0)} \\ \gamma_3^{(0)} \end{bmatrix} + \begin{bmatrix} \gamma_{11}^{(1)} & \gamma_{12}^{(1)} & \gamma_{13}^{(1)} \\ \gamma_{21}^{(1)} & \gamma_{22}^{(1)} & \gamma_{23}^{(1)} \\ \gamma_{31}^{(1)} & \gamma_{32}^{(1)} & \gamma_{33}^{(1)} \end{bmatrix} \begin{bmatrix} Y_t^{(1)} \\ Y_t^{(2)} \\ Y_t^{(3)} \end{bmatrix} \tag{4.1.10}$$

in order to also preserve the affine dynamics for the $\mathbb{P}$-measure (see Duffee, 2002).

The relationship described in (4.1.10) suggests that the $\mathbb{Q}$-measure dynamics of the AFNS model does not impose any restrictions on the parameters of the $\mathbb{P}$-measure dynamics, so there are an infinite number of possibilities in choosing the parameters of the $\mathbb{P}$-measure that can still satisfy the dynamics of the $\mathbb{Q}$-measure. Let therefore choose conveniently so that the state variables is independent and are described by the following SDEs under the $\mathbb{P}$-measure, that is,

$$dY_t = K^{\mathbb{P}}[\theta^{\mathbb{P}} - Y_t]dt + \Sigma dW_t^{\mathbb{P}} \ , \tag{4.1.11}$$

or equivalently,

$$\begin{bmatrix} dY_t^{(1)} \\ dY_t^{(2)} \\ dY_t^{(3)} \end{bmatrix} = \begin{bmatrix} k_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & k_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & k_{33}^{\mathbb{P}} \end{bmatrix} \left[\begin{bmatrix} \theta_1^{\mathbb{P}} \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{bmatrix} - \begin{bmatrix} Y_t^{(1)} \\ Y_t^{(2)} \\ Y_t^{(3)} \end{bmatrix} \right] dt + \begin{bmatrix} \sigma_{Y^{(1)}} & 0 & 0 \\ 0 & \sigma_{Y^{(2)}} & 0 \\ 0 & 0 & \sigma_{Y^{(3)}} \end{bmatrix} \begin{bmatrix} dW_{t,1}^{\mathbb{P}} \\ dW_{t,2}^{\mathbb{P}} \\ dW_{t,3}^{\mathbb{P}} \end{bmatrix} . \tag{4.1.12}$$

This dynamic is the path of each of the factor under the real-world probability measure $\mathbb{P}$ such that the state variables remain affine. The parameter $K^{\mathbb{P}}$ represents the mean reversion speed, $\theta^{\mathbb{P}}$ the mean reversion level, and Σ the volatility.

Finally, the yield curves for different time of maturities τ are as follows:

$$y_t = A + BY_t + \epsilon_t \ , \tag{4.1.13}$$

or equivalently,

$$\begin{bmatrix} y_t(\tau_1) \\ y_1(\tau_2) \\ \vdots \\ y_1(\tau_{N_y}) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_{N_y}}}{\lambda\tau_{N_y}} & \frac{1-e^{-\lambda\tau_{N_y}}}{\lambda\tau_{N_y}} - e^{-\lambda\tau_{N_y}} \end{bmatrix} \begin{bmatrix} Y_t^{(1)} \\ Y_t^{(2)} \\ Y_t^{(3)} \end{bmatrix} - \begin{bmatrix} \frac{A(\tau_1)}{\tau_1} \\ \frac{A(\tau_2)}{\tau_2} \\ \vdots \\ \frac{A(\tau_{N_y})}{\tau_{N_y}} \end{bmatrix} + \begin{bmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \vdots \\ \epsilon_t(\tau_3) \end{bmatrix}, \quad (4.1.14)$$

where $\tau_1, ...\tau_{N_y}$ corresponds to different maturities.

4.2 Estimation

The parameter estimation is done through the Kalman-filter maximum-likelihood estimation.

4.2.1 The Kalman Filter Design

The Kalman filtering algorithm is based on making predictive inference on unobserved values of the state variable Y based on past information and then making corrections afterwards on these predictions in order to estimate its parameters. More details will be given in Section 4.2.2. The current section focuses on defining the predictive equations and the forecasts to be used in the algorithm.

For the purpose of this thesis, the Kalman Filter is used to estimate the parameters of the AFNS model, based on fitting the yield curve over different maturities. Estimates for the yield curve are given by (4.1.13), which will later be compared with the actual observed yield curve. Let $A = A(\Theta)$ and $B = B(\Theta)$, where Θ represents the set of all parameters, such that A and B are functions of these parameters. The first step of the Kalman filtering algorithm is to forecast the yield curve Y for different maturities. The forecast equations are derived from conditional expected values and for the continuous-time AFNS models, given information at time t, the conditional mean and the conditional

covariance matrix of the state variable Y are

$$E^{\mathbb{P}}[Y_T \mid \mathcal{F}_t] = (I - e^{-K^{\mathbb{P}}\Delta t})\theta^{\mathbb{P}} + e^{-K^{\mathbb{P}}\Delta t}Y_t \ , \tag{4.2.1}$$

and

$$V^{\mathbb{P}}[Y_T \mid \mathcal{F}_t] = \int_0^{\Delta t} e^{-K^{\mathbb{P}}s}\Sigma\Sigma' e^{-(K^{\mathbb{P}})'s}ds \ , \tag{4.2.2}$$

where $\Delta t = T - t$ for $T > t$. These results are from the assumption that Y is assumed to be a multi-factor Vasicek process. See Christensen et al. (2011).

By reducing the time between the observations and by assuming that (4.2.1) gives predictive values for Y, define then the following state transition equation:

$$Y_{t_i|t_{i-1}} = E^{\mathbb{P}}[Y_{t_i} \mid \mathcal{F}_{t_{i-1}}] + \eta_{t_i} = (I - e^{-K^{\mathbb{P}}\Delta t_i})\theta^{\mathbb{P}} + e^{-K^{\mathbb{P}}\Delta t_i}Y_{t_{i-1}} + \eta_{t_i},$$

where $\Delta t_i = t_i - t_{i-1}$ is the time between observations and $Y_{t_i|t_{i-1}}$ represents the forecast given information at time t_{i-1}. Based on (4.2.2), the noises η_{t_i} have conditional covariance matrix given by

$$Q = \int_0^{\Delta t_i} e^{-K^{\mathbb{P}}s}\Sigma\Sigma' e^{-(K^{\mathbb{P}})'s}ds \ .$$

In fact, noises η_{t_i} and ϵ_{t_i} correspond respectively to the process and to the measurement noise vectors, and they are given by

$$\begin{bmatrix} \eta_{t_i} \\ \epsilon_{t_i} \end{bmatrix} \sim \mathcal{N}\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & H \end{bmatrix}\right] \ , \text{ where } H = \begin{bmatrix} \sigma_\epsilon^2(\tau_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_\epsilon^2(\tau_{N_y}) \end{bmatrix} \ .$$

A diagonal H matrix implies that the deviation of yields of different maturities from the yield curve are uncorrelated.

