On Basis Risk in Mortality CAT Bonds

by

Ruiyun Long

A thesis submitted to
The Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements
of the degree of

Master of Science

Department of Management I.H. Asper School of Business
The University of Manitoba
Winnipeg, Manitoba, Canada
May 2015

Thesis advisor

Xuemiao Hao

Author

Ruiyun Long

On Basis Risk in Mortality CAT Bonds

Abstract

Life re-insurers are exposed to mortality catastrophe risk. Mortality CAT bonds are a tool that can mitigate this risk. However, a key disadvantage of this tool is the existence of population basis risk, which occurs whenever there are differences between reference and insured populations. In this thesis, we propose a method to measure population basis risk of mortality CAT bonds. We consider a fictitious mortality CAT bond based on the mortality rates of two regional populations. We first obtain mortality change indexes by calibrating the MBMM model on these two regional populations. Then we use copula-based semi-parametric models to simulate the serial dependence and interdependence structure simultaneously between two regional mortality change indexes. Finally, we analyze the hedge effectiveness of the bond, from which we are able to quantify the population basis risk. We find that population basis risk decreases under certain circumstances.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my gratitude to my advisor Xuemiao Hao for the useful comments, remarks and engagement through the learning process of this master thesis. Furthermore, I would like to thank my committee, Professors Elif Acar, Jeffrey Pai and Rui Zhou, for their insights, expertise as well for their support on the way. Also, I would like to thank my parents and all the people who have supported me along the way.

Chapter 1

Introduction

This chapter includes the overall introduction of the thesis. First, the motivation of this thesis will be evaluated and will be introduced. After this, some definitions we use will be discussed. Finally, the structure of this thesis is briefly summarized.

1.1 Motivation

There has seen significant improvements in mortality rates in the past century. Figures 1.1 and 1.2 demonstrate these improvements for selected ages for England and Wales (EW), and Scotland (SL). From these plots we can summarize the following facts which may apply to most developed countries (for example, Australia, New Zealand, North America and Western Europe):

1. Mortality rates have declined remarkably at all ages.

2. At particular ages, mortality rates seem to have changed over time with

significant improvements in some decades and almost no improvements in other decades. For instance, for age 25-29, mortality rates for both EW and SL declined remarkably before 1960 (except for year 1918) and then stopped declining.

3. Mortality rates in specific years (such as year 1918 which may be associated with Flu pandemic) has an unexpected extreme increase.

4. Mortality rates for both EW and SL have been related to each other.

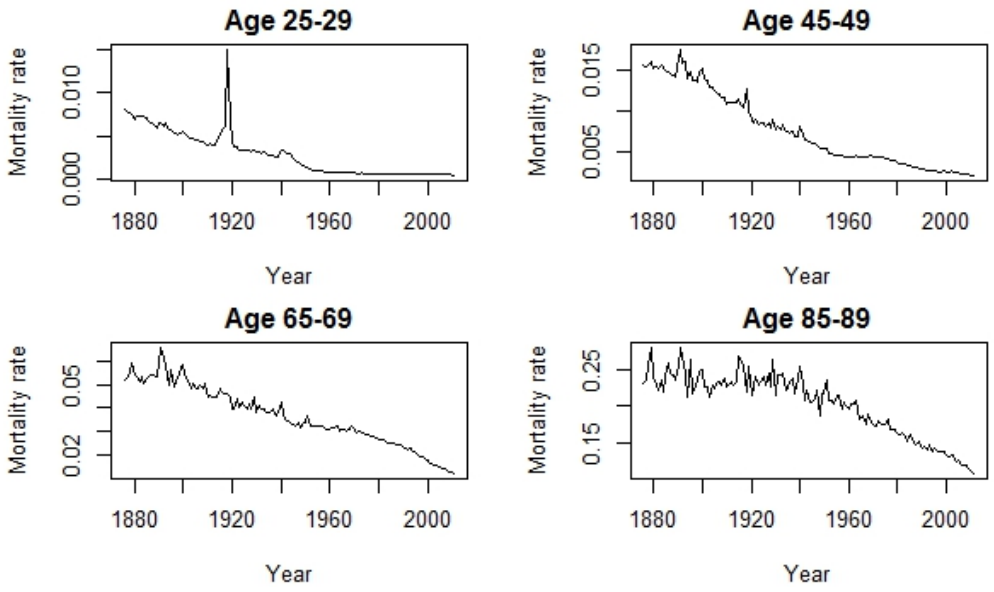


Figure 1.1: England and Wales: Mortality rates at ages 25-29, 45-49, 65-69 and 85-89 during 1876-2011.

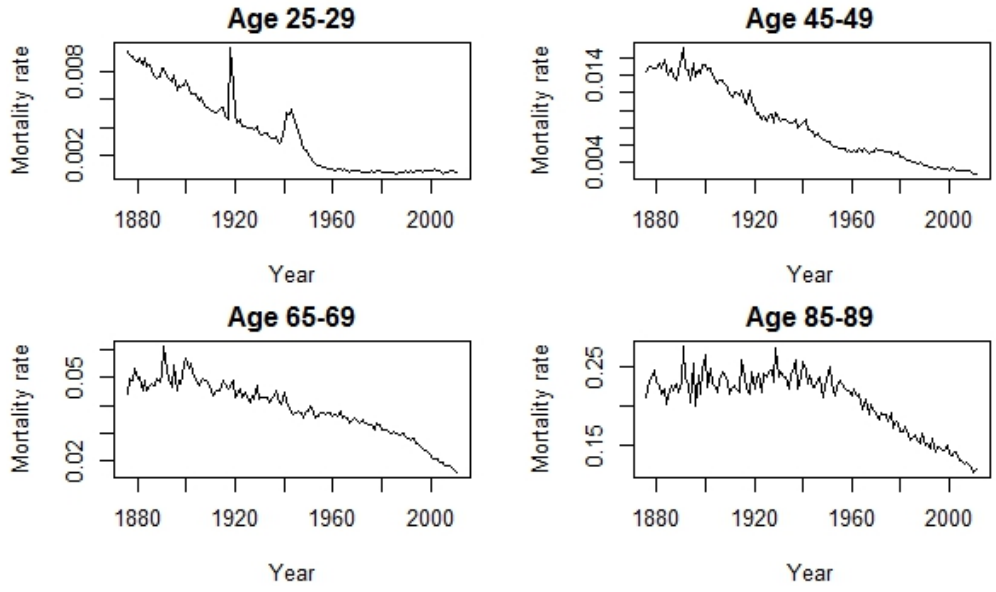


Figure 1.2: Scotland: Mortality rates at ages 25-29, 45-49, 65-69 and 85-89 during 1876-2011.

Many of these facts suggest that there will be considerable uncertainty in the risk exposure of life insurers and re-insurers since the mortality rates vary over time and may have unexpected extreme value under specific situations. In that case, mortality catastrophe bond is an effective method to mitigate mortality risk of a life insurance or reinsurance organization.

However, a key disadvantage of mortality catastrophe bond is the existence of basis risk, which happens because the bond payoff trigger is index based and the actual loss suffered may not be entirely matched by the bond payoff. This thesis aims to quantify the population basis risk of the mortality catastrophe bond using the hedge effectiveness. Population basis risk is the difference between reference and insured populations due to initial or emerging mismatches in age, gender, geographical location and socioeconomic class.

1.2 Definitions

Basis Risk

Offsetting vehicles are usually similar in structure to the investments being hedged, however, they still has differences that cause concern. For example, in the attempt to hedge against a five-year bond with the purchase of Treasury bill futures, there is a risk that the Treasury bill and the bond will not fluctuate identically.

Hedge Effectiveness

Hedge effectiveness is the extent to which a hedge transaction results in offsetting excess losses in mortality claims that the transaction was intended to provide. For instance, a hedge is considered to be highly effective if the excess losses in mortality claims and the hedging derivative (mortality catastrophe bond in this case) offset each other to a significant extent (80% to 125%).

Mortality Catastrophe Bond

Mortality catastrophe bonds (also known as mortality CAT bonds) are risk-linked securities that transfer mortality catastrophe risks from a sponsor to investors.

Mortality Rate

Mortality rate $m(x,t)$ is a measure of the number of deaths. It represents the underlying probability that an individual aged exactly x at time t will die before time $t+1$. The period t to $t+1$ will also be referred to as year t. In this thesis, $m(x,t)$ is the central mortality rate that is simply the ratio of deaths to exposure-to-risk in matched intervals of age x and time t.

1.3 Structure of The Thesis

The analysis of this thesis starts by briefly introducing the mortality catastrophe bond in Chapter 2. In particular, some risks that may cause unexpected extreme mortality rates and the history of mortality catastrophe bond are discussed. Next, Chapter 3 provides an overview of mortality models. Furthermore, the characteristics of the data are discussed in this chapter. Chapter 4 introduces the copula-based semi-parametric models that can take into account interdependence and serial dependence structure of time series simultaneously. This model is based on the assumptions that time series are stationarity and have Markovian structure. In Chapter 5 the forecasting of mortality rates and the population basis risk measuring are introduced. We quantify the population basis risk by analyzing the hedge effectiveness of the mortality catastrophe bond. Finally, in Chapter 6 conclusions are drawn and some possible topics for further research are given.

Chapter 2

Mortality CAT Bonds

This chapter is structured as follows. Section 2.1 briefly summarizes the risks that may cause catastrophic events for the purpose of explaining reasons for a bond issuance and discusses the life insurers' and re-insurers' management of catastrophic mortality risk causing by these events. Section 2.2 provides an overall literature review of mortality catastrophe bonds. Section 2.3 provides an overview of the mortality catastrophe bond markets and the mortality catastrophe bond structures.

2.1 Catastrophic Mortality Risk

A catastrophic event can be defined as "any natural or man-made incident, including terrorism, which results in extraordinary levels of mass casualties, damage, or disruption severely affecting the population, infrastructure, environment,

economy, national morale, and/or government functions."[1] In the context of life insurance and particularly for the purposes of reinsurance, a catastrophic event has a particular meaning or definition that one event causes claims more than an agreed number of lives insured within a given period, usually 24-72 hours. These events include diseases, natural disasters, terrorist attacks, industrial, transport, and other accidents, etc.

2.1.1 Catastrophic Mortality Risk

Disease

Diseases, such as Severe Acute Respiratory Syndrome (SARS) and Acquired Immunodeficiency Syndrome (AIDS), have occurred throughout history and caused a substantial number of deaths. SARS is a respiratory disease induced by the SARS coronavirus, which is an animal virus that crossed the species barrier to humans, see Peiris et al. (2003). The first known cases of SARS happened in Guangdong Province, China, in November 2002, and the last human transmission of SARS happened on July, 2003; see Guan et al. (2003). Over this short period there were 8,096 SARS cases in 26 countries inducing 774 deaths. While it seems that the spread of SARS has been controlled, the possibility of resurgence of a SARS epidemic remains as long as SARS coronavirus-like viruses are still present [2].

Another disease, AIDS, infects cells of the immune system in human, weakening the system and making infected individual more susceptible to other infections. AIDS has been classified as a pandemic by the World Health Organization

[1] `http://www.definitions.net/definition/catastrophicevent`, accessed June 15, 2012.
[2] `http://www.who.int/csr/resources/publications/WHO_CDS_CSR_ARO_2004_1.pdf?ua=1`

(WHO) since its explosion more than four decades ago. Over these four decades, more than 60 million people have been infected and nearly 30 million people have died from it.

Influenza Pandemic

Among these risks that can cause catastrophic events, influenza pandemic is the most common one. An influenza pandemic is an epidemic of an influenza virus that spreads on a worldwide scale and infects a large proportion of the human population. In contrast to the regular seasonal epidemics of influenza, which mostly happen during the fall and winter in temperate regions and all the year round for tropical and sub-tropical regions, the occurrence of influenza pandemics is not constrained by season, see Nguyen-Van-Tam and Hampson (2003), and it is possible that influenza pandemics may occur at any time during the year. Historically, influenza pandemics occurred irregularly, with the 1918 Spanish flu the most serious pandemic in recorded history. Pandemics can induce high levels of mortality, with the 1918 Spanish influenza pandemic estimated as inducing the deaths of approximately 50 million people across the world.

Natural Disasters

A natural disaster is an event induced by nature, the scale of which results in remarkable disruption and loss of human lives [3]. These natural disasters include cyclones, droughts, earthquakes, extreme temperatures, floods, landslides, tsunamis, volcanic activity, and wildfires. Table 2.1 depicts the 20 worst natural disasters with the highest death tolls since 1900.

Terrorist Attacks

[3]Centre for Research on the Epidemiology of Disasters, 2011

Year	**Location**	**Type of Natural Disaster**	**Estimated Number of Deaths**
1931	China	Flood	3,700,000
1928	China	Drought	3,000,000
1959	China	Flood	2,000,000
1943	Bangladesh	Drought	1,900,000
1942	India	Drought	1,500,000
1965	India	Drought	1,500,000
1900	India	Drought	1,250,000
1921	Soviet Union	Drought	1,200,000
1939	China	Flood	500,000
1920	China	Drought	500,000
1970	Bangladesh	Tropical cyclone	300,000
1983	Ethiopia	Drought	300,000
1976	China	Earthquake	242,000
2004	South East Asia	Earthquake and tsunami	230,000
2010	Haiti	Earthquake	222,570
1927	China	Earthquake	200,000
1920	China	Earthquake	180,000
1983	Sudan	Drought	150,000
1923	Japan	Earthquake	143,000
1935	China	Flood	142,000

Table 2.1: Twenty Worst Natural Disasters Since 1900
Source: Center for Research on the Epidemiology of Disasters

There is no agreement on the definition of terrorism, see Zeidan (2004), but one definition from the United States is that terrorism is premeditated, politically motivated violence perpetrated against noncombatant targets by sub-national groups or clandestine agents [4].

Although the data on terrorist attacks are mostly incomplete, we can still find that the deadliest terrorist attack in history (September 11 terrorist attacks in 2001) caused the deaths of nearly 3,000 people. There have been many other terrorist attacks, such as hijackings and bombings, which have resulted in hundreds of deaths. Table 2.2 lists the ten deadliest terrorist attacks in recent times.

Obviously, more severe terrorist attacks such as those involving the use of biological weapons have the potential to cause a substantial number of deaths.

2.1.2 Catastrophic Mortality Risk Management

For most life insurers, death benefit products constitute the majority of their risk business and in result, they are possibly to suffer a remarkable loss from those catastrophic events. Therefore, these events have the potential to severely affect the life insurers and may lead to breaches of solvency requirements.

In those cases, life insurers can transfer mortality risk through reinsurance, which is a form of agreement that can reduce the life insurers' possibility of having to pay a large obligation resulting from insurance claims. Generally, the reinsurance company receives part of a larger potential obligation in exchange for specific amount of the money. However, this agreement may pose unexpected extreme

[4]National Counterterrorism Center

Date	**Location**	**Terrorist Group**	**Estimated Number of Deaths**
Sep-11-2001	United States: NYC and Washington	Al-Qaeda	3,000
Apr-13-1994	Rwanda: Gikoro	Hutus	1,180
Mar-21-2004	Nepal: Bedi	Communist Party of Nepal-Maoist	518
Aug-14-2007	Iraq: Sinjar	Islamic State of Iraq/Mujahidin	430
Aug-19-1978	Iran: Abadan	Mujahideen-I-Khalq	430
Jan-17-2009	Congo: Tora	Lords Resistance Army	400
Mar-23-1994	Burundi: Bujumbura	Tutsi	400
Jul-18-1987	Mozambique: Homoine	Mozambique National Resistance Movement	386
May-23-1996	Burundi: Kivyuka	Tutsi	375
Dec-14-2009	Congo: Makombo Tapili	Lords Resistance Army	345

Table 2.2: Ten Deadliest Terrorist Attacks
Source: Kean and Lee (2004), National Consortium for the Study of Terrorism and Responses to Terrorism (2010), and National Counterterrorism Center

loss to reinsurance companies who need to find a way to hedge the risk.

Specifically, there are many ways in which a re-insurer company can mitigate extreme mortality risk. The simplest approach is to self-insure this risk. A re-insurer company may choose to set aside a portion of profit each year until it builds a meaningful contingency reserve to cover an extreme mortality event. The advantages of this method are that it can match up fully with the losses, it is the most cost efficient approach, and there is flexibility in how much to save each year. The problems of this method are that it takes long time to build a meaningful reserve and the reserve will most likely reduce return on equity.

Another method is to buy a simple high-limit stop-loss cover. This tool also has the benefit of covering actual losses with reimbursements. Additionally, it is very simple to negotiate a deal and to manage it. The disadvantages to this tool are that it is pretty costly, it usually has exclusions (such as terrorism, nuclear, biological), and the ceding company is exposed to the credit risk.

The final approach is a Mortality Catastrophe Bond. Mortality catastrophe bonds offer several advantages and disadvantages, which are discussed in Huynh et al. (2013) along with an introduction to the market and key characteristics of these instruments. This tool has the benefit of essentially eliminating the credit risk exposure for the issuer; see Bagus (2007). Additionally, it allows the issuer to secure fixed cost multi year coverage, typically ranging from three to five years, which also allows the issuer to spread the fixed cost of issuance over several years. A key disadvantage is the existence of basis risk, which happens because the bond payoff trigger is index based and the actual loss suffered is not entirely matched

by the bond payoff. We will discuss the basis risk in detail in Chapter 5.

2.2 Mortality Catastrophe Bond

2.2.1 Mortality Catastrophe Bond History

Mortality catastrophe bond is a recent market innovation that provides an alternative risk management to address the mortality catastrophe risk that life insurers and re-insurers expose to. The goal of this type of transaction is to transfer extreme risk into the capital markets as a risk mitigation tool. Catastrophe bonds in other industries such as earthquake bonds and windstorm bonds have been available in the market for years. However, the first Mortality Catastrophe Bond in the life insurance industry appeared in late 2003. In December 2003, Swiss Re issued a three-year life catastrophe bond, maturing on 1 January 2007, which serves as reducing Swiss Re's exposure to a catastrophic mortality deterioration, such as a repeat of the 1918 Spanish Flu pandemic. The issue size was $400 million. Investors receive quarterly coupons set at three-month U.S. dollar LIBOR135 basis points. However, the principal is unprotected and depends upon what happens to a specifically constructed index of mortality rates across five countries: France, Italy, Switzerland, the United Kingdom and the United States of America. Specifically, the principal is repayable in full if the mortality index does not exceed 1.3 times the 2002 base level during any of the three years of the bond's life. The principal is reduced by 5% for every 0.01 raise in the mortality index above this threshold and is entirely exhausted if the index exceeds 1.5 times the base level. The payoff schedule of the

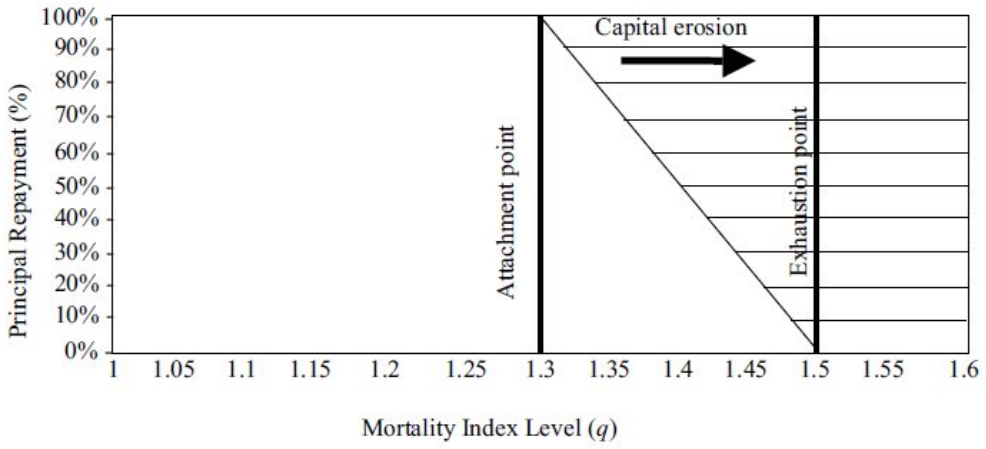


Figure 2.1: Terminal payoff of Swiss Re mortality bond to investors
Source: Blake et al. (2006)

bond is shown in Figure 2.1.

This bond was issued via a special purpose vehicle (SPV) called Vita Capital (VC). VC invests the $400 million principal in high-quality bonds and swaps the income stream on these for a LIBOR-linked cash flow. VC distributes the quarterly income to investors and any principal repayment at maturity. This structure is shown in Figure 2.2.

In April 2005, Swiss Re announced that it had issued a second life catastrophe bond with a principal of $362 million, using a new SPV called Vita Capital II. The maturity date is 2010 and the bond was issued in three tranches: Class B ($62 million), Class C ($200 million) and Class D ($100 million). The principal is at risk if, for any two consecutive years before maturity, the combined mortality index exceeds specified percentages of the expected mortality level (120% for Class B, 115% for Class C, and 110% for Class D).

