

**A MODIFIED LEE-CARTER APPROACH FOR
MORTALITY MODELLING AND FORECASTING
UNDER A NON-GAUSSIAN INNOVATION**

OLUWASEUN ENIOLAOLADIPUPO (111925)

PDS, B.Sc., M.Sc., (Statistics) (Ibadan)

November, 2019

**A MODIFIED LEE-CARTER APPROACH FOR
MORTALITY MODELLING AND FORECASTING
UNDER A NON-GAUSSIAN INNOVATION**

BY

OLUWASEUN ENIOLAOLADIPUPO (111925)

PDS, B.Sc., M.Sc., (Statistics) (Ibadan)

**A THESIS IN THE DEPARTMENT OF STATISTICS
SUBMITTED TO THE FACULTY OF SCIENCE IN
PARTIAL FUFILMENT OF THE REQUIREMENT
OF THE AWARD OF DOCTOR OF PHILOSOPHY**

November, 2019

ABSTRACT

The Lee-Carter (LC) model was primarily designed for modelling the mortality pattern with Gaussian error structure in most developed countries. Attempts at using LC for describing the pattern of mortality in developing countries have resulted in the violation of the normality assumption, due to its inability to accommodate variability across age-groups. Therefore, this study was aimed at developing a modified LC model to accommodate variability across all ages.

The LC model, $\ln m_{xt} = a_x + b_x k_t + \varepsilon_{xt}$ was modified using a gamma generator given as: $Gam(y) = \frac{\{-\ln [1-F(y)]\}^{\alpha-1}}{\Gamma\alpha} f(y)$; where m_{xt} is a random variable, $\ln m_{xt} = y$, a_x , b_x and k_t are location parameters, ε_{xt} is an error term, α is a shape parameter, $\Gamma\alpha$ is $(\alpha - 1)!$, $f(y)$ and $F(y)$ are the density and distribution functions of the normal random variable, respectively. The Kolmogorov Smirnov (KS) and Shapiro Wilks' (SW) procedures were used to test for normality. The Gamma-Normal Lee-Carter (GNLC) model, the LC model with normal error structure and the Brouhns' (BR) model with poisson error structure were fitted to Nigeria male and female all-cause mortality datasets obtained from the Global Health Observatory for age-groups < 1, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84 and 85⁺ for the period 2000-2015. Forecast of the mortality Index k_t was done in an Auto-regressive Integrated Moving Average (ARIMA) framework for a 20 year-period. Simulation study using data of sizes $n; t = (15; 5), (20; 20)$ and $(20; 30)$ was set up, where n and t are the sizes of the age-groups and periods, respectively. Model performance was evaluated using Bayesian Information Criterion (BIC) and the Corrected Akaike Information Criterion (CAIC). Significance was determined at 0.01 level.

The Probability Density Function (PDF), Cummulative Distribution Function (CDF) and Hazard Function (HF) of the GNLC model were derived. The PDF, CDF and HF of the GNLC were of the forms:

$$g(y) = \frac{1}{\Gamma\alpha} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left\{\frac{[y-(a_x+b_xk_t)]^2}{\sigma^2}\right\}} \left(-\ln\left\{1 - \Phi\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]\right\}\right)^{\alpha-1},$$

$$G(y) = \frac{1}{\Gamma\alpha} \gamma\left(\alpha, -\ln\left\{1 - \Phi\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]\right\}\right) \text{ and}$$

$$h(y) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left\{\frac{[y-(a_x+b_xk_t)]^2}{\sigma^2}\right\}} \left(-\ln\left\{1 - \Phi\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]\right\}\right)^{\alpha-1}}{\Gamma\left(\alpha, -\ln\left\{1 - \Phi\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]\right\}\right)}, \text{ respectively. The KS and SW}$$

procedures were significant at 0.01, confirming non-normality with P-values less than or equal to $(4.35 \times 10^{-13}, 2.22 \times 10^{-5})$ and $(4.18 \times 10^{-13}, 1.58 \times 10^{-13})$ for males and females, respectively. The BIC and CAIC for the forecast period were: -0.14, -0.55 (LC), -17.19, -17.60 (BR) and -40.99, -41.40 (GNLC) for the male and -7.05, -7.39 (LC), -14.74, -15.04 (BR) and -56.76, -57.06 (GNLC) for the female. For the simulated data, the obtained BIC and CAIC for the LC and GNLC models were for sample size $n; t = (15; 5)$: 285.6405, 263.8951; 281.3217, 257.4478, $n; t = (20; 20)$: 744.1807, 532.7462; 738.1579, 529.9777 and $n; t = (20; 30)$: 997.1916, 726.5509; 990.7245, 713.258, respectively.

The modified Lee-Carter model was able to accommodate variability across all age-groups better than the referenced classical Lee-Carter. Therefore, the modified Lee-Carter model is recommended for modelling mortality data from developing countries.

Keywords: Lee-Carter model, Mortality pattern, Non-Gaussian error, Gamma generator.

Word count: 465

CERTIFICATION

I certify that this work was carried out

By

Oluwaseun E. Oladipupo

In the Department of Statistics,

University of Ibadan

.....

(Supervisor)

Dr. Angela U. Chukwu

B.Sc. (Calabar), M.Sc. (Ibadan), PhD (Ibadan)

Department of Statistics, University of Ibadan

ACKNOWLEDGEMENT

My profound gratitude goes to my supervisor Dr. (Mrs.) A. U. Chukwu, for her timely counsel, guidance and encouragement when I lost focus and couldn't make progress on this study. Dr. O.S. Ewemooje kept encouraging me for months. It amazed me when I discovered that she was actually the one behind the camera. Ma, you have been more than a supervisor, you tutored and nurtured me like your own. My mentor per excellence, thank you for believing in me. May the Lord grant your heart desires over your children and continually distinguish you and your family for excellence in Jesus name (Amen).

I sincerely appreciate my Parents, Mr. and Mrs. Oladipupo and Mrs. Funmilayo Adeniyi for your support of so many years and how you stood in the gap for me; your joy will know no bounds over us in Jesus name (Amen). Also to my siblings; Mrs. Damilola Ojo, OluwaSegun Oladipupo, Oluwaseyifunmi Oladipupo, Rev. Femi Adeniyi and Mr. Opeyemi Adeniyi for your prayers and words of encouragement, I am grateful, God bless you all abundantly.

I acknowledge all the Lecturers in the department of Statistics: Prof. Iyaniwura, Prof. Bamiduro, Prof. Sodipo, Prof. G.N. Amahia, Prof. O.I. Shittu, Prof. Olubusoye, Dr. (Mrs.) A.A. Adepoju, Dr. J.F. Ojo, Dr. O.I. Osowole, Dr. (Mrs.) Alaba, Dr. C.G. Udomboso, Dr. Obisesan, Dr. (Mrs.) Otekunrin and Dr. S.O. Oyamakin. I am very grateful for all your corrective contributions.

This acknowledgement will not be complete if I fail to appreciate my spiritual father and mother at home; Venerable Prof. and Dr. (Mrs.) Amu. Thank you Sir and Ma for all your prayers and encouragement. God bless your ministry exceedingly.

DEDICATION

This work is dedicated to the Love of my life, the reason behind my existence, the Almighty God. In Him I live, in Him I move and in Him I have my being.

I am grateful to my husband and Head (Adegbilero-Iwari Idowu), for his love and support in every ramification. Id mi, your extraordinary encouragement and prayers all through these years made this great achievement a reality. Thank you for the sacrifices you made on my behalf, may God reward you exceedingly. To my Trinity boys, TiOluwa, UyiOluwa and EwaOluwa, thank you for those days you endured my absence in the home and our travelling to and fro. We did the work together! Yes, indeed, you are for signs and wonders. Team Iwari, this degree is yours. To God be the glory!

CONTENTS

Abstract	i
Certification	iii
Acknowledgement	iv
Dedication	v
Table of Contents	vi
List of Figures	xi
List of Tables	xii
1.0 Introduction	
1.1 Background to the Study	1
1.2 Statement of Problem	3
1.3 Objectives of the Study	4
1.4 Motivation	4
1.5 Thesis Layout	5
2.0 Literature Review	
2.1 Mortality Projections	6
2.2 The Lee and Carter (1992) Model and its Development	6
2.3 Application of the Lee-Carter Model	13
2.4 The Zografos–Balakrishnan-G Family of Distributions	18
2.5 The Lee and Carter (1992) Method	19
2.5.1 The Lee and Carter (1992) Model Assumptions	20

2.5.2	The Lee and Carter (1992) Model Parameter Estimation	20
2.5.3	Forecasting the Mortality Index	22
2.6	The Li <i>et al.</i> (2002) Method	22
2.6.1	Forecast Based on Data at few time points	22
2.7	Distribution of the Mortality Models	23
2.7.1	The Lee and Carter (1992) Model	23
2.7.1.1	The Normal Distribution	24
2.7.1.2	Properties of the Normal Distribution	24
2.7.1.3	The Central Limit Theorem	25
2.7.2	The Brouhns (2002) Model	25
2.7.2.1	The Poisson Distribution	25
2.7.2.2	Properties of the Poisson Distribution	26
2.7.3	Renshaw and Haberman (2006) Model	27
2.8	Normality Test	27
2.8.1	The Kolmogorov-Smirnov Test (K-S)	28
2.8.2	The Shapiro-Wilk Test	28
2.9	Information Criteria for Model Selection	29
2.9.1	Akaike Information Criterion (AIC)	29
2.9.2	Bayesian Information Criterion (BIC)	30
2.9.3	Corrected Akaike Information Criterion (CAIC)	30
2.9.4	Hannan-Quinn Information Criterion (HQC)	31
2.10	Forecast	32
2.10.1	Formulation of the Autoregressive Moving Average Model	32
2.11	Proposed Method	32
2.12	Reason for Proposed Method	33
2.13	The Generalized Gamma Distribution	33
2.13.1	Characterization using shape k and scale θ	33
2.13.2	Characterization using shape α and rate β	34
2.13.3	The Gamma and Incomplete Gamma Functions	34
2.14	The Zografos and Balakrishnan Gamma Generator	34

3.0	Methodology	
3.1	Developing the Gamma-Normal Lee-Carter Model.....	36
3.1.1	Probability Density Function of Gamma-normal Lee-Carter	36
3.2	Statistical Properties of the Proposed Model	37
3.2.1	Cumulative Distribution Function	37
3.2.2	Verification of Proper Probability Density Function	37
3.2.3	Asymptotic Behaviour	39
3.2.4	The Hazard Function	39
3.3	Estimation	41
3.3.1	The Likelihood and Log-likelihood Functions	41
3.3.2	Estimation of Parameters	42
3.3.3	Components of the Score Vector.....	53
4.0	Analysis of Data and Discussion	
4.1	Analysis of Data	55
4.2	Data Choice and Source	55
4.3	Data Structure and Notation	55
4.4	Results and Discussion	56
4.4.1	Descriptive Analysis for Age-specific Mortality Data	59
4.4.2	Boxplots for Mortality and Log Mortality Rates.....	64
4.4.3	Test for Normality using Kolmogorov Smirnov and Shapiro-Wilks Procedures.....	69
4.4.4	General Pattern of Mortality by Age-Lee-Carter Model.....	72
4.4.5	Relative Pace of Change in Mortality by Age-Lee-Carter Model.....	74
4.4.6	Mortality Time Trend - Lee-Carter Model.....	76
4.4.7	General Pattern of Mortality by Age- Brouhns Model.....	78
4.4.8	Relative Pace of Change in Mortality by Age-Brouhns Model.....	80
4.4.9	Mortality Time Trend - Brouhns Model.....	82

4.4.10	General Pattern of Mortality by Age- Renshaw- Haberman Model.....	84
4.4.11	Relative Pace of Change in Mortality by Age- Renshaw- Haberman.....	86
4.4.12	Mortality Time Trend -Renshaw Haberman	88
4.4.13	General Pattern of Mortality by Age- Gamma-Normal Lee-Carter Model	90
4.4.14	Relative Pace of Change in Mortality by Age-Gamma-Normal Lee-Carter Model.....	92
4.4.15	Mortality Time Trend-Gamma-Normal Lee-Carter.....	94
4.4.16	Comparison of MLEs of Parameter a_x Across all the Models	97
4.4.17	Comparison of MLEs of Parameter b_x Across all the Models	100
4.4.18	Comparison of MLEs of Parameter k_t Across all the Models	103
4.4.19	Estimates of Parameters α under the Gamma-Normal Lee-Carter Model	105
4.4.20	Measures of Goodness of Fit for Male and Female when $h = 16$	108
4.4.21	Forecast Estimates for the Lee-Carter Model.....	111
4.4.22	Forecast Estimates for the Brouhns Model.....	114
4.4.23	Forecast Estimates for the Gamma-Normal Lee-Carter Model	117
4.4.24	Comparison of Mean Forecast Estimates Across the Models	

.....	117
4.4.25 Measures of Goodness of Fit for the Forecast when $h = 20$..	119
4.4.26 Measures of Goodness of Fit results for the Simulated Data.	121
5.0 Summary, Conclusion, Contribution to Knowledge and Suggestion for Further Study	
5.1 Summary	122
5.2 Conclusion	122
5.3 Contribution to Knowledge	122
5.4 Suggestion for Further Study	123
References	124
Appendix A	131
Appendix B	227

LIST OF FIGURES

4.1:	Boxplots for Female Mortality Rates	60
4.2:	Boxplots for Male Mortality Rates	61
4.3:	Boxplots for Female Log Mortality Rates	62
4.4:	Boxplots for Male Log Mortality Rates Across Time-period	63
4.5:	General Pattern of Mortality by Age (Lee-Carter)	71
4.6:	Relative Pace of Change in Mortality by Age (Lee-Carter)	73
4.7:	Mortality Time Trend (Lee-Carter)	75
4.8:	General Pattern of Mortality by Age (Brouhns)	77
4.9:	Relative Pace of Change in Mortality by Age (Brouhns)	79
4.10:	Mortality Time Trend (Brouhns)	81
4.11:	General Pattern of Mortality by Age (Renshaw-Haberman).....	83
4.12:	Relative Pace of Change in Mortality by Age (Renshaw-Haberman)....	85
4.13:	Mortality Time Trend (Renshaw-Haberman)	87
4.14:	General Pattern of Mortality by Age (Gamma-Normal Lee-Carter Model)..	89
4.15:	Relative Pace of Change in Mortality by Age (Gamma-Normal Lee-Carter Model).....	91
4.16:	Mortality Time Trend (Gamma-Normal Lee-Carter Model)	93

LIST OF TABLES

4.1	Descriptive Analysis for Male Mortality Data	57
4.2:	Descriptive Analysis for Female Mortality Data	58
4.3:	Kolmogorov Smirnov Normality Test Across Ages	65
4.4:	Kolmogorov Normality Test Across Time Period	66
4.5:	Shapiro Wilks Normality Test Across Age-groups	67
4.6:	Shapiro Wilks Normality Test Across Time Period	68
4.7:	Comparison of the MLEs of Parameter a_x under the Four Models for Nigeria Female Mortality Data.	95
4.8:	Comparison of the MLEs of Parameter a_x under the Four Models for Nigeria Male Mortality Data	96
4.9:	Comparison of the MLEs of Parameter b_x under the Four Models for Nigeria Female Mortality Data.....	98
4.10:	Comparison of the MLEs of Parameter b_x under the Four Models for Nigeria Male Mortality Data.....	99
4.11:	Comparison of the MLEs of Parameter k_t under the Four Models for Nigeria Female Mortality Data	101
4.12:	Comparison of the MLEs of Parameter k_t under the Four Models for Nigeria	

Male Mortality data	102
4.13: Estimates of Parameter α under the Gamma-Normal Model	104
4.14: Measures of Goodness of Fit for Male and Female when h=16	106
4.15: Measures of Goodness of Fit Results for the Mortality Model when h=16....	107
4.16: Forecast Estimates for the Lee-Carter Model (Male)	109
4.17: Forecast Estimates for the Lee-Carter Model (Female)	110
4.18: Forecast Estimates for the Brouhns Model (Male)	112
4.19: Forecast Estimates for the Brouhns Model (Female)	113
4.20: Forecast Estimates for the Gamma-Normal Model Lee-Carter (Male).....	115
4.21: Forecast Estimates for the Gamma-Normal Model Lee-Carter (Female)....	116
4.22: Measures of Goodness of Fit for the Forecast when h=20.....	118
4.23: Measures of Goodness of Fit Results for the Simulated Data	120

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Whether in developed or developing countries, mortality models play a very significant role in demographic projections as they help to provide a sufficient understanding of mortality trends and the uncertainty of future mortality rates.