4.2.2 The Kalman Filter Algorithm

The Kalman Filter is an optimization algorithm that implements a predictor-corrector type estimator in order to minimize the estimated error covariance. The algorithm functions as a feedback control by first predicting *a priori* predictor and then making respectively *a posteriori* adjustments on the estimated predictor based on feedback from actual information. The Kalman Filter thus consists of two sets of equations defined as the *time update* equations and the *measurement update* equations. The former is responsible for projecting forward in time the predictor and its associated error covariance. The latter is responsible for correcting afterwards the estimated predictor. Define

$$\begin{aligned}\phi_{t_i}^{(0)} &= (I - e^{-K^{\mathbb{P}}\Delta t_i})\theta^{\mathbb{P}} \ , \\ \phi_{t_i}^{(1)} &= e^{-K^{\mathbb{P}}\Delta t_i} \ .\end{aligned}$$

Based on (4.2.1) and (4.2.2), the *time update* equations from time t_{i-1} to t_i are described in the following. The estimated predictors, given information at time t_{i-1}, are given by

$$\hat{Y}_{t_i} = \phi_{t_i}^{(0)}(\Theta) + \phi_{t_i}^{(1)}(\Theta)\dot{Y}_{t_{i-1}} \ ,$$

with mean square error covariance matrix

$$\hat{\Sigma}_{t_i} = \phi_{t_i}^{(1)}(\Theta)\dot{\Sigma}_{t_{i-1}}\phi_{t_i}^{(1)}(\Theta)' + Q(\Theta) \ ,$$

where Θ represents the vector of all parameters to estimate, i.e. $\Theta = \{K^{\mathbb{P}}, \theta^{\mathbb{P}}, \Sigma, \lambda\}$.

At time t_i, $\hat{Y}_{t_i}$ is updated by using additional information provided by the actual observed yields y_{t_i}. The corrected variables are thus given by the *measurement update* equations:

$$\begin{aligned}\dot{Y}_{t_i} &= \hat{Y}_{t_i} + \hat{\Sigma}_{t_i}B(\Theta)'\hat{F}_{t_i}^{-1}v_{t_i} \ , \\ \dot{\Sigma}_{t_i} &= \hat{\Sigma}_{t_i} - \hat{\Sigma}_{t_i}B(\Theta)'\hat{F}_{t_i}^{-1}B(\Theta)\hat{\Sigma}_{t_i} \ ,\end{aligned}$$

where

$$v_{t_i} = y_{t_i} - E[y_{t_i} \mid \mathcal{F}_{t_{i-1}}] = y_{t_i} - A(\Theta) - B(\Theta)\hat{Y}_{t_i} \ ,$$
$$\hat{F}_{t_i} = cov(v_{t_i}) = B(\Theta)\hat{\Sigma}_{t_i}B(\Theta)' + H(\Theta) \ ,$$

are *a postepriori* error and error covariance matrix respectively, based on the measurement matrix y_{t_i}. In other words, v_{t_i} represents the out-of sample forecast error given information at time t_i and F_{t_i} is the resulting covariance matrix error. Moreover, it is assumed that at time zero,

$$\dot{Y}_0 = \theta^{\mathbb{P}} \ ,$$
$$\dot{\Sigma}_0 = \int_0^\infty e^{-K^{\mathbb{P}}s}\Sigma\Sigma' e^{-(K^{\mathbb{P}})'s}ds \ .$$

Once the Kalman filter has completed its algorithm and has produced all the required values, the AFNS parameters are then estimated based on maximizing the likelihood of the prediction-error decomposition with respect to Θ. With N_y being the number of yields, the Gaussian log likelihood is given as follows:

$$\log l(y_{t_1}, \ldots, y_{t_T}; \Theta) = \sum_{t=1}^{T} \left(-\frac{N_y}{2} log(2\pi) - \frac{1}{2} log|\hat{F}_{t_i}| - \frac{1}{2} v'_{t_i} \hat{F}_{t_i}^{-1} v_{t_i} \right) \ .$$

The Kalman-Filter algorithm works by minimizing the prediction error.

4.2.3 AFNS estimated parameters

The AFNS parameters are estimated using the described above Kalman Filter algorithm and they are based on daily Canadian yield curves obtained from Canadian zero-coupon bonds from 1991-01-02 till 2017-10-31. The estimates are obtained by considering yield curves from different maturities: 3 months, 6 months, 9 months, 1 year, 18 months, 2 years, 3 years, 4 years, 5 years, 7 years, 8 years, 9 years, 10 years, 15 years,

20 years and 30 years. The data was extracted from the Bank of Canada website[1]. The estimated parameters are summarized below:

Table 4.1: Estimated parameters for the AFNS model

$k_{11}^{\mathbb{P}}$	0.011059202
$k_{22}^{\mathbb{P}}$	0.344903793
$k_{33}^{\mathbb{P}}$	0.792906078
$\theta_1^{\mathbb{P}}$	0.072266054
$\theta_2^{\mathbb{P}}$	-0.027039646
$\theta_3^{\mathbb{P}}$	-0.017515206
$\sigma_{Y^{(1)}}^{\mathbb{P}}$	0.005513094
$\sigma_{Y^{(2)}}^{\mathbb{P}}$	0.011397127
$\sigma_{Y^{(3)}}^{\mathbb{P}}$	0.019603447
λ	0.359347873

The estimated parameters for the independent state variables AFNS model provide a good fit to the data. The boxplots of the forecast errors v_t show that the errors tend to be null which is evidence of good fit. Shown in Figure 4.2 are the boxplots of the errors for a few maturities. As observed for all maturities, the median of the errors are near zero, and the range is also small. Moreover, the boxplot is symmetric so that errors revolve around zero. As shown in Figure 4.1, outliers come mostly from the first observations because the Kalman Filter Algorithm assumes an initial value for the yield curve and it then adapts itself to subsequent values for the observed yields. As more observations are fed to the algorithm, prediction errors diminish and stabilize themselves.