In brief, mortality catastrophe bond is one of numerous ways that insurers and

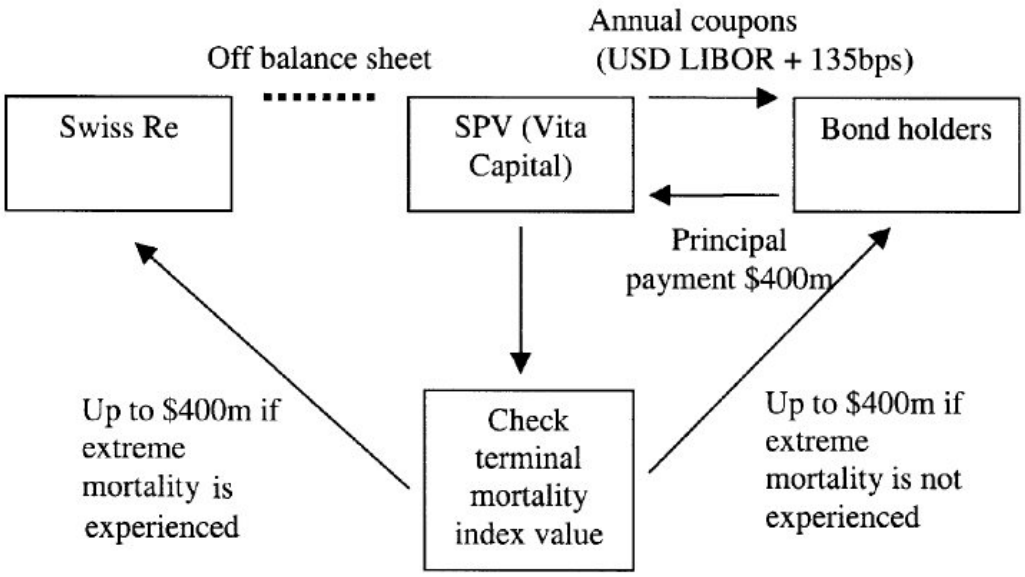


Figure 2.2: The structure of Swiss Re mortality bond Source: Blake et al. (2006)

re-insurers can do to mitigate extreme mortality risk. In particular, investors purchase bonds and receive a return on their investment. If an event does not happen, investors receive their principal back at the end of the term (usually three to five years). If the event does happen, the investors will lose part or all of their investment, which is paid to the insurance company to offset some or all of its loss.

2.2.2 Literature Review

Life insurers and re-insurers are exposed to the future mortality uncertainty risk. Catastrophic mortality events pose a significant threat to the life insurance industry as they induce an unexpected growth in mortality over a short period of time, which may lead to a substantial rise in claims and the potential for severe adverse financial consequences, such as breaches in regulatory solvency and capital requirements, see Cox and Yungui (2004).

While there are a range of catastrophic mortality events that may impact the life insurance industry, an influenza pandemic, such as Spanish Flu pandemic happened in 1918, is considered the most serious threat. The exposure to catastrophic mortality events such as influenza pandemics has been hard for life insurers and re-insurers to manage since the possibility of such events happening in any year is low while the potential for devastating losses is high once they happen. In contrast to the inherent credit risk associated with traditional reinsurance, mortality catastrophe bonds bear no credit risk for sponsors, see Bagus (2007). According to Cowley and Cummins (2005), mortality catastrophe bonds are based on a specified mortality index, which is calculated as a weighted average of general population mortality rates. Therefore, the issue of basis risk comes out, resulting in imperfect hedge effectiveness as the possibility of gains or losses existing in the hedged position. In particular, the sponsor is concerned that the bond payoff will be inadequate to cover the actual loss suffered, see Coughlan et al. (2011).

2.3 Structure of Mortality Bonds

Once the base mortality is set, the bond issuer must determine a trigger point (or attachment point). The trigger point is usually set at 100 percent + X of expected mortality. The larger X is, the lower the chance of an event will happen and, hence, the investor can expect a lower possibility of losing money. At some point 100 percent + Y (so-called exhaustion level or detachment point), there is a total loss to the investor. Between 100 percent + X and 100 percent + Y, there will be a grading of loss to the investor (See Figure 2.1 where X=30% and Y=50%). A bond issuer

may issue different levels of risk to reach investors with different risk appetites. Theses different levels of risk are called tranches. Tranches closer to the actual mortality will cover most extreme mortality losses and tranches further from the actual mortality will cover less losses.

The advantages of this type of approach to risk mitigation are that there is absolutely no credit risk (see Bagus (2007)), and the bond issuer may be able to release some risk capital. The disadvantages are cost, complication and the risk that insured life mortality will be poor that even the population mortality index used will not cover this risk. This is called basis risk and it is a mismatch risk. Basis risk does not necessarily invalidate the effect of hedging because it can be minimized by properly structuring and calibrating the hedging instrument to ensure high hedge effectiveness. If the basis risk is small relative to the risk of the initial unhedged position, hedging strategy is beneficial, see Coughlan et al. (2011). We will explain this in details in Chapter 5.

Although we consider a fictitious mortality CAT bond in this thesis, one should be advised that it very much shares the same payoff structure as that of most mortality bonds in practice. See Vita and Tartan bonds for example. We assume that the bond is issued at the end of year t_0 with maturity $T = 5$ years. The principal of the bond is P. At maturity, the principal repayment of the bond is at risk and depends on a mortality index of a population (the *reference population*). Precisely, let q_t represent the weighted average mortality rate at year t, i.e.

$$q_t = \sum_{\text{all } x} w_x m_{x,t}, \tag{2.1}$$

where $m_{x,t}$ is the central mortality rate for age group x in calendar year t, and w_x is the weight applied to $m_{x,t}$. The weights for the weighted average mortality rate are displayed in Table 2.3. Then the mortality index at time t, denoted by i_t, is defined as

$$i_t = \frac{q_t + q_{t-1}}{q_{t_0} + q_{t_0-1}}. \tag{2.2}$$

At the end of year $t = t_0 + j$, $j = 2, 3, 4, 5$, if i_t exceeds a certain level, the so-called *trigger level* or *attachment point*, a, investors will lose principal. If the index exceeds the so-called *exhaustion level* or *detachment point*, d, their complete principal will be lost. Accordingly, define l_t, the *accumulated loss ratio* at the end of year t, as

$$l_t = \min\left\{\max\left\{l_{t-1}, \frac{i_t - a}{d - a}\right\}, 100\%\right\}, \tag{2.3}$$

with $l_{t_0+1} = 0$. Therefore, by maturity the total amount of Pl_{t_0+5} from the bond principal will be used to cover mortality losses of the bond issuer. However, the mortality index of the reference population may not reflect the mortality experience of the bond issuer.

Age Groups (x)	**Age Weights (w_x)**
0	1.27%
1-4	4.95%
5-9	5.59%
10-14	5.84%
15-19	6.32%
20-24	6.73%
25-29	6.84%
30-34	6.54%
35-39	6.73%
40-44	7.34%
45-49	7.30%
50-54	6.39%
55-59	5.69%
60-64	6.01%
65-69	4.77%
70-74	3.89%
75-79	3.18%
80-84	2.38%
85-89	1.48%
90-94	0.59%
95-99	0.15%
100-104	0.02%
105-109	0.00%
110+	0.00%

Table 2.3: Weights for Weighted Average Mortality Rate

Chapter 3

Mortality Models

Modeling and forecasting mortality are two of the primary concerns of actuaries, life insurance companies, pension plans, national governments, etc. Specifically, corporations providing defined benefit pension plans must assure right funding of future liabilities, which highly depend on the future mortality rates forecasts; National governments use forecasts of mortality rates to plan social security and health care programs.

Mortality modeling has a very long history. Many models have been proposed since Gompertz (1825) published his law of mortality in 1825. Gompertz (1825) shows that the mortality rate stands for the body's propensity to succumb to death and that its inverse is the body's ability to withstand death. This law is based on an assumption that a person's resistance to death decreases as his years increase. Makeham (1867) modifies the original Gompertz formula, mentioning that while the Gompertz formula interprets adequately the age-dependent component of mortality, it does not properly reflect the environmental forces of mortality that

are unrelated to those associated with aging. Therefore, Makeham's Law states that the human death rate is the sum of an age-independent component and an age-dependent component. However, a difficulty of Makeham law of mortality is to identify which elements of its patterns are really age related risks and which are non-age related risks. In order to approach this difficulty, Carnes and Olshansky (1997) try to separate extrinsic causes of death so that one can find out a population pattern that reflects an intrinsic mortality schedule, interpreting the increasing vulnerability of individuals with increasing age.

The above mortality models, however, are static, which are unsatisfactory since no death is ever entirely one or the other. Indeed, actual mortality is stochastic and evolves over time. Therefore, while the above mortality models can fit data at a fixed point in time, the parameters must be re-fit periodically to accommodate changes in mortality patterns. In order to approach this problem, an alternative approach is explored by Lee and Carter (1992) to model mortality rates with sets of age-specific constants, $\{a_x\}$ and $\{b_x\}$, and time-varying index k_t. In this case, mortality at each age changes as k_t changes, which is an advantage for forecasting. Most models, then, are developed based on the Lee-Carter model, considering both age patterns and trends over time. For example, Renshaw and Steven (2005) extend the Lee-Carter modeling framework through the introduction of a wiser class of generalized, parametric, non-linear models. This allows the modeling and extrapolation of age-specific cohort effects and the more familiar age-specific period effects. Booth et al. (2006) change the Lee-Carter model by optimally selecting the time period over which to fit the model and adjust the state variable, k_t to

match the total number of deaths in each year. Jong and Tickle (2006) address the issue with the Lee-Carter model where k_t time series have an independent impact at each age and simulate mortality rates as a smoothed state space model. Cairns et al. (2006) present a two-factor stochastic model for the development of mortality curve through time. The first factor affects mortality rate at all ages in the same way, whereas the second factor affects mortality rate at higher ages much more than at lower ages. However, the above models all suffer from the same problem as Lee and Carter (1992), in which they all model the level of log mortality rates and therefore misrepresent the temporal dependence structure of mortality rates by age group.

Another model based on the Lee-Carter model is proposed by Mitchell et al. (2013), which avoids the common problem of modeling log mortality levels by modeling changes in log mortality rates (hereinafter referred to as the "MBMM mortality model"). Mitchell et al. (2013) propose a model that expresses the log mortality rate changes as an age group dependent linear transformation of a mortality change index. This model simulates better human mortality rates by modeling the time series of mortality rate changes rather than mortality rate levels, see Section 3.2 for more detail. Given this appropriate model, one can forecast the level and age distribution of mortality for the next few decades. The rest of this chapter is organized as follows.

In Section 3.1, we describe the Lee-Carter model in more detail and provide information to fit this model. In Section 3.2 we discuss the MBMM mortality model we will use and compare this model to the Lee-Carter model. An overview of data

on England and Wales, and Scotland and the model calibration are provided in Section 3.3.

3.1 The Lee-Carter Model

One of the first papers to model mortality rates considering both current age and year was written by Lee and Carter (1992). For a single population, let $m(x,t)$ be the central death rate for age x in year t. Lee and Carter (1992) model the log central mortality rate of age group x at time t, $\ln m_{x,t}$, as a bi-linear model

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

$$\text{s.t.} \sum_{\text{all x}} b_x = 1, \sum_{\text{all t}} k_t = 0 \tag{3.2}$$

where a_x describes the general shape of the mortality curve, k_t is the temporal mortality time-varying index that captures the evolution of rates over time, b_x describes how age group x responds to the temporal mortality time-varying index, and $\epsilon_{x,t}$ is the error term.

Specifically, a_x are averages over time of the $\ln m_{x,t}$ and e^{a_x} is the general shape across age of the mortality schedule. The b_x shows us which rates go down rapidly and which rates go down slowly in response to changes in k_t. The b_x could be negative for some ages, indicating that mortality at those ages tends to move up when falling at other ages; in reality this does not seem to exist over the long time. When k_t is linear in time, mortality at each age varies at its own constant exponential rate. The advantage of this model is that negative death rate cannot occur because of its

exponential property; as k_t goes to negative infinity, each age-specific rate goes to 0.

The error term, $\epsilon_{x,t}$, with mean 0 and variance σ^2, reflects specific age historical influences not captured by the Lee-Carter model. The variances σ^2 over time of age-specific components of $\epsilon_{x,t}$ should not differ greatly because they are deviations from the logs of the rates.

Lee and Carter (1992) model mortality rates for different ages over time by obtaining an unobserved state variable from the historical data on mortality rates. This state variable represents a single temporal mortality time-varying index applicable for each age group in the entire population. Since it is possible that each age group reacts differently to the temporal mortality time-varying index, the mortality rate of each group is a linear process of the mortality time-varying index.

To estimate the unknown index k_t, the singular value decomposition method should be applied to find a least squares solution after the averages over time of the logarithms of the rates $\ln[m(x,t)]$ has been subtracted from the matrix of the logarithms of the rates. That is to say, we first need to build the matrix, $M_{x,t}$, of log mortality rate $\ln[m(x,t)]$. Then from this matrix we must obtain the mean over time for each age group, a_x, and create a demeaned matrix $\hat{M}_{x,t}$ such that $\hat{M}_{x,t} = M_{x,t} - a_x$. From this matrix we can use the singular value decomposition to obtain b_x and k_t. The singular value decomposition of $\hat{M}$ finds matrices U, S and V such that $\hat{M} = USV'$, where U and V are orthogonal matrices and S is a non-negative diagonal matrix, with the same dimension as $\hat{M}$, and $S_{1,1}$ is largest element of S. The U and V are called the left-singular vectors and right-singular

vectors of $\hat{M}$, respectively. This means that k_t is a scale multiple times the first column of V times $S_{1,1}$ and b_x is the first column of U divided by that same scalar. See Jolliffe (1986) for more detail on singular value decomposition.

On the other hand, since mortality rates have been trending downwards over at least the last 100 years for all age groups, the covariance matrix of mortality rates in the Lee-Carter model immensely overestimates dependence. For example, the covariance over 100 years between the log mortality rates of people aged 1-4 and people aged 70-74 is very high because early in the time series the mortality rates for both age groups were relatively high, and later in the history the mortality rates for both age groups were relatively low. Even so, this does not imply that mortality rates for these two groups are highly dependent. Factors that can reduce the mortality rates for these two age groups are different.

3.2 The MBMM mortality model

Building upon the idea of bi-linear modeling of age and time from the Lee-Carter model, Mitchell et al. (2013) propose a new model of mortality rates that is based on log mortality rate changes rather than levels. The MBMM mortality model is

$$\ln m_{x,t+1} - \ln m_{x,t} = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t}, \tag{3.3}$$

where α_x describes the average change in log mortality rate for age group x, κ_t describes a mortality change index at time t that is the same for all age groups, β_x describes how age group x responds to the mortality change index, and $\epsilon_{x,t}$ is the

error term, which we assume to be normally distributed with zero mean and to be uncorrelated across time and age groups.

This new model proposed by Mitchell et al. (2013) considers the mortality data from a different perspective. The trend in the data derives from previous years' mortality rates rather than a trend in the unknown mortality time-varying index. In this way we firstly perform a simple transformation of the data prior to modeling, the subsequent modeling can better replicate the dynamics of mortality rates through time, which avoids the common problem of modeling log mortality levels as we discussed before. That is to say, by considering changes of mortality rates we are able to capture the dependence structure between ages of mortality more precisely.

This means that the change in log mortality rate for each age group x from year t to year $t+1$ is a linear transformation of a time indexed variable κ_t plus some error. Mitchell et al. (2013)'s interpretation of this model is that in each year we find a random vector of log mortality rate changes and the log mortality rate from the next year for each age group is the log mortality rate from this year plus the random vector. Although it looks similar to the Lee-Carter model, Mitchell et al. (2013) model changes in log mortality rates rather than levels of log mortality rates. As said in Mitchell et al. (2013), "by considering the changes we are able to more accurately capture the dependence structure between ages of mortality and use this to construct a more encompassing model" (p. 276).

The values of α_x, β_x and κ_t are not unique in this representation of the model. To address this identifiability problem we need to impose restrictions on the model to

guarantee uniqueness. We require α_x to be the mean change in log mortality rates, which forces κ_t to have a mean of zero. We also require the sum of β_x across age groups to be equal to 1. Then we apply the singular value decomposition to obtain unique estimates for β_x and κ_t.

The difference between fitting this model and the Lee-Carter model is that we first need to build the matrix $M_{x,t}$ of log mortality rate changes. Specifically, we take the matrix of log mortality rates from the second year to last year under consideration and then subtract from the matrix of log mortality rates from the first year to the penultimate year. In this way we can get $M_{x,t} = \ln[m(x,t+1)] - \ln[m(x,t)]$. Then from this matrix we apply the singular value decomposition method as described above.

The comparisons of the MBMM mortality model versus the Lee-Carter model are presented in Figures 3.1 and 3.2 for both EW and SL data from 1876 to 2011. Figure 3.1 shows the actual log England and Wales mortality rate for two different age groups 0 and 25-29 and the Lee-Carter prediction of the log mortality rate and the MBMM mortality model's prediction of the log mortality rate. We can see that the MBMM mortality model is much better than the Lee-Carter model which calculates the mortality rates for next year without considering the mortality rates this year. We choose age groups 0 and 25-29 at random because all age groups show better levels of accuracy for the MBMM mortality model than the Lee-Carter model. We also show similar plots for Scotland in age groups 0 and 25-29 in Figure 3.2.

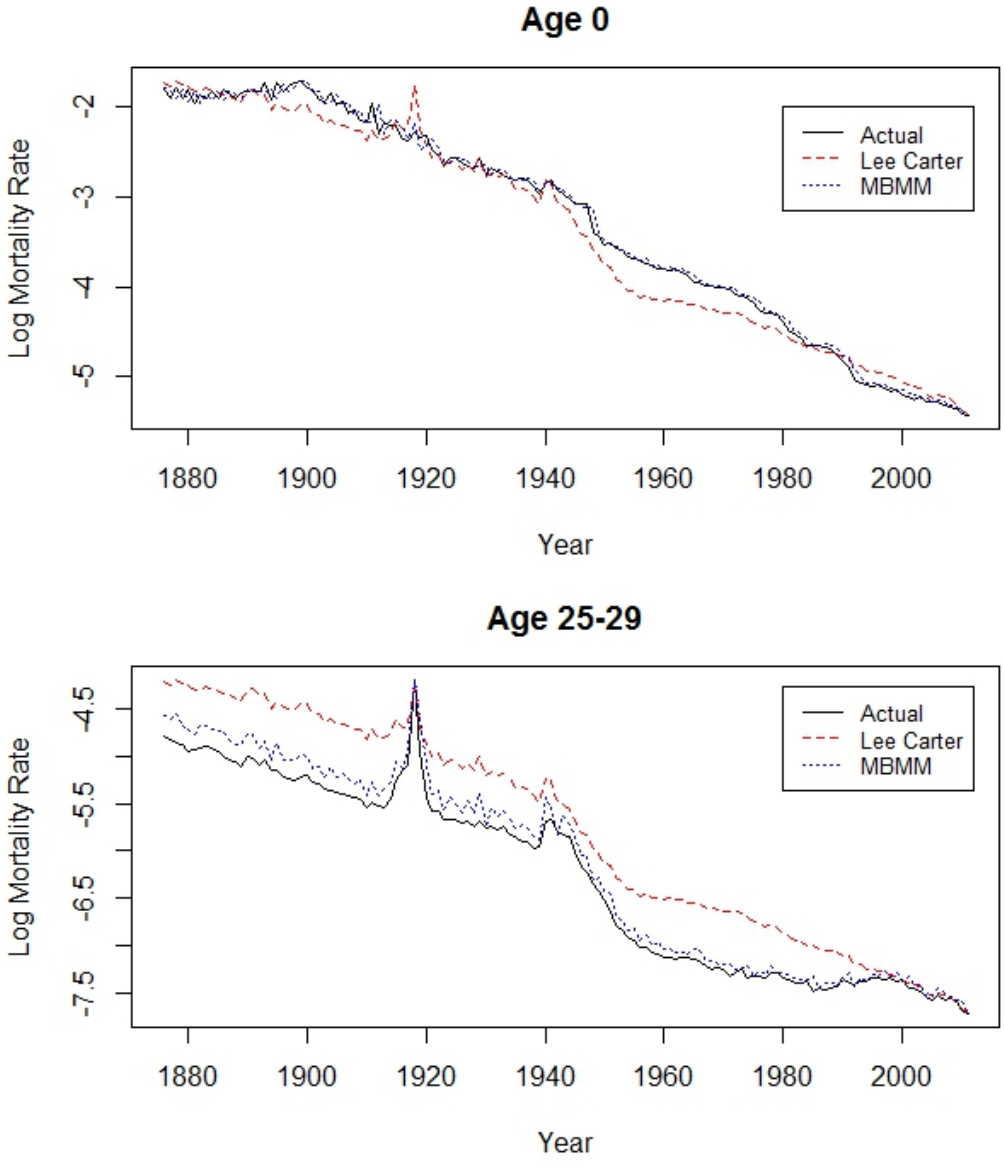


Figure 3.1: Comparison of log mortality rate predictions for England and Wales data during 1876-2011