Mortality data remain a great challenge for observing, evaluating and projecting the health condition of a large part of the world's population. Due to scarcely available resources, many developing countries especially in Sub-Saharan Africa lack vital registration systems that could consistently and continuously collect information on mortality. Hence the need for mortality models which are of utmost importance in population studies.

Many industrialized countries experienced an upward surge in expected life expectancy during the past century. Improvements in medicine, hygiene, sanitary conditions and living standard led to rapidly decreasing infant mortality rates in the early part of the century and afterwards to first increasing and then decreasing mortality in adults in the concluding part of the twentieth century. In Africa, however, the level of mortality varies distinctly across countries. In contrast to other major parts of the world, mortality is higher in the African continent.

According to Keyfitz and Caswell (2005), "to represent a population as a number changing in time, and ignoring its age composition, is comparable to treating the earth as a point in space". In the same vein, the mortality pattern of an area cannot be studied without considering the peculiarity in different age-groups that makes up its population.

Unlike their counterparts in the developed World who enjoy the privilege of being a 'She', from infancy to adulthood, the female gender in many developing countries is

faced with challenges: from low-socio cultural values to poor socio-economic growth and a frail health care system.

Maternal mortality affects the female gender, specifically, women in the child-bearing age-group (15-49 years) and it has been a serious concern in many developing countries. About a decade ago, the world Health survey reported that the risk of a woman in a developing country dying from a pregnancy-related cause during her lifetime was about thirty-six times higher compared to a woman living in a developed country. Nigeria for instance has been confirmed as having the second highest number of maternal mortality rate in the world after India in a recent United Nations report.

With respect to infants and children, mortality still remains high in many developing countries. Compared to other parts of the world, the rate of mortality drop has been slow in Africa. According to Timeous (1999), an increase in the number of mothers who went to school in the 90's slackened in many African countries and this had a direct negative impact on infant and childhood mortality.

Adult mortality (15-60 years) rate has risen sharply since HIV became prevalent. 'HIV and AIDS' have become one of the most devastating epidemics in the world. Its impact is concentrated among adults in the age range that usually forms new sexual partnerships. Thus, the HIV epidemic is likely to have a dramatic effect on the age pattern of mortality in Africa as well as on its overall level.

Recent changes in infectious disease have also had a negative influence on mortality in sub-Saharan Africa. In Africa, the impact of malaria on the population of both young and old is quiet overwhelming. Before the 1980s, malaria fever was majorly treated with chloroquine and this was thought to have had a significant impact on mortality. However, between the early 1980's and early 1990's, malaria resistant to chloroquine spread across the continent. As a consequence, according to Bradley (1991) and Ewbank and Gribble (1993), chloroquine- resistant malaria exerted a significant upward influence on mortality during that time period. Furthermore, about three to four decades ago, non-communicable diseases like hypertension, diabetes and cancer was nothing to worry about since they

were extremely rare occurrences. But recent figures show an upward trend in the prevalence of these diseases for some specific age-groups.

Since the 1970s, economic difficulties have distressed many African countries. This is another major determinant which suggests that mortality decline in Africa may have decelerated further in recent times. In many developing countries, Government's budget on health services is usually too meager and in many instances insufficient to cater for the people's health needs.

Furthermore, a long history of military conflict has left countries and areas such as southern Somalia, Sudan and Ethiopia among the most deprived parts of Africa. Just as what obtains in other regions, warfare and political crisis in developing countries often bring with it substantial mortality from famine and disease.

Difference of opinion in Religion is also a major determinant in the level of mortality in Africa. In Nigeria for instance, clashes between Muslims and Christians usually occur in the North. Islamic Jihadist group like the Boko Haram have not made matters any better. The United Nations and some other humanitarian agencies have had to intervene to aid an extremely awful aftermath.

1.2 Statement of Problem

Many developed countries experienced an appreciable reduction in mortality rates during the past century. However, in many developing countries, the story is not the same as mortality levels varies markedly across age groups and between regions within particular countries. The Lee and Carter (1992) model was mainly designed for the United States mortality and it has been used by many industrialized countries as a reference point to model their mortality pattern. Hanna (2007), Chukwu and Oladipupo (2012) and Taruvinga *et al.* (2017) used the Lee and Carter (1992) model for describing the pattern of mortality for some developing countries with defective data situation. These attempts have resulted in the violation of the normality assumption. This is due to the model's

inability to accommodate variability across age-groups. The model and some of its variants assume a normal distribution for its error term and this is not the situation for many developing countries.

Therefore, this study was aimed at developing a modified Lee and Carter model under a non-Gaussian error innovation to accommodate variability across all ages.

1.3 Objectives of the Study

The main focus of this research work is to develop a modified Lee-Carter approach with improved modelling properties with the aim of capturing variability across all ages

The specific objectives are:

- i. to derive some statistical properties of the proposed model such as its probability density function, its cumulative distribution function, its hazard function, its asymptotic behaviour and the verification of proper probability density function.
- ii. to use the maximum likelihood estimation procedure in estimating the model's parameters in order to achieve better results.
- iii. to determine future mortality trend using the proposed model.
- iv. to make a statistical comparison of the proposed model and other Lee and Carter variants using the Akaike information criterion, the Bayesian information criterion, the Corrected Akaike information criterion and the Hannan-Quinn information criterion.

1.4 Motivation

The Lee-Carter model is the most popular approach to mortality modelling Reese (2015). It has been applied by many developed countries to model their mortality pattern with satisfactory results. “Surprisingly, however, the basic theoretical properties of the Lee-Carter model have never been thoroughly investigated” (Reese 2015).

Attempts in using the model for describing the pattern of mortality in developing Countries have led to the violation of the normality assumption, due to its inability to

accommodate variability across all age-groups. The consequence of this is mortality curves crossing each other at different ages. Hence, because of our own peculiar exigencies, assumptions underlining the classical Lee and Carter (1992) model and its variants should be modified when it comes to applying them to developing Countries.

1.5 Thesis Layout

In addition to this introductory chapter, which contains the background of the study, statement of problem, objectives and motivation for the study, this research work contains four other chapters stated as follows;

In chapter two, existing works on mortality models, the development and application of the Lee-Carter model and a brief overview on the Zografos–Balakrishnan-GFamily of distributions were reviewed. The methodology of Zografos and Balakrishnan (2009) and Lima *et al* (2015) framework were looked into.

In chapter three, the methodology of the proposed Gamma-Normal Lee-Carter distribution was discussed. The empirical derivations of the parameters of the proposed distribution and its statistical properties were considered.

In chapter four, the proposed model was applied to a combination of all- cause mortality data. The results obtained were compared to some Lee-Carter variants.

In chapter five, summary and conclusion were drawn while contribution to knowledge and suggestions for further research were also given.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

This chapter begins with a review of existing mortality models followed by a review of research works on the Lee and Carter (1992) model which will be in two sections; works on the development of the Lee and Carter model followed by related works based on its application. Finally a review of methods which includes works based on the Zografos–Balakrishnan-GFamily of distributions and some other relevant statistical tools.

2.1 Mortality Projections

The theory and practice of modelling mortality have evolved rapidly in recent decades and there are many ways to mortality forecast (Bongaarts 2004). However, deMoivre (1725), Gompertz (1825) and Makeham (1860) are the most famous in the history of mortality modelling (Wang 2007). In recent times, quite a number of new methods have been established for forecasting mortality using stochastic models, such as, Pollard (1987), Olshansky (1988), Alho and Spencer (1985), Keyfitz (1991), Alho (1990) and Lee and Carter (1992).

2.2 The Lee and Carter (1992) Model and its Development

In 1992, Ronald Lee and Lawrence Carter published a stochastic model for the forecasts of the level and age pattern of mortality. The method was based on a blend of statistical time series techniques and a simple approach to dealing with the age distribution of mortality. The methodology defines the logarithm of a time series of age-specific death rates as the addition of an age-specific component that is time independent and another component which is the product of a time-varying parameter and an age-specific component that represents how rapidly or slowly mortality at each age varies when the general level of mortality changes. The resulting estimate of the time-varying parameter is then modeled and forecast as a stochastic time series using standard methods. Suffice to say that although the Lee and Carter model was specially proposed for the United

States mortality pattern, it has been modified and applied beyond the scope of its original creation.

Instead of the singular value decomposition procedure used by Lee and Carter (1992), Wilmoth (1993) proposed two alternatives in estimating the Lee-Carter model's parameters; the weighted least squares and the maximum likelihood approaches. Wilmoth's intuition was that the reason the observed number differs from the fitted number of deaths is that the estimates of the mortality time trend k_t are computed by minimizing the least square error over log-mortality, rather than mortality. As a result, age groups with small numbers of deaths receive the same weight as age groups with large numbers, even though they contribute very little to the totals. Both techniques have the significant advantage over the original Lee-Carter formulation in that they deal naturally with the case in which the observed number of deaths is zero, which occurs when analyzing cause-specific data or when dealing with small Countries.

Lee and Miller (2001) found that a major problem with the Lee-Carter model is the assumption that the age component is invariant over time. They proposed a simple and satisfactory solution to this problem by stating that the mortality experience of the world's leading industrial nations suggests the substantial age-time interaction. As a result, they assumed that the age component which applies to only half the century is invariant. They stated three modifications. First, the fitting period commences in 1970. Second, the adjustment involves fitting to error term in year t ; third, the jump-off rates are taken to be the actual rates in the jump-off year. In Lee and Miller model, the parameter k_t is adjusted by fitting a Poisson regression model as follows: $\ln D(x, t) = \ln N(x, t) + \ln m'(x, t) + \varepsilon'(x, t)$, where $D(x, t)$ denotes the annual number of deaths for aged x at time t ; $N(x, t)$ denotes the population aged x at time t ; $[\ln m'(x, t)] = a(x) + b(x) + k'(t)$; $k'(t)$ refers to an adjusted parameter for k_t ; $\varepsilon'(x, t)$ refers to the residuals after adjustment of k_t .

Booth *et al.* (2002) observed that the original Lee-Carter method makes use of only the first term of the singular value decomposition, but in principle the second and higher order terms could be incorporated in the base model. Their work was later corroborated

by Girosi and King (2007) who stated that there is an information loss with only one principal component. The full expanded model is of the form: $\ln [m(x, t)] = a(x) + b_1(x)k_1(t) + b_2(x)k_2(t) + \dots + b_n(x)k_n(t)$ where $b_i(x)k_i(t)$ is referred to as the i^{th} term of the rank n approximation. Their work helped to gain a better understanding of age-time interactions hence contributing to the validity and reliability of the Lee-Carter model.

In the spirit of Wilmoth (1993), Brouhns *et al.* (2002) implemented an alternative procedure for a log-bilinear formulation of the Lee-Carter model based on Poisson error structures. Specifically, they switched from a classical linear model to a generalized linear model, substituting Poisson random variation for the number of deaths for an additive error term on the logarithm of mortality rates. Furthermore, as pointed out by the researchers, the main drawback of the OLS estimation through the Singular Value Decomposition (SVD) procedure is that the errors are assumed to be homoscedastic. Instead of resorting to SVD for estimating a_x , β_x and k_t they determined these parameters by maximizing the log-likelihood based on model $D_{xt} \sim \text{poisson}(E_{xt}\mu_x(t))$ with $\mu_x(t) = \exp(a_x + \beta_x k_t)$. It was discovered that the Poisson model was able to account for slightly more variability than its SVD counterpart when applied on Belgian mortality data.

A major challenge with the Lee and Carter (1992) model framework is the long-time span data requirement. Li *et al.* (2002) tried to overcome this challenge by extending the Lee-Carter approach to situations in which mortality data are available at only a few points in time or at unevenly spaced intervals; situations often encountered in statistics for Third World Countries. They proposed that the improved Lee-Carter method can provide accurate mean mortality forecasts for Countries with historical data at only a few time points, if the earliest and latest points are sufficiently far apart in time. The model; $\ln m_{xu(t)} = a_x + b_x k_{u(t)} + \varepsilon_{xu(t)}$ with $\hat{k}_{u(t)} = \hat{k}_{u(t-1)} + \hat{\theta}[u(t) - u(t-1)] + [(\varepsilon_{u(t-1)} + 1) + \dots + \varepsilon_{u(t)}]$ was used on China's sex- combined mortality data for the years 1974, 1981 and 1990. It is important to state here that Li *et al.*'s major contribution is on the mortality time trend (parameter k_t) built in the model. Their method was able to account

for extra variability arising from fewer time points when forecasting the k_t parameter. However, their method does not account for the extra variability which will also be present in the age-specific (α_x and β_x) parameters.

Renshaw and Haberman (2006) extension of the Lee-Carter (LC) method bears some resemblance with that of Booth *et al.* (2002). However, the former proposed the incorporation of a cohort effect in their model. This was motivated by the lack of goodness of fit of the LC model to the death rates of England and Wales and the fact that its application led to a cohort effect of decreasing death rates. Furthermore the duo investigated the feasibility of constructing mortality forecasts on the basis of the first two sets of Singular Value Decomposition vectors, rather than just on the first set of such vectors, as obtained in the Lee-Carter approach. Their model is stated as: $\ln [m(x, t)] = a(x) + b_1(x)k_1(t) + b_2(x)l_{t-x} + \varepsilon_{xt}$ under the Gaussian framework or equivalently $\ln [m(x, t)] = a(x) + b_1(x)k_1(t) + b_2(x)l_{t-x} + \log e_{xt}$ under the Poisson setting.

Pedroza (2006) faulted the assumption made by Lee and Carter (1992) and several of its variants that the main source of error in the Lee-Carter model comes from the forecasting error of the mortality index k_t . Hence, he proposed a Bayesian approach to the Lee-Carter model which incorporates all sources of variation in the model when forming forecasts. When applied on U.S. male mortality data, the resulting prediction intervals from the Bayesian approach were much wider than those obtained by the original Lee-Carter method with prediction intervals which were too narrow for short-term forecasts, especially for young adults and older age group. Also, the forecast error associated with the model reflected more accurately under their approach. Moreover, their work shows how multiple imputations can be readily incorporated into the model to handle missing data.