[1]https://www.bankofcanada.ca/rates/interest-rates/bond-yield-curves/

Figure 4.1: Prediction errors for yields of maturity 30-years

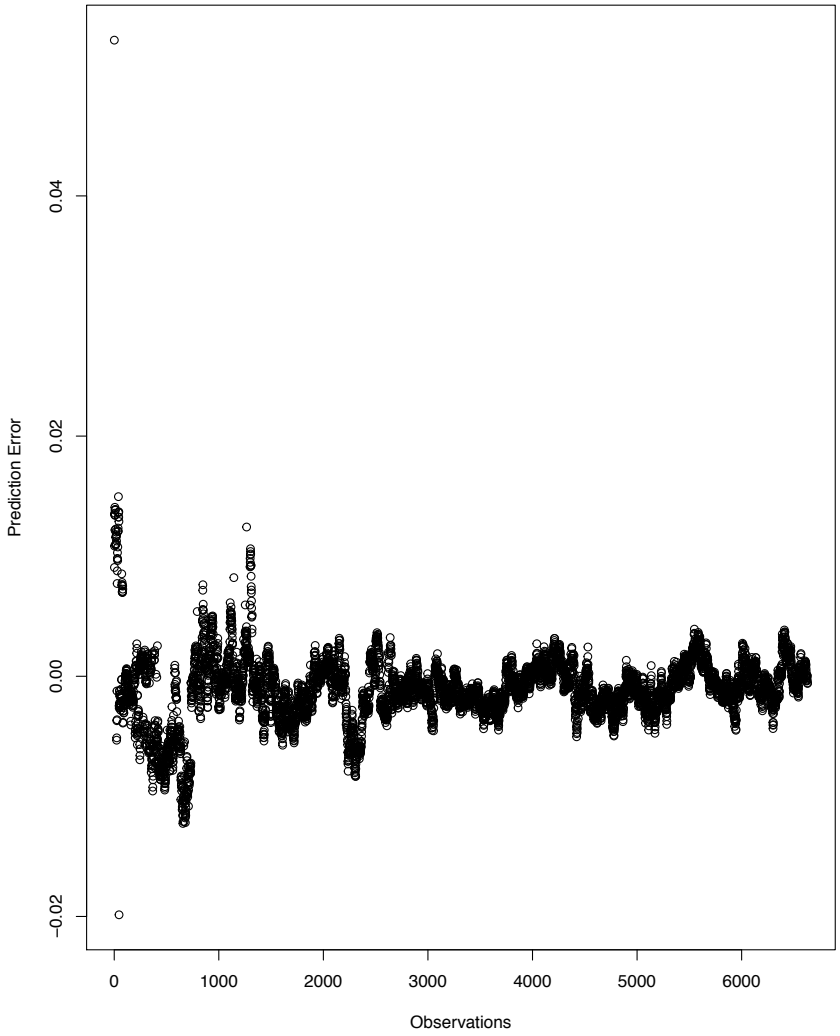
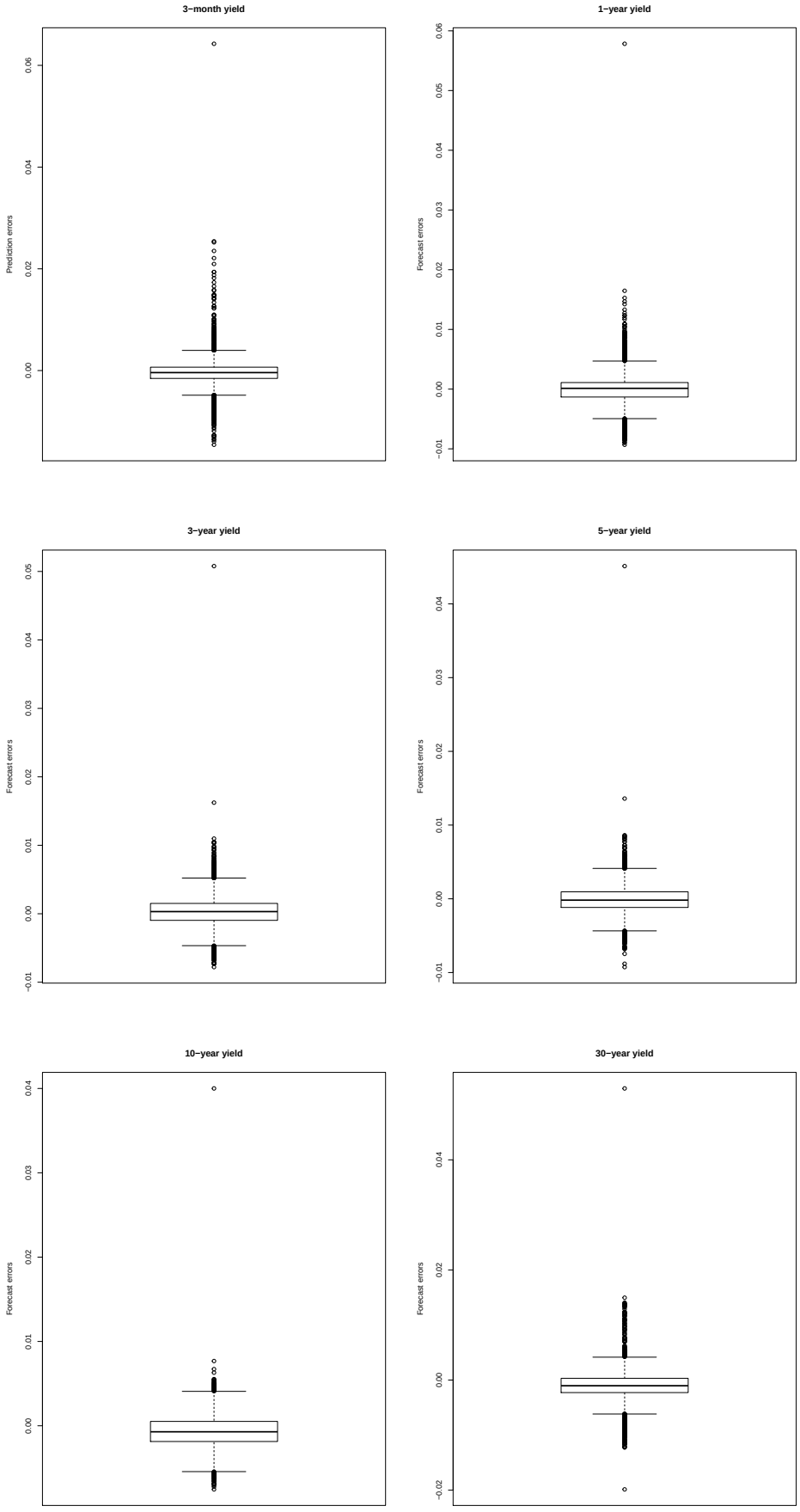


Figure 4.2: Boxplots of the forecast errors for yields of maturities 3 months, 1 year, 3 years, 5 years, 10 years and 30 years



Chapter 5

Numerical Results

In Chapter 2, the MRT risk decomposition was introduced as a meaningful risk decomposition and it has been derived under an insurance setting for an annuity and life insurance portfolios. It is assumed that the risks are generated by the state variable X associated with the sub-filtration $\mathbb{G}$ and, by the random death process associated with the sub-filtration $\mathbb{I}$. $\mathbb{G}$ represents the information generated by the fluctuation of the implicated rates, so that $X(t) = \left(Y_t^{(1)}, Y_t^{(2)}, Y_t^{(3)}, m(t)\right)$ correspond respectively to the level of the yield curve, the slope of the yield curve, the curvature of the yield curve, and finally the systematic mortality rate. The yield curve is assumed to follow the AFNS model and, the mortality rate is assumed to follow the Lee-Carter model. This chapter presents the numerical results obtained by applying the MRT risk decomposition to the annuity and the insurance portfolios.

5.1 MRT Decomposition

In Chapter 2, the MRT Decomposition was derived for both an annuity portfolio and an insurance portfolio. For a given portfolio L, the total risk represented by $L_0 - E[L_0]$ can be decomposed in the following manner:

$$L_0 - E[L_0] = R_{Y^{(1)}} + R_{Y^{(2)}} + R_{Y^{(3)}} + R_m + R_N \ , \tag{5.1.1}$$

where the risks $\{R_{Y^{(1)}}, R_{Y^{(2)}}, R_{Y^{(3)}}, R_m, R_N\}$ correspond respectively to the risks associated with the level of the yield curve, the slope of the yield curve, the curvature of the yield curve, the systematic mortality rate, and the unsystematic mortality rate. Their values are summarized below according to the given portfolio.

5.1.1 Annuity portfolio

For an annuity portfolio of the form $L_0 = \sum_{j=1}^{T}(P_0 - N(t_j))e^{-\int_0^{t_j} r(s)ds}$ as in 1.1.1, the risks are given by:

$$R_{Y^{(1)}} = \sum_{j=1}^{T}\int_0^{t_j}(P_0 - N(t-))e^{-\int_0^t r(s,X(s))ds}\frac{\partial f^{A_2}}{\partial Y^{(1)}}(t,X(t))dM_{Y^{(1)}}(t)\ , \tag{5.1.2}$$

$$R_{Y^{(2)}} = \sum_{j=1}^{T}\int_0^{t_j}(P_0 - N(t-))e^{-\int_0^t r(s,X(s))ds}\frac{\partial f^{A_2}}{\partial Y^{(2)}}(t,X(t))dM_{Y^{(2)}}(t)\ , \tag{5.1.3}$$

$$R_{Y^{(3)}} = \sum_{j=1}^{T}\int_0^{t_j}(P_0 - N(t-))e^{-\int_0^t r(s,X(s))ds}\frac{\partial f^{A_2}}{\partial Y^{(3)}}(t,X(t))dM_{Y^{(3)}}(t)\ , \tag{5.1.4}$$

$$R_m = \sum_{j=1}^{T}\int_0^{t_j}(P_0 - N(t-))e^{-\int_0^t r(s,X(s))ds}\frac{\partial f^{A_2}}{\partial m}(t,X(t))dM_m(t)\ , \tag{5.1.5}$$

and

$$R_N = -\sum_{j=1}^{T}\int_0^{t_j} e^{-\int_0^t r(s,X(s))ds} f^{A_2}(t,X(t))dM^N(t)\ . \tag{5.1.6}$$