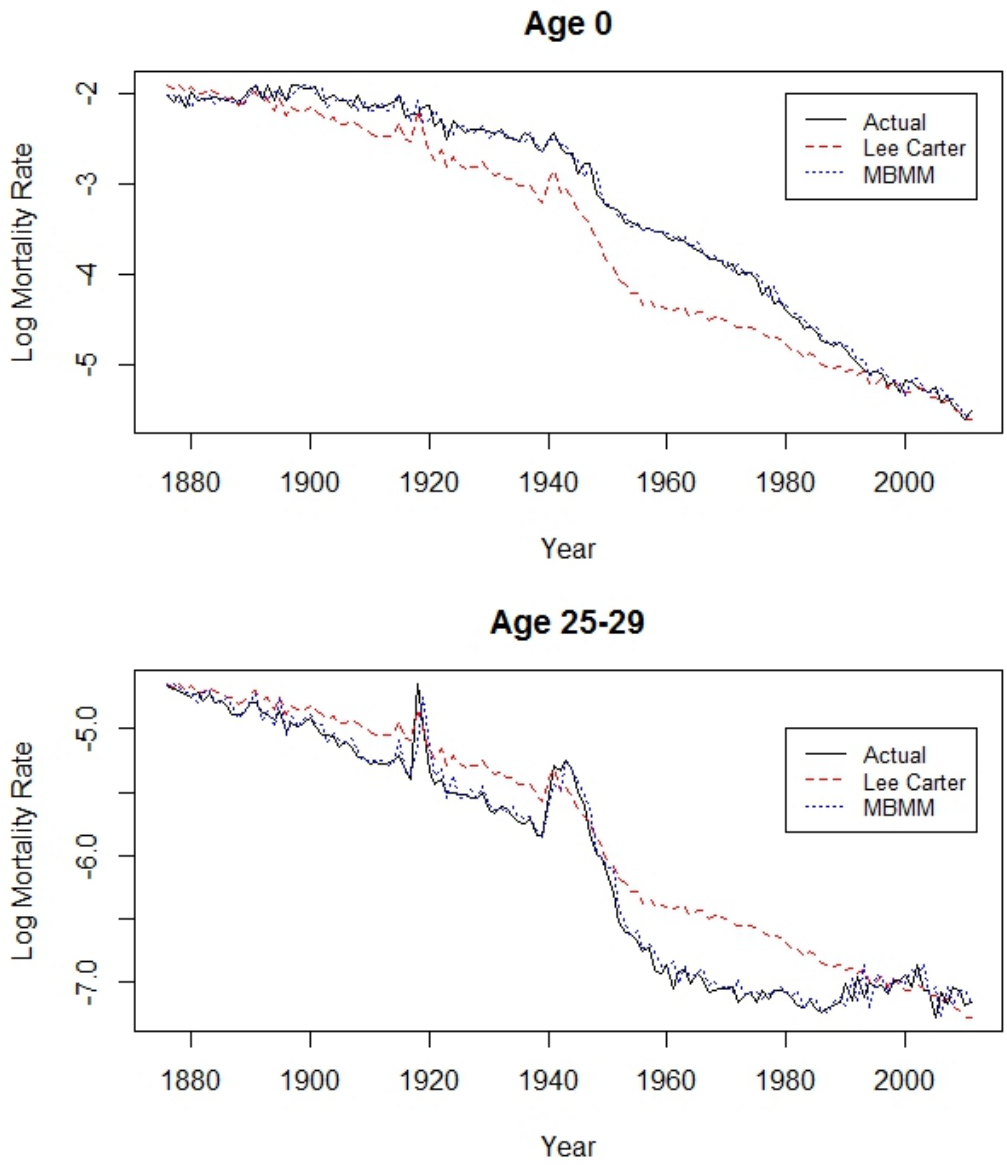


Figure 3.2: Comparison of log mortality rate predictions for Scotland data during 1876-2011

3.3 Model Calibration

3.3.1 Data Description

In this thesis, we use the England and Wales (EW) civilian population as the reference population and the Scotland (SL) civilian population as the insured pop-

ulation. The mortality data we use is from the Human Mortality Database[1] (HMD) (2013). The data separates the population into 24 different age groups: (0), (1-4), (5-9), then every 5 years until (105-109) and finally (110+). We use data that combines both sex and race categories to consider civilian population in EW and SL in each age group. The HMD provides central mortality rates for the EW civilian population back to year 1841 and that for the SL civilian population back to year 1855. However, according to the U.K. Office for National Statistics birth registration was incomplete until 1874 and therefore the quality of the mortality data in 1875 is assumed to be lower than in later years. Hence, in this thesis we use mortality rates from 1876 to 2011 for both populations.

One should know that basis risks remain between reference and insured populations no matter how similar characteristics these two populations share because the two populations are not the same people. A study on the 1957-1958 and 1968-1969 influenza pandemics from Woolnough et al. (2007) suggests that insured population has around 12% lower excess death rates in standard ordinary policyholders compared with age and gender matched general population. Therefore, we analyze population basis risks using the difference between the two population.

We have the following reasons to use the two regional populations. First of all, their mortality rates are available in the HMD while it is hard to find mortality data for a real insured population. Second, civilian population excludes the military, also, death counts exclude military deaths that occurred abroad, which is very much like a typical insured population. Last, a regional civilian population is part of the total population. This relation provides a good example to analyze

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

population basis risk.

We treat the EW population as reference population instead of the UK population because there are only 89 years data available for UK in the HMD which is limited for mortality modeling. Also, according to 2011 Census [2], the population of EW accounted for 88.8% of the population of UK, therefore we use the EW population as the total population.

3.3.2 Data Analysis: Parameter Estimation

The estimates of α_x and β_x for EW and SL from 1876 to 2011 are given in Tables 3.1 and 3.2, respectively. For $\epsilon_{x,t}$, we simulate it as the age and year specific error which has an expected value of zero and is assumed to be uncorrelated across time and age group. The estimates of α_x, β_x, $\epsilon_{x,t}$ and κ_t (described below) will be used in Chapter 5 to construct forecasts of age-specific central mortality rates. Since α_x and β_x are age specific and do not vary over time, the mortality rates forecasting will focus on forecasting $(\kappa_{t,1}, \kappa_{t,2})$, where $\kappa_{t,1}$ and $\kappa_{t,2}$ are mortality change indexes for EW and SL, respectively.

Figure 3.3 shows the time series of $\kappa_{t,1}$ and $\kappa_{t,2}$ from 1876 to 2011. We observe both the dependence between $\kappa_{t,1}$ and $\kappa_{t,2}$ series and their own serial dependence. Therefore, we further investigate the dependence structure of $\kappa_{t,1}$ and $\kappa_{t,2}$ in the rest of this section.

Dependence Measures

Correlation plays a central role in describing the structure between time series

[2]Source: 2011 Census: Population Estimates for the United Kingdom, 27 March 2011

Age group	α_x	β_x
0	-0.02684562	0.029090873
1-4	-0.03762468	0.092754210
5-9	-0.03224814	0.074965005
10-14	-0.02728377	0.073761473
15-19	-0.02314139	0.085833844
20-24	-0.02193603	0.092506388
25-29	-0.02151008	0.098072547
30-34	-0.02029481	0.085866332
35-39	-0.01879237	0.061082562
40-44	-0.01721324	0.047381748
45-49	-0.01518214	0.038600279
50-54	-0.01358755	0.034789897
55-59	-0.01181602	0.025508727
60-64	-0.01154772	0.023428148
65-69	-0.01066173	0.022973729
70-74	-0.00984461	0.019745118
75-79	-0.00869981	0.018599972
80-84	-0.00751580	0.015439109
85-89	-0.00568088	0.014350693
90-94	-0.00404974	0.011537777
95-99	-0.00267211	0.011084219
100-104	-0.00143431	0.009271685
105-109	-0.00052833	0.007386185
110+	-0.00003048	0.005969481

Table 3.1: Fitted values of α_x and β_x of England and Wales data for 1876-2011 using singular value decomposition method

Age group	α_x	β_x
0	-0.02589783	0.03741179
1-4	-0.03906638	0.09110817
5-9	-0.03261521	0.06486377
10-14	-0.02972871	0.05250511
15-19	-0.02197736	0.05273089
20-24	-0.02019697	0.05239806
25-29	-0.01846976	0.05915137
30-34	-0.01634217	0.05195447
35-39	-0.01451149	0.04147394
40-44	-0.01369569	0.03256099
45-49	-0.01228445	0.03230266
50-54	-0.01180972	0.03151236
55-59	-0.00947882	0.02907023
60-64	-0.00972131	0.02510831
65-69	-0.00743909	0.02883687
70-74	-0.00752351	0.02979919
75-79	-0.00688912	0.03054010
80-84	-0.00563015	0.02957460
85-89	-0.00414092	0.03895919
90-94	-0.00305955	0.04089462
95-99	-0.00155231	0.04155089
100-104	-0.00046002	0.03963308
105-109	0.00032563	0.03519760
110+	0.00075404	0.03086172

Table 3.2: Fitted values of α_x and β_x of Scotland data for 1876-2011 using singular value decomposition method

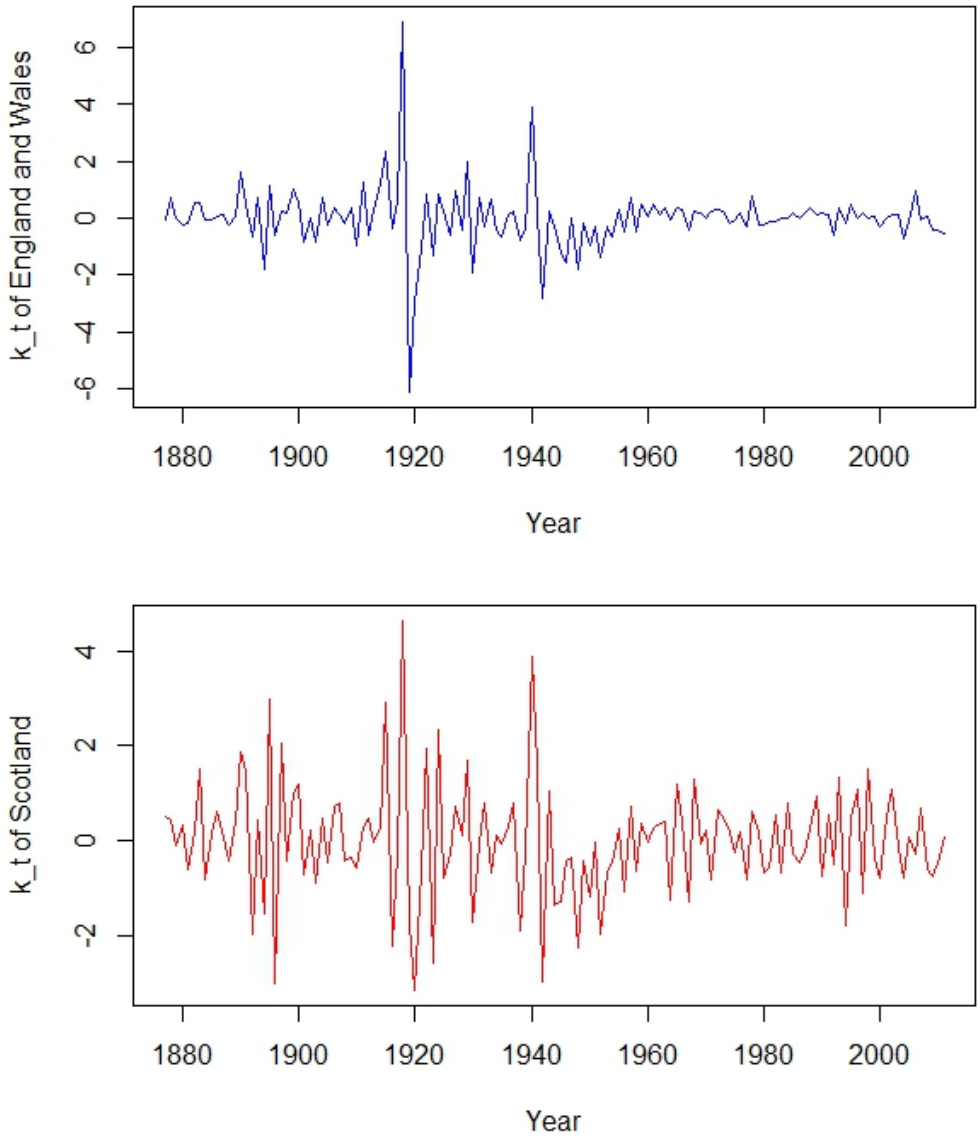


Figure 3.3: The mortality change indexes $\kappa_{t,1}$ (England and Wales) and $\kappa_{t,2}$ (Scotland) from 1876 to 2011

$\kappa_{t,1}$ and $\kappa_{t,2}$. For interdependence between time series of $\kappa_{t,1}$ and $\kappa_{t,2}$, we consider rank correlations - the Spearman's Rho and Kendall's Tau - which are copula-based dependence measures. In contrast to linear correlation, these two measures are simple scalar measures of correlation that depend only upon the copula but not

	Spearman's Rho	**Kendall's Tau**
EW vs. SL	0.70	0.51

Table 3.3: Spearman's Rho and Kendall's Tau of pseduo-observations of mortality change indexes $\kappa_{t,1}$ (England and Wales) and $\kappa_{t,2}$ (Scotland) from 1876 to 2011

upon the marginal distributions, thus can be used in the parametrization of copulas. The correlations are calculated by looking at the ranks of the data alone, thus we replaced κ_t by its rank R_t and defined the sequence

$$\hat{U}_t = R_t/(n+1). \tag{3.4}$$

The Spearman's Rho and Kendall's Tau test results of $\hat{U}_t$ are showed in Table 3.3. Specifically, the Spearman's Rho of the two $\hat{U}_t$ series is 0.70 while the Kendall's Tau is 0.51 (both range from -1 to 1), which are close to 1, indicating that two series are positively correlated.

As to the serial dependence, we observe that there was a big jump down right after a big jump up in 1918. This phenomenon was also observed by others, e.g. Mitchell et al. (2013). Therefore, we test for the presence of serial dependence in the uni-variate time series κ_t using Ljung-Box statistics test, whose null hypothesis is that the datum are independently distributed. Table 3.4 shows that the p-values of Ljung-Box test are 3.05% for EW and 0.02% for SL (both less than 5% level), rejecting the null hypothesis of independence. The reasons for serial dependence are numerous, for example, as shown in Strómberg (2012), there are lots of factors, such as weather, forcing mortality rates to be serial correlated. Therefore, we

Ljung-Box Test			
	Lag	***p*-value (%)**	**Test Results**
EW	1	3.05	serial dependent
SL	1	0.02	serial dependent

Table 3.4: p-values for testing serial dependence of $\kappa_{t,1}$ (England and Wales) and $\kappa_{t,2}$ (Scotland) from 1876 to 2011

need to find a model that can analyze both interdependence and serial dependence structure of time series $\kappa_{t,1}$ and $\kappa_{t,2}$ simultaneously.

Chapter 4

Copula-based Semi-parametric Models

As shown in Chapter 3, the key to forecast future mortality rates is to simulate the time series of vector $(\kappa_{t,1}, \kappa_{t,2})$. Since there exists an embedded dependence structure, we use the copula-based semi-parametric models proposed by Rémillard et al. (2012) that take into account both interdependence and serial dependence in $(\kappa_{t,1}, \kappa_{t,2})$ simultaneously. While Rémillard et al. (2012)studied $(d_1 + d_2)$-dimensional random vector $(X, Y)^{\mathrm{T}}$, this thesis only considers 4-dimensional vector $(X_{t-1}^{\mathrm{T}}, X_t^{\mathrm{T}})^{\mathrm{T}}$ in which $d_1 = d_2 = 2$ and $X_t = (\kappa_{t,1}, \kappa_{t,2})^{\mathrm{T}}$. Specifically, we fit copulas on $(X_{t-1}^{\mathrm{T}}, X_t^{\mathrm{T}})^{\mathrm{T}}$ to model the dependence among $\kappa_{t-1,1}, \kappa_{t-1,2}, \kappa_{t,1}, \kappa_{t,2}$. By doing this, we combine both interdependence and serial dependence and model them simultaneously. The other advantage of this approach is that it is not necessary to model univariate time series.

The rest of this chapter is structured as follows. In Section 4.1, we introduce

copula-based Markovian models for time series. In Section 4.2 we discuss two preliminary hypotheses that need to be assumed before using the copula-based semi-parametric models. Parameter estimation is introduced in Section 4.3 using maximum pseudo likelihood method and goodness-of-fit test. Model calibration and goodness-of-fit test results are provided in Section 4.4. An overview of copulas is given in Appendix A and some general properties of conditional copulas are provided in Appendix A.4.

4.1 Copula Models

Over the past two decades, copulas have gained popularity to be an essential modeling tool in describing the dependence between random variables. Trivedi and Zimmer (2005) state that "The copula approach is a useful method for deriving joint distributions given the marginal distributions, especially when the variables are non normal." Broadly, copulas can be used to simulate a correlation structure independently from the marginal distributions of these variables. Moreover, copulas provide a more flexible approach to model correlation structures compared to other traditional methods. Copulas are especially popular in the high dimensional statistical applications as they can be easily modeled and the distributions of random vectors can be determined by estimating the marginals and copulas separately. See Appendix A for a summary of some useful results in copula theory that are essentially useful in this thesis.

In our application, we aim to model the dependence in 2-dimensional time series $X_t = (\kappa_{t,1}, \kappa_{t,2})^{\mathrm{T}}$ through a copula-based semi-parametric model proposed by

Rémillard et al. (2012). The model we are going to use only require that the process X_t be Markovian and stationary, and all marginal distributions be continuous and independent of time. Precisely, we consider C to be the copula associated with vector $(X_{t-1}^{\mathrm{T}}, X_t^{\mathrm{T}})^{\mathrm{T}}$. The copula Q of X_t is then the same as the copula X_{t-1} as the copula property mentioned in AppendixA, i.e., if $\mathbf{1} = (1,1)^{\mathrm{T}}$ then, for all $u \in [0,1]^2$,

$$Q(u) = C(u, \mathbf{1}) = C(\mathbf{1}, u). \tag{4.1}$$

Let $\mathbf{F}$ denote the transformation of X_t, i.e.,

$$\mathbf{x} = (x_1, x_2)^{\mathrm{T}} \mapsto \mathbf{F}(x) = (F_1(x_1), F_2(x_2))^{\mathrm{T}}, \tag{4.2}$$

where F_1, F_2 are marginal distribution functions. Denote $U_t = \mathbf{F}(X_t)$. Then U_t is a 2-dimensional time series such that $(U_{t-1}, U_t) \sim C$ and $U_t \sim Q$. Because F is unknown, the Markovian stationary time series U is not observable.

To calculate the copula parameters or to simulate observations of U_t, we need to calculate the conditional distribution of U_t given U_{t-1}. Calculating conditional copula enables us to combine interdependence and serial dependence and model them simultaneously. In Appendix A.4, we analyze the properties of the conditional copula in a general context. It is then applied to multivariate time series.

4.1.1 Markovian Models with Meta-elliptic Copulas

Meta-elliptical copulas are copulas associated with elliptical distributions through equation (A.2). Two meta-elliptic copulas that are popular are Gaussian and Student-

t copulas. According to Rémillard et al. (2012), a vector Y has a elliptical distribution with generator g, location parameter μ and positive definite symmetric dispersion matrix Σ, denoted $Y \sim \epsilon(g, \mu, \Sigma)$, if its density is given, for all $y \in \mathbb{R}^d$, by

$$h(y) = \frac{1}{|\Sigma|^{1/2}} g\left\{(y-\mu)^{\mathrm{T}}\Sigma^{-1}(y-\mu)\right\}, \tag{4.3}$$

where, for arbitrary $r \in (0, \infty)$,

$$\frac{\pi^{d/2}}{\Gamma(d/2)} r^{(d-2)/2} g(r) \tag{4.4}$$

is a density on $(0, \infty)$. In fact it is the density of $\xi = (Y-\mu)^{\mathrm{T}}\Sigma^{-1}(Y-\mu)$.

In order to generate Y, simply set $Y = \mu + \xi^{1/2} A^{\mathrm{T}}\zeta$, where $A^{\mathrm{T}}A = \Sigma$, ξ has density as equation (4.4) and is independent of ζ, with ζ uniformly distributed on $S_d = \left\{y \in \mathbb{R}^d : \|y\| = 1\right\}$. Since copulas are invariant by increasing transformations, the underlying copula of $Y \sim \epsilon(g, \mu, \Sigma)$ depends only on g and the correlation matrix R associated with Σ.

For instance, the Gaussian distribution is a special case of elliptic distribution with generator $g(r) = e^{-r/2}/(2\pi)^{d/2}$ for $r \in (0, \infty)$. Another interesting family of elliptic distribution is the Pearson type VII, with generator $g(r) = \Gamma(\alpha + d/2)(1 + r/v)^{-\alpha-d/2} / \left\{(\pi v)^{d/2}\Gamma(\alpha)\right\}$ for $r \in (0, \infty)$, where $\alpha, \nu > 0$. The case $\alpha = \nu/2$ corresponds to the multivariate Student-t, while if $\alpha = 1/2$ and $\nu = 1$, one gets the multivariate Cauchy distribution.

Suppose that $X = (X_1, X_2)^{\mathrm{T}} \sim \epsilon(g, 0, R)$ with correlation matrix

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{4.5}$$

Set $\Omega = R_{22} - R_{21}R_{11}^{-1}R_{12}$ and $B = R_{21}R_{11}^{-1}$. Then $|R| = |R_{11}|\,|\Omega|$ and

$$x^{\mathrm{T}}R^{-1}x = (x_2 - Bx_1)^{\mathrm{T}}\Omega^{-1}(x_2 - Bx_1) + x_1^{\mathrm{T}}R_{11}^{-1}x_1. \tag{4.6}$$

Accordingly, $X_1 \sim \epsilon(g_1, 0, R_{11})$ and the conditional distribution of X_2 given $X_1 = x_1$ is $\epsilon(g_2, Bx_1, \Omega)$, where g_1 and g_2 are respectively given by

$$g_1(r) = \int_{\mathbb{R}^{d_2}} g(\|x_2\|^2 + r)\mathrm{d}x_2 = \frac{2\pi^{d_2/2}}{\Gamma(d_2/2)} \int_0^\infty s^{d_2-1} g(s^2 + r)\mathrm{d}s \tag{4.7}$$

and

$$g_2(r) = g(r + x_1^{\mathrm{T}}R_{11}^{-1}x_1)/g_1(x_1^{\mathrm{T}}R_{11}^{-1}x_1). \tag{4.8}$$

The following result is a consequence of these facts.