Delwarde *et al.* (2007) pointed out the fact that the Poisson version of the Lee-Carter model with the assumption; $D_{xt} \sim \text{poisson}(E_{xt}\mu_x(t))$ as proposed by Brouhns *et al.* (2002) induces equi-dispersion with the mean and variance of D_{xt} equal to each other i. e. $E[D_{xt}] = V[D_{xt}]$. With reference to the documentation of

Brown (2003) on the US population, which states that aside gender, age and time, mortality rates are also influenced by other exogenous factors such as race, ethnicity, income, wealth, marital status and educational attainment and the correlations work in the direction that individuals with higher socio-economic status live longer than those in lower socio-economic groups; Delwarde *et al.* (2007) felt that this heterogeneity tends to increase the variance compared to the mean (a phenomenon termed as over-dispersion), ruling out the Poisson specification in favor of a mixed Poisson model. He therefore proposed a negative binomial regression model for estimating the parameters a_x , b_x and k_t .

Hyndman and Ullah (2007) proposed a robust version of the Lee-Carter model to avoid difficulties with outlying years and to avoid problems associated with data grouped into age intervals. His methodology can be considered a successor to Lee and Carter because it involves the use of non-parametric smoothing and principal component analysis to estimate the model's parameters. However, it differs from the Lee-Carter model in some important respects. First, it uses the functional data paradigm which immediately leads to the use of non-parametric smoothing to reduce some of the inherent randomness in the observed data. Second, it proposes a robust version of principal components to avoid difficulties with outlying years. They first define the problem as a function $y_t(x)$, which denotes the log of the observed mortality or fertility rate for age x in year t . Then suppose there is an underlying smooth function $f_t(x)$, which implies observing with error at discrete points of x . The observations are $\{x_i, y_t(x_i)\}$, $t = 1, \dots, n$, $i = 1, \dots, p$ which satisfy $y_t(x) = f_t(x_i) + \sigma_t(x_i) \varepsilon_{t,i}$, where $\varepsilon_{t,i}$ is an independent and identically distributed standard normal random variable and $\sigma_t(x_i)$ denotes the amount of noise varying with x . Typically, $\{x_1, \dots, x_p\}$ are single years of age ($x_1 = 0, x_2 = 1, \dots$) or denote 5-year age groups. Their approach is aimed in forecasting $y_t(x)$ for x belonging to $[x_1, x_p]$ and $t = n + 1, \dots, n + h$.

Li and Chan (2007) developed an outlier-adjusted version of the Lee-Carter model after performing a systematic outlier analysis of the mortality index built in the Lee-Carter model. Through outlier detection, they found out that mortality levels in the United States

and Canada were vulnerable to events like pandemics and wars; for instance, World War I in 1916, “Spanish flu” pandemic in 1918, Korean War in 1950, improved social condition in 1975 and so on. By incorporating the effect of outliers, an outlier-adjusted Lee-Carter model was created. They concluded that the outlier-adjusted version seemed to be an attractive alternative to the original model in predicting the long-term trend of mortality rates because of the better fit and the enhancement in forecast efficiency.

Girosi and King (2007) observed that the Lee-Carter model is a special case of the random walk with drift model. They proved that the age profile forecast from the model will always become less smooth and unrealistic after a point (when forecasting forward or backwards in time) and will eventually deviate from any given baseline. Furthermore, the researchers pointed out the fact that parameters β_x and k_t in the Lee-Carter model can be estimated through maximum likelihood estimation. However the multiple maxima or constraints will make standard optimization program work poorly. They concluded that whether one uses the singular value decomposition or eigen values, the theoretical justification of the procedure remains maximum likelihood and that there is information loss with one principal component.

Inspired by a linear re-parameterisation of Kannisto (1992) model, Doray (2008) proposed a new model for the old-age population, where the logit of μ_x is modeled as a function analogous to the one in the Lee-Carter model, that is, $\text{logit } \mu_{x,t} = \log\left(\frac{\mu_{x,t}}{1-\mu_{x,t}}\right) = a_x + b_x k_t + \varepsilon_{xt}$. In a comparative block study of the two variants of the Lee-Carter model and Doray (2008), Doray and Tang (2011) discovered that the latter yielded the smallest mean relative absolute errors and also had the best fit when applied to Canadian mortality data from 1976 to 2005 for ages 80 to 105 for women and 70 to 99 for men.

Koissi and Shapiro (2008) proposed a modification of the Lee-Carter model which incorporates variations in the age-specific parameters without violating the normality assumption made by Lee and Carter (1992). The parameters were computed for different periods. The time series obtained from the process were modeled using polynomial interpolation. Such model was said to be a realistic alternative to the standard Lee-Carter assumption of constant age parameters, although it is less simple than the original model.

In a comparative study of the simulation strategies applied under the log-bilinear Poisson setting, Renshaw and Haberman (2008) highlighted that Monte Carlo simulation approach should not be used for risk measurement in the Lee-Carter modelling, since different choices of constraints result in widely differing simulated confidence and prediction interval widths. On the basis of these findings, D'Amato *et al.* (2009) worked out a two-stage procedure combining the stratified sampling technique together with the Bootstrap procedure on the Poisson Lee-Carter model. The bootstrap simulation procedure was implemented using stratified sampling technique. The efficiency indicator showed that the method is more robust for obtaining forecasts.

Wang and Preston (2009) discovered that apart from age and time, smoking histories also influenced adult mortality in the United States. This they achieved by incorporating a variable representing the intensity of smoking within a cohort into the original Lee-Carter projection model. The introduction of this variable accounted for important anomalies in the age-sex pattern of mortality change and enabled the use of a common temporal trend of mortality change for the two sexes. They projected the age-specific mortality rates for men and women at ages 50–84 between 2004 and 2034 in the United States. From their findings, mortality was projected to decline much faster when smoking was introduced into the model

In the light of Li *et al.* (2002) and Doray (2008), Zhao (2012) introduced a modified Lee-Carter model for analysing short-base-period mortality data, for which the original Lee-Carter model produces severely fluctuating predicted age-specific mortality. Approximating the unknown parameters (a_x and b_x) in the modified model by linearized cubic splines and other additive functions, the researcher pointed out that the model can be simplified into a logistic regression when fitted to binomial data. The proposed model is of the form: $\ln\left(\frac{p(x,t)}{1-p(x,t)}\right) = a(x) + b(x)k(t)$ where the expected death rate $p(x, t)$ and the functions of age x and time t ; $a(x)$, $b(x)$ and $k(t)$ are all unknown. The method assumes a binomial distribution for the number of deaths; $d(x, t) \sim \text{Bin}[n(x, t), p(x, t)]$, the relation $p(x, t) = \frac{\exp(a_x + \beta_x k_t)}{1 + \exp(a_x + \beta_x k_t)}$ is assumed. The expected death rate estimated from

the modified model is smooth, not only over ages but also over years. The analysis of mortality data in China (2000-2008) demonstrates the advantages of the new model over some existing models.

Previous studies had suggested the possibility of cohort effect in Japanese death rate. Taking this into account, Igawa (2013) analysed the residual structure involved in the application of the Lee-Carter model to Japanese death rates. After observing a time series correlation and undulation in the residual phase on the age and year axes, He went ahead to propose the LC-VAR (Lee-Carter Vector Autoregressive) model to address the problem. The application of Lee-Carter Vector Autoregressive on Japanese mortality data improved the fit of the model. Keeping in mind the condition that a cohort effect may vary over time, they submitted that the LC-VAR model is useful not only for short- to middle-term mortality projections and the evaluation of uncertainties but also for the description of cohort and period effects.

According to Reese (2015), the superficial design of the Lee-Carter framework, leads to inconsistencies that impair the usefulness of the estimated parameters. He observed that the trend estimate obtained from the Singular Value Decomposition used in the Lee-Carter model and some of its variants was not consistent, implying that, in contrast to common practice, estimation uncertainty should not be disregarded. He proposed that taking an asymptotic perspective allows one to avoid a number of restrictive assumptions. From his contribution, asymptotic distributions were derived and variance estimators were provided for the parameter estimators of the model.

2.3 Application of the Lee-Carter Model

Lee and Carter (1992) applied their model on sex-combined United States age-specific population data from 1933-1987 and projections were made up to the year 2065. The authors were concerned that extrapolating sex differentials in mortality would produce implausible differentials in the future. Results obtained from their work confirm that the method did well over the period of data availability: age patterns were quite stable and the trend in the fundamental time-varying parameter was surprisingly linear. They however

stated that if one were to go back much before 1900, the linearity of change would be violated because sustained mortality decline is chiefly a twentieth century phenomenon in the developed countries.

Booth *et al.* (2005) applied the Lee and Carter (1992) and two other variants; the Lee and Miller (2001) and the Booth *et al.* (2002) to mortality data from ten developed countries (both male and female populations) from the time period 1900-1985 while forecast was done from 1986 to 2001 depending on the country. For short-term forecast, the Booth *et al.* (2002) outperformed the other two variants with narrower prediction intervals and forecast errors.

Wang (2007) focused on a long-term study of mortality rates for Swedish population data and how a forecaster could get a better performance of the Lee-Carter model by the selection of an optimal time period to fit the model. In his research work, three different estimation periods of 1800-1900, 1850-1900 and 1875-1900 were used to predict the mortality rates for 1901-2004. Four estimation periods of 1850-1950, 1900-1950, 1925-1950 and 1940-1950 were used to predict the mortality rates for 1951-2004. The results showed that the long estimation periods of 1850-1900 and 1900-1950 yielded the best forecasting performances for prediction series of 1901-2004 and 1951-2004 respectively. While the prediction with short estimation period like 1875-1900 and 1940-1950 did not work well. The efficiency of the Lee-Carter model was studied by examining the residual term on a logarithmic scale. It was discovered that most of the residual terms showed a similar lack of systematic pattern but they could not be described as random as marked clusters occurred. He stated that their occurrence was probably as a result of infectious diseases, which were the leading cause of death during certain periods. This indicates that the selection of an appropriate estimation period is important for forecasting mortality and issues such as infectious diseases and medical factors are always very important contributors to the trend of mortality

Hanna (2007) applied the Lee-Carter model to four Countries in Eastern Europe; Bulgaria [1947-2003], Czech Republic [1950-2004], Hungary [1950-2001], Slovakia [1950-2005] and five Countries in the former Soviet Union; the Baltic States, Ukraine and Russia

[1959-2005]. For the Countries where the model fitted very well, the estimated time trend $\hat{\kappa}_t$ was linearly decreasing with time and the \hat{b}_x values (relative pace of change) were decreasing with age implying that the mortality rates at younger ages decline more rapidly than for older ages. The research work further showed that for Bulgaria, in contrary to the female data, the male data created a problem with the ages between 40 and 60 years. When compared with the female data, the Lee-Carter model could not be applied to the Bulgarian male mortality data with substantial results. The study also showed that Hungarian data had a slightly systematic effect in the male and female models. For some ages the estimated trend was increasing while the actual trend was decreasing. Another problem with the Hungarian data pointed out by this study was the deviation from the age profile for the male data in the broad age span 30 to 75. The \hat{b}_x values were negative at the same time as the $\hat{\kappa}_t$ values were negative. The $\hat{\kappa}_t$ time trend after the year 1970 was negative. This means that if the \hat{b}_x values are negative, the estimated mortality: $\ln m_x(t) = \hat{a}_x + \hat{b}_x \hat{\kappa}_t$ would increase for each year. The $\hat{b}_x \hat{\kappa}_t$ term was positive for x greater than 35 and t greater than 1975. This may result in mortality curves crossing each other with higher mortality for a 60 year old male than for an 80 year old male in the future. However, the problem with negative \hat{b}_x values was not found for the female data.

Koissi *et al.* (2004) used the methods proposed by Wilmoth (1993) to study the efficiency of the Lee-Carter model. The model's parameters were estimated using the Singular Value Decomposition (SVD), the Maximum Likelihood Estimate (MLE) and the Weighted Least Square (WLS) methods. The computations were made using the vital rates (1955 to 1999) from Finland, Sweden, Denmark and Norway. The results of their analysis showed that under an appropriately chosen estimation period, the model fit the observed death rates quite well. The estimates for the parameters \hat{a}_x (general pattern of mortality by age) and \hat{b}_x (relative pace of change in mortality by age as $\hat{\kappa}_t$ varies) were almost alike with the SVD, the MLE or the WLS methods. The mortality index $\hat{\kappa}_t$ had a

common (almost) linear decreasing trend in the four Countries, although some variations existed with the three methods. A bootstrap simulation indicated that the use of the MLE results in smaller mean squared errors for the parameters \hat{a}_x and \hat{b}_x than the use of the other two methods. However, the Singular Value Decomposition method was the best alternative for the mortality index $\hat{\kappa}_t$.

Chukwu and Oladipupo (2012) applied the Li *et al.*(2002) to model adult mortality in Nigeria. The data set used was the age-specific mortality rates of males and females aged 15-84 years for the time periods 1990, 2000 and 2009. The model's parameters were estimated using the singular value decomposition technique. They observed that the model follows the mortality pattern very well for most of the ages except that the fit of the model was better for the male data than the female data due to the latter's heterogeneous nature. This result coincides with that of Hanna (2007) as obtained in the case of Russian male and female mortality data.

Emilio (2012) projected the mortality pattern of Italy using the lee-carter model. They fitted the model to Italian mortality rates for the period 1965-1999 and generated forecasts for nine subsequent years. In their findings, they observed that the model performs very well in explaining mortality for ages 40 to 100 years. However, the power of the Lee-Carter model significantly drops as the observations on mortality rates become more volatile at very old ages.

Eglé *et al.* (2012) employed the Lee-carter method by comparing the suitability of the model for different Countries namely France, Lithuania and Belarus for the time period 1970-2005. By using their age specific mortality rates, they discovered that the model almost accurately describes and predicts mortality for France. However, only short-term (1–2 years) forecasts for the Belarusian and Lithuanian mortality were accurate, while the accuracy of longer-term (4–5 years) forecasts is lower, especially for men. From their results, they concluded that the Lee–Carter method is most suitable for populations with a clear upward or downward mortality trend over time.

Bamidele and Adejumo (2014) tried to explain mortality improvements for males aged 40-65 for the time period 1994-2004 using Nigerian mortality data. A combination of the

Lee Carter model and its variant (Renshaw-Haberman model) was utilized. On the Bayesian information ranking criterion, the Renshaw-Haberman model for the data dominates. However, taking into account the robustness of the parameter estimates, they stated that the Lee Carter model is preferred for the dataset. With respect to the age groups studied, they concluded that there has been approximately linear improvement over time in mortality rates at all ages, but that the improvements have been greater at lower ages than at higher ages.

As a follow-up to their previous work in 2012, Chukwu and Adegbilero-Iwari (2015) desired to see the performance of the Li *et al.* (2002) on sex-combined adult mortality data. Unlike the results obtained in many developed Countries, it was discovered that the parameter \hat{b}_x (relative pace of change in mortality by age) violated the normal descending trend with age. Instead, it exhibited a fluctuating increase and decrease within the age-groups especially with respect to the middle ages. This is also similar to the results obtained with the female data in Chukwu and Oladipupo (2012). The reason behind this is not far-fetched. In Countries where the model has fitted very well, the estimated time trend \hat{k}_t decreased linearly with time (which indicates a decreasing mortality pattern overtime) while the \hat{b}_x values (relative pace of change) were decreasing with age.