Since the state process $X(t)$ is assumed to possess interdependent components with the volatility matrix being

$$\sigma\left(t, Y^{(1)}, Y^{(2)}, Y^{(3)}, m(t)\right) = diag\{\sigma_{Y^{(1)}}, \sigma_{Y^{(2)}}, \sigma_{Y^{(3)}}, b_x\sigma m(t)\}$$

and recalling that $dM(t) = \sigma(t)dW(t)$, then

$$dM_{Y^{(1)}}(t) = \sigma_{Y^{(1)}}dW_{Y^{(1)}}(t)\ , \tag{5.1.7}$$

$$dM_{Y^{(2)}}(t) = \sigma_{Y^{(2)}} dW_{Y^{(2)}}(t) \ , \tag{5.1.8}$$

$$dM_{Y^{(3)}}(t) = \sigma_{Y^{(3)}} dW_{Y^{(3)}}(t) \ , \tag{5.1.9}$$

$$dM_m(t) = b_x \sigma m(t) dW_m(t) \ . \tag{5.1.10}$$

Recalling Proposition 1,

$$dM^N(t) = dN(t) - (P_0 - N(t))m(t)dt \ . \tag{5.1.11}$$

In addition, developing (2.2.19) gives

$$f^{A_2}(t,\omega) = E[e^{-\int_t^{t_j} m(s,X(s))ds} e^{-\int_t^{t_j} r(s,X(s))ds} \mid X(t) = \omega]$$

$$= E[e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega] \ E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid X(t) = \omega] \ ,$$

since $r(s, X(s))$ and $m(s, X(s))$ are independent. Using (4.1.5) gives

$$f^{A_2}(t,\omega) = e^{B^{(1)}(t,t_j)Y^{(1)}(t) + B^{(2)}(t,t_j)Y^{(2)}(t) + B^{(3)}(t,t_j)Y^{(3)}(t) + A(t,t_j)} \ E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid X(t) = \omega] \ .$$

However, since there is no analytical solution for $E[e^{-\int_t^{t_j} m(s,X(s))ds} | X(t) = \omega]$, Monte-Carlo simulations are used for the purpose of its calculation.

Moreover,

$$\frac{\partial f^{A_2}}{\partial Y^{(i)}}(t,\omega)) = B^{(i)}(t,t_j) \ e^{B^{(1)}(t,t_j)Y^{(1)}(t) + B^{(2)}(t,t_j)Y^{(2)}(t) + B^{(3)}(t,t_j)Y^{(3)}(t) + A(t,t_j)}$$
$$\times E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid X(t) = \omega] \ ,$$

for $i = 1, 2, 3$, and

$$\frac{\partial f^{A_2}}{\partial m}(t,\omega)) = e^{B^{(1)}(t,t_j)Y^{(1)}(t) + B^{(2)}(t,t_j)Y^{(2)}(t) + B^{(3)}(t,t_j)Y^{(3)}(t) + A(t,t_j)}$$
$$\times \frac{\partial E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid X(t) = \omega]}{\partial m} \ .$$

Furthermore, approximating with the finite difference method yields

$$\frac{\partial E[e^{-\int_t^{t_j} m(s,X(s))ds}|X(t)=\omega]}{\partial m} = \frac{\partial E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid m(t,X(t))=\omega_m]}{\partial m}$$

$$\approx \lim_{h\to 0} \frac{E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid m(t,X(t))=\omega_m+h] - E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid m(t,X(t))=\omega_m]}{h}. \tag{5.1.12}$$

5.1.2 Life insurance portfolio

For a life insurance portfolio of the form $L_0 = \sum_{j=1}^{T}(P_0 - N(t_{j-1}))e^{-\int_0^{t_j} r(s)ds} - \sum_{j=1}^{T}(P_0 - N(t_j))e^{-\int_0^{t_j} r(s)ds}$, the risks are given by:

$$\begin{aligned} R_{Y^{(1)}} &= \sum_{j=1}^{T}\left(\int_0^{t_{j-1}} (P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_1}}{\partial Y^{(1)}}(t,X(t))dM_{Y^{(1)}}(t)\right) \\ &+ \sum_{j=1}^{T}\left(\int_{t_{j-1}}^{t_j} (P_0 - N(t_{j-1}))e^{-\int_0^t r(s)ds}\frac{\partial f^{B_1}}{\partial Y^{(1)}}(t,X(t))dM_{Y^{(1)}}(t)\right) \\ &- \sum_{j=1}^{T}\left(\int_0^{t_j} (P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_2}}{\partial Y^{(1)}}(t,X(t))dM_{Y^{(1)}}(t)\right), \end{aligned} \tag{5.1.13}$$

$$\begin{aligned} R_{Y^{(2)}} &= \sum_{j=1}^{T}\left(\int_0^{t_{j-1}} (P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_1}}{\partial Y^{(2)}}(t,X(t))dM_{Y^{(2)}}(t)\right) \\ &+ \sum_{j=1}^{T}\left(\int_{t_{j-1}}^{t_j} (P_0 - N(t_{j-1}))e^{-\int_0^t r(s)ds}\frac{\partial f^{B_1}}{\partial Y^{(2)}}(t,X(t))dM_{Y^{(2)}}(t)\right) \\ &- \sum_{j=1}^{T}\left(\int_0^{t_j} (P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_2}}{\partial Y^{(2)}}(t,X(t))dM_{Y^{(2)}}(t)\right), \end{aligned} \tag{5.1.14}$$

$$\begin{aligned} R_{Y^{(3)}} &= \sum_{j=1}^{T}\left(\int_0^{t_{j-1}} (P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_1}}{\partial Y^{(3)}}(t,X(t))dM_{Y^{(3)}}(t)\right) \\ &+ \sum_{j=1}^{T}\left(\int_{t_{j-1}}^{t_j} (P_0 - N(t_{j-1}))e^{-\int_0^t r(s)ds}\frac{\partial f^{B_1}}{\partial Y^{(3)}}(t,X(t))dM_{Y^{(3)}}(t)\right) \\ &- \sum_{j=1}^{T}\left(\int_0^{t_j} (P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_2}}{\partial Y^{(3)}}(t,X(t))dM_{Y^{(3)}}(t)\right), \end{aligned} \tag{5.1.15}$$

$$
\begin{aligned}
R_m = &\sum_{j=1}^{T}\left(\int_0^{t_{j-1}}(P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_1}}{\partial m}(t,X(t))dM_m(t)\right)\\
&+\sum_{j=1}^{T}\left(\int_{t_{j-1}}^{t_j}(P_0 - N(t_{j-1}))e^{-\int_0^t r(s)ds}\frac{\partial f^{B_1}}{\partial m}(t,X(t))dM_m(t)\right) \qquad (5.1.16)\\
&-\sum_{j=1}^{T}\left(\int_0^{t_j}(P_0 - N(t-))e^{-\int_0^t r(s)ds}\frac{\partial f^{A_2}}{\partial m}(t,X(t))dM_m(t)\right)\ ,
\end{aligned}
$$

$$
R_N = -\sum_{k=1}^{T}\int_{0^+}^{t_{j-1}} e^{-\int_0^t r(s)ds} f^{A_1} dM^N(t) + \sum_{k=1}^{T}\int_{0^+}^{t_j} e^{-\int_0^t r(s)ds} f^{A_2} dM^N(t)\ . \qquad (5.1.17)
$$

The compensated processes $dM(t) = \{dM_{Y^{(1)}}(t), dM_{Y^{(2)}}(t), dM_{Y^{(3)}}(t), dM_m(t), dM^N(t)\}$ are the same as described in 5.1.1. Moreover,

$$
\begin{aligned}
f^{A_1}(t,\omega) &= E[e^{-\int_t^{t_{j-1}} m(s,X(s))ds} e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega]\\
&= e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)}\\
&\times E[e^{-\int_t^{t_{j-1}} m(s,X(s))ds}]\ ,
\end{aligned}
$$

$$
\begin{aligned}
f^{A_2}(t,\omega) &= E[e^{-\int_t^{t_j} m(s,X(s))ds} e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega]\\
&= e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)}\\
&\times E[e^{-\int_t^{t_j} m(s,X(s))ds}]\ ,
\end{aligned}
$$

$$
\begin{aligned}
f^{B_1}(t,\omega) &= E[e^{-\int_t^{t_j} r(s,X(s))ds} | X(t) = \omega]\\
&= e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)}\ .
\end{aligned}
$$