Let R be a correlation matrix of the form (4.5) and suppose that $C_{g,R}$ is the copula associated with the elliptic distribution $\epsilon(g, 0, R)$. Then the conditional copula of V given $U = u$ is $C_{g_2,\hat{\Omega}}$, where $\hat{\Omega}$ is the correlation matrix built from $\Omega = R_{11} - R_{12}R_{22}^{-1}R_{21}$, and g_2 is defined by Equation (4.8).

Therefore if g is the generator of the 2 dimensional Pearson type VII with parameters (α, ν), then g_1 is the generator of the 2 dimensional Pearson type VII with parameters (α, ν), and g_2 is the generator of the 2 dimensional Pearson type VII

(α', ν'), with $\alpha' = \alpha + 2/2$ and $\nu' = \nu + x_1^{\mathrm{T}} R_{11}^{-1} x_1$. Therefore if the joint distribution of (X_1, X_2) is a Student-t with parameter (ν, R), then the conditional distribution of X_2, given $X_1 = x_1$, is $\epsilon\left\{g, Bx_1, (\nu + x_1^{\mathrm{T}} R_{11}^{-1} x_1)\Omega/(\nu + 2)\right\}$, where g is the generator of a Student-t with $\nu + 2$ degrees of freedom. It follows that the conditional copula of a Student-t copula with parameter (ν, R) is a Student-t copula with parameters $(\nu + 2, \tilde{\Omega})$; see Rémillard et al. (2012).

The following algorithm is to generate Markovian time series having a joint meta-elliptic copula. Let F_1 be the distribution of function associated with $\epsilon(g_1, 0, R_1)$, where g_1 is defined by Equation (4.7).

Algorithm 1. Let g be the generator of a 4-dimensional elliptic distribution F. To generate a time series $U_0, \ldots, U_n$, with stationary distribution C_{g_1,R_1} such that $(U_{t-1}, U_t) \sim C_{g,R}$ with correlation matrix R of the form (4.5), proceed as follows:

1. Generate $X_0 = (X_{01}, X_{02}) \sim \epsilon(g, 0, R)$.
2. Set $U_0 = (F_1(X_{01}), F_1(X_{02}))$.
3. For each $t \in \{1, \ldots, n\}$,
   (a) generate $V_t \sim \epsilon(g_2, 0, \Omega)$;
   (b) set $X_t = V_t + B \times X_{t-1}$ and $U_t = (F(X_{t1}), F(X_{t2}))$.

4.1.2 Markovian Models with Archimedean Copulas

Archimedean copulas were first presented in statistics in Genest and Mackay (1986a) and Genest and Mackay (1986b). A copula C is said to be Archimedean with generator ϕ when it can be expressed, for all $u = (u_1, \ldots, u_d)$, in the form

$C(u) = \phi^{-1}\{\phi(u_1) + \ldots + \phi(u_d)\}$ for some choice of bijection $\phi : (0,1] \to [0,\infty)$ which is unique up to a scale factor. As proved in Mcneil and Neŝlehovà (2009), sufficient conditions on ϕ are that $\phi(1) = 0$ and that, for all $s > 0$ and $k \in \{1,\ldots,d\}$,

$$h_k(s) = (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}x^k}\phi^{-1}(s) > 0. \tag{4.9}$$

Here, suppose C is a 4-dimensional Archimedean copula with generator ϕ, and set $A_u = C(u, \mathbf{1})$, where $u \in (0,1)^2$. So, for arbitrary $t \in (0,1]$,

$$F_{j,u}(t) = \Pr(V_j \leq t \,|U = u\,) = \frac{h_2\{\phi(A_u) + \phi(t)\}}{h_2\{\phi(A_u)\}} \tag{4.10}$$

and thus the associated quantile function is

$$Q_u(s) = \phi^{-1}\left[h_2^{-1}\left[sh_2\{\phi(A_u)\}\right] - \phi(A_u)\right]. \tag{4.11}$$

This leads to the following result, which is already described in Mesfioui and Quessy (2008).

If $(U,V) \sim C_{4,\phi}$ then the copula associated with the conditional distribution of V given $U = u \in (0,1)^2$ is Archimedean with generator defined, for all $t \in (0,1]$, by

$$\psi_u(t) = h_2^{-1}\left[th_2\{\phi(A_u)\}\right] - \phi(A_u), \tag{4.12}$$

where $A_u = C_{4,\phi(u,\mathbf{1})} = C_{2,\phi(u)}$ and h_2 is defined in equation (4.9).

The following example is showed in Rémillard et al. (2012).

$\phi_\theta(t) = (t^{-\theta} - 1)/\theta$ with $\theta > 0$ generates the Clayton copula with positive

association, and in this case one has

$$h_k(s) = (1 + s\theta)^{-k-1/\theta} \prod_{j=0}^{k-1} (1 + j\theta) \tag{4.13}$$

for arbitrary $s > 0$ and integer $k \geq 0$. As presented in Mesfioui and Quessy (2008), the conditional Clayton copula is also Clayton, with parameter $\theta/(1 + d_1\theta)$. One also can evaluate h_d for Frank and Gumbel-Hougaard copulas; see Barbe et al. (1996) for details on the evaluation.

Using the previous calculations, one can now propose a general algorithm to model a Markovian time series with joint copula $C_{4,\phi}$.

Algorithm 2. To get a Markov chain $U_0, \ldots, U_n$ with stationary distribution $C_{2,\phi}$ and joint distribution of $(U_{t-1}, U_t) \sim C_{4,\phi}$, proceed as follows:

1. Generate $U_o \sim C_{2,\phi}$.
2. For $t \in \{1, \ldots, n\}$,
   (a) set $A_{U_{t-1}} = C_{4,\phi}(U_{t-1}, \mathbf{1}) = C_{2,\phi}(U_{t-1})$;
   (b) generate $V_t \sim C_{2,\psi_{U_{t-1}}}$, where ψ_u is defined by equation (4.12);
   (c) set $U_t = (U_{t,1}, U_{t,2})$, where $U_{t,k} = Q_{U_{t-1}}(V_{t,k})$ for both $k = 1, 2$ and Q_u is defined by equation (4.11).

4.2 Preliminary Assumptions

To use the above models, we only ask that the process $\kappa_{t,1}$ and $\kappa_{t,2}$ be Markovian and stationary, and that their marginal distributions F_1 and F_2 be continuous and

independent of time. In this section, we will discuss the assumption of stationarity and Markovian structure.

4.2.1 Stationarity

A stationary process is a stochastic process whose joint probability distribution does not vary when shifted in time. Therefore, parameters such as the mean value, variance and auto-correlation structure also do not vary over time and do not follow any trends.

Formally, let X_t be a stochastic process and let $F_x(x_{t_1+\tau}, \ldots, x_{t_d+\tau})$ be the cumulative distribution function of the joint distribution of X_t at times $t_1 + \tau, \ldots, t_d + \tau$. So, X_t is supposed to be stationary if, for all d, for all τ, and for all $t_1, \ldots, t_d$,

$$F_x(x_{t_1+\tau}, \ldots, x_{t_d+\tau}) = F_x(x_{t_1}, \ldots, x_{t_d}). \tag{4.14}$$

Difference Stationary

We test for the presence of difference stationary using augmented DickeyFuller (ADF) test. In statistics and econometrics, an ADF test is a test for a unit root in a time series sample. A unit root is a feature of processes that evolve through time that can cause problems in statistical inference involving time series models. A linear stochastic process has a unit root if 1 is a root of the process's characteristic equation. Such a process is non-stationary. The ADF test is an augmented version of the DickeyFuller test for a larger and more complicated set of time series models. The ADF statistic, used in the test, is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit root at some level of

Augmented Dickey-Fuller Test			
	Dickey-Fuller	***p*-value (%)**	**Test Results**
EW	-5.52	1	stationary
SL	-4.22	1	stationary

Table 4.1: Estimates of augmented DickeyFuller test for mortality change index $\kappa_{t,1}$ (England and Wales) and $\kappa_{t,2}$ (Scotland) over 135 years (1876-2011)

confidence. See Kirchgàssner et al. (2012) for details on ADF test.

The results of ADF test for both series over 135 years are listed in Table 4.1. According to those results, ADF statistic values are both negative (-5.52 and -4.22 for EW and SL, respectively), strongly reject the hypothesis that there is a unit root. The rejection of the null hypothesis also can be seen by the p-values (both less than 5%), showing that the difference stationary properties exist.

Trend Stationary

We test for the presence of trend stationary using Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. In econometrics, KPSS tests are used for proving a null hypothesis that an observable time series is stationary around a deterministic trend. In 1991, Kwiatkowski et al. (1991) proposed a test of the null hypothesis that an observable series is trend stationary. KPSS type tests are intended to complement unit root tests, such as the Dickey-Fuller tests. By executing both ADF test and KPSS test, one can distinguish if series are stationary.

The results of KPSS test for both series over 135 years are listed in Table 4.2. According to those results, p-values are both equal to 10% and greater than significant level (5%), showing that trend stationary properties exist.

Kwiatkowski-Phillips-Schmidt-Shin Test		
	p**-value (%)**	**Test Results**
EW	10	stationary
SL	10	stationary

Table 4.2: Estimates of Kwiatkowski-Phillips-Schmidt-Shin test for mortality change index $\kappa_{t,1}$ (England and Wales) and $\kappa_{t,2}$ (Scotland) over 135 years (1876-2011)

4.2.2 Markovian Structure

In probability theory and statistics, a Markovian structure or Markov process, is a stochastic process that satisfies the Markov property. A Markovian structure can be considered as 'memoryless'. Generally speaking, a process satisfies the Markov property if one can make predictions for the future of the process depending only on its current state just as well as one depending on the process's full history. In other words, let a stochastic process be $S_t, t \geq 0$, S_t has a Markovian structure if, for all $u > v > 0$, the conditional probability distribution S_u, given the whole history of the process up to and including time v, depends only on the value of S_v.

Markovian structures are usually applied in probability and statistics in one of the following two ways. The first way is a stochastic process. A stochastic process may be proved mathematically to have the Markov property, and as a result to have the properties that can be deduced from this for all Markov processes. The second is in modeling a process. One may assume the process to be Markov, and take this as the basis for a construction. In modeling process, assuming that the Markov property holds is one of a limited number of simple ways of presenting

statistical dependence into a model for a stochastic process.

The reasons for our assumption of Markovian structure are numerous. For example, as shown in Chen and Cox (2009), the mortality factor κ_t is modeled with transitory jump effects such that

$$\kappa_t = \kappa_{t-1} + \mu + \sigma Z_t + Y_t N_t - Y_{t-1} N_{t-1}, \tag{4.15}$$

where μ and σ are constants, Y_t is the the jump severity variable in year t, N_t is the number of jumps occurring in year t, and Z_t is a standard normal random variable that is independent of Y_t and N_t. Equation 4.15 shows clearly that the mortality factor κ_t is a random walk with drift $\kappa_{t-1} + \mu + \sigma Z_t$ and can be predicted depending only on its current state.

Also, given the fact that mortality jumps, like the 1918 Spanish flu and the 2004 earthquake and tsunami, are induced by short-term catastrophic events, the mortality rates will decrease right after these events. We believe the mortality factor κ_t this year only depends on the factor last year and has a Markovian structure.

4.3 Estimation and Goodness-of-fit in Copula Models

In order to describe copula models in Section 4.1.1 in such a way that their values can predict the time series data, we need to estimate parameter of each copula. Start with a time series of 2-dimensional vectors $X_t = (\kappa_{t,1}, \kappa_{t,2})^{\mathrm{T}}$ with $t \in \{1, \ldots, 135\}$, and a copula C_θ which is associated with $(X_{t-1}^{\mathrm{T}}, X_t^{\mathrm{T}})^{\mathrm{T}}$. The goal is to calculate $\theta \in \vartheta \subset \mathbb{R}^p$ without any prior knowledge of the margins.

As proposed by Rémillard et al. (2012), first, given that the margins are unknown, X_t is substituted for its rank R_t. Next, define the sequence

$$\hat{U}_t = R_t/(n+1) \tag{4.16}$$

of normalized ranks. These pseudo-observations are then close to being uniformly distributed over $[0,1]$, when n is large enough. Set $Q_\theta(u) = C_\theta(u, \mathbf{1})$ for all $u \in [0,1]^2$ and recall that by hypothesis, $C_\theta(\mathbf{1}, v) = Q_\theta(v)$ for all $v \in [0,1]^2$.

4.3.1 Estimation by The Maximum Pseudo Likelihood Method

An extension of the maximum pseudo likelihood method to the Markovian case proposed by Genest (1995) is to define the maximum pseudo likelihood estimator $\hat{\theta}_n$ by

$$\hat{\theta}_n = \arg\max_{\tilde{\theta}\in\vartheta} \sum_{t=2}^{n} \log\left\{ \frac{c_{\tilde{\theta}}(\hat{U}_{t-1}, \hat{U}_t)}{q_{\tilde{\theta}}(\hat{U}_{t-1})} \right\}, \tag{4.17}$$

where $c_{\tilde{\theta}}$ is the density of $C_{\tilde{\theta}}$, assumed to be non-vanishing on $(0,1)^4$, and $q_{\tilde{\theta}}$ is the density of $Q_{\tilde{\theta}}$.

Under additional assumptions listed below, Rémillard et al. (2012) prove that the maximum pseudo likelihood estimator behaves nicely.

Assumptions. Suppose that,

(A1) c_θ is positive on $(0,1)^4$ and continuously differentiable as a function of θ; the gradient of c_θ with respect to θ is denote as $\dot{c}_\theta$.

(A2) For all $t > 2$,

$$\Delta M_t = G_\theta(U_{t-1}, U_t) = \frac{\dot{c}_\theta(U_{t-1}, U_t)}{c_\theta(U_{t-1}, U_t)} - \frac{\dot{q}_\theta(U_{t-1}, U_t)}{q_\theta(U_{t-1}, U_t)} \tag{4.18}$$

is square integrable.

(A3) $G_\theta(u, v)$ is continuously differentiable with respect to (u, v).

(A4) $(\mathbb{F}_{1,n}, \mathbb{F}_{2,n}) \to (\mathbb{F}_1, \mathbb{F}_2)$ as $n \to \infty$ in the Skorohod space $\mathcal{D}([0,1])^{\otimes 2}$, where, for $k = 1, 2$ and $u_k \in \mathbb{R}$,

$$\mathbb{F}_{k,n}(u_k) = n^{1/2} \left\{ F_{k,n}(u_k) - u_k \right\}, F_{k,n}(u_k) = \frac{1}{n} \sum_{t=1}^{n} \mathbb{I}(U_{k,t} \leq u_k). \tag{4.19}$$

Note that by (A1), the sequence U_t is ergodic [Bradley (2005), Theorem 3.5].

Assume conditions (A1)-(A4) and let

$$\theta_n = \arg\max_{\tilde{\theta} \in \vartheta} \sum_{t=2}^{n} \log \left\{ \frac{c_{\tilde{\theta}}(U_{t-1}, U_t)}{q_{\tilde{\theta}}(U_{t-1})} \right\}. \tag{4.20}$$

Then $\Theta_n = n^{1/2}(\hat{\theta}_n - \theta) \to \Theta \sim \mathcal{N}_p(0, \ell^{-1})$ as $n \to \infty$, where

$$\ell = \int_{(0,1)^4} \frac{\dot{c}_\theta(u, v) \dot{c}_\theta(u, v)^{\mathrm{T}}}{c_\theta(u, v)} \mathrm{d}v \mathrm{d}u - \int_{(0,1)^2} \frac{\dot{q}_\theta(u) \dot{q}_\theta(u)^{\mathrm{T}}}{q_\theta(u)} \mathrm{d}u, \tag{4.21}$$

and where $\dot{f}$ denotes the gradient with respect to θ. In addition, for the maximum pseudo likelihood estimator $\hat{\theta}_n$ defined by equation (4.17), one has $\hat{\Theta}_n = n^{1/2}(\hat{\theta}_n - \theta) \to \hat{\Theta} = \Theta + \tilde{\Theta} \sim \mathcal{N}_p(0, J)$, for some covariance matrix J, the joint law of Θ and

$\tilde{\Theta}$ is Gaussian, and

$$\begin{aligned}\tilde{\Theta} =& \ell^{-1} \int \nabla_u C_\theta(u,v) \left\{\mathbb{F}_1(u_1), \mathbb{F}_2(u_2)\right\}^{\mathrm{T}} \mathrm{d}C_\theta(u,v) \\ &+ \ell^{-1} \int \nabla_u C_\theta(u,v) \left\{\mathbb{F}_1(v_1), \mathbb{F}_2(v_2)\right\}^{\mathrm{T}} \mathrm{d}C_\theta(u,v).\end{aligned} \tag{4.22}$$

Because the covariance of $\mathbb{F}_1, \mathbb{F}_2$ is given by an infinite series, it would be extremely hard to obtain a direct estimation of the covariance matrix J of $\hat{\Theta} = \Theta + \tilde{\Theta}$. However, according to Rémillard (2011), using the results on parametric bootstrap for dynamic models, one could generate N samples of time series with copula $C_{\hat{\theta}_n}$ and estimate θ for each sample.

For each $k \in \{1, \ldots, 135\}$, let $\hat{\theta}_n^{(k)}$ denote the estimate of θ and write

$$z_k = n^{1/2}(\hat{\theta}_n^{(k)} - \hat{\theta}_n). \tag{4.23}$$

The random vectors $z_1, \ldots, z_{135}$ then converge to independent copies of $\hat{\Theta}$. Hence, one could estimate J by the empirical covariance $z_1, \ldots, z_{135}$.

4.3.2 Goodness-of-fit

Following on Genest et al. (2009), we use the Rosenblatt transform to construct goodness-of-fit tests for serial dependent data $(X_{t-1}^{\mathrm{T}}, X_t^{\mathrm{T}})^{\mathrm{T}}$, where $X_t = (\kappa_{t,1}, \kappa_{t,2})^{\mathrm{T}}$. Since we have substituted X_t for its pseudo-observations U_t and recall that Rosenblatt's transform of a 2-variate copula C is the mapping $\Re$ from $(0,1)^2 \to (0,1)^2$ so

that $u = (u_1, u_2) \mapsto \Re(u) = (e_1, e_2)$ with $e_1 = u_1$ and

$$e_k = \frac{\partial C(u_1, u_2, 1, 1)}{\partial u_1} / \frac{\partial C(u_1, 1)}{\partial u_1}. \tag{4.24}$$

A key property of Rosenblatt's transform is that $U \sim C$ if and only if $E = \Re(U) \sim \Pi$, i.e., E is uniformly distributed on $[0,1]^2$.

The use of Rosenblatt transforms based on a parametric bootstrap was proposed in Rémillard et al. (2010) and Rémillard and Papageorgiou (2008) for a multivariate regime switching Gaussian model. The validity of the parametric bootstrap approach for dynamic models is demonstrated in Rémillard (2011).

Recall that in the present setting, (U_t) is a stationary Markov process with $(U_{t-1}, U_t) \sim C$. The object is to prove the null hypothesis that C belongs to a given parametric family, i.e., $H_0 : C \in \{C_\theta : \theta \in \vartheta\}$. Let

$$\Re_\theta(u, v) = (\Re_\theta^{(1)}(u), \Re_\theta^{(2)}(u, v)) \tag{4.25}$$

be the Rosenblatt transform associated with the 4-dimensional copula C_θ, where $\Re_\theta^{(1)}$ is the Rosenblatt transform associated with the 2-dimensional copula Q_θ, and $\Re_\theta^{(2)}$ is the Rosenblatt transform associated with the conditional distribution of U_2 given $U_1 = u$. Accordingly, it follows that under the null hypothesis, the 2-dimensional time observations defined, for all $t \geq 2$, by

$$E_1 = \Re_\theta^{(1)}(U_1), E_t = \Re_\theta^{(2)}(U_{t-1}, U_t) \tag{4.26}$$

are independent and uniformly distributed over $[0,1]^2$.