Due to mortality data availability of 21 years, Taruvunga *et al.* (2017) applied the Lee-Carter (1992) model to Zimbabwe mortality data for the age groups 0-84 years between the time periods 1983 to 2004. Forecast was done from 2005 to 2014. Results from their empirical analysis showed that the Lee- Carter model provided a better fit for thirteen age groups but failed to fit for four age groups namely 50-54, 55-59, 75-79 and 80-84.

2.4 The Zografos–Balakrishnan-GFamily of Distributions

Current literature has recommended some other ways of extending well-known distributions. The earliest are the class of distributions generated by Kumaraswamy (1980). The more recent ones are the class of distributions generated by a standard beta random variable introduced by Eugene *et al.*(2002); generalized gamma-generated

distributions by Zografos and Balakrishnan (2009); generalized beta random variable introduced by Alexander *et al.* (2012); the class of beta random variable introduced by Pescim *et al.* (2012).

In distribution theory, the Gamma distribution is a two-parameter family of continuous probability distributions and is the most popular model for analyzing skewed data.

Manton (2009) pointed out that the Gamma distribution is usually used in mortality models for heterogeneous population especially in situations where the distribution of risk levels is flatter and has thicker tails.

Zografos and Balakrishnan (2009) presented two generalized gamma-generated distributions with an extra positive parameter for any continuous baseline cdf $G(x)$, $x \in R$. Their distribution springs from the generalized gamma distribution as introduced by Stacy (1962). The number of parameters of the Zografos-Balakrishnan and Ristic-Balakrishnan-Gamma generated family of distributions is equal to that of the Gamma distribution plus an additional shape parameter $\alpha > 0$. For $\alpha = 1$, the generalized gamma distribution is a basic exemplar of the Zografos-Balakrishnan-G and Ristic-Balakrishnan-G distributions with a continuous crossover towards cases with different shapes. The Zografos-Balakrishnan-G family of distributions is particularly useful for heterogeneous data because it allows for greater flexibility of its tails.

Nadarajah *et al.* (2015) presented some special cases of the Zografos-Balakrishnan-G family which includes the Gamma-Normal, the Gamma-Weibull, the Gamma-Gumbel, Gamma-log-normal and the Gamma-log-logistic distributions.

Lima *et al.* (2015) took the work of Zografos-Balakrishnan (2009) further by considering a three-parameter distribution called the gamma-normal distribution. Various structural properties of the new distribution were derived and studied. Based on three criteria (AIC, CAIC and BIC), their work presented the fact that the Zografos-Balakrishnan-Normal distribution provided a better fit than the skew-normal (normal, beta-normal and Kumaraswamy normal) distributions.

2.5 The Lee and Carter (1992) Method

The Lee-Carter methodology for forecasting mortality rates is a simple bi-linear model in the variables x (age) and t (calendar year). The model is defined as;

$$\ln m_{xt} = a_x + b_x k_t + \varepsilon_{xt} \quad (2.1)$$

Where;

m_{xt} : is the matrix of the observed age-specific death rate at age x during year t . It is obtained from observed deaths divided by population exposed to risk. It is subject to random fluctuation.

$\ln m_{xt}$: is the logarithm of the matrix of observed age-specific death rate at age x during year t .

a_x : is the average of $\ln m_{xt}$ over time t . It describes the (average shape of the age profile) general pattern of mortality by age.

k_t : is the time trend for the general mortality. It captures the main time trend on the logarithmic scale in mortality rates at all ages. k_t is also referred to as the mortality index.

b_x : indicates the relative pace of change in mortality by age as k_t varies. It describes the pattern of deviations from the age profile when the parameter k_t varies. It modifies the main time trend according to whether change at a particular age is faster or slower than the main trend.

ε_{xt} : is the error term at age x and time t . It reflects the age specific influences not captured by the model.

2.5.1 The Lee and Carter (1992) Model Assumptions

1. The Lee-Carter model assumes that parameter b_x is invariant over time for all x .
2. It assumes that parameter k_t is fixed over age-groups for all t .
3. The practical application of the model assumes that the disturbances ε_{xt} are normally distributed i.e. $\varepsilon_{xt} \sim N(0, \sigma^2)$

2.5.2 The Lee and Carter (1992) Model Parameter Estimation

The model in (2.1) above is underdetermined (not unique). For the purpose of insuring a unique solution, the following constraints were used by Lee and Carter (1992):

$$\sum_{x=1}^m b_x^2 = 1 \quad \sum_{t=1}^n \kappa_t = 0 \quad (2.2)$$

Where $x = 1 \dots m$ age groups and $t = 1 \dots n$ calendar years.

To estimate the parameters, the values that minimize Q are chosen;

$$Q = \sum_{x,t} (a_x + b_x \kappa_t - q_{xt})^2, \quad (2.3)$$

Where $q_{xt} = \ln m_{xt}$, q_{xt} and Q are subject to the constraints above. To find values that minimize Q , Lagrange's multipliers; α and β are introduced such that;

$$R = Q - \alpha \sum_t \kappa_t - \beta \sum_x b_x^2. \quad (2.4)$$

Substituting equation (2.3) into (2.4) and then minimizing R ;

$$R = \sum_{x,t} (a_x + b_x \kappa_t - q_{xt})^2 - \alpha \sum_t \kappa_t - \beta \sum_x b_x^2$$

Taking the derivative of R with respect to a_x , b_x and κ_t respectively;

$$\frac{dR}{da_x} = 2 \sum_t (a_x + b_x \kappa_t - q_{xt}) \quad \text{for all } x \quad (2.5)$$

$$\frac{dR}{db_x} = 2 \sum_t \kappa_t (a_x + b_x \kappa_t - q_{xt}) - 2\beta b_x \quad \text{for all } x \quad (2.6)$$

$$\frac{dR}{d\kappa_t} = 2 \sum_x b_x (a_x + b_x \kappa_t - q_{xt}) - \alpha \quad \text{for all } t \quad (2.7)$$

Setting the derivatives equal to zero, the parameters of the Lee-Carter model are finally estimated as;

$$\hat{a}_x = \frac{1}{n} \sum_{t=1}^n \ln m_{xt} \quad \hat{b}_x = \frac{\sum_t k_t z_{xt}}{\sqrt{\sum_x (\sum_t k_t z_{xt})^2}} \quad k_t = \sum_x b_x z_{xt} \quad (2.8)$$

Where $z_{xt} = \ln m_{xt} - \hat{a}_x$

In the original paper written by Lee and Carter (1992), k_t is re-estimated by the method of the “second stage estimation”. They noticed that once b_x and k_t have been estimated, the observed total number of deaths D_t is not guaranteed to be equal to the fitted number of deaths. Therefore, they made a second stage estimation of k_t by finding a value that makes the observed number of deaths equal to the predicted number of deaths. That is, they searched for k_t such that,

$$D_t = \sum_x \exp(a_x + b_x k_t) N_{x,t}$$

Where:

D_t is the total number of deaths in year t

$N_{x,t}$ is the population of age x in year t .

No analytic solution is available for k_t so it can only be found by using an iterative numerical scheme such as the Newton Raphson method. However, this second stage of estimation does not have any impact on whether the Lee-Carter model could be fitted to the data or not.

2.5.3 Forecasting the Mortality Index

A peculiar feature of the Lee-Carter model and its variant is that once the data are fitted to the model and the values of the vectors \hat{a}_x , \hat{b}_x and \hat{k}_t are found, only the mortality index \hat{k}_t needs to be predicted. The mortality trend \hat{k}_t is modeled using standard univariate time series model. In most applications to date, it has been found that a random walk (Arima 0, 1, 0) with drift fits very well. After obtaining the predicted values of the mortality index \hat{k}_t , the values are then plugged back into the original Lee-Carter model to obtain forecast of the mortality rates. The forecast of the mortality index is modeled as;

$$k_t = k_{t-1} + \hat{\theta} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \quad (2.9)$$

Where the drift parameter $\hat{\theta}$ is expressed as;

$$\hat{\theta} = \frac{1}{T-1} \sum_{t=1}^{T-1} [\hat{k}_{t+1} - \hat{k}_t] \quad (2.10)$$

The standard error of the random fluctuations is estimated as:

$$S.e(\varepsilon_t) = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T-1} [\hat{k}_{t+1} - \hat{k}_t - \hat{\theta}]^2} \quad (2.11)$$

The variance of the drift parameter is given by:

$$\text{Var}(\hat{\theta}) = \frac{\sigma_{rw}^2}{T-1} \quad (2.12)$$

While its standard error is:

$$S.e(\hat{\theta}) = \sqrt{\frac{\sigma_{rw}^2}{T-1}} = \sqrt{\frac{\frac{1}{T-1} \sum_{t=1}^{T-1} [\hat{k}_{t+1} - \hat{k}_t - \hat{\theta}]^2}{T-1}} \quad (2.13)$$

2.6 The Li *et al.* (2002) Method

The model framework of the Li *et al.* (2002) bears essentially the same features as that of Lee and Carter (1992) except that the former is designed for situations where mortality data is available at few time points or unequally spaced interval of time. Their model is;

$$\ln m_{xu(t)} = a_x + b_x k_{u(t)} + \varepsilon_{xu(t)} \quad (2.14)$$

$$\varepsilon_{xu(t)} \sim N(0, \sigma^2)$$

2.6.1 Forecast Based on Data at few Time Points

Now let data be collected at times $u(0), u(1), \dots, u(T)$, the forecast of the mortality index in equation (2.9) is expressed as:

$$\hat{k}_{u(t)} = \hat{k}_{u(t-1)} + \hat{\theta}[u(t) - u(t-1)] + [(\varepsilon_{u(t-1)} + 1) + \dots + \varepsilon_{u(t)}] \quad (2.15)$$

The drift parameter $\hat{\theta}$ is the average rate of decline in the mortality time trend k_t , both for forecasting and for describing history. Li *et al.* (2002) felt that just as the average speed of linear movement depends only on the initial and terminal positions and their times, so $\hat{\theta}$ is determined only by the first and last values of $\hat{k}_{u(t)}$ and $u(t)$, and is independent of other values of $\hat{k}_{u(t)}$. Thus, the mean forecasts of $\hat{k}_{u(t)}$ depend mainly on

the death rates at starting and ending points of the historical period, and mortality data at years between the two points do not matter much.

The unbiased estimate of $\hat{\theta}$ is estimated as;

$$\hat{\theta} = \frac{\sum_{t=1}^T [\hat{k}_{u(t)} - \hat{k}_{u(t-1)}]}{\sum_{t=1}^T [u(t) - u(t-1)]} = \frac{\hat{k}_{u(T)} - \hat{k}_{u(0)}}{u(T) - u(0)} \quad (2.16)$$

$$\text{Var of } \varepsilon_{u(t)} = \frac{\sum_{t=1}^T [\hat{k}_{u(t)} - \hat{k}_{u(t-1)} - \hat{\theta}[u(t) - u(t-1)]]^2}{[u(T) - u(0) - \frac{\sum_{t=1}^T [u(t) - u(t-1)]^2}{u(T) - u(0)}]} \quad (2.17)$$

$$\begin{aligned} \text{Var}(\hat{\theta}) &= \frac{\text{var}\{\sum_{t=1}^T [(\varepsilon_{u(t-1)} + 1) + \dots + \varepsilon_{u(t)}]\}}{[u(T) - u(0)]^2} \\ &= \frac{\sigma^2}{u(T) - u(0)} \approx \frac{\text{Var}(\varepsilon_{u(t)})}{u(T) - u(0)} \end{aligned} \quad (2.18)$$

2.7 Distribution of the Mortality Models

2.7.1 The Lee and Carter (1992) Model

$$\text{Model Structure: } \ln m_{xt} = a_x + b_x k_t + \varepsilon_{xt} \quad (2.19)$$

The Lee-Carter approach assumes that its error term follows a normal distribution.

2.7.1.1 The Normal Distribution

The Normal distribution is a continuous distribution and it is the most important distribution in statistics. Abraham De Moivre (1667-1754) was first person to publish the formula for this distribution. Also, another prominent contributor is Carl Friedrich Gauss (1777-1855). Thus the distribution is frequently called the Gaussian distribution in recognition of Gauss' contribution.

Density function of the Normal distribution

The normal distribution has its probability density function given as:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-\mu}{\sigma}\right]^2}, \quad \sigma > 0, \quad -\infty < y < \infty \quad (2.20)$$

Where μ and σ are the two parameters of the distribution. Y is a normally distributed random variable and is continuous taking values between $-\infty$ and $+\infty$. Hence the probability density function of the error term embedded in the Lee and Carter (1992) model is given as:

$$\text{Pdf: } f(y, a_x, b_x, k_t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_x k_t)}{\sigma}\right]^2}, \quad \sigma > 0, \quad -\infty < y < \infty : y = \ln m_{xt} \quad (2.21)$$

2.7.1.2 Properties of the Normal Distribution

- The normal distribution is symmetrical about its mean. This implies that the curve on either side of μ is a mirror image of the other side. Its graph produces a bell-shaped curve.
- The mean, median and mode are all equal
- The total area under the curve about the X-axis is one square unit. This is based on the fact that the normal distribution is a probability distribution.
- The normal distribution is completely determined by the parameters μ and σ . Different values of μ shift the graph of the distribution along the X-axis. Also, different values of σ determine the degree of flatness or peakedness of the graph of the distribution. Because of this characteristic, μ is often referred to as the location parameter while σ is called the scale parameter. Hence the Normal distribution has no shape parameter.

2.7.1.3 The Central Limit Theorem

The Central limit theorem is the second fundamental theorem in probability. This theorem states that if S_n is the sum of n mutually independent random variables, then the distribution function of S_n is well-approximated by a certain type of continuous function known as a normal density function, which is given by equation (2.20) above. In other words, this theorem implies that irrespective of the underlying population distribution model, as the sample size increases, the sample mean tends to be normally distributed around the population mean, and its standard deviation shrinks as the sample size increases.

2.7.2 The Brouhns *et al.* (2002) Model

The structure of the Brouhns *et al.* (2002) model is given as:

$$\text{Model Structure: } D_{xt} = e^{a_x + b_x k_t} E_{xt} + e^{\varepsilon_{xt}} \quad (2.22)$$

The parameters a_x , b_x and k_t are basically the same as in the classical Lee-Carter model.

D_{xt} denotes the number of deaths at age-group x in year t .

E_{xt} is the population exposed to risk at age-group x in year t .

The Brouhns model assumes that its random component follows a Poisson distribution.

2.7.2.1 The Poisson Distribution

The Poisson distribution (a discrete distribution) is associated with counts of number of occurrences of a relatively rare event over a specified interval of time. In 1837, Simeon Poisson published the derivation of this distribution which bears his name. It was originally derived as an approximation to the binomial distribution.