The respective derivatives are given by

$$
\begin{aligned}
\frac{\partial f^{A_1}}{\partial Y^{(i)}}(t,\omega)) =& B^{(i)}(t,t_j)\ e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)}\\
&\times E[e^{-\int_t^{t_{j-1}} m(s,X(s))ds}]\ ,
\end{aligned}
$$

$$\frac{\partial f^{A_1}}{\partial m}(t,\omega)) = e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)} \times \frac{\partial E[e^{-\int_t^{t_{j-1}} m(s,X(s))ds} \mid X(t)=\omega]}{\partial m} ,$$

$$\frac{\partial f^{A_2}}{\partial Y^{(i)}}(t,\omega)) = B^{(i)}(t,t_j)\ e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)} \times E[e^{-\int_t^{t_j} m(s,X(s))ds}] ,$$

$$\frac{\partial f^{A_2}}{\partial m}(t,\omega)) = e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)} \times \frac{\partial E[e^{-\int_t^{t_j} m(s,X(s))ds} \mid X(t)=\omega]}{\partial m} ,$$

$$\frac{\partial f^{B_1}}{\partial Y^{(i)}}(t,\omega)) = B^{(i)}(t,t_j)\ e^{B^{(1)}(t,t_j)Y^{(1)}(t)+B^{(2)}(t,t_j)Y^{(2)}(t)+B^{(3)}(t,t_j)Y^{(3)}(t)+A(t,t_j)} ,$$

and

$$\frac{\partial f^{B_1}}{\partial m}(t,\omega)) = 0 ,$$

for $i = 1, 2, 3$, where $\frac{\partial E[e^{-\int_t^{t_j} m(s,X(s))ds} | X(t)=\omega]}{\partial m}$ is given by (5.1.12), for $j \geq 1$.

5.2 Numerical results

5.2.1 Euler's principle

Once the decomposition of the risk has been achieved and the corresponding empirical distribution functions of each factor have been obtained, the next step then consists of quantifying and allocating a homogeneous risk measure ρ to the different sources of risk and determining the contribution of each risk factor with respect to the total risk. Under

the Euler principle, the contributions are estimated using the following equation:

$$\rho(\bar{L}) = \sum_{i=1}^{5} \frac{\partial \rho}{\partial h} \left(\bar{L} + \bar{R}_i h \right) \ , \tag{5.2.1}$$

where $\rho(\bar{L})$ represents the risk measure for the total risk $\bar{L}$, and each summand is interpreted as the risk contribution of the respective risk factor with respect to the total risk. See Tasche (2008) for more details. Three risk measures will be used: the standard deviation, the 99% Value-at-Risk ($VaR_{0.99}$), and the 99% Tail-Value-at-Risk ($TVaR_{0.99}$).

5.2.2 Simulations

The steps in producing the simulated results are outlined here. First, set the number of simulations to 100,000 and the number of steps within a year is equal to 10. The latter is to be used within an Euler scheme context to project the rates $r(t)$ and $m(t)$ as well as to approximate the stochastic integrals. As basic scenario for point of reference, let the initial number of policyholders to be 100 and their age to be 65. Notice that as a result from the above described Lee-Carter model design, the age limit of the policyholders is set to be 109. Consider also instead the risk proportional to the number of policyholders, i.e. $\bar{L} = \frac{L}{P_0}$. Each simulation is proceeded as follows:

1. First, simulate the paths of the mortality index $\kappa(t)$ in order to simulate $m(t)$ using (3.4.2) and (3.4.6). Once $m(t)$ is simulated, ${}_tp_{x_i}$ can then be calculated using (1.0.2). The counting process of the number of deaths within each year, $N(t) - N(t-1)$, can be simulated as well since $N(t) - N(t-1) \sim binomial(p_{t-1} \ , \ q_{x+t-1})$. Similarly, the paths of the yield curve factors $(Y_t^{(1)}, Y_t^{(2)}, Y_t^{(3)})$ are simulated using (4.1.11), with r_t being, under the real world probability $\mathbb{P}$, the sum of the first two factors as in (4.1.3). Moreover, bond prices are given in (4.1.5) with parameters $(B^{(1)}(t,T), B^{(2)}(t,T), B^{(3)}(t,T), A(t,T))$ being deterministic and given in detail in Subsection 4.1.3.

2. All required variables are now simulated and compute next the MRT decomposition according to the desired portfolio. Section 5.1 gives detail to the formulas to compute

with Subsections 5.1.1 and 5.1.2 corresponding respectively to the annuity portfolio and the life insurance portfolio.

5.2.3 Numerical results: Annuity portfolios

Consider the base scenario to be an annuity portfolio for $P_0 = 100$ policyholders of age 65. This scenario is typical of a pension plan context. The resulting empirical distributions of the total risk along with its respective risk factors are shown in Figure 5.1. The plots indicate that the level of the yield curve is central to the total risk since the distribution of its risk factor approaches and represents the most to the distribution of the total risk. In fact, this is logical because the level of the yield curve is responsible for the long term yield factor and annuities span over long term periods.

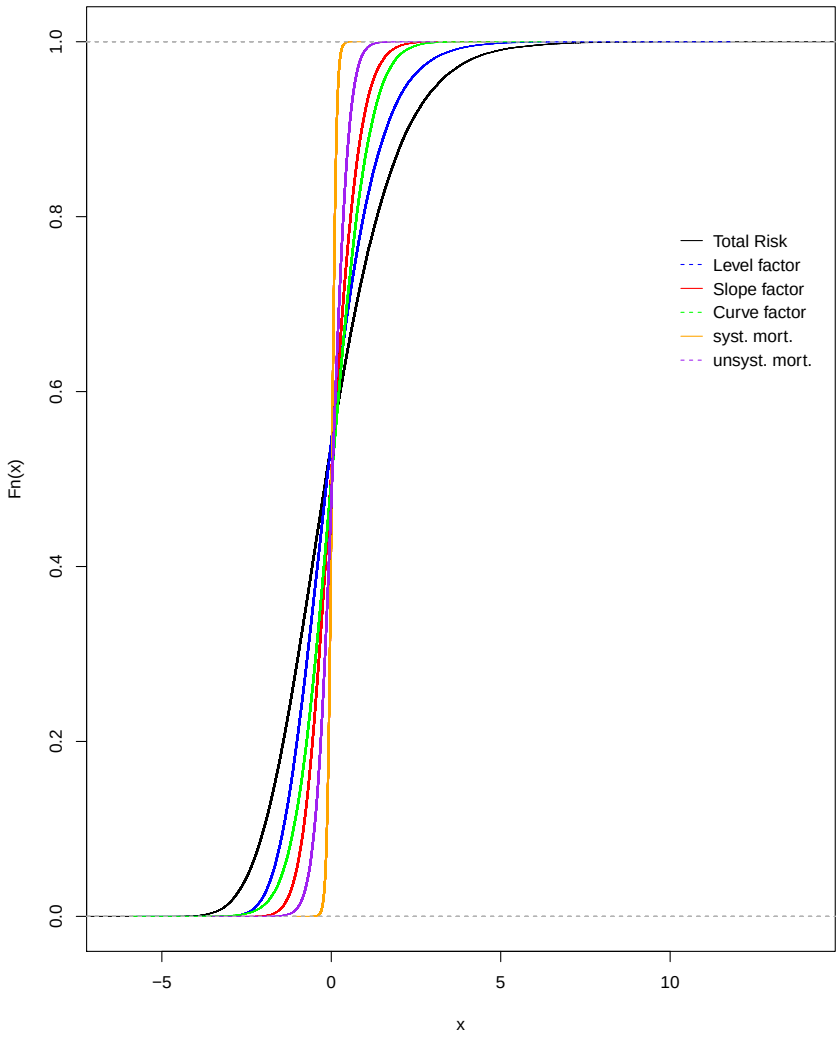
Using Euler's principle, the allocated risk contributions of each factor are shown in Table 5.1, along with their percentage of their contributions to the total risk. L^* represents the total risk as simulated directly using (1.1.1), whereas $\bar{L}$ is obtained using the MRT decomposition as in (2.2.18). Since both values are similar for the three risk measures, then the MRT decomposition does indeed add up to the total risks and the aggregation property of a meaningful risk decomposition is thus satisfied. The differences might be explained by error approximations such as the calculations of stochastic integrals.