Since θ is unknown and U_t is unobservable, θ must be estimated and U_t has to be replaced by a pseudo-observation $\hat{U}_t$. Suppose that $\hat{\theta}$ is a "regular" estimator of θ, in the sense of Genest and Rémillard (2008) and Rémillard (2011), based on the pseudo sample $\hat{U}_1, \ldots, \hat{U}_t$, and for all $t \geq 2$, set

$$\hat{E}_1 = \Re_{\hat{\theta}}^{(1)}(\hat{U}_1), \hat{E}_t = \Re_{\hat{\theta}}^{(2)}(\hat{U}_{t-1}, \hat{U}_t). \tag{4.27}$$

Under H_0, the empirical distribution function defined, for all $u \in [0,1]^2$, by

$$G_t(u) = \frac{1}{t}\sum_{i=1}^{t} \mathbb{I}(\hat{E}_i \leq u) \tag{4.28}$$

should be "close" to Π, the 2-dimensional independence copula. Mimicking Genest et al. (2009), one can prove the null hypothesis with a Cramér-von Mises type statistic

$$\begin{aligned} S_t = T(\mathbb{G}_t) &= \int_{[0,1]^2} \mathbb{G}_t^2(u)\mathrm{d}u = t\int_{[0,1]^2} \{G_t(u) - \Pi(u)\}^2 \,\mathrm{d}u \\ &= \frac{t}{3^2} - \frac{1}{2}\sum_{i=1}^{t}\prod_{k=1}^{2}(1-\hat{E}_{ik}^2) + \frac{1}{t}\sum_{i=1}^{t}\sum_{j=1}^{t}\prod_{k=1}^{2}\{1-\max(\hat{E}_{ik}, \hat{E}_{jk}),\} \end{aligned} \tag{4.29}$$

where $\mathbb{G}_t = t^{1/2}(G_t - \Pi)$. Using tools described Ghoudi and Rémillard (2004) together with the convergence results of the empirical processes described in Rémillard et al. (2012), one can determine that $\mathbb{G}_t$ converges to a (complicated) continuous centered Gaussian process $\mathbb{G}$. Regarding goodness-of-fit, the results of Genest and Rémillard (2008) on the parametric bootstrap can be extended to a Markovian set-

ting, showing that p-values for tests of goodness-of-fit based on the empirical copula or the Rosenblatt transform can be estimated by Monte Carlo methods. The proof of the validity of that extension is given in Rémillard (2011).

4.4 Model Calibration and Goodness-of-fit Results

Having identified the models that can estimate both serial dependence and interdependence structure simultaneously, the next step is to attempt to fit this copula-based Markovian models to our data. We choose to test the adequacy of two families of copulas: meta-elliptic copulas, such as Gaussian and Student-t; and Archimedean copulas, such as Clayton, Frank, Gumbel and Joe.

Gaussian copula

The Gaussian copula is denoted as

$$C_\rho^N(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \tag{4.30}$$

where $-1 < \rho < 1$ is the correlation coefficient. The Gaussian copula does not exhibit any tail dependence.

t Copula

The t Copula is derived from the t-distribution as:

$$C_{\nu,\rho}^t(u,v) = t_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v)), \tag{4.31}$$

where $t_{\nu,\rho}$ is the joint distribution function of a bi-variate Student-t distribution

with degree of freedom ν and correlation ρ. The t copula have both lower and upper tail dependence.

Clayton copula

The Clayton copula is given by

$$C_\theta^C(u,v) = (u^{-\theta} + v^\theta - 1)^{-1/\theta}, \tag{4.32}$$

where $0 < \theta < \infty$ is a parameter determining dependence structure of the copula. The Clayton copula is an asymmetric copula, which exhibits lower tail dependence only.

Gumbel copula

The Gumbel copula is given by

$$C_\theta^G(u,v) = \exp\left(-\left((-\ln(u))^\theta + (-\ln(v))^\theta\right)^{1/\theta}\right) \tag{4.33}$$

where $1 < \theta < \infty$ and the Gumbel copula exhibits upper tail dependence only.

Frank copula

The Frank copula is a symmetric copula, which is given by

$$C_\theta^F(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right), \tag{4.34}$$

where $\theta \in (-\infty, \infty)\setminus\{0\}$. It exhibits no lower or upper tail dependence. For $\theta \to \pm 0$, it implies independence structure.

Joe copula

The Joe copula takes the following form:

$$C_\theta^J(u,v) = 1 - \Big((1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta\Big)^{1/\theta}, \tag{4.35}$$

where $1 < \theta < \infty$. The Joe copula exhibits upper tail dependence. It is similar to Gumbel copula, but the upper tail dependence is stronger. Actually, it is closer to being the reverse of the Clayton copula.

The corresponding goodness-of-fit test results for those copulas are shown in Tables 4.3 and 4.4. The null hypothesis is then tested at the 5% level using the goodness-of-fit test described in Section 4.3. Genest and Rémillard (2008) show that an approximate p-value is given by

$$\frac{1}{2}\sum_{k=1}^{135} \mathbb{I}(S_t > S_t^*) \tag{4.36}$$

where S_t is given in Equation (4.29) and S_t^* is the asymptotic null distribution depending on the unknown value of θ. If the p-value is equal to or smaller than the 5% level, it suggests that the observed data are inconsistent with the null hypothesis that the observed data follows the specific copula, and thus that copula model must be rejected.

Therefore, the zero p-value for Joe copula indicates that it is strongly rejected. As other five copulas all have p-values higher than significant level (5%), we look for the copula with largest p-value. As we can see in Table 4.4, the Clayton copula exhibits the largest p-value, however, it is not the best copula for data prediction in this case. As we discussed above, Clayton copula is an asymmetric copula, which

exhibits lower tail dependence only. The form of copula is given by Equation (4.32). In this case, $\theta = 0.04$ which is very close to 0, implies independence structure and is against our assumption of dependence. See Chapter 3 for details on tests of dependence structure.

As we can see in Table 4.3, the Gaussian and Student-t copula models have p-values of 26% and 29%, respectively. Therefore, they are the best in estimating mortality change indexes. Also, we test copula models using the Akaike Information Criterion (AIC), which is also a measure of the quality of a statistical model. AIC chooses models by both maximizing value of the log pseudo-likelihood and minimizing the number of free parameters.

According to Akaike (1981),

$$\begin{aligned} AIC &= -2(\text{maximum log likelihood}) + 2(\text{number of free parameters}) \\ &= -2\hat{\ell}(\theta \,|X_i) + 2q \end{aligned} \tag{4.37}$$

where $\hat{\ell}(\theta \,|X_i)$ is the maximized value of the log pseudo-likelihood function $\ell(\theta \,|X_1, X_2, \ldots, X_d)$ defined as

$$\ln \ell(\theta \,|X_1, X_2, \ldots, X_d) = \sum_{i=1}^{d} \ln f(X_i \,|\theta), \tag{4.38}$$

and q is the number of estimated parameters.

As we can see from (4.37), AIC not only is a measure related to goodness-of-fit test, but also includes a penalty as the number of estimated parameters increases. The penalty of increasing parameters discourages over-fitting (increasing the number of free parameters in the model so that one can fit the data better).

	ρ				ν	p-value(%)	AIC
Gaussian	1.000	0.747	-0.197	-0.129	NA	26	-235.58
	0.747	1.000	-0.189	-0.335			
	-0.197	-0.189	1.000	0.747			
	-0.129	-0.335	0.747	1.000			
Student-t	1.000	0.691	-0.205	-0.133	3.47	29	-245.94
	0.691	1.000	-0.138	-0.334			
	-0.205	-0.138	1.000	0.691			
	-0.133	-0.334	0.691	1.000			

Table 4.3: Estimate of ρ, ν and p-values of the goodness-of-fit test and AIC values for the fitted Gaussian and Student-t copulas.

The AIC values in Tables 4.3 and 4.4 confirm that the Gaussian and the Student-t copulas are the best in estimating mortality change indexes because they have lowest AIC values. These two meta-elliptical copulas are also more advantageous in application because there are six parameters in each of these two copulas and each parameter exhibits different dependence structure. The Gaussian and student-t copulas allow greater flexibility in simulating the dependence structure of mortality rates. On the other hand, Archimedean copulas only depend on one parameter to describe the dependence structure between two mortality change indexes, forcing both interdependence and serial dependence follow the same dependence structure, which is not reasonable to us.

Also, the Gaussian and Student-t copula results confirm the positive correlation (0.747 for Gaussian and 0.691 for Student-t) between two mortality change indexes $(U_{1,t}, U_{2,t})$, negative correlation (-0.197 for Gaussian and -0.205 for Student-

	$\boldsymbol{\theta}$	p**-value(%)**	AIC		$\boldsymbol{\theta}$	p**-value(%)**	AIC
Clayton	0.040	63	-17.71	**Frank**	0.011	11	-7.42
Gumbel	1.000	16	-7.66	**Joe**	1.000	0	-3.24

Table 4.4: Estimate of θ and p-values of the goodness-of-fit test and AIC values for the fitted Clayton, Frank, Gumbel and Joe copulas.

t) between mortality change indexes of the two consecutive years $(U_{1,t-1}, U_{1,t})$ for the population of EW, and negative correlation (-0.335 for Gaussian and -0.334 for Student-t) between mortality change indexes of the two consecutive years $(U_{2,t-1}, U_{2,t})$ for the population of SL, which can not be modeled by the Clayton copula because the Clayton copula can just model positive correlation. In this thesis, we choose to use only the Student-t copula to predict the mortality rates because of three reasons. Firstly, Student-t copulas have been proved to be widely used in financial applications because of its tail dependence properties. Secondly, the AIC value of the Student-t copula is higher than that of the Gaussian copula. Last, the main difference between Gaussian and Student-t copulas are that Student-t copulas exhibit both lower and upper tail dependence while Gaussian copulas exhibit no tail dependence. Given the mortality data we got, both lower and upper tail dependence structure are obvious.

Specifically, Figure 4.1 shows a scatter plot of population EW and SL. In Figure 4.2 we transform Figure 4.1 using the quantile function of the standard normal distribution as marginal distributions to get realizations from it. We can see that in Figure 4.2, our data shows both lower and upper tail dependence structure as the Student-t copula. The comparison of the Gaussian and Student-t copulas is shown

in Figure 4.3, in which we show 2000 simulated points from Gaussian and Student-t copulas and transform these two copula data using the quantile function of the standard normal distribution as marginal distributions to get realizations from two different meta distributions. The Gaussian copula is simulated with parameter $\rho = 0.747$ and the Student-t copula is simulated with parameters $\nu = 3.47$ and $\rho = 0.691$, which are parameters fitted with time series $(U_{1,t}, U_{2,t})$ using the maximum pseudo-likelihood method. Therefore, we choose to use only the Student-t copula to do the forecasting.

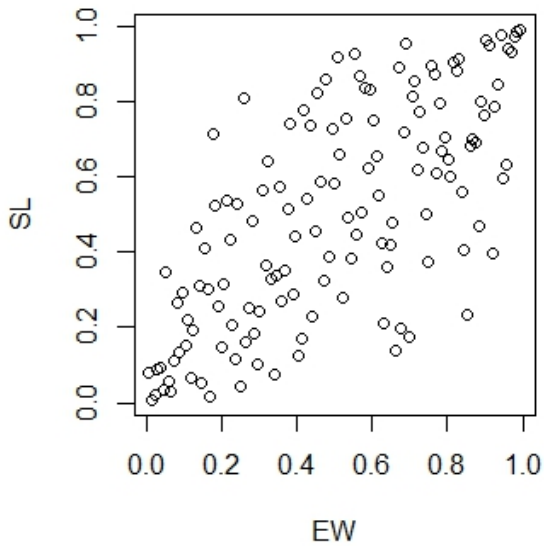


Figure 4.1: Scatter plot of England and Wales, and Scotland in uniform scale

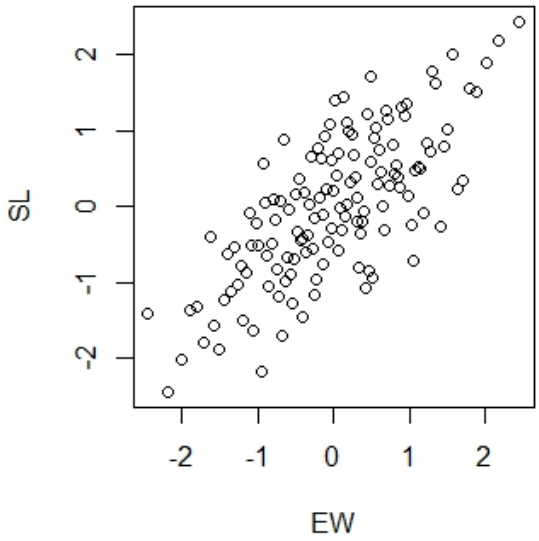


Figure 4.2: Scatter plot of England and Wales, and Scotland data using the quantile function transformation of the standard normal distribution

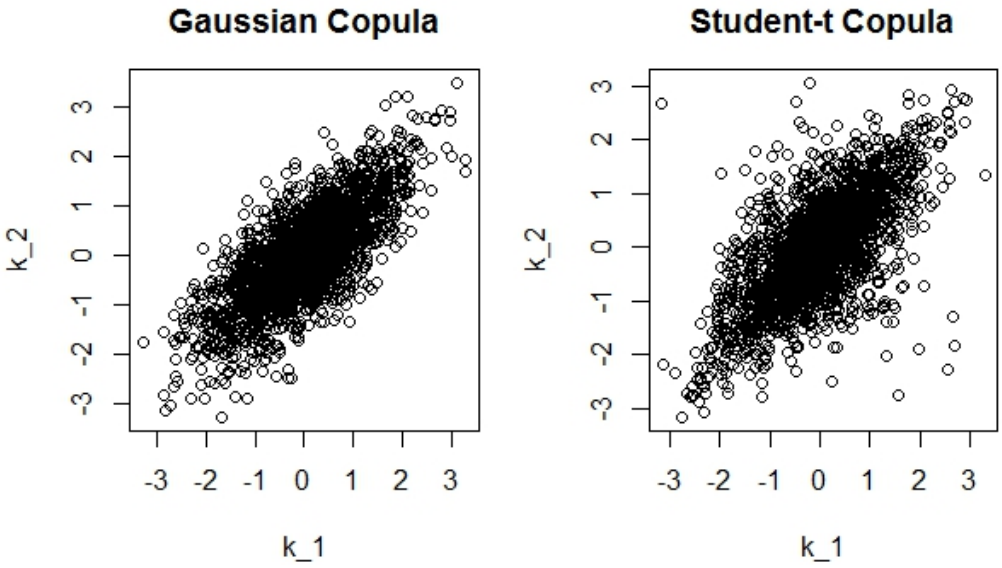


Figure 4.3: 2000 simulated points from two distributions with standard normal margins, constructed using the Gaussian (parameter $\rho = 0.747$) and Student-t (parameters $\nu = 3.47$ and $\rho = 0.691$) copulas data

Chapter 5

Forecasting and Measuring Population Basis Risk

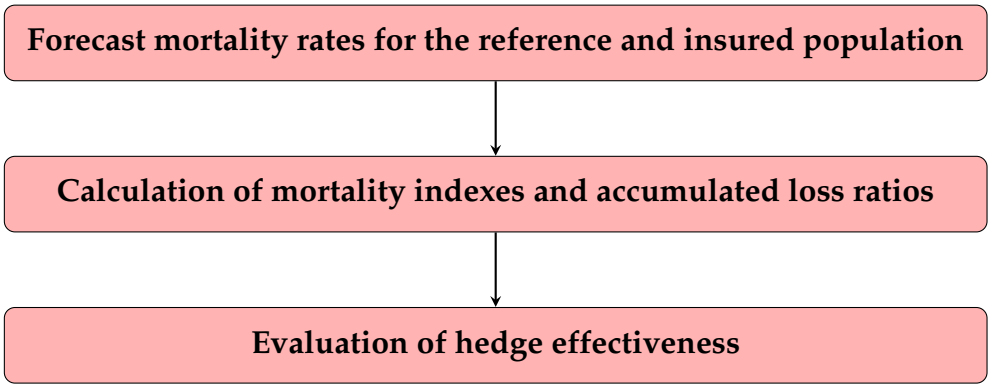
In Chapter 4 the Student-t copula has been proved to be the best copula for mortality change indexes $(\kappa_{t,1}, \kappa_{t,2})$. Hence, to quantify the population basis risk of the mortality catastrophe bond we are going to simulate future mortality change indexes $(\kappa_{t,1}, \kappa_{t,2})$ using Markovian models with the Student-t copula. The algorithm of the simulation has been introduced in Chapter 4.

We consider a typical life insurer's portfolio of individuals, which contains 100 thousand policyholders with average sum insured of £100 thousand British pounds. Suppose this portfolio comes from a so-called *insured population*. We ignore sampling risk, i.e., we assume the realized mortality is exactly the same as the intrinsic mortality of the insured population.

The life insurer's liability is to pay death benefits for this portfolio. Let X denote the amount in British pounds of the total claims from this portfolio in the five

years period of the bond contract. Then $X = 100,000N$ where N is the portfolio's total number of claims in the five years period. Since N is determined by the mortality rates of the insured population, the insurer is exposed to risk of unexpected mortality deterioration. In practice, insurers usually buy reinsurance to protect them from huge unexpected death benefit payments. We assume the life insurer is willing to pay up to an amount R, the so-called *ceding company's retention level*, and buys a stop-loss reinsurance from a re-insurer to cover the excess over R. Through this contract of reinsurance the extreme mortality risk is transferred from the insurer to the re-insurer and the re-insurer is obliged to pay the insurer $(X - R)_+$ no matter how large X would be, where $x_+ = \max(x, 0)$. The re-insurer can hedge its extreme mortality risk by issuing a mortality CAT bond.

The framework for forecasting mortality rates and assessing population basis risk is shown below. The life re-insurer's total extra costs are determined by the insured population (SL) mortality rates while the mortality catastrophe bond payoffs are determined by the reference population (EW) mortality rates. The estimation of hedge effectiveness uses calculations of weighted average mortality rates, mortality indexes and accumulated loss ratios, which are calculated by forecast mortality rates.



The remainder of this chapter is structured as follows. Section 5.1 outlines the methodology adopted for the forecasting. Section 5.2 shows quantification of population basis risk for the mortality catastrophe bond through hedge effectiveness.

5.1 Forecasting

After we have chosen the model we would like to forecast the $(\kappa_{t,1}, \kappa_{t,2})$ for five years in the future for each age group and convert them to central mortality rates using model (3.3). To do this we assume that the error term $\epsilon_{x,t}$ for each age group is Gaussian with a mean of zero and the standard deviation of σ_x. The simulations of future mortality rates are used to calculate weighted average mortality rates, mortality indexes and accumulated loss ratios using equations (2.1), (2.2) and (2.3).

Specifically, for each forecast scenario, we first simulate pseudo-observations of $(\kappa_{t,1}, \kappa_{t,2})$ by using Markovian models of the Student-t copula. Note that the parameters are given in Table 4.3 and the simulation algorithm is introduced in Section 4.1.1.

After we have simulated pseudo-observations of $(\kappa_{t,1}, \kappa_{t,2})$ we transform them back to mortality change indexes $\kappa_{t,1}$ and $\kappa_{t,2}$. In this thesis, we use linear interpo-

lation method to get $\kappa_{t,1}$ and $\kappa_{t,2}$. Specifically, for $i = 1, 2$ we first order the pseudo-observation obtained in Section 4.3 as $\hat{U}_{(1),i} \leq \ldots \leq \hat{U}_{(135),i}$ and correspondingly $\kappa_{(1),i} \leq \ldots \leq \kappa_{(135),i}$. If the pseudo-observation $U_{t,i}$ satisfies $\hat{U}_{(j-1),i} \leq U_{t,i} \leq \hat{U}_{(j),i}$ for some $j \in \{2, 3, \ldots, 135\}$, then the value $\kappa_{t,i}$ corresponds to $U_{t,i}$ is given by the equation

$$\kappa_{t,i} = \frac{U_{t,i} - \frac{j-1}{136}}{1/136} * \kappa_{(j),i} + \frac{\frac{j}{136} - U_{t,i}}{1/136} * \kappa_{(j-1),i}. \tag{5.1}$$

If $U_{t,i} < \hat{U}_{(1),i}$, then

$$\kappa_{t,i} = \kappa_{(1),i} + \ln\left(\frac{U_{t,i}}{1/136}\right). \tag{5.2}$$

On the other hand, if $U_{t,i} > \hat{U}_{(135),i}$, then

$$\kappa_{t,i} = \kappa_{(135),i} - \ln\left(\frac{1 - U_{t,i}}{1/136}\right). \tag{5.3}$$

Equations (5.2) and (5.3) try to approximate a value of $\kappa_{t,i}$ at a value outside the range of $\left\{\kappa_{(1),i}, \ldots, \kappa_{(135),i}\right\}$ using the exponential extrapolation method. Using the exponential extrapolation method, we assume that $\kappa_{t,i}$ goes to negative infinity as $U_{t,i}$ goes to 0, and that $\kappa_{t,i}$ goes to positive infinity as $U_{t,i}$ goes to 1. However, equations (5.2) and (5.3) can lead to a value of $\kappa_{t,i}$ that forces the mortality rate $m(x,t)$ greater than 1, which is not reasonable. Therefore, we add a constraint that

$$m_{x,t} \leq 1 \tag{5.4}$$

during the exponential extrapolation process.

Life re-insurers total extra losses

The life re-insurers total extra losses simulate from aggregate annual claims for a typical life re-insurers portfolio of individual fully underwritten yearly renewable term insurance. Since we suppose that the insurer reserves up to 99 percentile of next year's total claims, we calculate value-at-risk of the weighted average mortality rate q_t at 99 percentile. That is to say, the life re-insurer will expose to the extreme 1% mortality risk. Therefore, we define Excess Losses as

$$\text{Excess Losses}_i = (\text{TOT}_i - \text{VaR}_{0.99}(\text{TOT}))_+ \tag{5.5}$$

where TOT_i is the total insurance claims for scenario i, $i = 1, 2, \ldots, 100,000$,

$$\text{TOT}_i = 100,000 \times 100,000 \times \sum_{t=1}^{5} q_{i,t}. \tag{5.6}$$

The Excess Losses describes the amount of money that the reinsurance company need to pay to the life insurance company.