Definition

Given a continuous interval and assume that discrete events occur randomly throughout the interval. If the interval can be partitioned into sub-intervals of small enough length such that;

- (i) the probability of more than one occurrence in a sub-interval is zero;
- (ii) the probability of one occurrence in a sub-interval is the same for all subintervals and proportional to the length of the subinterval;
- (iii) the occurrence of an event in one subinterval has no effect on the occurrence or non-occurrence in another non-overlapping subinterval,

If the mean number of occurrences in the interval is λ , the random variable X that equals the number of occurrences in the interval is said to have the Poisson distribution with parameter λ .

Density Function of the Poisson Distribution

The probability density function of the Poisson distribution is stated as:

$$f(y) = \frac{e^{-\lambda} \lambda^y}{y!}, \quad y = 0, 1, 2, \dots, n \quad (2.23)$$

The density function of the random component in the Brouhns model is given as:

$$\text{Pdf: } f(y, a_x, b_x, k_t) = \frac{e^{-\mu_{xt} E_{xt}} \mu_{xt}^{D_{xt}}}{D_{xt}!}, D_{xt} = 1, 2, 3, \dots : \mu_{xt} = e^{a_x + b_x k_t}$$

(2.24) 2.7.2.2 Properties of the Poisson Distribution

- The expected value (mean) of the Poisson distribution is λ
- Its variance is λ
- Its third central moment is λ
- Its Cumulative distribution function (CDF) is $f(x) = \sum_{y=0}^x \frac{e^{-\lambda} \lambda^y}{y!}$
- The binomial probability density function approaches the Poisson density function under certain limiting conditions.

2.7.3 Renshaw and Haberman (2006) Model

$$\text{Model Structure: } m_{xt} = e^{a_x + b_x k_t + l_{t-x} + \varepsilon_{xt}}$$

$$\text{Pdf: } f(y, a_x, b_x, k_t, l_{t-x}) = \frac{e^{-[\exp(a_x + b_x k_t + l_{t-x}) E_{xt}]} [\exp(a_x + b_x k_t + l_{t-x}) E_{xt}]^{D_{xt}}}{D_{xt}!}, D_{xt} = 1, 2, 3, \dots \quad (2.25)$$

In this study, the error term of the Renshaw and Haberman model is also considered under a Poisson distribution framework.

2.8 Normality Test

A normality test is a statistical process used to determine if a sample or any group of data fits a standard normal distribution. The assumption of normality is particularly common in classical statistical tests. Tests for normality calculate the probability that the sample was drawn from a normal population. A normality test can be performed statistically or graphically.

According to Chambers *et al.* (1983) ‘Graphical methods provide powerful diagnostic tools for confirming assumptions, or, when the assumptions are not met, for

suggesting corrective actions. Without such tools, confirmation of assumptions can be replaced only by hope. Graphical methods include the use of;

- a. Histograms,
- b. Stem-and-leaf plots,
- c. Quartile-Quartile (Q-Q) probability plots,
- d. Cumulative frequency (P-P) plots.
- e. Box-and-whisker plots

The benefit of statistical tests over the graphical test for normality is that statistical tests are more precise since actual probabilities are calculated.

Statistical tests include;

- a. Anderson Darling test (Stephens, 1974)
- b. Jarque-Bera test (Bowman and Shenton ,1975)
- c. Shapiro-Wilks test (Shapiro and Wilk, 1965)
- d. Kolmogorov-Smirnov test (Kolmogorov and Smirnov , 1933,1948)
- e. D'Agostino test (D'Agostino, 1986)

It is important to state that for grouped data, each group must be tested for normality.

The hypotheses used in normality tests are:

Ho: The sample data are not significantly different than a normal population.

H₁: The sample data are significantly different than a normal population.

2.8.1 The Kolmogorov-Smirnov Test (K-S)

The Kolmogorov-Smirnov test was first derived by Kolmogorov (1933) and later modified and proposed as a test by Smirnov (1948). The (K-S) test is based on the empirical distribution function (ECDF). An attractive feature of this test is that the distribution of the K-S test statistic itself does not depend on the underlying cumulative distribution function being tested. Another advantage is that it is an exact test (the chi-square goodness-of-fit test depends on an adequate sample size for the approximations to be valid). Despite these advantages, the K-S test has some important limitations:

- i. It only applies to continuous distributions.
- ii. It tends to be more sensitive near the center of the distribution than at the tails.

The test statistic is given as;

$$D = \sup_x |F_n(X) - F(X, \mu, \sigma)| \quad (2.26)$$

Where,

$F(X, \mu, \sigma)$ is the theoretical cumulative distribution function of the normal distribution function and $F_n(X)$ is the empirical distribution function of the data. If it gives large values of D then it indicates the data are not normal. When population parameters (μ and σ) are unknown then sample estimates are used instead of parameter values.

2.8.2 The Shapiro-Wilk Test

Given a sample X_1, X_2, \dots, X_n , of n real-valued observations, the Shapiro-Wilk test developed by Shapiro and Wilk (1965) is a test of the composite hypothesis that the data are independent and identically distributed and normal, i.e. $N(\mu, \sigma^2)$ for some unknown real μ and some $\sigma > 0$. It is one of the most popular tests for normality assumption diagnostics. Despite its popularity, the Shapiro-Wilk test has some limitations. They are

- i. It depends on availability of values of a_i , and for large sample cases their computation may be much more complicated.
- ii. It was designed for small sample sizes (≤ 50)

The test statistic is of the form:

$$w = \frac{(\sum a_i y_{(i)})^2}{\sum (y - \bar{y})^2} \quad (2.27)$$

Where,

$y_{(i)}$ is the i th order statistics while a_i is the i th expected value of normalized order statistics. For independently and identically distributed observations, the values of a_i can be obtained from the table presented by Shapiro and Wilk (1965) for sample sizes up to 50. w can be expressed as a square of the correlation coefficient between a_i and $y_{(i)}$. So w is location and scale invariant and is always less than or equal to 1. In the plot of $y_{(i)}$ against a_i , an exact straight line would lead to w very close to 1. Hence, if w is significantly less than 1, the hypothesis of normality will be rejected.

2.9 Information Criteria for Model Selection

Information criteria have found a wide range of practical applications in empirical work. Their use extends to choosing explanatory variables in regression models and selecting lag lengths in time series models. Frequently used information criteria are the Akaike information criterion (AIC), Bayesian information criterion (BIC) and the Hannan-Quinn information criterion (HQC).

2.9.1 Akaike Information Criterion (AIC)

The Akaike information criterion was developed by Akaike (1973) to estimate the expected Kullback-Leibler discrepancy between the model generating the data and a fitted candidate model. It is grounded in the concept of entropy, in effect offering a relative measure of the information lost when a given model is used to describe reality. The AIC is not a test of the model in the sense of hypothesis testing; rather it is a tool for model selection. Given a data set, several competing models may be ranked according to their AIC, with the one having the lowest AIC being the best. In the general case, the Akaike information criterion is expressed as:

$$\text{AIC} = 2K - 2\log [L(\hat{\theta})] \quad (2.28)$$

Where,

K is the number of parameters in the statistical model,

L is the maximized value of the likelihood function under the estimated model.

2.9.2 Bayesian Information Criterion (BIC)

The Bayesian information criterion (BIC) or Schwarz Criterion derived by Schwarz (1978) is a criterion for model selection among a class of parametric models with different numbers of parameters. BIC arises from a Bayesian viewpoint with equal prior probability on each model and very vague priors on the parameters, given the model. The assumed purpose of the BIC-selected model was often simple prediction; as opposed to scientific understanding of the process or system under study. The BIC is very closely related to AIC except that the former takes into cognizance the logarithm of the sample size i.e. $\log(n)$. According to Burnham and Anderson (2002), $\log(n)$ is needed for

idealized asymptotic consistency. In BIC, the penalty for additional parameters is stronger than that of the AIC. The formula for the BIC is given as;

$$BIC = K \log n - 2 \log \{L(\hat{\theta})\} \quad (2.29)$$

Where,

n is the sample size ; K and L are as defined earlier in the Akaike information criterion.

2.9.3 Corrected Akaike Information Criterion (CAIC)

Hurvich and Tsai (1989) proposed the Corrected Akaike information criterion (CAIC). The rationale behind their proposal is that in some settings, the AIC may be characterized by a large negative bias which limits its effectiveness as a model selection criterion. The advantage of the CAIC over the AIC is that in small-sample applications, CAIC estimates the expected discrepancy with less bias than the AIC. However, the benefit of the Akaike information criterion over the corrected Akaike information criterion is that the AIC is more universally applicable, since its derivation is more general while the derivation of the CAIC relies on the form of the candidate model. Hurvich and Tsai (1989) justified CAIC for linear, non-linear regression and auto-regressive modelling. Since then, it has been extended to a number of additional frameworks which includes auto-regressive moving-average modelling (Hurvich *et al.* 1990), vector auto-regressive modelling (Hurvich and Tsai, 1993) and multivariate regression modelling (Bedrick and Tsai, 1994). In the general case the corrected Akaike information criterion is expressed as:

$$CAIC = AIC + \frac{[2K(K+1)]}{(n-K-1)} \quad (2.30)$$

Where, K and n are as defined earlier in the Akaike information criterion.

2.9.4 Hannan-Quinn Information Criterion (HQC)

Hannan and Quinn (1979) derived a criterion Hannan-Quinn information criterion (HQC) for model selection. The HQC is a measure of the goodness of fit of a statistical model, and is often used as a criterion for model selection among a finite set of models. Similar

to AIC, the HQC introduces a penalty term for the number of parameters in a model but the penalty is larger than the one in the AIC. The penalty term is:

$$c \log\{\log(n)\}$$

Where,

n is the sample size and c is a constant greater than 2.

Burnham and Anderson (2002) point out that the HQC seems to have seen little use in practice. Furthermore, HQC, like BIC, but unlike AIC, is not asymptotically efficient. Given any two estimated models, the model with the lower value of HQC is preferred.

The formula for the HQC is expressed as:

$$\text{HQC} = 2K \log\{\log(n)\} - 2 \log\{L(\hat{\theta})\} \quad (2.31)$$

Where, K , L , and n are as defined earlier in the Akaike information criterion.

2.10 Forecast

Forecast is the estimation of the value of a variable (or set of variables) at some future point in time. A forecasting exercise is usually carried out in order to provide an aid to decision-making and in planning the future. A population forecast is a projection in which certain assumptions are considered to yield a realistic picture of the future development of a population.

2.10.1 Formulation of the Autoregressive Moving Average Model

Box and Jenkins (1976) developed a method for analyzing stationary univariate time series data. The Box–Jenkins (ARIMA) method differences a series to stationarity and then combines the moving average with autoregressive parameters to yield a comprehensive model amenable to forecasting. The model developed serves not only to explain the underlying process generating the series, but as a basis for forecasting.

Yaffee and McGee (1999) states that the basic processes of the Box–Jenkins ARIMA (p, d, q) model include the autoregressive process, the integrated process, and the moving average process. With respect to the order of integration of a model which is indicated by the $I(d)$ distribution designation. If a series is $I(0)$, then it is stationary and has an ARIMA $(p, 0, q)$ designation. If a series requires first differencing to render it stationary,

then $d=1$ and it is distributed as $I(1)$, and is given an ARIMA($p,1,q$) designation. p is the order of the autoregressive process while q is the order of the moving average process. Once the process has been transformed into stationarity, one can proceed with the analysis.

2.11 Proposed Method

The distribution arising from order statistics using the gamma generator as introduced by Zografos and Balakrishnan (2009) and Lima *et al.* (2015) will be considered for this study.

2.11.1 Reason for proposed method

The importance and role of order statistics in real life phenomenon cannot be over-emphasized. In recent years, there has been a deliberate focus on new techniques for building meaningful distributions which includes the use of the logit of beta, the gamma generator and so on. Manton (2009) pointed out that the gamma distribution is the most popular model for analyzing skewed data and is usually used in mortality models for heterogeneous population. The Zografos–Balakrishnan- G family of distributions is particularly useful for heterogeneous data because it allows for greater flexibility of its tails.

Some fundamental properties of the gamma distribution useful in the empirical derivations will be considered briefly in the next section.

2.12 The Generalized Gamma Distribution

It is worthy to state that the parameterizations of the gamma distribution discussed in this section emanates from the Stacy (1962) generalized gamma density whose distribution function is expressed as;

$$f(x, \alpha, \beta, \rho) = \frac{\left(\frac{\rho}{\alpha\beta}\right)x^{\beta-1}e^{-\left(\frac{x}{\alpha}\right)^\rho}}{\Gamma\left(\frac{\beta}{\rho}\right)}, \quad x \in R; \quad \alpha, \beta, \rho > 0 \quad (2.32)$$

For non-negative values of x and positive values of α, β and ρ , the familiar gamma, exponential, chi-squared and Weibull variates are special cases.

2.12.1 Characterisation using Shape k and Scale θ

This is more common where for example the gamma distribution is used to model waiting times. For instance, in life testing, the waiting time until death is a random variable that is frequently modeled with a gamma distribution. A random variable X that is gamma-distributed with shape k and rate θ is denoted as;

$$X \sim \Gamma(k, \theta) \equiv \text{Gamma}(k, \theta)$$

The probability density function using the shape-scale parameterisation is;

$$f(x) = \frac{e^{-\frac{x}{\theta}} x^{k-1}}{\Gamma k \theta^k}, \quad x > 0, k, \theta > 0 \quad (2.33)$$

2.12.2 Characterisation using Shape α and Rate β

The parameterization with α and β is more common in situations where the gamma distribution is used as a conjugate prior distribution for various types of inverse scale (rate) parameters, such as the λ of an exponential distribution or a Poisson distribution or for that matter, the β of the gamma distribution itself. A random variable X that is gamma-distributed with shape α and rate β is denoted as;

$$X \sim \Gamma(\alpha, \beta) \equiv \text{Gamma}(\alpha, \beta)$$

The corresponding probability density function using the shape-rate parameterisation is;

$$f(x) = \frac{e^{-\beta x} x^{\alpha-1} \beta^\alpha}{\Gamma \alpha}, \quad x > 0, \alpha, \beta > 0 \quad (2.34)$$

2.12.3 The Gamma and Incomplete Gamma Functions

The gamma function is defined as;

$$\Gamma \alpha = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad (2.35)$$

The lower incomplete gamma function is defined as;

$$\gamma(\alpha, x) = \int_0^x x^{\alpha-1} e^{-x} dx \quad (2.36)$$

While the upper incomplete gamma function is defined as;

$$\Gamma(\alpha, x) = \int_x^\infty x^{\alpha-1} e^{-x} dx \quad (2.37)$$

2.12.4 The Zografos and Balakrishnan Gamma Generator

Consider a continuous distribution F with density f , and further Stacy's generalized gamma density; $f(x) = \frac{e^{-x\beta} x^{\alpha-1} \beta^\alpha}{\Gamma\alpha}$ for $x > 0$ and positive values of the parameters α and β . Based on this density, by replacing x with $-\ln [1 - F(x)]$, Zografos and Balakrishnan (2009) introduced the family with pdf;

$$g(x, \alpha, \beta) = \frac{\beta^\alpha}{\Gamma\alpha} [-\ln [1 - F(x)]]^{\alpha-1} e^{-[-\ln [1 - F(x)]] \beta} \frac{f(x)}{1 - F(x)}, \quad x \in R, \quad \alpha, \beta > 0 \quad (2.38)$$

This is the generalized gamma (or Stacy)-generated distribution by parent F .