Table 5.1: The risk contributions for an annuity portfolio issued at age 65

	L^*	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	1.7276	1.7153	0.87	0.2726	0.4645	0.008	0.1002
			50.7%	15.9%	27.1%	0.5%	5.8%
$\text{VaR}_{0.99}$	5.0384	5.0304	3.3002	0.8929	0.9375	-0.1175	0.0173
			65.6%	17.8%	18.6%	-2.3%	0.3%
$\text{TVaR}_{0.99}$	6.2077	6.0833	3.405	0.9524	1.298	0.0341	0.3938
			56.0%	15.7%	21.3%	0.6%	6.5%

Table 5.1 confirms the results given by Figure 5.1. The allocated risk contributions show that in total, the interest rate makes up for around 93% of the total risk, with the level factor accounting for half of the risk. This means that changes in the interest

Figure 5.1: Empirical CDF of the risks associated with a whole life annuity portfolio issued at age 65



rate are critical because they affect directly the premium pricing of the annuity and the discounting effect takes even greater importance over long periods. The systematic mortality risk suggests that it has a small impact on the total risk and that this risk is rather deterministic. This may be explained by the fact that even with the improvement of life expectancy, human lifespan is by default limited and the decrease in the systematic mortality rate cannot compensate for the high unsystematic risk presented in the old ages range. Let now compare these results as the age of the policyholders varies for the annuity portfolio.

5.2.4 Ages in the annuity portfolio

Table 5.2: The risk contributions for an annuity portfolio issued at age 55

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	2.6570	1.585	0.3429	0.6622	0.0058	0.061
		59.7%	12.9%	24.9%	0.2%	2.3%
$\text{VaR}_{0.99}$	7.8395	5.5615	0.6986	1.6663	-0.1624	0.0755
		70.9%	8.9%	21.3%	-2.1%	1.0%
$\text{TVaR}_{0.99}$	10.5360	7.0185	1.2363	1.9631	0.0356	0.2825
		66.6%	11.7%	18.6%	0.3%	2.7%

Table 5.3: The risk contributions for an annuity portfolio issued at age 75

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.9715	0.3707	0.1761	0.2482	0.0105	0.166
		38.2%	18.1%	25.5%	1.1%	17.1%
$\text{VaR}_{0.99}$	2.5942	1.0345	0.4659	0.5859	0.034	0.474
		39.9%	18.0%	22.6%	1.3%	18.3%
$\text{TVaR}_{0.99}$	3.1648	1.3243	0.5779	0.6776	0.034	0.551
		41.8%	18.3%	21.4%	1.1%	17.4%

Tables 5.2, 5.3, and 5.4 show that as the ages of policyholders in the portfolio increases, the level factor takes less importance in the risk contributions. On the one hand, this is intuitive because as age increases, the coverage period for the policyholders also decreases before their eventual death and annuities span over shorter periods. On the other hand, the unsystematic risk also takes more importance in the contribution of the risk because as age increases, policyholders are also more prone to die and thus, they are more exposed to that risk. Notice that the overall risk also decreases because the event of their death and consequently the end of their annuity policy is also more certain and less variable. However, the systematic mortality risk does not seem again to have any impact on the portfolio risk. This might be explained by the same argument as in Subsection 5.2.3.

Table 5.4: The risk contributions for an annuity portfolio issued at age 85

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.4915	0.0994	0.0752	0.0794	0.0067	0.2308
		20.2%	15.3%	16.2%	1.4%	47.0%
$\text{VaR}_{0.99}$	1.3035	0.3216	0.1355	0.2258	0.0031	0.6176
		24.7%	10.4%	17.3%	0.2%	47.4%
$\text{TVaR}_{0.99}$	1.4996	0.3396	0.2447	0.2293	0.024	0.662
		22.6%	16.3%	15.3%	1.6%	44.1%

5.2.5 Number of policyholders in an annuity portfolio

Let increase the number of policyholders in a given portfolio to 1,000 in order to see the effects of the number of policyholders on the risks. Tables 5.5 and 5.6 show the results. Just as expected, the unsystematic risk is diversifiable and decreases as the number of policyholders increases.

Table 5.5: The risk contributions for an annuity portfolio issued at age 65 for 1,000 policyholders

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	1.6761	0.8969	0.2812	0.4788	0.0069	0.0123
		53.5%	16.8%	28.6	0.4%	0.7%
$\text{VaR}_{0.99}$	4.7026	3.2163	0.7195	0.9826	-0.1225	-0.0932
		68.4%	15.3%	20.9%	-2.6%	-2.0%
$\text{TVaR}_{0.99}$	5.8448	3.4101	0.9425	1.3934	0.0402	0.0586
		58.3%	16.1%	23.8%	0.7%	1.0%

5.2.6 Deferred annuities

Another typical example of annuities is the deferred annuity. This type of annuities is especially common in a pension context where participants can usually start to receive their benefits at age 65. Tables 5.7 and 5.8 show results for a 15-year deferred annuity issued at age 50, and a 35-year deferred annuity issued at age 30, respectively. As the deferred period increases, the contribution of the level factor also increases since payments

Table 5.6: The risk contributions for an annuity portfolio issued at age 85 for 1,000 policyholders

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.3727	0.1299	0.0976	0.1051	0.0084	0.0317
		34.9%	26.2%	28.2%	2.3%	8.5%
$\text{VaR}_{0.99}$	0.9737	0.3307	0.2808	0.252	0.0218	0.0885
		34.0%	28.8%	25.9%	2.2%	9.1%
$\text{TVaR}_{0.99}$	1.1333	0.4089	0.2847	0.2974	0.0267	0.1155
		36.1%	25.1%	26.2%	2.4%	10.2%

start at a later date for the same period of coverage. In response, the slope factor and the curve factor, which are responsible for shorter terms, has less importance to the overall risk. Mortality risks also decrease because as the deferred period is extended, the number of policyholders that die during this period is also more probable. However, those are not covered yet and when the annuity actually starts, fewer policyholders will remain and the total number of policyholders exposed to the unsystematic risk will, therefore, become smaller.

Table 5.7: The risk contributions for a 35-year deferred annuity portfolio issued at age 30

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	3.3064	2.7275	0.1711	0.3827	0.005	0.0201
		82.5%	5.2%	11.6%	0.2%	0.6%
$\text{VaR}_{0.99}$	12.782	10.2636	0.9827	1.6422	-0.0842	-0.0223
		80.3%	7.7%	12.8%	-0.7%	-0.2%
$\text{TVaR}_{0.99}$	21.5002	18.0892	1.2013	2.0291	0.064	0.1167
		84.1%	5.6%	9.4%	0.3%	0.5%

5.2.7 Numerical results: Life insurance

Next, let consider life insurance portfolios. Similarly as before, set the reference portfolio to be such that the number of policyholders is $P_0 = 100$, aged 65 as well. The resulting empirical distributions of the total risk along with its respective risk factors are shown in Figure 5.2. Like the annuity portfolio, the plots indicate that the level of the

Table 5.8: The risk contributions for a 15-year deferred annuity portfolio issued at age 50

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	2.5141	1.7091	0.2278	0.5338	0.004	0.0394
		68.0%	9.1%	21.2%	0.2%	1.6%
$\text{VaR}_{0.99}$	8.3896	6.0098	0.8154	1.4415	-0.0742	0.1971
		71.6%	9.7%	17.2%	-0.9%	2.3%
$\text{TVaR}_{0.99}$	11.3055	8.1495	0.9474	1.9786	0.0265	0.2036
		72.1%	8.4%	17.5%	0.2%	1.8%

yield curve is central to the total risk. Again, the systematic mortality risk appears to have a very minimal impact on the total risk. Notice that the risk is much less than that of an annuity portfolio because the life insurance portfolio does not require periodic payments over time but consists only of one single benefit premium. In an annuity portfolio, the risk is accumulated with each periodic payment. Table 5.9 validates the observation made on Figure 5.2 and shows the respective risk contributions. L^* represents the total risk as simulated directly through (1.1.2) and with some approximation errors, the MRT decomposition does add up to the total risk.