Mortality catastrophe bond payoff

The mortality catastrophe bond payoff is triggered based on a specified mortality index (2.2), as described in the Chapter 2. In reality, the mortality catastrophe bond must be adjusted according to the life re-insurer's hedging objectives, which we assume in this thesis is to protect against significant mortality claims arising from unexpected circumstance causing the reference population's mortality rates excess a specific mortality rate level. In other words, we do not expect the re-insurer to pay for protection against every mortality claim arising, they will only require protection once mortality claims reach a financially stressful level (or so-

called *trigger level* or *attachment point*).

As discussed in Chapter 2, the bond issuer must determine a trigger point (or attachment point), a, and the exhaustion level or detachment point, d. In this thesis, we assume a bond issuer may issue different levels of tranche to reach investors with different risk appetites. Accumulated loss ratios can be calculated according to different tranches using equation (2.3).

Therefore, we define bond payoff as

$$\text{Bond Payoff} = P * l_{t_0+5} \tag{5.7}$$

where P is the bond principal and l_{t_0+5} is the accumulated loss ratio (2.3) in year 5.

5.2 Measuring Population Basis Risk

Basis risk occurs whenever there are differences between an underlying hedged portfolio and the associated hedging instrument. Its existence means imperfect hedge effectiveness due to the possibility of gains or losses in the hedged position. Basis risk does not necessarily invalidate the effect for hedging since it can be minimized by properly structuring and calibrating the hedging instrument to ensure high hedge effectiveness. If the basis risk is small compared to the risk of the initial unhedged position, it means the hedging strategy is beneficial, see Coughlan et al. (2011).

Questions arise from basis risk have been examined for several index-based insurance linked securities (ILS). In life insurance industry, the existing literature

has primarily concentrated on longevity linked securities. For example, Cairns et al. (2014) prove that, at least for medium and large pension plans, longevity risk can be substantially hedged using index hedges as an alternative to customized longevity hedges; Coughlan et al. (2011) suggest that high levels of hedge effectiveness should be achieved with appropriately calibrated, static, index-based longevity hedges; Ngai and Sherris (2011) investigate the effectiveness of static hedging strategies for longevity risk management using longevity bonds and derivatives (q-forwards) for the retail products: life annuity, deferred life annuity, indexed life annuity, and variable annuity with guaranteed lifetime benefits.

In other industries, catastrophe derivative contracts, catastrophic loss index securities and industry loss warranties have been examined. For example, Harrington and Niehaus (1999) suggest that basis risk is not likely to be a significant problem with state-specific catastrophe derivative contracts; Cummins et al. (2004) analyze the effectiveness of catastrophic-loss index options in hedging hurricane losses for Florida insurers; Zeng (2000) presents an alternative measure of the industry loss warranty (ILW) basis risk, specifically the conditional probability that the ILW policy does not pay out, given an actual loss suffered by the policyholder that exceeds some critical level.

According to Coughlan et al. (2011), for mortality catastrophe bonds, basis risk could be differences between reference and insured populations due to initial or emerging mismatches in age, gender, geographical location and socioeconomic class, so-called population basis risk. In this thesis, we only consider the basis risk arises from these situations.

In this section we use the life re-insurer's total extra losses and mortality catastrophe bond payoffs introduced in Section 5.1 to quantify the population basis risk. We propose the following *hedge effectiveness* (HE) as risk measure:

$$\text{HE}(x) = \frac{\text{Bond payoff}}{\text{Excess losses}} \Bigg| \text{Excess losses} > x. \tag{5.8}$$

A similar HE definition can be found in Huynh et al. (2014), where HE is defined as the ratio of the bond payoff over the excess claims incurred as that in (5.8) but without the condition of the excess claim is greater than x.

HE indicates that the bond payoff is insufficient to cover the excess losses incurred when it is less than 100%; On the other hand, when HE is greater than 100%, it means that the bond payoff exceeds the excess losses incurred, thereby fully hedging. HE cannot be less than 0 because the bond payoff cannot be negative.

5.2.1 When Excess losses level is equal to 0

Assume that 80% of excess losses for the insured population can be hedged. Under this specified hedging objective, the life re-insurer intends to issue specific amount of the bond principal P to cover the excess losses caused by extreme mortality events. After determining the amount of the bond principal P, hedge effectiveness ratios of reference population (EW) bond payoff versus insured population (SL) excess losses are calculated. The difference between the hedge effectiveness using EW as reference population and that using SL as reference population

is caused by the population basis risk. The start of the bond period is assumed to be January 1st 2012, the maturity of the bond is five years.

We set he principal amount at £95,620,479 British pounds so that 80% of excess losses for the insured population can be fully hedged ($HE \geq 100\%$). Here, we use 113% as the attachment point at which the bond payoff is always triggered whenever there is excess losses, given that the reference population is also SL population, i.e. there is no population basis risk.

We run 100,000 simulations on the mortality rates forecasting for both EW and SL populations. Each simulation generates values for the actual aggregate claims and bond's accumulated losses, which are used to calculate the value of the HE. We conduct analysis based on two difference tranches to analyze the population basis risk in different conditions.

Table 5.1 shows hedge effectiveness using SL as reference population versus that using EW as reference population with two tranches ($a = 102\%$, $d = 132\%$; $a = 120\%$, $d = 150\%$). When the attachment point is equal to 102% and the detachment point is equal to 132%, the probability of HE greater than 0 is 100% for SL versus 82.9% for EW; when the attachment point is equal to 120% and the detachment point is equal to 150%, the probability of HE greater than 0 is 63.1% for SL versus 44.6% for EW, showing relatively high population basis risk. While the estimated mean HE are always greater than 100% for both populations under two different levels of tranches, the estimated median HE are fully-hedged when attachment point and detachment point are low and under-hedged when attachment point and detachment point are high. In the respect of probability of HE

Reference Population	**Attachment Point (%)**	**Detachment Point (%)**	**Mean (%)**	**Median (%)**	$\Pr(HE > 0)$ **(%)**
SL	102	132	2876	435	100.00
EW	102	132	556	150	82.90
SL	120	150	194	90	63.10
EW	120	150	125	0	44.60

Table 5.1: Estimated HE when excess losses level is 0 (Principal: £95,620,479)

greater than 0, the results suggest that the population basis risk is relatively low when tranches are low; on the other hand, In the respect of estimated mean HE, the results suggest that the population basis risk is relatively low when tranches are high. These results are not in conflict with each other. Overall, this suggests that bond issuers should synthesize each kind of situation when it comes to population basis risk.

Figure 5.1 shows that the estimated density of HE using EW as reference population is more spread, more positively skewed and less peaked compared to that using SL as reference population when tranches are low. On the other hand, when tranches are high, the estimated density of HE using EW as reference population is more spread, more positively skewed and more peaked compared to that using SL as reference population. The difference between HE using EW as reference population and that using SL as reference population is population basis risk. Overall, this proves the potential beneficial effect for re-insurers of choosing different tranches when dealing with different situations.

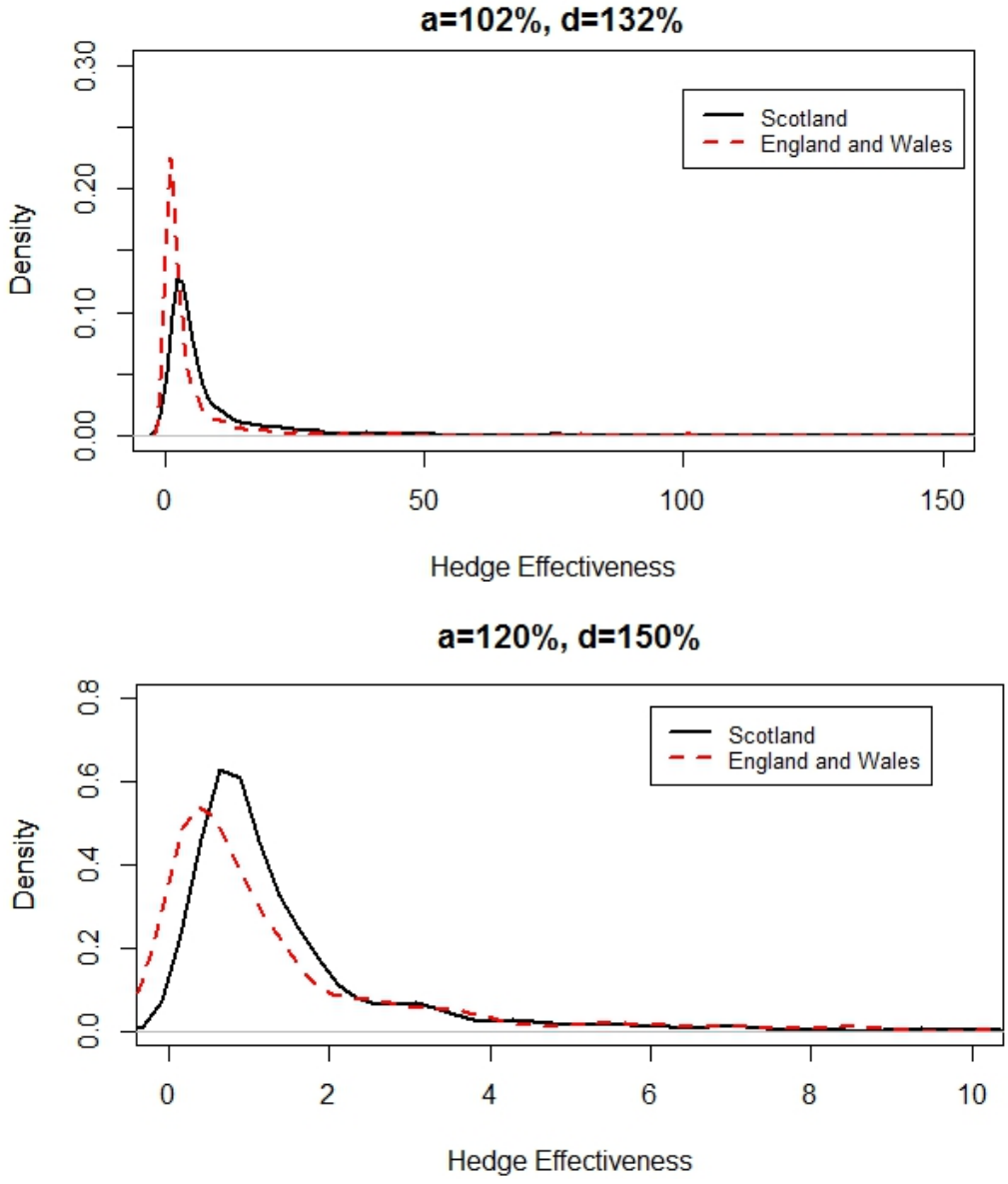


Figure 5.1: Estimated HE when excess losses level is 0 (Principal: £95,620,479)

5.2.2 When Excess losses level is greater than 0

As life re-insurers focus mainly on extreme mortality events, a key measure is the effectiveness of the hedge under more extreme circumstances. Hence, different levels of excess losses are examined. Table 5.2 reports the estimated HE when

varying the excess losses levels. In the respect of the probability of HE greater than 0, as the excess losses levels increase, the population basis risk becomes smaller when the tranches are low, and the population basis risk becomes higher when the tranches are high. The reason for these phenomenon is that the first group of tranches ($a = 102\%$, $d = 132\%$) is in the central part of the mortality indexes while the second group of tranches ($a = 120\%$, $d = 150\%$) is at the tail of the mortality indexes.

On the other hand, in the respect of the estimated mean HE, the population basis risk decreases under two different levels of tranches as the excess losses levels increase. These two results are not in conflict with each other. These results just confirm that bond issuers should synthesizes each kind of situation when it comes to population basis risk.

Figures 5.2, 5.3, 5.4 and 5.5 demonstrate that the estimated density distribution of the HE using EW as reference population becomes closer to that using SL as reference population as the excess losses values increase from £5,000,000 to £10,000,000, than £20,000,000, and finally £30,000,000. Overall, these figures show that the population basis risk decreases when extreme circumstances are under consideration.

Reference Population	**Attachment Point (%)**	**Detachment Point (%)**	**Mean (%)**	**Median (%)**	$\Pr(HE > 0)$ **(%)**
		x=£5,000,000, Figure 5.2			
SL	102	132	411	334	100.00
EW	102	132	193	128	84.96
SL	120	150	111	86	73.20
EW	120	150	51	0	48.42
		x=£10,000,000, Figure 5.3			
SL	102	132	307	275	100.00
EW	102	132	155	117	88.08
SL	120	150	97	84	81.00
EW	120	150	43	5	52.33
		x=£20,000,000, Figure 5.4			
SL	102	132	217	214	100.00
EW	102	132	112	102	92.61
SL	120	150	86	80	92.88
EW	120	150	33	11	57.26
		x=£30,000,000, Figure 5.5			
SL	102	132	182	181	100.00
EW	102	132	93	88	93.78
SL	120	150	84	78	99.17
EW	120	150	48	12	60.58

Table 5.2: Estimated HE when varying excess losses levels (Principal: £95,620,479)

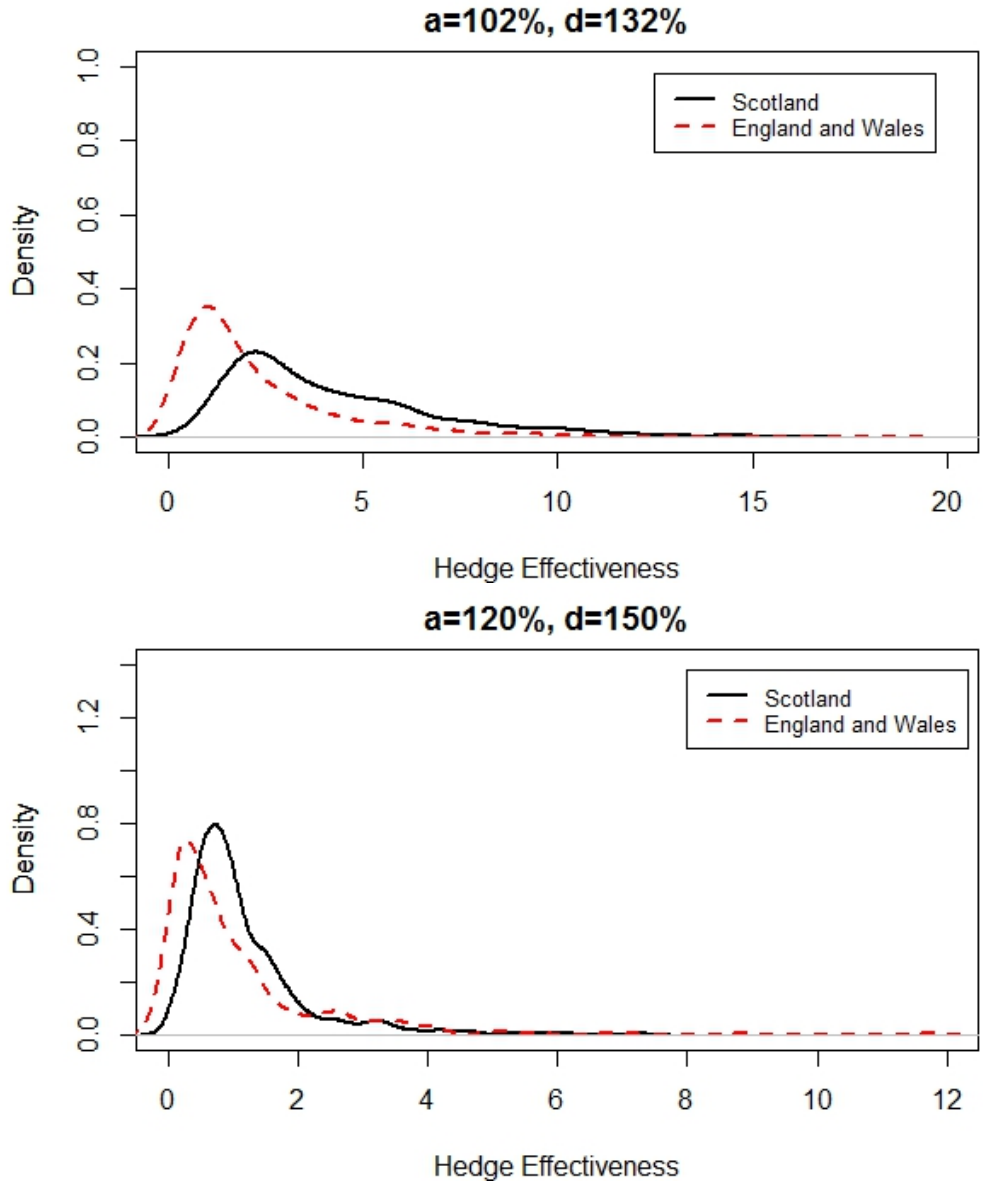


Figure 5.2: Estimated density of HE when excess losses level is £5,000,000 (Principal: £95,620,479)

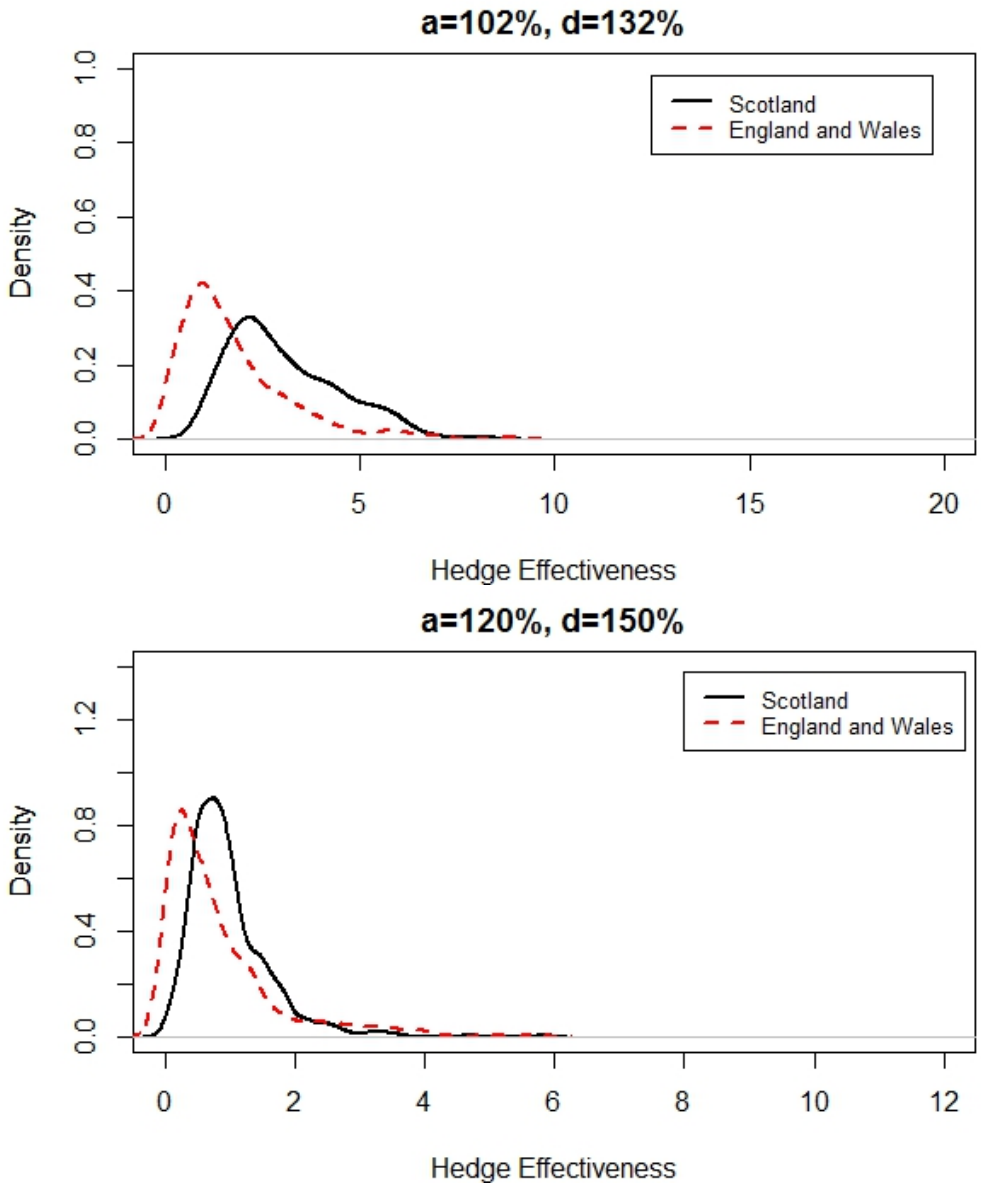


Figure 5.3: Estimated density of HE when excess losses level is £10,000,000 (Principal: £95,620,479)