If $\beta = 1$ in (3.7), it corresponds to the gamma-generated distribution by parent F . This family of distributions has its pdf as;

$$g(x, \alpha) = \frac{1}{\Gamma\alpha} [-\ln [1 - F(x)]]^{\alpha-1} e^{-[-\ln [1 - F(x)]]} \frac{f(x)}{1 - F(x)}$$

$$g(x, \alpha) = \frac{1}{\Gamma\alpha} [-\ln [1 - F(x)]]^{\alpha-1} f(x), \quad x \in R, \quad \alpha > 0. \quad (2.39)$$

Remark

Suppose $X_{U(1)}, X_{U(2)}, \dots, X_{U(n)}$ are upper record values arising from a sequence of i.i.d. continuous random variables from a population with cdf $F(x)$ and pdf $f(x)$. Then, the pdf of the n th upper record value $X_{U(n)}$ is given by;

$$f(x_{U(n)}) = \frac{[-\ln [1 - F(x)]]^{n-1}}{(n-1)!} f(x), \quad -\infty < x < \infty \text{ for } n = 1, 2, \dots \quad (2.40)$$

Converting the positive integral parameter n to a positive real parameter say; α , the family of densities with pdf which is precisely the class of gamma-generated densities is obtained. It is given by;

$$g(x) = \frac{[-\ln [1 - F(x)]]^{\alpha-1}}{\Gamma\alpha} f(x), \quad -\infty < x < \infty, \quad \alpha > 0. \quad (2.41)$$

While its Cumulative distribution function is given by;

$$G(x) = \frac{1}{\Gamma\alpha} \int_0^{-\ln(1-F(x))} t^{\alpha-1} e^{-t} dt \quad (2.42)$$

CHAPTER THREE

METHODOLOGY

In chapter two, a review of works based on the Lee-Carter model, the Zografos–Balakrishnan-Gfamily of distributions and some other aspects was done. In this section, the methodology of the proposed Gamma-Normal Lee-Carter distribution is discussed. The empirical derivations of the parameters of the proposed distribution and its statistical properties are also presented.

3.1 Developing the Gamma-Normal Lee-Carter Model

The Lee-carter approach assumes that its error term follows a normal distribution whose probability density function must satisfy the condition below.

$$f(y) = \frac{1}{\sigma} \phi \left[\frac{y-(a_x+b_x k_t)}{\sigma} \right] \quad (3.1)$$

$$\phi \left[\frac{y-(a_x+b_x k_t)}{\sigma} \right] = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y-(a_x+b_x k_t)}{\sigma} \right]^2}$$

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y-(a_x+b_x k_t)}{\sigma} \right]^2}, \quad \sigma > 0, b_x > 0, k_t > 0 \quad (3.2)$$

Its cumulative distribution function is given by;

$$F(y) = \Phi \left(\frac{y-(a_x+b_x k_t)}{\sigma} \right) \quad (3.3)$$

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{1}{2} \left[\frac{t-(a_x+b_x k_t)}{\sigma} \right]^2} dt \quad (3.4)$$

3.1.1 Probability Density Function of Gamma-Normal Lee-Carter Model

The probability density function of the proposed model is derived by substituting equations (3.1) and (3.3) into (2.41) to give;

$$g(y) = \frac{1}{\Gamma\alpha} \frac{1}{\sigma} \phi \left[\frac{y-(a_x+b_x k_t)}{\sigma} \right] \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_x k_t)}{\sigma} \right) \right] \right]^{\alpha-1} \quad (3.5)$$

$$= \frac{1}{\Gamma\alpha} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y-(a_x+b_x k_t)}{\sigma} \right]^2} \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_x k_t)}{\sigma} \right) \right] \right]^{\alpha-1}, \quad \alpha > 1, \sigma > 0 \quad (3.6)$$

When $\alpha = 1$, equation (3.6) coincides with the normal distribution.

3.2 Statistical Properties of the Proposed Model

In this section some statistical properties of the proposed distribution will be considered. These include; cumulative distribution function, verification of proper probability density function, hazard function and asymptotic behaviour.

3.2.1 Cumulative Distribution Function

The cumulative distribution function is derived by substituting equation (3.3) into (2.42) to give;

$$\begin{aligned} p(X \leq x) &= \frac{1}{\Gamma\alpha} \int_0^{-\ln \left\{ 1 - \Phi \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right] \right\}} t^{\alpha-1} e^{-t} dt \\ &= \frac{1}{\Gamma\alpha} \int_0^{-\ln \left\{ 1 - \Phi \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right] \right\}} t^{\alpha-1} e^{-t} dt \end{aligned} \quad (3.7)$$

From the definition of the lower incomplete gamma function in equation (2.36),

$$\begin{aligned} \int_0^{-\ln \left\{ 1 - \Phi \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right] \right\}} t^{\alpha-1} e^{-t} dt \text{ reduces to; } \gamma \left(\alpha, -\ln \left\{ 1 - \Phi \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right] \right\} \right) \\ \text{Hence, } G(y) = \frac{1}{\Gamma\alpha} \gamma \left(\alpha, -\ln \left\{ 1 - \Phi \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right] \right\} \right), \quad \alpha > 1, \sigma > 0 \end{aligned} \quad (3.8)$$

3.2.2 Verification of Proper Probability Density Function

Here we seek to verify whether the proposed distribution integrates to unity.

$$\text{We need to show that the } \int_0^{\infty} g(y) dy = 1 \quad (i)$$

$$\text{From (3.16), } g(y) = \frac{1}{\Gamma\alpha} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right]^2} \right] \left[-\ln \left[1 - \Phi \left(\frac{y - (a_x + b_x k_t)}{\sigma} \right) \right] \right]^{\alpha-1}$$

Substituting equation (3.9) into equation (i), we have;

$$\int_0^{\infty} \frac{1}{\Gamma\alpha} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y - (a_x + b_x k_t)}{\sigma} \right]^2} \right] \left[-\ln \left[1 - \Phi \left(\frac{y - (a_x + b_x k_t)}{\sigma} \right) \right] \right]^{\alpha-1} dy \quad (ii)$$

We set $\Phi\left(\frac{y-(a_x+b_xk_t)}{\sigma}\right) = z$

$$\frac{dz}{dy} = \frac{1}{\sigma} \Phi\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right] = \left[\frac{e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2}}{\sigma\sqrt{2\pi}} \right]$$

$$dy = \frac{\sigma\sqrt{2\pi} dz}{e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2}}$$

Equation (ii) becomes;

$$\begin{aligned} & \int_0^{\infty} \frac{1}{\Gamma\alpha} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \right] [-\ln [1 - z]]^{\alpha-1} \frac{\sigma\sqrt{2\pi} dz}{e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2}} \\ &= \frac{1}{\Gamma\alpha} \int_0^{\infty} [-\ln [1 - z]]^{\alpha-1} dz \end{aligned} \quad (iii)$$

Applying the log function rule to (iii), we have

$$\begin{aligned} &= \frac{1}{\Gamma\alpha} \int_0^{\infty} [\ln [1 - z]^{-1}]^{\alpha-1} dz \\ &= \frac{1}{\Gamma\alpha} \int_0^{\infty} [\ln [1 - z]^{-1}]^{\alpha-1} dz \\ &= \frac{1}{\Gamma\alpha} \int_0^{\infty} \left[\ln \left[\frac{1}{1-z} \right] \right]^{\alpha-1} dz \end{aligned} \quad (iv)$$

Applying Euler's integral to equation (iv);

$$\int_0^{\infty} \left[\ln \left[\frac{1}{1-z} \right] \right]^{\alpha-1} dz \text{ reduces to a one parameter gamma function: } \int_0^{\infty} z^{\alpha-1} e^{-z} dz$$

$$\text{Where } \int_0^{\infty} z^{\alpha-1} e^{-z} dz = \Gamma\alpha \quad (v)$$

Equation (iv) becomes $\frac{\Gamma\alpha}{\Gamma\alpha} = 1$.

Hence the proposed distribution integrates to 1.

3.2.3 Asymptotic Behaviour

We seek to investigate the behaviour of the model in equation (3.6) as $y \rightarrow 0$

From (3.6), $g(y) = \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi\left(\frac{y-(a_x+b_xk_t)}{\sigma}\right)\right]\right]^{\alpha_x-1}$

i. If $\alpha_x = 1$, we have

$$\lim_{y \rightarrow 0} g(y) = \lim_{y \rightarrow 0} \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi\left(\frac{y-(a_x+b_xk_t)}{\sigma}\right)\right]\right]^{\alpha_x-1} \quad (\text{vi})$$

From equation (vi),

$$\begin{aligned} \Gamma(\alpha_x) &= \Gamma(1) = \int_0^{\infty} e^{-y} y^{1-1} dy = [-e^{-y}]_0^{\infty} \\ &= \lim_{y \rightarrow \infty} [-e^{-y}] - [-e^{-0}] = 0 - (-1) = 1 \end{aligned}$$

Equation (vi) becomes $\lim_{y \rightarrow 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2}$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(a_x+b_xk_t)}{\sigma}\right]^2} \quad (\text{vii})$$

$$= \frac{1}{\sigma} \Phi\left[\frac{(a_x+b_xk_t)}{\sigma}\right] \quad (\text{viii})$$

ii. If $\alpha_x < 1$,

$$\lim_{y \rightarrow 0} g(y) = \lim_{y \rightarrow \infty} \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi\left(\frac{y-(a_x+b_xk_t)}{\sigma}\right)\right]\right]^{\alpha_x-1} \quad (\text{ix})$$

$$= \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi\left(\frac{-(a_x+b_xk_t)}{\sigma}\right)\right]\right]^{\alpha_x-1}$$

$$= \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi\left(\frac{-(a_x+b_xk_t)}{\sigma}\right)\right]\right]^{\alpha_x-1} \quad (\text{x})$$

This is a function of the gamma-normal distribution

3.2.4 The Hazard Function

The hazard function is a conditional density, given that the event in question has not yet occurred prior to time t .

For the proposed distribution, the hazard rate function is obtained using;

$$h(y) = \frac{g(y)}{1-G(y)} \quad (3.9)$$

$$g(y) = \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right] \right]^{\alpha_x-1}$$

$$G(y) = \frac{1}{\Gamma\alpha_x} \gamma \left[\alpha, -\ln \left(1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right) \right], \quad (3.10)$$

$$h(y) = \frac{\frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right] \right]^{\alpha_x-1}}{1 - \frac{\gamma[\alpha, -\ln \left(1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right)]}{\Gamma\alpha_x}}, \quad (3.11)$$

$$h(y) = \frac{\frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right] \right]^{\alpha_x-1}}{\Gamma\alpha_x - \gamma[\alpha, -\ln \left(1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right)] / \Gamma\alpha_x}, \quad (3.12)$$

This reduces to;

$$h(y) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right] \right]^{\alpha_x-1}}{\Gamma\alpha - \gamma[\alpha, -\ln \left(1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right)]}, \quad (3.13)$$

Recall from equations (2.35), (2.36) and (2.37), the following definitions;

$$\Gamma\alpha = \int_0^{\infty} x^{\alpha-1} e^{-x} dx ;$$

$$\gamma(\alpha, x) = \int_0^x x^{\alpha-1} e^{-x} dx ;$$

$$\Gamma(\alpha, x) = \int_x^{\infty} x^{\alpha-1} e^{-x} dx$$

Inferring from the above definitions, the denominator of equation (3.13) reduces to an upper incomplete gamma function given as;

$\Gamma(\alpha_x, -\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right])$ Such that equation (3.13) becomes;

$$h(y) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma}\right]^2} \left[-\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right] \right]^{\alpha_x-1}}{\Gamma(\alpha_x, -\ln \left[1 - \Phi \left(\frac{y-(a_x+b_xk_t)}{\sigma} \right) \right])} \quad (3.14)$$

3.3 Estimation

In the Lee and Carter (1992) model and several of its variants, the model's parameters were estimated using the singular value decomposition. We deviate from this approach by using the maximum likelihood estimation to estimate the parameters of the proposed method in order to achieve better results.

3.3.1 The Likelihood and Log-likelihood Functions

The likelihood function is defined as the density function regarded as a function of θ . Let X_1, X_2, \dots, X_n be a random sample of size n distributed according to the likelihood function of vector of parameters, $\theta = (a_x, b_x, k_t, \alpha_x \text{ and } \sigma_i)$

From equation (3.5) the probability density function of the proposed model is given by:

$$g(y) = \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y-(a_x+b_xk_t)}{\sigma_i}\right]^2} \left[-\ln\left[1 - \Phi\left(\frac{y-(a_x+b_xk_t)}{\sigma_i}\right)\right]\right]^{\alpha_x-1}$$

The likelihood function is given by:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^n \frac{1}{\Gamma\alpha_x} \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{y_i-(a_x+b_xk_t)}{\sigma_i}\right]^2} \left\{-\ln\left[1 - \Phi\left(\frac{y_i-(a_x+b_xk_t)}{\sigma_i}\right)\right]\right\}^{\alpha_x-1} \\ &= \prod_{i=1}^n e^{-\frac{1}{2\sigma_i^2}[y_i-(a_x+b_xk_t)]^2} \left\{-\ln\left[1 - \Phi\left(\frac{y_i-(a_x+b_xk_t)}{\sigma_i}\right)\right]\right\}^{\alpha_x-1} (\sigma_i\Gamma\alpha_x)^{-n} (2\pi)^{-\frac{n}{2}} \quad (3.15) \end{aligned}$$

The log-likelihood function is expressed as:

$$\begin{aligned} \ln L(\theta) &= -n(\ln\sigma_i + \ln\Gamma\alpha_x) - \frac{n}{2}(\ln 2\pi) - \frac{1}{2\sigma_i^2} \sum_{i=1}^n [y_i - (a_x + b_xk_t)]^2 + \\ &(\alpha_x - 1) \sum_{i=1}^n \ln \left\{-\ln\left[1 - \Phi\left(\frac{y_i-(a_x+b_xk_t)}{\sigma_i}\right)\right]\right\} \quad (3.16) \end{aligned}$$

Having obtained the log-likelihood function in (3.16) above, we proceed with the estimation of parameters.