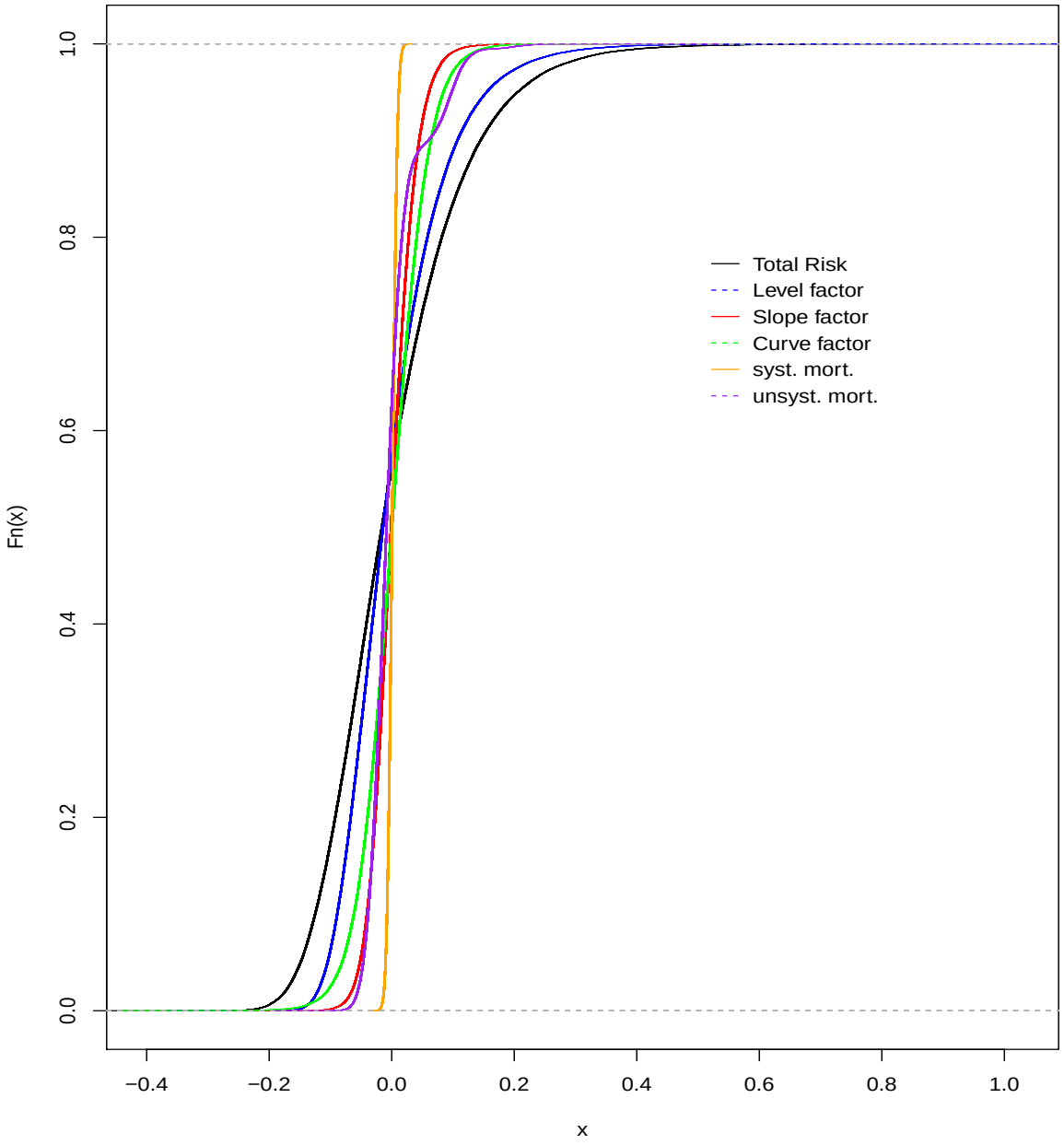
Table 5.9: The risk contributions for a life insurance portfolio issued at age 65

	L^*	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.1168	0.1120	0.0619	0.0109	0.0232	0.0004	0.0157
			55.3%	9.7%	20.7%	0.4%	14.0%
$\text{VaR}_{0.99}$	0.3726	0.3447	0.2252	0.0365	0.0612	0.0003	0.0216
			65.3%	10.6%	17.8%	0.1%	6.3%
$\text{TVaR}_{0.99}$	0.4707	0.4382	0.2858	0.0478	0.0779	0.0003	0.0265
			65.2%	10.9%	17.8%	0.1%	6.0%

5.2.8 Ages in the life insurance

Similarly as in the life annuity portfolio, when age increases, the risk from the level factor is less present in the total risk because the coverage period takes shorter terms since

Figure 5.2: Empirical CDF of the risks associated with a whole life insurance portfolio issued at age 65



the lifespan of the policyholders reduces. Consequently, the slope factor also takes slightly more weight in the total risk. Moreover, an increase in ages also makes the policyholders more prone to death and they are therefore more exposed to the unsystematic mortality risk. Tables 5.10, 5.11 and 5.12 show such trends for portfolios of life insurance for policyholders with age of 55, 75 and 85 years old respectively, where an increase in ages shifts the risks from the level factor toward the unsystematic mortality rate.

Table 5.10: The risk contributions for a life insurance portfolio issued at age 55

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.1193	0.0803	0.0088	0.0207	0.0003	0.0093
		67.3%	7.4%	17.4%	0.3%	7.8%
$\text{VaR}_{0.99}$	0.3834	0.2993	0.0239	0.0687	-0.0048	-0.0037
		78.1%	6.2%	17.9%	-1.3%	-1.0%
$\text{TVaR}_{0.99}$	0.588	0.4427	0.0495	0.0848	0.0003	0.0107
		75.3%	8.4%	14.4%	0.1%	1.8%

Table 5.11: The risk contributions for a life insurance portfolio issued at age 75

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.1006	0.0419	0.0129	0.0233	0.0003	0.0222
		41.7%	12.8%	23.2%	0.3%	22.1%
$\text{VaR}_{0.99}$	0.2783	0.1284	0.0233	0.0619	0.0004	0.0643
		46.1%	8.4%	22.2%	0.1%	23.1%
$\text{TVaR}_{0.99}$	0.3362	0.1567	0.0439	0.0709	0.0004	0.0643
		46.6%	13.1%	21.1%	0.1%	19.1%

5.2.9 Number of policyholders in the life insurance

Again, the unsystematic mortality risk proves to be diversifiable as it tends toward zero as the number of policyholders increases. As a result, the risks associated with the yield curve factors make up for most of the risk contributions. Tables 5.13 and 5.14 show such trends for an increase of policyholders to 1000.