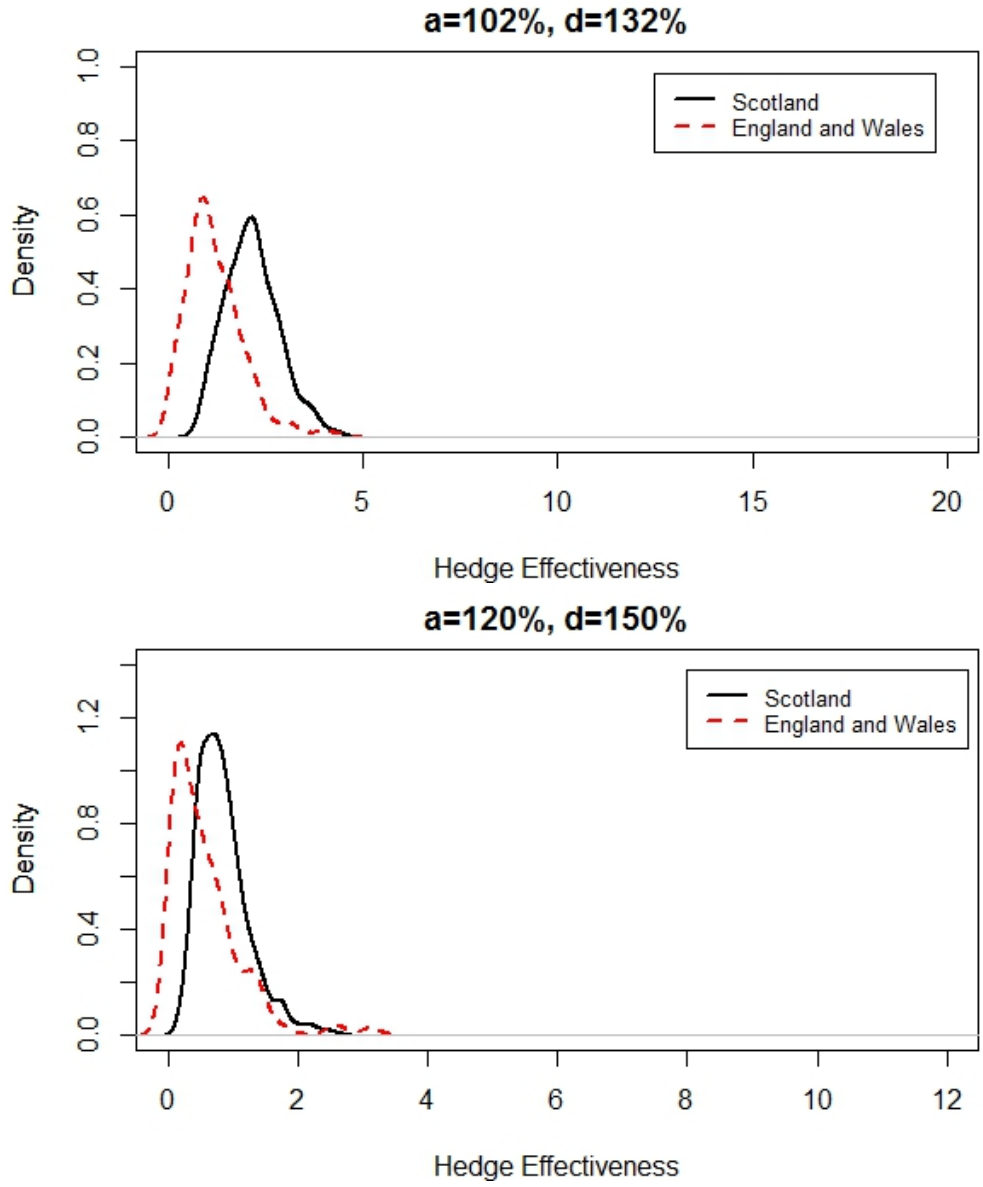


Figure 5.4: Estimated density of HE when excess losses level is £20,000,000 (Principal: £95,620,479)

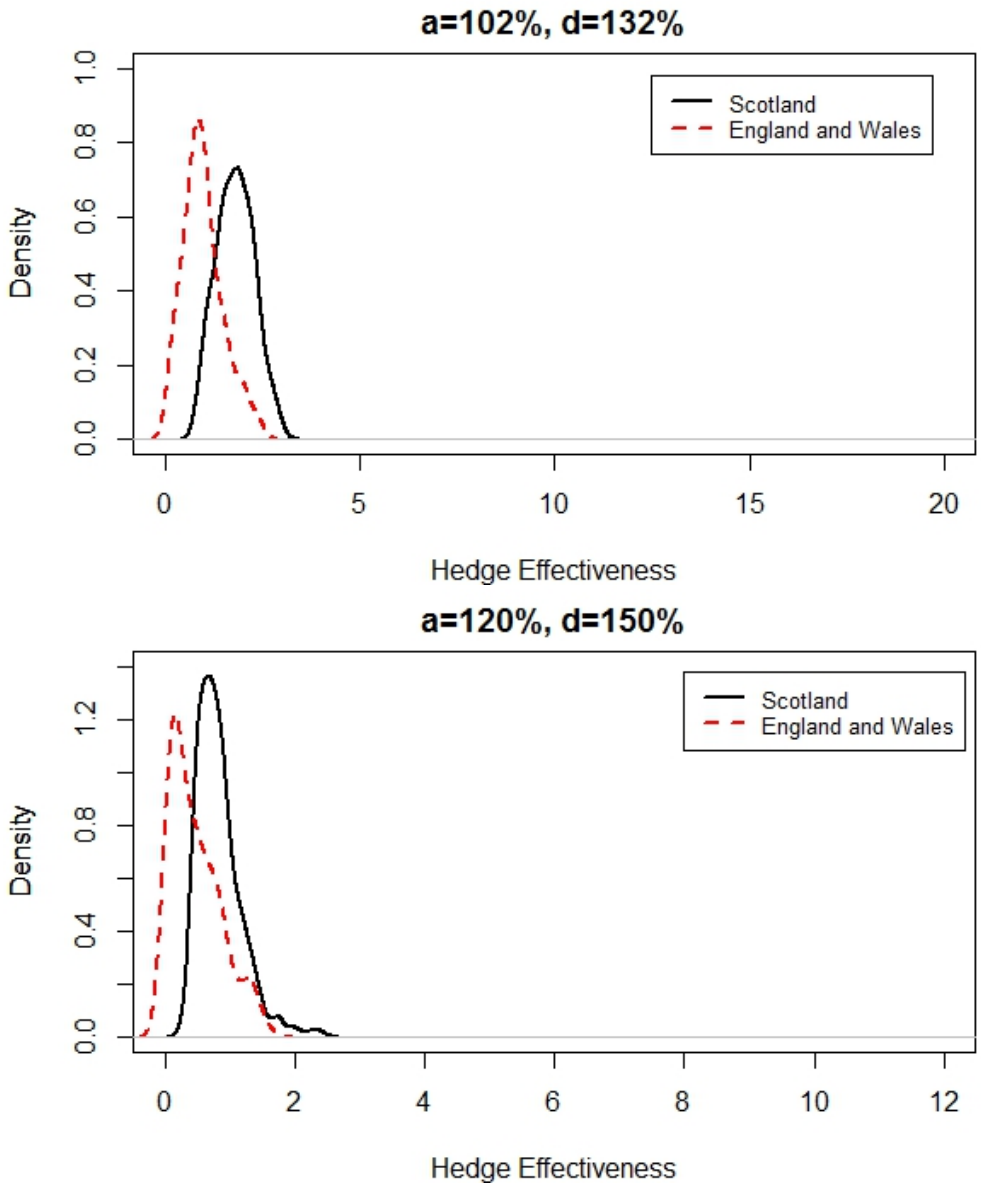


Figure 5.5: Estimated density of HE when excess losses level is £30,000,000 (Principal: £95,620,479)

Chapter 6

Conclusions

This chapter concludes the analysis of this thesis. First, the results that were derived in the analysis will be evaluated. After this, some possibilities to extend the analysis will be discussed.

6.1 Summary of Results

The analysis of this thesis concentrated on the population basis risk of a fictitious mortality catastrophe bond which shares the same payoff structure as that of most mortality bonds in practice. Normally, basis risk occurs whenever there are differences between an underlying hedged portfolio and the associated hedging instrument. In this thesis, we only considered basis risk arises between reference and insured populations due to initial or emerging mismatches in age, gender, geographical location and socioeconomic class, so-called population basis risk.

To analyze the population basis risk of this mortality catastrophe bond, we first

obtained the mortality change indexes $\kappa_{t,1}$ and $\kappa_{t,2}$ by calibrating the MBMM mortality model on England and Wales (EW), and Scotland (SL) populations. In Chapter 3 the comparisons of the MBMM mortality model versus the Lee-Carter model are showed. We can see that the MBMM mortality model is much better than the Lee-Carter model which calculates the mortality rates for next year without considering the mortality rates this year. After obtaining the mortality change indexes $\kappa_{t,1}$ and $\kappa_{t,2}$, we tested the interdependence and serial dependence between them. For the interdependence, The Spearmans Rho and Kendalls Tau are tested to be 0.70 and 0.51, respectively, indicating that two series are positively correlated. For serial dependence, since the null hypothesis of Ljung-Box test are both rejected, two mortality change indexes series show serial dependence structure. Therefore, the use of copula-based semi-parametric models was investigated.

The copula-based semi-parametric models can tackle simultaneously serial dependence and interdependence between time series $\kappa_{t,1}$ and $\kappa_{t,2}$. It differs from the usual approach to time series copula modeling in which the series are first modeled individually and copulas are used to model the dependence. We executed parameter estimation and goodness-of-fit tests for the models, considering meta-elliptical copulas and Archimedean copulas. The results showed that Gaussian and Student-t copulas are best copulas for data prediction, with p-values 26% and 29% respectively. By using Gaussian or Student-t copulas they allow greater flexibility in simulating the dependence structure of mortality rates, because there are six parameters in each of these two copulas and each parameter exhibits different dependence structure; see Chapter 4. In this thesis, we chose to use only the

Student-t copula to predict the mortality rates because of three reasons. Firstly, Student-t copulas have been proved to be widely used in financial applications because of its tail dependence properties. Secondly, the AIC value of the Student-t copula is higher than that of the Gaussian copula. Last, the main difference between Gaussian and Student-t copulas are that Student-t copulas exhibit both lower and upper tail dependence while Gaussian copulas exhibit no tail dependence. Given the mortality data we got, both lower and upper tail dependence structure are obvious.

Finally, we proposed a method to measure the population basis risk of mortality CAT bonds. The analysis found that in the respect of the probability of HE greater than 0, as the excess losses levels increase, the population basis risk becomes smaller when the tranches are low, and the population basis risk becomes higher when the tranches are high. On the other hand, in the respect of the estimated mean HE, the population basis risk decreases under two different levels of tranches as the excess losses levels increase. These two results are not in conflict with each other. These results just confirm that bond issuers should synthesizes each kind of situation when it comes to population basis risk.

6.2 Discussions

To derive the results of this thesis, many choices had to be made. Some of these choices were related to model selection. While other choices could have been made that probably would have resulted in different outcomes. Hence, this section discusses some alternatives that could have been chosen and that can be used in

further research.

The analysis of this thesis concentrated solely on the mortality catastrophe bond without paying any coupon. In reality however, the mortality catastrophe bond may pay quarterly coupons or be constructed to be a more complicated structure. When constructing the mortality index, we considered a ratio of any two consecutive years before maturity to two consecutive years basic levels. One can also construct the mortality index considering only one year weighted average mortality rate.

To analyze and forecast the mortality rates, the MBMM mortality model was used. However, other complicated models can be used as long as they model changes of log mortality rates rather than levels of log mortality rates, such as a two factor model. Making forecasting of future mortality rates using two factor model is very similar to using the one factor model. In the case of two factors we need a model to fit $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$. Because $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are obtained using the singular value decomposition they are necessarily uncorrelated, we assume that $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are independent. This implies that they may be fit independently to each other.

In this thesis the Student-t copula was used in the analysis of the dependence structure between time series $(\kappa_{t,1}, \kappa_{t,2})$. However, one also can use the Gaussian copula if they observe higher Gaussian copula p-value and no tail dependence structure.

Finally, the choice of pricing the bond principle was somewhat arbitrary, because a wide range of pricing principles exists.

Appendix A

Introduction to Copulas

The well-known Sklar's Theorem says that any multivariate joint distribution can be written in terms of the uni-variate marginal distribution function and a copula that depicts the correlation structure between the variables (Sklar (1959)). This means that every multivariate cumulative distribution function

$$F(x_1, \ldots, x_d) = \Pr[X_1 \leq x_1, \ldots, X_d \leq x_d] \tag{A.1}$$

of a random vector $(X_1, X_2, \ldots, X_d)$ with marginals $F_i(x) = \Pr[X_i \leq x]$ can be written as

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \tag{A.2}$$

where C here is a copula.

In 1981, Schweizer and Wolff (1981) present the first paper relating copulas to the study of correlation among random variables and used copulas to specify various natural nonparametric measures of correlation for pairs of random vari-

ables. Subsequently, several papers follow this work using copulas to connect multivariate distributions to their one-dimensional margins, see Joe (1997) and Nelsen (2006).

There are many desirable features of copulas, which make them popular in simulating models. Some of these features include non-linear dependence, ability to quantify correlation for heavy tail distributions, very flexible, etc. and are described in more details below.

Non-linear dependence

Imagine that we have two data sets and we want to find out the correlation between them. Obviously, there are lots of methods that we can use. Usually people will assume that the relation between them is linear, however, what if it is not? In that case, the correlation coefficient might be zero although there is strong spatial correlation between data sets. The most common method that can be used to overcome this shortcoming is copula, which is a function that enables us to combine uni-variate distributions to get a joint distribution with a specific correlation structure.

Ability to quantify correlation for heavy tail distributions

Since many distributions are non normal (there are deviations from normality), it can be difficult to calculate the joint distribution models. For example, if there is tail dependence in the distribution of asset prices. In this case, copula is an attractive method, see Bouye and Roncalli (2000). In fact, even when the correlations remain the same, some extreme cases are much more remarkable under copulas. For example, if we use the same correlation coefficient such as ρ=0.3, the copu-

las under Gaussian copulas or under Student-*t* copulas (copulas whose properties focus on issues related to the correlation of extreme values) will show different behaviors. The fact that copulas are able to model extreme cases that due to the tail-dependence is examined by many studies, see Schmidt (2007).

Very flexible: parametric, semi-parametric or non-parametric

Copulas are very flexible and specifying marginal distributions using them is very straightforward. The copula divides the marginal properties and its dependence structure, therefore giving a more flexible model. Also, the copula approach has a useful function when one needs to obtain joint distributions given marginal distributions, especially when variables are non normal. For example, a copula can be used to analyze how exchangeable categorical data can be incorporated into two sets of correlated data (parents and offspring), each of which is characterized by an aggregation parameter, see Trégouét et al. (1999). Another example is given that copula models can be considered to analyze the dependence between the three pairs of exchange rates (Euro, Japanese Yen and Great Britain Sterling Pound) against the US dollar using non-parametric and semi-parametric methods, see Boero et al. (2011).

Quicker and more stable computation procedure

Copula parameters can be fitted by many methods, one of which is called Monte Carlo experiment. Large scale Monte Carlo experiment can be simulated in a short period of time, confirming that the computation procedures of copulas can provide an efficient and much faster method to test the dependence structure.

A.1 Basic Properties

Property 1. A d-dimensional copula is a distribution function on $[0,1]^d$ with standard uniform marginal distributions.

Let $C(\mathbf{u}) = C(u_1, \ldots, u_d)$ for the multivariate dfs that are copulas. Therefore C is a mapping of the form $C : [0,1]^d \to [0,1]$, i.e. a mapping of the unit hypercube into the unit interval. The following three properties must hold.

(1.1) $C(u_1, \ldots, u_d)$ is increasing in each component u_i.

(1.2) $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $i \in \{1, \ldots, d\}$, $u_i \in [0,1]$.

(1.3) For all $(a_1, \ldots, a_d), (b_1, \ldots, b_d) \in [0,1]^d$ with $a_i \leq b_i$ we have

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\ldots+i_d} C(u_{1i_1}, \ldots, u_{di_d}) \geq 0, \tag{A.3}$$

where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j \in \{1, \ldots, d\}$.

The property (1.1) is obviously required of any multivariate df and the property (1.2) is the requirement of uniform marginal distributions. The property (1.3) is less obvious, however the so-called rectangle inequality in (A.3) ensures that if the random vector $(U_1, \ldots, U_d)'$ has df C, then $\Pr(a_1 \leq U_1 \leq b_1, \ldots, a_d \leq U_d \leq b_d)$ is non-negative. These three properties characterize a copula; if a function C meets all of them, it is a copula. Note that, for $2 \leq k < d$, the k-dimensional margins of a d-dimensional copula are themselves copulas.

In working with copulas we must be familiar with the operations of probability and quantile transformation, as well as the properties of generalized inverses. The

following property can be found in numerous probability texts.

Property 2. Let G be a distribution function and let $G^{\leftarrow}$ denote its generalized inverse, i.e. the function $G^{\leftarrow}(y) = \inf\{x : G(x) \geq y\}$.

(2.1) **Quantile transformation.** If $U \sim U(0,1)$ has a standard uniform distribution, accordingly $\Pr(G^{\leftarrow}(U) \leq x) = G(x)$.

(2.2) **Probability transformation.** If Y has df G, where G is a continuous univariate df, accordingly $G(Y) \sim U(0,1)$.

Proof. Let $y \in \mathbb{R}$ and $u \in (0,1)$. For the first part use the fact that

$$G(y) \geq u \Leftrightarrow G^{\leftarrow}(u) \leq y \tag{A.4}$$

from which it follows that

$$\Pr(G^{\leftarrow}(U) \leq y) = \Pr(U \leq G(y)) = G(y). \tag{A.5}$$

For the second part we conclude that

$$\Pr(G(Y) \leq u) = \Pr(G^{\leftarrow} \circ G(Y) \leq G^{\leftarrow}(u)) = \Pr(Y \leq G^{\leftarrow}(u)) = G \circ G^{\leftarrow}(u) = u, \tag{A.6}$$

where the first inequality follows from the fact that $G^{\leftarrow}$ is strictly increasing.

Property (2.1) is the key to stochastic simulation. If we can generate a uniform variate U and calculate the inverse of a df G, accordingly we can sample from that df. Properties (2.1) and (2.2) together mean that we can transform risks with a

particular continuous df to have any other continuous distribution. For instance, if Y has a standard normal distribution, accordingly $\Phi(Y)$ is uniform by Property (2.1), and, because the quantile function of a standard exponential df G is $G^{\leftarrow}(y) = -\ln(1-y)$, the transformed variable $Z := -\ln(1-\Phi(Y))$ has a unit exponential distribution by Property (2.2).

A.2 Elliptical Copulas

Elliptical copulas are simply the copulas of elliptically contoured (or elliptical) distributions. Consider, for example, building a distribution with the Gaussian copula C_P^N but arbitrary margins; such a model is known as a *meta-Gaussian* distribution. Another example is that *meta-t_ν* distribution has the copula $C_{\nu,P}^t$ and arbitrary margins.

Tail dependence property

Coefficients of tail dependence. Consider a pair of uniform rvs (U_1, U_2) whose distribution $C(u_1, u_2)$ is a normal variance mixture copula. Because of the radial symmetry of C (see McNeil et al. (2005)), it suffices to consider the formula for the lower tail-dependence coefficient to calculate the coefficient λ

$$\lambda = \lim_{q\to 0^+} \frac{\mathrm{d}C(q,q)}{\mathrm{d}q} = \lim_{q\to 0^+} \Pr(U_2 \le q | U_1 = q) + \lim_{q\to 0^+} \Pr(U_1 \le q | U_2 = q). \quad \text{(A.7)}$$

Since C is exchangeable we have

$$\lambda = 2 \lim_{q\to 0^+} \Pr(U_2 \le q \,|U_1 = q\,). \quad \text{(A.8)}$$

The below shows the difference between Gaussian and Student-t copulas that Student-t copulas have tail dependence, while the Gaussian copulas are asymptotically independent in the tail.

Example (asymptotic independence of the Gaussian copula). To evaluate the tail-dependence coefficient for the Gaussian copula C_ρ^N, let $(X_1, X_2) := (\Phi^{-1}(U_1), \Phi^{-1}(U_2))$, so that (X_1, X_2) has a bi-variate normal distribution with standard margins and correlation ρ. It follows equation (A.8) that

$$\begin{aligned}\lambda &= 2 \lim_{q \to 0^+} \Pr(\Phi^{-1}(U_2) \leq \Phi^{-1}(q) | \Phi^{-1}(U_1) = \Phi^{-1}(q)) \\ &= 2 \lim_{x \to -\infty} \Pr(X_2 \leq x | X_1 = x).\end{aligned} \tag{A.9}$$

Using the fact that $X_2 | X_1 = x \sim N(\rho x, 1 - \rho^2)$, it can be calculated that

$$\lambda = 2 \lim_{x \to -\infty} \Phi(x\sqrt{1-\rho}/\sqrt{1+\rho}) = 0, \tag{A.10}$$

provided $\rho < 1$. Therefore, the Gaussian copula is asymptotically independent in both tails.

Example (asymptotic dependence of the Student-t copula). To evaluate the tail-dependence coefficient for the Student-t copula $C_{\nu,\rho}^t$, $(X_1, X_2) := (t_\nu^{-1}(U_1), t_\nu^{-1}(U_2))$, where t_ν denotes the distribution function of a uni-variate Student-t distribution with ν degrees of freedom. Therefore, $(X_1, X_2) \sim t_2(\nu, \mathbf{0}, P)$, where P is correlation matrix with off-diagonal element ρ. By calculating the conditional density from the joint and marginal densities of a bi-variate Student-t distribution, it can be checked

that, conditional on $X_1 = x$,

$$\left(\frac{\nu+1}{\nu+x^2}\right)^{1/2} \frac{X_2 - \rho x}{\sqrt{1-\rho^2}} \sim t_{\nu+1}. \tag{A.11}$$

Using an argument similar to Example above we find that

$$\lambda = 2t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right). \tag{A.12}$$

Provided that $\rho > -1$, the copula of the bi-variate Student-t distribution is asymptotically dependent in both the upper and lower tail.