3.3.2 Estimation of Parameters

The following derivatives of equation (3.16) with respect to $\alpha_x, \sigma_i, a_x, b_x$ and k_t are taken, respectively.

a. Parameter α_x (the shape parameter)

$$\frac{\partial \ln L(\theta)}{\partial \alpha_x} = \frac{-n\Gamma\alpha_x'}{\Gamma\alpha_x} + \sum_{i=1}^n \ln \left\{ -\ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}$$

Since $\frac{\Gamma\alpha_x'}{\Gamma\alpha_x}$ is a di-gamma function, $\psi(\alpha_x)$

$$\frac{\partial \ln L(\theta)}{\partial \alpha_x} = -n\psi(\alpha_x) + \sum_{i=1}^n \ln \left\{ -\ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}$$

$$\frac{\partial \ln L(\theta)}{\partial \alpha_x} = \sum_{i=1}^n \ln \left\{ -\ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\} - n\psi(\alpha_x) \quad \forall x \quad (3.17)$$

b. Parameter σ_i (scale parameter)

The next target is to obtain $\frac{\partial \ln L(\theta)}{\partial \sigma_i}$;

For ease of differentiation, equation (3.16) is split into smaller units;

$$W = -n(\ln \sigma_i + \ln \Gamma \alpha_x)$$

$$X = -\frac{n}{2}(\ln 2\pi)$$

$$Y = -\frac{1}{2\sigma_i^2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2 = -\frac{1}{2}\sigma_i^{-2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

$$Z = (\alpha_x - 1) \sum_{i=1}^n \ln \left\{ -\ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}$$

$$\frac{\partial W}{\partial \sigma_i} = -\frac{n}{\sigma_i} \quad (3.18)$$

$$\frac{\partial X}{\partial \sigma_i} = 0 \quad (3.19)$$

$\frac{\partial Y}{\partial \sigma_i} = \sigma_i^{-3} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$ simplifies to;

$$\frac{\partial Y}{\partial \sigma_i} = \frac{1}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]^2 \quad (3.20)$$

Using logarithmic rule, $(-b \log a = \log a^{-b})$, Z is re-expressed as;

$$Z = (\alpha_x - 1) \sum_{i=1}^n \ln \left\{ \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}^{-1} \quad (3.21)$$

A combination of the product rule and the log of log derivative in equation (3.22) is used to obtain the derivative of Z ,

$$[\log (\log u)]' = \frac{\log u'}{\log u} = \frac{u'}{u \log} \quad (3.22)$$

$$\text{So that } \frac{\partial Z}{\partial \sigma_i} = (\alpha_x - 1) \sum_{i=1}^n \frac{\left\{ \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \right\}'}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}} \quad (3.23)$$

We continue with the numerator of (3.23)

$$\text{Let } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] = U$$

$$\text{And } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} = R$$

Such that $U^{-1} = R$

$$\text{Now } \frac{\partial R}{\partial U} = -U^{-2}, \quad \frac{\partial U}{\partial \sigma_i} = - \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2} \right] - \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\frac{\partial U}{\partial \sigma_i} = \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2} \right] \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\text{We seek } \frac{\partial R}{\partial \sigma_i} = \frac{\partial R}{\partial U} \times \frac{\partial U}{\partial \sigma_i}$$

$$= -U^{-2} \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2} \right] \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$= - \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-2} \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2} \right] \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \quad (3.24)$$

Substituting (3.24) back into the numerator of (3.23) leads to;

$$\begin{aligned}
\frac{\partial Z}{\partial \sigma_i} &= (\alpha_x - 1) \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2}\right] \phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2}\right] \phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} - \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i^2}\right] \phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{\frac{1}{\sigma_i} \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right] \left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)}{\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]}
\end{aligned}$$

This further reduces to;

$$\frac{\partial Z}{\partial \sigma_i} = \frac{(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right] \phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right] \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \quad (3.25)$$

Bringing together the results obtained in (3.18), (3.19), (3.20) and (3.25), We have,

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial \sigma_i} &= \\
&= -\frac{n}{\sigma_i} + \frac{1}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]^2 + \frac{(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right] \phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right] \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&\quad \forall x, t
\end{aligned} \quad (3.26)$$

c. Parameter α_x (the general pattern of mortality by age)

To obtain $\frac{\partial \ln L(\theta)}{\partial \alpha_x}$, recall W, X, Y and Z as used in previous section.

$$W = -n(\ln \sigma_i + \ln \Gamma \alpha_x)$$

$$X = -\frac{n}{2}(\ln 2\pi)$$

$$Y = -\frac{1}{2\sigma_i^2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

$$= -\frac{1}{2} \sigma_i^{-2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

We proceed as follows;

$$\frac{\partial W}{\partial a_x} = 0 \quad (3.27)$$

$$\frac{\partial X}{\partial a_x} = 0 \quad (3.28)$$

$\frac{\partial Y}{\partial a_x}$; we apply the product of function of a function stated below;

$$\text{Let } -\frac{1}{2\sigma_i^2} = U \quad \text{and} \quad \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2 = V$$

Then,

$$\frac{\partial Y}{\partial a_x} = U \frac{\partial V}{\partial a_x} + V \frac{\partial U}{\partial a_x} \quad (3.29)$$

$$\text{Let } y_i - (a_x + b_x k_t) = P$$

$$\text{It implies that } \sum_{i=1}^n P^2 = V$$

$$\frac{\partial V}{\partial P} = 2 \sum_{i=1}^n P, \quad \frac{\partial P}{\partial a_x} = -1$$

$$\text{Now } \frac{\partial V}{\partial a_x} = \frac{\partial V}{\partial P} \times \frac{\partial P}{\partial a_x} = -2 \sum_{i=1}^n P$$

$$\frac{\partial V}{\partial a_x} = -2 \sum_{i=1}^n [y_i - (a_x + b_x k_t)], \quad \frac{\partial U}{\partial a_x} = 0$$

It follows from equation (3.29) that;

$$\frac{\partial Y}{\partial a_x} = \left(-\frac{1}{2\sigma_i^2}\right) \{-2 \sum_{i=1}^n [y_i - (a_x + b_x k_t)]\}$$

$$\frac{\partial Y}{\partial a_x} = \frac{1}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \quad (3.30)$$

Next is $\frac{\partial Z}{\partial a_x}$;

Recall from (3.32), $Z = (\alpha_x - 1) \sum_{i=1}^n \ln \left\{ \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}^{-1}$

Leaving $(\alpha_x - 1)$ constant and applying the log derivative rule in (3.33)

$$\frac{\partial Z}{\partial a_x} = (\alpha_x - 1) \sum_{i=1}^n \frac{\left\{ \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \right\}'}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}} \quad (3.31)$$

We proceed with the numerator of (3.31),

Recall;

$$\text{Let } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] = U$$

$$\text{And } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} = R$$

Such that $U^{-1} = R$,

$$\frac{\partial R}{\partial U} = -U^{-2}$$

$$\frac{\partial U}{\partial a_x} = \frac{1}{\sigma_i} \phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right)$$

We seek $\frac{\partial R}{\partial a_x} = \frac{\partial R}{\partial U} \times \frac{\partial U}{\partial a_x}$

$$= \frac{-U^{-2}}{\sigma_i} \phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right)$$

Thus the derivative of the numerator (3.31) w. r. t a_x yields:

$$- \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-2} \frac{1}{\sigma_i} \phi \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \quad (3.32)$$

Substituting (3.32) into the numerator of (3.31) leads to;

$$\begin{aligned}
\frac{\partial Z}{\partial a_x} &= (\alpha_x - 1) \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \frac{1}{\sigma_i} \phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \frac{1}{\sigma_i} \phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} - \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{-\frac{1}{\sigma_i} \phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right] \left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}}{-\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{-\frac{1}{\sigma_i} \phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{-\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right] \left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&= (\alpha_x - 1) \sum_{i=1}^n \frac{\frac{1}{\sigma_i} \phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right] \left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\
&= \frac{(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right] \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \tag{3.33}
\end{aligned}$$

Bringing together the results obtained in (3.27), (3.28), (3.30) and (3.33),

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial a_x} &= \\
\frac{1}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right] &+ \frac{(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right] \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \quad \forall x \tag{3.34}
\end{aligned}$$

d. Parameter b_x (the relative pace of change in mortality by age)

To obtain $\frac{\partial \ln L(\theta)}{\partial b_x}$, recall W, X, Y and Z as used in previous sections.

$$W = -n(\ln \sigma_i + \ln \Gamma \alpha_x)$$

$$X = -\frac{n}{2} (\ln 2\pi)$$

$$Y = -\frac{1}{2\sigma_i^2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2 = -\frac{1}{2} \sigma_i^{-2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

We proceed as follows;

$$\frac{\partial W}{\partial b_x} = 0 \quad (3.35)$$

$$\frac{\partial X}{\partial b_x} = 0 \quad (3.36)$$

$$\frac{\partial Y}{\partial b_x} = U \frac{\partial V}{\partial b_x} + V \frac{\partial U}{\partial b_x}$$

Recall equations U, V and P used earlier in previous section.

$$U = -\frac{1}{2\sigma_i^2}$$

$$V = \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

$$P = y_i - (a_x + b_x k_t)$$

It implies that $\sum_{i=1}^n P^2 = V$

$$\frac{\partial V}{\partial P} = 2 \sum_{i=1}^n P, \quad \frac{\partial P}{\partial b_x} = -k_t$$

$$\text{Now } \frac{\partial V}{\partial b_x} = \frac{\partial V}{\partial P} \times \frac{\partial P}{\partial b_x} = -2k_t \sum_{i=1}^n P$$

$$\frac{\partial V}{\partial b_x} = -2k_t \sum_{i=1}^n [y_i - (a_x + b_x k_t)],$$

$$\frac{\partial U}{\partial b_x} = 0$$

$$\text{It follows that } \frac{\partial Y}{\partial b_x} = U \frac{\partial V}{\partial b_x} + V \frac{\partial U}{\partial b_x}$$

$$= \left(-\frac{1}{2\sigma_i^2}\right) \{-2k_t \sum_{i=1}^n [y_i - (a_x + b_x k_t)]\}$$

$$\frac{\partial Y}{\partial b_x} = \frac{k_t}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \quad (3.37)$$

Next is $\frac{\partial Z}{\partial b_x}$;

Recall from (3.32); $Z = (\alpha_x - 1) \sum_{i=1}^n \ln \left\{ \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}^{-1}$

Leaving $(\alpha_x - 1)$ constant and applying the log derivative rule in (3.22)

$$\frac{\partial Z}{\partial b_x} = (\alpha_x - 1) \sum_{i=1}^n \frac{\left\{ \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \right\}'}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}} \quad (3.38)$$

We begin with the numerator of (3.38),

Recall equations U and R used previously;

$$\text{Let } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] = U$$

$$\text{And } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} = R$$

Such that $U^{-1} = R$,

$$\frac{\partial R}{\partial U} = -U^{-2}$$

$$\frac{\partial U}{\partial b_x} = -\frac{k_t}{\sigma_i} \left[-\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\frac{\partial U}{\partial b_x} = \frac{k_t}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

We seek $\frac{\partial R}{\partial b_x} = \frac{\partial R}{\partial U} \times \frac{\partial U}{\partial b_x}$

$$\frac{\partial R}{\partial b_x} = -U^{-2} \frac{k_t}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\frac{\partial R}{\partial b_x} = - \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-2} \frac{k_t}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \quad (3.39)$$

Substituting (3.39) into the numerator of (3.38) gives;

$$\frac{\partial Z}{\partial b_x} = (\alpha_x - 1) \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \frac{k_t}{\sigma_i} \left[\Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}}$$

This can be re-expressed as;

$$\begin{aligned} \frac{\partial Z}{\partial b_x} &= \frac{k_t(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-2} \left[\Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]}{\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}} \\ &= \frac{k_t(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \left[\Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]}{\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}} \\ &= \frac{k_t(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{-\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1} \left[\Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]}{-\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\ &= \frac{k_t(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\Phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right] \left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]^{-1}}{\ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \\ \frac{\partial Z}{\partial b_x} &= \frac{k_t(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\Phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right) \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \end{aligned} \quad (3.40)$$

Finally bringing together all the results obtained in (3.35), (3.36), (3.37) and (3.40),

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial b_x} &= \frac{k_t}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \\ &\quad + \frac{k_t(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\Phi\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right]}{1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right) \ln\left[1 - \Phi\left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i}\right)\right]} \forall x \end{aligned} \quad (3.41)$$

e. Parameter k_t (mortality index)

To obtain $\frac{\partial \ln L(\theta)}{\partial k_t}$, recall W, X, Y and Z as used in previous sections.

$$W = -n(\ln \sigma_i + \ln \Gamma \alpha_x)$$

$$X = -\frac{n}{2}(\ln 2\pi)$$

$$Y = -\frac{1}{2\sigma_i^2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2 = -\frac{1}{2} \sigma_i^{-2} \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

We proceed as follows;

$$\frac{\partial W}{\partial k_t} = 0 \tag{3.42}$$

$$\frac{\partial X}{\partial k_t} = 0 \tag{3.43}$$

$$\frac{\partial Y}{\partial k_t} = U \frac{\partial V}{\partial k_t} + V \frac{\partial U}{\partial k_t}$$

Recall equations U, V and P used earlier in previous section.

$$U = -\frac{1}{2\sigma_i^2}$$

$$V = \sum_{i=1}^n [y_i - (a_x + b_x k_t)]^2$$

$$P = y_i - (a_x + b_x k_t)$$

It implies that $\sum_{i=1}^n P^2 = V$

$$\frac{\partial V}{\partial P} = 2 \sum_{i=1}^n P, \quad \frac{\partial P}{\partial k_t} = -b_x$$

$$\text{Now } \frac{\partial V}{\partial k_t} = \frac{\partial V}{\partial P} \times \frac{\partial P}{\partial k_t} = -2b_x \sum_{i=1}^n P$$

$$\frac{\partial V}{\partial k_t} = -2b_x \sum_{i=1}^n [y_i - (a_x + b_x k_t)],$$

$$\frac{\partial U}{\partial k_t} = 0$$

It follows that $\frac{\partial Y}{\partial k_t} = U \frac{\partial V}{\partial k_t} + V \frac{\partial U}{\partial k_t}$

$$= \left(-\frac{1}{2\sigma_i^2}\right) \{-2b_x \sum_{i=1}^n [y_i - (a_x + b_x k_t)]\}$$

$$\frac{\partial Y}{\partial k_t} = \frac{b_x}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \quad (3.44)$$

Next is $\frac{\partial Z}{\partial k_t}$:

$$\text{Recall from (3.21), } Z = (\alpha_x - 1) \sum_{i=1}^n \ln \left\{ \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\}^{-1}$$

Leaving $(\alpha_x - 1)$ constant and applying the log derivative rule in (3.22)

$$\frac{\partial Z}{\partial k_t} = (\alpha_x - 1) \sum_{i=1}^n \frac{\left\{ \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \right\}'}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}} \quad (3.45)$$

We begin with the numerator of (3.45),

Recall equations U and R used previously;

$$\text{Let } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] = U$$

$$\text{And } \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} = R$$

Such that; $U^{-1} = R$,

$$\frac{\partial R}{\partial U} = -U^{-2}$$

$$\text{Now } \frac{\partial U}{\partial k_t} = -\frac{b_x}{\sigma_i} \left[-\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\frac{\partial U}{\partial k_t} = \frac{b_x}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\text{We seek } \frac{\partial R}{\partial k_t} = \frac{\partial R}{\partial U} \times \frac{\partial U}{\partial k_t}$$

$$\frac{\partial R}{\partial k_t} = -U^{-2} \frac{b_x}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]$$

$$\frac{\partial R}{\partial k_t} = - \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-2} \frac{b_x}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \quad (3.46)$$

Substituting (3.46) into the numerator of (3.45) gives;

$$\frac{\partial Z}{\partial k_t} = (\alpha_x - 1) \sum_{i=1}^n \frac{- \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-2} \frac{b_x}{\sigma_i} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}}$$

This can be re-expressed as;

$$\begin{aligned} \frac{\partial Z}{\partial k_t} &= \frac{b_x(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{- \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-2} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}} \\ &= \frac{b_x(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{- \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]}{\ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}} \\ &= \frac{b_x(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{- \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1} \left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]}{- \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \\ &= \frac{b_x(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\left[\Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]^{-1}}{\ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \\ \frac{\partial Z}{\partial k_t} &= \frac{b_x(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\Phi \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]}{1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \end{aligned} \quad (3.47)$$

Finally bringing together all the results obtained in (3.42), (3.43), (3.44) and (3.47),

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial k_t} &= \frac{b_x}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \\ &\quad + \frac{b_x(\alpha_x - 1)}{\sigma_i} \sum_{i=1}^n \frac{\Phi \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]}{1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \forall t \end{aligned} \quad (3.48)$$

3.3.3 Components of the Score Vector

Therefore, from section 3.3.2, the components of the score vector $U(\theta)$ of the Gamma-Normal Lee-Carter are of the forms:

$$U_{a_x}(\theta) = \frac{1}{\sigma_i} \sum_t \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] + \frac{(\alpha - 1)}{\sigma_i} \sum_t \frac{\phi \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \quad (3.49)$$

$$U_{\alpha}(\theta) = \sum_t \ln \left\{ - \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \right\} - n \psi(\alpha) \quad (3.50)$$

$$U_{\sigma_i}(\theta) = \frac{n}{\sigma_i} + \frac{1}{\sigma_i} \sum_{i=1}^n \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]^2 + \frac{(\alpha - 1)}{\sigma_i} \sum_{i=1}^n \frac{\left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] \phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right)}{\left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right] \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \quad (3.51)$$

$$U_{b_x}(\theta) = \frac{k_t}{\sigma_i} \sum_t \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] + \frac{k_t(\alpha - 1)}{\sigma_i} \sum_t \frac{\phi \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]}{1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \quad (3.52)$$

$$U_{k_t}(\theta) = \frac{b_x}{\sigma_i} \sum_x \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right] + \frac{b_x(\alpha - 1)}{\sigma_i} \sum_x \frac{\phi \left[\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right]}{1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \ln \left[1 - \Phi \left(\frac{y_i - (a_x + b_x k_t)}{\sigma_i} \right) \right]} \quad (3.53)$$

Setting these expressions to zero and solving them simultaneously yields the maximum likelihood estimates (MLEs) of the five parameters. The package *bbmle* by Bolker (2017) was used to solve for the parameters in equations 3.49 to 3.53 above.