5.2.10 Term Life Insurance

Notice that since the age limit is assumed to be 109 years old, a whole life insurance issued to a policyholder aged 65 years old can then also be considered as a 44 years term life insurance. In Subsection 5.2.8, ages have been varied across the policyholders in order to test the lifespan factor on the MRT decomposition. Indirectly, varying ages also tested the effect of the different coverage period on the MRT decomposition as well. Here, let test the coverage period factor directly by testing the number of years itself covered by

Table 5.12: The risk contributions for a life insurance portfolio issued at age 85

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.0802	0.0221	0.0125	0.0168	0.0001	0.0286
		27.6%	15.6%	20.9%	0.1%	35.7%
$\text{VaR}_{0.99}$	0.2078	0.0579	0.0306	0.0417	0.0001	0.0774
		27.9%	14.7%	20.1%	0.0%	37.2%
$\text{TVaR}_{0.99}$	0.2386	0.0668	0.0354	0.0419	0.0003	0.0941
		28.0%	14.8%	17.6%	0.1%	39.4%

Table 5.13: The risk contributions for a life insurance portfolio issued at age 65 for 1000 policyholders

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.105	0.0664	0.0116	0.0249	0.0005	0.0017
		63.2%	11.0%	23.7%	0.5%	1.6%
$\text{VaR}_{0.99}$	0.3428	0.2264	0.0458	0.0747	-0.005	0.0009
		66.0%	13.4%	21.8%	-1.5%	0.3%
$\text{TVaR}_{0.99}$	0.4249	0.2927	0.0464	0.0837	0.0001	0.0019
		68.9%	10.9%	19.7%	0.0%	0.4%

a term life insurance, issued for 100 policyholders aged 65. Tables 5.15 and 5.16 show that as the number of years covered by the policy decreases, the interest rate exercises less and less influence on the total risk while the unsystematic risk accounts for most of the risk. This is explained because as the number of covered years decreases, the number of deaths also decreases since the time frame has been reduced and consequently, the policyholders are becoming younger and are less prone to die. Thus, the most important source of fluctuation consists of the actual event of deaths occurring. With the time frame decreasing, the discounting effect on the insurance amount becomes smaller and the interest rate factors have fewer impacts on the total risk to the point that they can become deterministic because as the discounting effect reduces, any fluctuations in the interest rate become less influential. Thus, fluctuations in the risk depend mainly on fluctuations of the events of death occurring.

Table 5.14: The risk contributions for a life insurance portfolio issued at age 85 for 1000 policyholders

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.0661	0.0269	0.0151	0.0204	0.0002	0.0036
		40.7%	22.8%	30.9%	0.3%	5.4%
$\text{VaR}_{0.99}$	0.1752	0.0784	0.0423	0.0543	-0.002	0.0021
		44.7%	24.1%	31.0%	-1.1%	1.2%
$\text{TVaR}_{0.99}$	0.1991	0.0891	0.046	0.0583	0.0003	0.0055
		44.8%	23.1%	29.3%	0.2%	2.8%

Table 5.15: The risk contributions for a 10-years term life insurance portfolio issued at age 65

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.0372	0.0007	0.0004	0.0006	0.0003	0.0351
		1.9%	1.1%	1.6%	0.8%	94.4%
$\text{VaR}_{0.99}$	0.1086	0.0012	0.0011	-0.0019	-0.0004	0.1085
		1.1%	1.0%	-1.7%	-0.4%	99.9%
$\text{TVaR}_{0.99}$	0.1352	0.0013	0.0012	0.001	0.0004	0.1313
		1.0%	0.9%	0.7%	0.3%	97.1%

Table 5.16: The risk contributions for a 20-years term life insurance portfolio issued at age 65

	$\bar{L}$ (Total)	$\bar{R}_{Y^{(1)}}$	$\bar{R}_{Y^{(2)}}$	$\bar{R}_{Y^{(3)}}$	$\bar{R}_m$	$\bar{R}_N$
Std	0.0598	0.0119	0.0042	0.0076	0.001	0.0351
		19.9%	7.0%	12.7%	1.7%	58.7%
$\text{VaR}_{0.99}$	0.1695	0.0263	0.0101	0.0141	0.0015	0.1175
		15.5%	6.0%	8.3%	0.9%	69.3%
$\text{TVaR}_{0.99}$	0.2083	0.0308	0.0115	0.016	0.0028	0.1472
		14.8%	5.5%	7.7%	1.3%	70.7%

Conclusion

This thesis provides an application of the MRT decomposition method to allocate the contribution of each source of uncertainty to the total risk in the insurance or annuity portfolios. The risk factors are assumed to be the level yield factor, the slope yield factor, the curve yield factor, the systematic mortality rate, and the unsystematic mortality rate.

Chapter 1 sets up the mathematical framework needed to apply the MRT decomposition to the different portfolios at hand. Chapter 2 introduces the concept of the MRT decomposition and derives explicitly the formulas for the risk decomposition of the annuity and the insurance portfolios.

Chapter 3 and chapter 4 introduces the models to be used and to be calibrated for the systematic mortality rate and for the interest rate respectively. Mortality is assumed to follow the Lee-Carter model since the model reflects the growing improvement in life expectancy. Interest rate is assumed to follow the AFNS model because the model allows the rate to be driven by three factors: the level, the slope, and the curvature factors. Moreover, the model imposes an arbitrage-free environment.

Chapter 5 demonstrates the applicability of the MRT decomposition through numerical examples. Outlined here are some main results. It is observed that for a whole life insurance and for a whole life annuity portfolio issued for policyholders at age 65, the most important source of risk is the level factor of the yield curve because these portfolios span over long periods. Notice that the slope factor is less influential because it is associated with short term yield factor. However, as the coverage period of the policies decreases, it is also observed that the unsystematic mortality risk gains on weight. That is because shorter periods mean that the discounting becomes minimal and any fluctuation

in the interest rate does not affect much to the overall expectation of the liabilities. Thus, risks depend mainly on the actual event of death. It is also observed to be diversifiable where it disappears as the number of policyholders increases. Surprisingly, the MRT risk decomposition shows that the impact of the systematic mortality rate is almost negligible. This may be explained by the fact that human lifespan is limited and that despite the improvement in life expectancy, one cannot simply live forever. In this thesis, the age is limited to 109 years old and the improvement in life expectancy only slows down the death process. Therefore, from a mortality point of view, the unsystematic mortality is much more important than the systematic mortality because the stakes are more about when the policyholders will actually die.

Bibliography

[1] T. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging.* Springer, Berlin. 2004

[2] H. Bühlmann. *Life insurance with stochastic interest rates.* In: Ottaviani G (ed) Financial Risk in Insurance. Springer, Berlin. 1995

[3] J. H. E. Christensen. The Affine Arbitrage-Free Class of Nelson-Siegel Term Strucutre Models, 2010

[4] M. Christiansen and M. Helwich. Some further ideas concerning the interaction between insurance and incestment risks. *Blätter der DGVFM*, 29(2):253-266, 2008.

[5] F. X. Diebold and C. Li. Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics* 130, 337-364, 2006.

[6] F. X. Diebold, G.D.Rudebusch, S.B. Aruoba. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* 131, 309-338, 2006.

[7] G.R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57, 405-443, 2002.

[8] S. H. Friedberg, A. J. Insel, and L. E. Spence. *Linear Algebra.* Pearson, 2003.

[9] J.M. Harrison and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications* 11, 215-260, 1981.

[10] M. Jeanblanc, M. Yor and M. Chesney. *Mathematical Methods for Financial Markets.* Springer, 2009.

[11] R. D. Lee and L. R. Carter. Modeling and forecasting US mortality. *Journal of the American statistical association*, 87(419):659-671, 1992.

[12] C. R. Nelson and A. F. Siegel. Parsimonious Modeling of Yield Curves. *Journal of Business*, 60:473-489, 1987.

[13] G. Parker. Stochastic Analysis of the Interaction Between Investment and Insurance Risks. *North American Actuarial Journal*, 55-84, 1997.

[14] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer Verlag, Berlin, 2005.

[15] K. Schilling, D. Bauer, M. C. Christiansen, and A. Kling. Decomposing life insurance liabilities into risk factors, 2015.

[16] D. Tasche. Capital allocation to business units and sub-portfolios: the euler principle. *arXiv preprint arXiv:0708.2542*, 2007.