Contour plot of the bi-variate Gaussian copula with $\rho = 0.7$ is shown on the left in Figure A.1, while contour plot for the Student-t copula with parameters $\rho = 0.7, \nu = 2$ is shown on the right in Figure A.1. Note that the marginals for both distributions are uniform and the only difference is the structure of the correlation. The contour plots comfirm that the Student-t copula produces more extreme events.

A.3 Archimedean Copulas

The Clayton copula (4.32), the Gumbel copula (4.33), the Frank copula (4.34) and the Joe copula (4.35) belong to the family of so-called Archimedean copulas, which has been very widely studied. In this section we look at the simple structure of these copulas and some properties that we need in the thesis.

Bi-variate Archimedean copulas

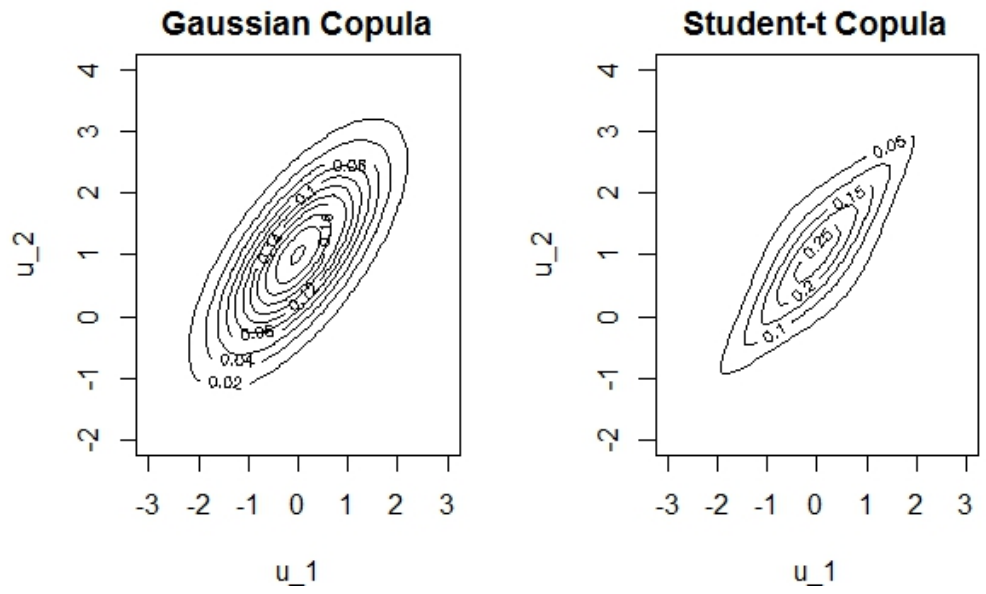


Figure A.1: Contour-plots of the Gaussian (parameter $\rho = 0.7$) and Student-t (parameters $\rho = 0.7, \nu = 2$) copulas, respectively.

All four examples that we have introduced have the form

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)), \tag{A.13}$$

where ϕ is a decreasing function from $[0,1]$ to $[0,\infty]$, satisfying $\phi(0) = \infty, \phi(1) = 0$, known as the generator of the copula, and ϕ^{-1} is its inverse.

To obtain a copula in a case when $\phi(0) < \infty$ we present a so-called pseudo-inverse of the generator and give a theorem that explains exactly when a construction resembling (A.13) yields a copula.

Definition (pseudo-inverse). Suppose $\phi : [0,1] \to [0,\infty]$ is continuous and strictly decreasing with $\phi(1) = 0$ and $\phi(0) \leq \infty$. We define a pseudo-inverse of ϕ

with domain $[0, \infty]$ by

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & 0 \leq t \leq \phi(0), \\ 0, & \phi(0) < t \leq \infty. \end{cases} \tag{A.14}$$

Theorem (bi-variate Archimedean copula). Let $\phi : [0,1] \to [0,\infty]$ be continuous and strictly decreasing with $\phi(1) = 0$ and $\phi^{[-1]}(t)$ as in (A.14). Then

$$C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2)) \tag{A.15}$$

is a copula if and only if ϕ is convex.

All copulas constructed according to (A.15) are called bi-variate Archimedean copulas. If $\phi(0) = \infty$ the generator is supposed to be strict and we may replace the pseudo-inverse $\phi^{[-1]}$ by the ordinary functional inverse ϕ^{-1} as in (A.13).

Archimedean Copulas have a wide range of applications for some reasons: firstly, it is easy to be constructed; secondly, there are many families of copulas belong to it and each of them show different dependence structure which can fit data widely; last, it has many nice properties.

Contour plot of the bi-variate Clayton copula with $\theta = 2$ is shown on the left in Figure A.2, while contour plot for the Gumbel copula with parameters $\theta = 2$ is shown on the right in Figure A.2. Note that the marginals for both distributions are uniform and the only difference is the structure of the correlation. They appear to have lower and upper tail dependence structure, respectively, as discussed above.

Multivariate Archimedean copulas

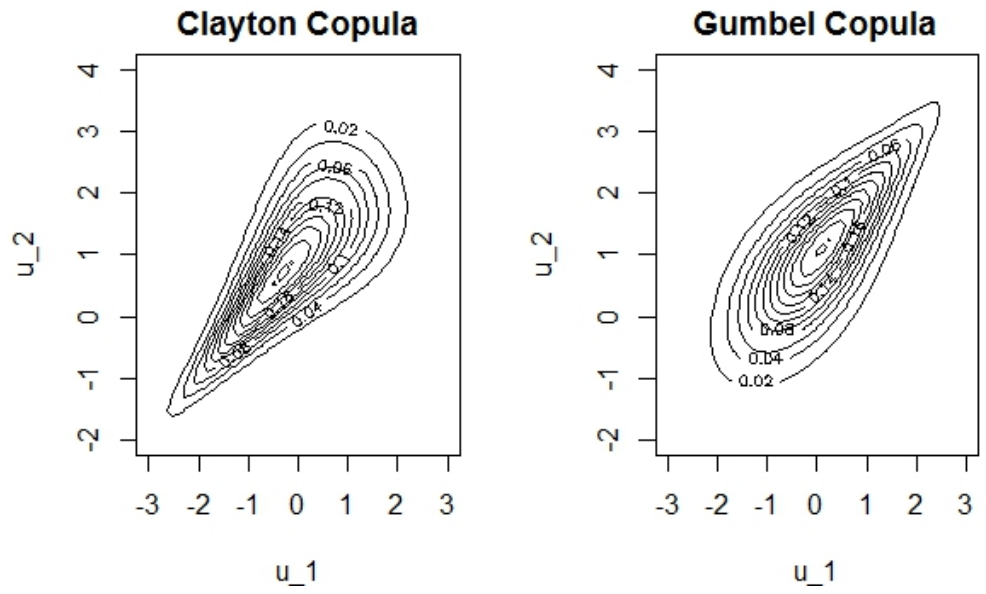


Figure A.2: Contour-plots of the Clayton (parameter $\theta = 2$) and Gumbel (parameter $\theta = 2$) copulas, respectively.

It seems natural to attempt to construct a higher-dimensional Archimedean copula according to $C(u_1, \ldots, u_d) = \phi^{[-1]}(\phi(u_1) + \ldots + \phi(u_d))$. While this structure may fail to define an appropriate distribution function for arbitrary dimension d. An example where this happens is obtained if we assume the generator $\phi(t) = 1 - t$, which is not strict. In this case we obtain the Fréchet lower bound for copulas, which is not itself a copula for $d > 2$.

A required condition for the d-dimensional construction to succeed in all dimensions is that ϕ should be a strict Archimedean copula generator, though this is not sufficient. It is proved by Kimberling (1974) that if $\phi : [0, 1] \to [0, \infty]$ is a strict Archimedean copula generator, accordingly

$$C(u_1, \ldots, u_d) = \phi^{-1}(\phi(u_1) + \ldots + \phi(u_d)) \tag{A.16}$$

gives a copula in any dimension d if and only if the generator inverse $\phi^{-1} : [0, \infty] \to [0, 1]$ is strictly monotonic. A decreasing function $f(t)$ is strictly monotonic on an interval $[a, b]$ if it satisfies

$$(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k} f(t) \geq 0, k \in \mathbb{N}, t \in (a, b). \tag{A.17}$$

All of the generators mentioned above have inverses which are strictly monotonic on $[0, \infty]$ (if we restrict to $\theta \geq 0$ for the Clayton copula) and all extend to arbitrary dimensions using the construction (A.16). For instance, a d-dimensional Clayton copula is

$$C_\theta^C(u) = (u_1^{-\theta} + \ldots + u_d^{-\theta} - d + 1)^{-1/\theta}, \theta \geq 0, \tag{A.18}$$

where the limiting case $\theta = 0$ should be interpreted as the d-dimensional independence copula.

Another method of describing these Archimedean copulas which extend to arbitrary dimensions is in terms of Laplace-Stieltjes transforms of dfs on $\mathbb{R}^+$, because every strictly monotonic function mapping from $[0, \infty]$ to $[0, 1]$ can be expressed in terms of such transforms. Let G be a df on $\mathbb{R}^+$ satisfying $G(0) = 0$ with Laplace-Stieltjes transform

$$\hat{G}(t) = \int_0^\infty \exp^{-tx} \mathrm{d}G(x), t \geq 0. \tag{A.19}$$

If we define $\hat{G}(\infty) := 0$, it is not hard to check on that $\hat{G} : [0, \infty] \to [0, 1]$ is a continuous, strictly decreasing function with the property of strict monotonicity (A.17).

A.4 The Conditional Copula

In this section, we analyze the properties of the conditional copula in a general context. It is applied to multivariate time series in Chapter 4.

Let H be the cumulative function of the joint distribution of the $(d_1 + d_2)$-dimensional random vector (X, Y), where X has continuous marginal distribution $F_1, \ldots, F_{d_1}$ and Y has continuous marginal distributions $G_1, \ldots, G_{d_2}$. Sklar (1959) states that the cumulative distribution function, for all $x_1, \ldots, x_{d_1}, y_1, \ldots, y_{d_2} \in \mathbb{R}$, can be written as

$$H(x, y) = C\left\{F_1(x_1), \ldots, F_{d_1}(x_{d_1}), G_1(y_1), \ldots, G_{d_2}(y_{d_2})\right\} \tag{A.20}$$

where C is a unique $(d_1 + d_2)$-dimensional copula.

Following Rémillard et al. (2012), we assume that copula C is absolutely continuous with density c and that densities f_i of F_i and g_i of G_i for all $i \in \{1, \ldots, d_1\}$ and $j \in \{1, \ldots, d_2\}$ can be written as

$$h(x_1, \ldots, x_{d_1}, y_1, \ldots, y_{d_2}) = c\left\{F_1(x_1), \ldots, F_{d_1}(x_{d_1}), G_1(y_1), \ldots, G_{d_2}(y_{d_2})\right\} \times \prod_{i=1}^{d_1} f_i(x_i) \prod_{j=1}^{d_2} g_j(y_j). \tag{A.21}$$

As we set $u = \mathbf{F}(x) = (F_1(x_1), \ldots, F_{d_1}(x_{d_1}))$ and $U_t = \mathbf{F}(X_t)$, we can derive from equations (A.20) and (A.21) that the density of X is $f_X(x) = c_U(u) \times f_1(x_1) \times \ldots \times f_{d_1}(x_{d_1})$, where c_U is the density of the copula $Q(u) = C(u, \mathbf{1})$. Similarly, we can set $v = \mathbf{G}(y) = (G_1(y_1), \ldots, G_{d_2}(y_{d_2}))$ and $V = \mathbf{G}(Y)$. Then the conditional density of

Y given $X = x$ can be written as

$$f_{Y|X}(y;x) = \frac{f(x,y)}{f_X(x)} = c_{V|U}(v;u)\prod_{j=1}^{d_2} g_j(y_j), \tag{A.22}$$

where

$$c_{V|U}(v;u) = \frac{c(u,v)}{c_U(u)} \tag{A.23}$$

is the conditional density of V given $U = u$. The density (A.23) is the density of $V = (G_1(Y_1), \ldots, G_{d_2}(Y_{d_2}))$ given $U = (F_1(X_1), \ldots, F_{d_1}(X_{d_1}))$ = u. Hence, the conditional copulas of V given U and Y given X are the same.

Sections 4.1.1 and 4.1.2 are a few examples of application in the general case where $X \in \mathbb{R}^{d_1}$ and $Y \in \mathbb{R}^{d_2}$. In the Markovian case, $d_1 = d_2 = d$; in other applications, e.g. p-Markov processes, one could have $d_1 \neq d_2$.

Bibliography

H. Akaike. Likelihood of a model and information criteria. *Journal of Econometrics*, 16, 1981.

G. Bagus. Taming the cat: securitizing pandemic losses offers life insurers a guaranteed way to fund potential claims. *Best's Review*, pages 92–94, 2007.

P. Barbe, C. Genest, K. Ghoudi, and B. Rémillard. On kendall's process. *Journal of Multivariate Analysis*, 58(2):197–229, 1996.

D. Blake, A. J. G. Cairns, and K. Dowd. Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, 12(1):153–228, 2006.

G. Boero, P. Silvapulle, and A. Tursunalieva. Modelling the bivariate dependence structure of exchange rates before and after the introduction of the euro: a semi-parametric approach. *International Journal of Finance & Economics*, 16(4):357–374, 2011.

H. Booth, P. D. Jong, R. J. Hyndman, and L. Tickle. Lee-carter mortality forecasting:

a multi-country comparison of variants and extensions. *Demographic Research*, 15:289–310, 2006.

A. N. A. R. G. Bouye, E.and Durrleman and T. Roncalli. Copulas for finance: A reading guide and some applications. *Groupe de Recherche Operationnalle, Credit Lyonnais, France.*, 2000.

R. C. Bradley. Basic properties of strong mixing conditions. a survey and some open questions. *Probab. Surveys 2*, pages 107–114, 2005.

A. J. Cairns, K. Dowd, D. Blake, and G. D. Coughlan. Longevity hedge effectiveness: a decomposition. *Quantitative Finance*, pages 217–235, 2014.

A. J. G. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *The Journal of Risk and Insurance*, 73(4):687–718, 2006.

B. A. Carnes and S. J. Olshansky. A biologically motivated partitioning of mortality. *Experimental Gerontology*, 325:615–631, 1997.

H. Chen and S. H. Cox. Modeling mortality with jumps: Applications to mortality securitization. *The Journal of Risk and Insurance*, 76(3):727–751, 2009.

G. Coughlan, M. Khalaf-Allah, Y. Ye, S. Kumar, A. Cairns, D. Blake, and K. Dowd. Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal*, 15:150–176, 2011.

A. Cowley and J. D. Cummins. Securitization of life insurance assets and liabilities. *Journal of Risk and Insurance*, 72(2):193–226, 2005.

S. H. Cox and H. Yungui. Modeling mortality risk from exposure to a potential future extreme event and its impact on life insurance. Technical report, Department of Risk Management and Insurance J. Mack Robinson College of Business, 2004.

J. Cummins, D. Lalonde, and R. D. Phillips. The basis risk of catastrophic-loss index securities. *Journal of Financial Economics*, 71(1):77–111, 2004.

C. Genest and R. J. Mackay. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 14(2):145–159, 1986a.

C. Genest and R. J. Mackay. The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician*, 40(4):280–283, 1986b.

C. Genest and B. Rémillard. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Inst. H. Poincaré Probab. Statist.*, 44(6): 1096–1127, 2008.

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance Mathematics and Economics*, 44(2):199–213, 2009.

K. Genest, C.and Ghoudi. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552, 1995.

K. Ghoudi and B. Rémillard. Empirical processes based on pseudo-observations. ii. the multivariate case. *Asymptotic methods in stochastics*, 44:381–406, 2004.

B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115:513–583, 1825.

Y. Guan, B. J. Zheng, Y. Q. He, X. L. Liu, Z. X. Zhuang, C. L. Cheung, S. W. Luo, P. H. Li, L. J. Zhang, Y. J. Guan, K. M. Butt, K. L. Wong, K. W. Chan, W. Lim, and K. F. Shortridge. Isolation and characterization of viruses related to the sars coronavirus from animals in southern china. *Science (New York, N.Y.)*, 302(5643): 276–278, 2003.

S. Harrington and G. Niehaus. Basis risk with pcs catastrophe insurance derivative contracts. *The Journal of Risk and Insurance*, 66(1):49–82, 1999.

A. Huynh, B. Browne, and A. Bruhn. A review of catastrophic risks for life insurers. *Risk Management and Insurance Review*, 16(2):233–266, 2013.

A. Huynh, A. Bruhn, and B. Browne. Catastrophic mortality bonds: Analysing basis risk and hedge effectiveness. *Australian Journal of Actuarial Practice*, 1:45–62, 2014.

H. Joe. *Multivariate models and dependence concepts*. London ; New York : Chapman & Hall, 1997.

I. T. Jolliffe. *Principal component analysis*. New York : Springer-Verlag, 1986.

P. D. Jong and L. Tickle. Extending lee-carter mortality forecasting. *Mathematical Population Studies*, 13:1–18, 2006.

T. H. Kean and H. Lee. The 9/11 commission report : final report of the national commission on terrorist attacks upon the united states. Technical report, Washington, D.C. : National Commission on Terrorist Attacks upon the United States : For sale by the Supt. of Docs., U.S. G.P.O., 2004.

C. Kimberling. A probabilistic interpretation of complete monotonicity. *aequationes mathematicae*, 10(2):152–164, 1974.

G. Kirchgàssner, J. Wolters, and U. Hassler. *Introduction to Modern Time Series Analysis*. Dordrecht : Springer, 2012.

D. Kwiatkowski, P. C. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54: 159–178, 1991.

R. D. Lee and L. R. Carter. Modeling and forecasting us mortality. *Journal of the American Statistical Association*, 87(419):659–671, 1992.

W. M. Makeham. On the law of mortality. *Journal of the Institute of Actuaries (1866-1867)*, 13:325–358, 1867.

A. J. Mcneil and J. Neŝlehovà. Multivariate archimedean copulas, d-monotone functions andl_1-norm symmetric distributions. *The Annals of Statistics*, 37(5): 3059–3097, 2009.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management : concepts, techniques and tools*. Princeton, N.J. : Princeton University Press, 2005.

M. Mesfioui and J.-F. Quessy. Dependence structure of conditional archimedean copulas. *Journal of Multivariate Analysis*, 99(3):372–385, 2008.

D. Mitchell, P. Brockett, R. Mendoza-Arriaga, and K. Muthuraman. Modeling and forecasting mortality rates. *Insurance: Mathematics & Economics*, 52(2):275–285, 2013.

R. B. Nelsen. *An introduction to copulas*. New York : Springer, 2006.

A. Ngai and M. Sherris. Longevity risk management for life and variable annuities: The effectiveness of static hedging using longevity bonds and derivatives. *Insurance Mathematics and Economics*, 49(1):100–114, 2011.

J. S. Nguyen-Van-Tam and A. W. Hampson. The epidemiology and clinical impact of pandemic influenza. *Vaccine*, 21(16):1762–1768, 2003.

J. Peiris, L. Poon, Y. Guan, J. Nicholls, K. Yuen, S. Lai, L. Yam, W. Lim, W. Yee, W. Yan, M. Cheung, V. Cheng, D. Tsang, R. Yung, T. Ng, and K. Chan. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*, 361: 1319–1325, 2003.

B. Rémillard. Validity of the parametric bootstrap for goodness-of-fit testing in dynamic models. Technical report, Technical report 3, SSRN Working Paper Series No 1966476, 2011.

B. Rémillard and N. Papageorgiou. Modelling asset returns with markov regime switching models. Technical report, Tech. Report 3, DGAM-HEC Alternative Investments Research, 2008.

B. Rémillard, A. Hocquard, and N. Papageorgiou. Option pricing and dynamic discrete time hedging for regime-switching geometric random walks models. Technical report, SSRN Working Paper Series No 1591146, 2010.

B. Rémillard, N. Papageorgiou, and F. Soustra. Copula-based semiparametric models for multivariate time series. *Journal of Multivariate Analysis*, 110:30–42, 2012.

A. Renshaw and H. Steven. Mortality reduction factors incorporating cohort effects. Technical report, Actuarial Research Paper, 2005.

T. Schmidt. *Coping with Copulas*. Risk Books, J. Rank (Ed.), Risk Books, 2007.

B. Schweizer and E. F. Wolff. On nonparametric measures of dependence for random variables. *The Annals of Statistics*, 9:879–885, 1981.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231, 1959.

D. Strómberg. Weather and infant mortality in africa. In *C.E.P.R. Discussion Papers*, 2012.

D.-A. Trégouét, P. Ducimetiére, V. Bocquet, S. Visvikis, F. Soubrier, and L. Tiret. A parametric copula model for analysis of familial binary data. *The American Journal of Human Genetics*, 64:886–893, 1999.

P. K. Trivedi and D. M. Zimmer. *Copula Modeling: An Introduction for Practitioners*, volume 1. Foundations and trends in econometrics, 2005.

K. Woolnough, B. Ivanovic, S. Kramer, and J. Busenhart. *Pandemic influenza: A 21st century model for mortality shocks*. Zurich : Swiss Re, Swiss Reinsurance Company, cop, 2007.

S. Zeidan. Desperately seeking definition: the international community's quest for identifying the specter of terrorism. *Cornell International Law Journal*, 36(3):491, 2004.

L. Zeng. On the basis risk of industry loss warranties. *The Journal of Risk Finance*, 1:27–32, 2000.