CHAPTER FOUR

ANALYSIS OF DATA AND DISCUSSION

4.1 Analysis of Data

In chapter three, the Gamma-normal Lee-Carter model was presented with derivations made for its parameters. This chapter entails data analysis and discussion.

4.2 Data Choice and Source

The data sets used for the study are the age-specific mortality data for both males and females in Nigeria, a developing country located in West Africa. The reason behind our choice of the two sexes instead of a combined sex mortality data is to avoid problems with implausible sex differentials raised by Li *et al.* (2000). The data set was previously studied by Chukwu and Oladipupo (2012) under the Li *et al.* (2002) framework. It was obtained from the Global Health Observatory (WHO 2017), an arm of the World Health Organisation Indicator and Measurement Registry (IMR). IMR is a central source of metadata of health-related indicators used by WHO and other organizations. The data sets are displayed in Tables 1.1 to 1.4 of the Appendix.

4.3 Data Structure and Notation

For the mortality data set, the age distribution of the population ranges between less than one year old to 85 years and above for the time period 2000 to 2015 for both males and females. Throughout the study, the number of deaths (d_{xt}), the central exposures or the population exposed to risk (E_{xt}) and the mortality rates (m_{xt}) are arranged in a rectangular array format comprising ages (on the row) $x = x_1, x_2, \dots, x_k$ and calendar years (on the columns) $t = t_1, t_2, \dots, t_n$.

4.4 Results and Discussion

The results in this study begin with a descriptive analysis of the dataset. Two non-parametric procedures are employed to test for normality. They are the Kolmogorov-Smirnov and Shapiro-Wilks normality tests.

Due to the inability to obtain closed-form expressions of the parameter estimates and the presence of the bilinear predictors present in the Lee-Carter model and its variants, the parameters of the models are estimated through an iterative minimization technique. The

proposed Gamma-Normal Lee-Carter (GNLC) is compared with three other variants of the Lee-Carter model. They are; the Lee and Carter (1992) with the abbreviation (LC), the Brouhns (2002) model - (BR), the Renshaw Haberman (2006) - (RH). The *R* packages; *BBMLE* by Bolker (2017), *StMoMo* by Villegas *et al.* (2015) and *ILC* by Haberman and Butt (2010) were used in the analysis of data.

Forecast of the mortality indexes was done under suitable Arima framework for a 20-year period within 80-95% confidence interval for each model. Simulation study was done using data of sizes $n ; t = (15; 5), (20; 20)$ and $(20; 30)$ where n and t are the sizes of the age-groups and time periods respectively.

To compare the proposed model with some other variants of the Lee-Carter, some criteria like, Akaike information criterion (AIC), Corrected Akaike information criterion (CAIC), Bayesian Information Criterion (BIC) and the Hannan Quinine Information Criterion (HQC) were considered.

Table 4.1: Descriptive Analysis for Male Mortality Data

	Mean	Median	Mode	Skewness	Kurtosis	Minimum	Maximum
<1	.104250	.103000	.0800	.1500	-1.281	.0800	.1320
1-4	.016188	.015500	.0110	.2180	-1.271	.0110	.0220
5-9	.006125	.006000	.0060	-.192	-.821	.0050	.0070
10-14	.003313	.003000	.0030	.8950	-1.391	.0030	.0040

15-19	.004250	.004000	.0040	1.278	-.440	.0040	.0050
20-24	.006313	.006000	.0060	.8950	-1.391	.0060	.0070
25-29	.007188	.007000	.0070	-.197	-.373	.0060	.0080
30-34	.008438	.008500	.0090	-.653	-.321	.0070	.0090
35-39	.010563	.011000	.0110	-.246	.249	.0090	.0120
40-44	.012375	.012000	.0120	-.421	-.454	.0110	.0130
45-49	.013875	.014000	.0130	-.146	-1.405	.0120	.0150
50-54	.017750	.017500	.0170	.114	-1.336	.0160	.0190
55-59	.022813	.023000	.0240	-.253	-1.355	.0210	.0240
60-64	.033813	.033500	.0360	.083	-1.573	.0310	.0360
65-69	.051250	.051000	.0540	.018	-1.657	.0480	.0540
70-74	.083438	.082500	.0880	.196	-1.635	.0790	.0880
75-79	.137063	.135500	.1320	.332	-1.602	.1310	.1440
80-84	.222125	.220000	.2150	.352	-1.675	.2150	.2310
85+	.350063	.347500	.3450	.450	-1.662	.3430	.3590

Table 4.2: Descriptive Analysis for Female Mortality Data

	Mean	Median	Mode	Skewness	Kurtosis	Minimum	Maximum
<1	.086375	.085500	.0660	.183	-1.224	.0660	.1100
1-4	.015938	.015500	.0110	.181	-1.388	.0110	.0220
5-9	.006313	.006000	.0060	-.205	-.377	.0050	.0070
10-14	.003438	.003000	.0030	.279	-2.219	.0030	.0040

15-19	.004000	.004000	.0040	-	-	.0040	.0040
20-24	.005125	.005000	.0050	.343	1.653	.0040	.0060
25-29	.006938	.007000	.0060	.129	-1.646	.0060	.0080
30-34	.008813	.009000	.0090	-.375	-.948	.0070	.0100
35-39	.010938	.011000	.0110	-1.474	2.273	.0080	.0120
40-44	.011375	.012000	.0120	-1.545	2.277	.0090	.0120
45-49	.011875	.012000	.0120	-.627	.754	.0100	.0130
50-54	.014313	.014500	.0150	-1.397	2.016	.0120	.0150
55-59	.018813	.019000	.0200	-.253	-1.355	.0170	.0200
60-64	.029563	.029500	.0310	-.075	-1.259	.0270	.0320
65-69	.047313	.047000	.0500	-.014	-1.555	.0440	.0500
70-74	.080063	.079500	.0840	.074	-1.608	.0750	.0850
75-79	.133563	.132500	.1290	.160	-1.562	.1260	.1410
80-84	.220563	.219000	.2290	.172	-1.645	.2110	.2300
85+	.340250	.338000	.3320	.294	-1.684	.3320	.3500

4.4.1 Descriptive Analysis for Age-specific Mortality Data

The descriptive analysis for the age-specific mortality data is shown in tables 4.1 and 4.2. On the average, age-group 10-14 has the lowest mortality rate for both gender with 0.003313 for the males and 0.003438 for the females. The highest mortality rates on the average are found in age-group 85⁺ with 0.35006 for the males and 0.34025 for the females. Furthermore, the results show that for the male data set, 12 age-groups exhibit

some degree of skewness to the right while 7 age-groups are left-skewed. However, for the female data, the ratio of left to right skewness is 9:9 with only one age-group (15-19) having no skew. This implies that only age-group 15-19 exhibits a Normal distribution.

Mortality Rate (Male)

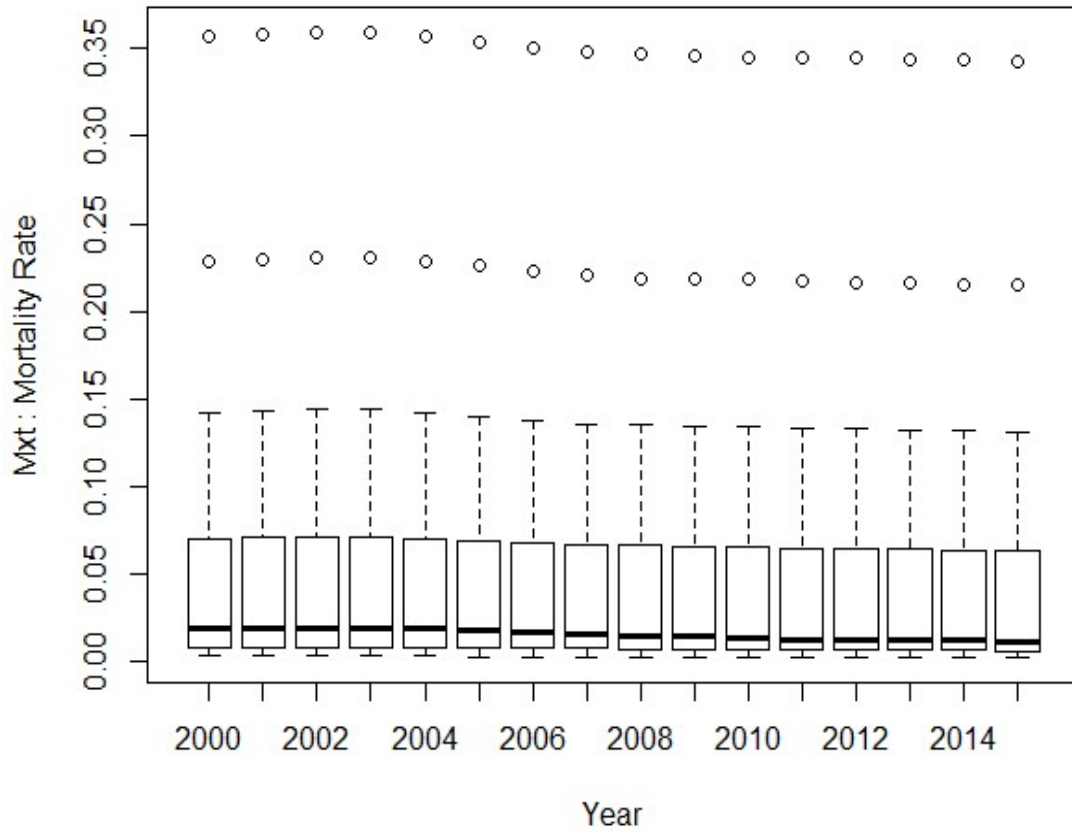


Figure 4.1: Boxplots for Male Mortality Rates

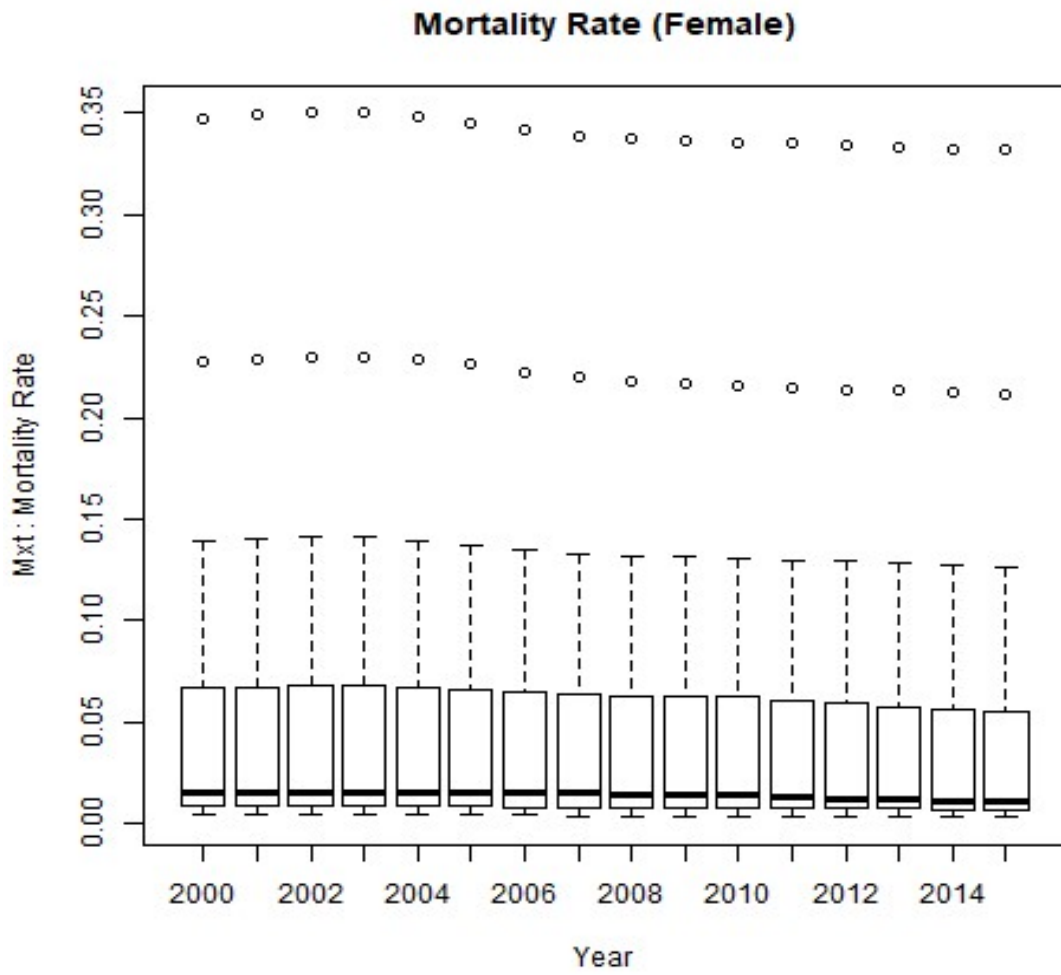


Figure 4.2: Boxplots for Female Mortality Rates

log Mortality Rate (Male)

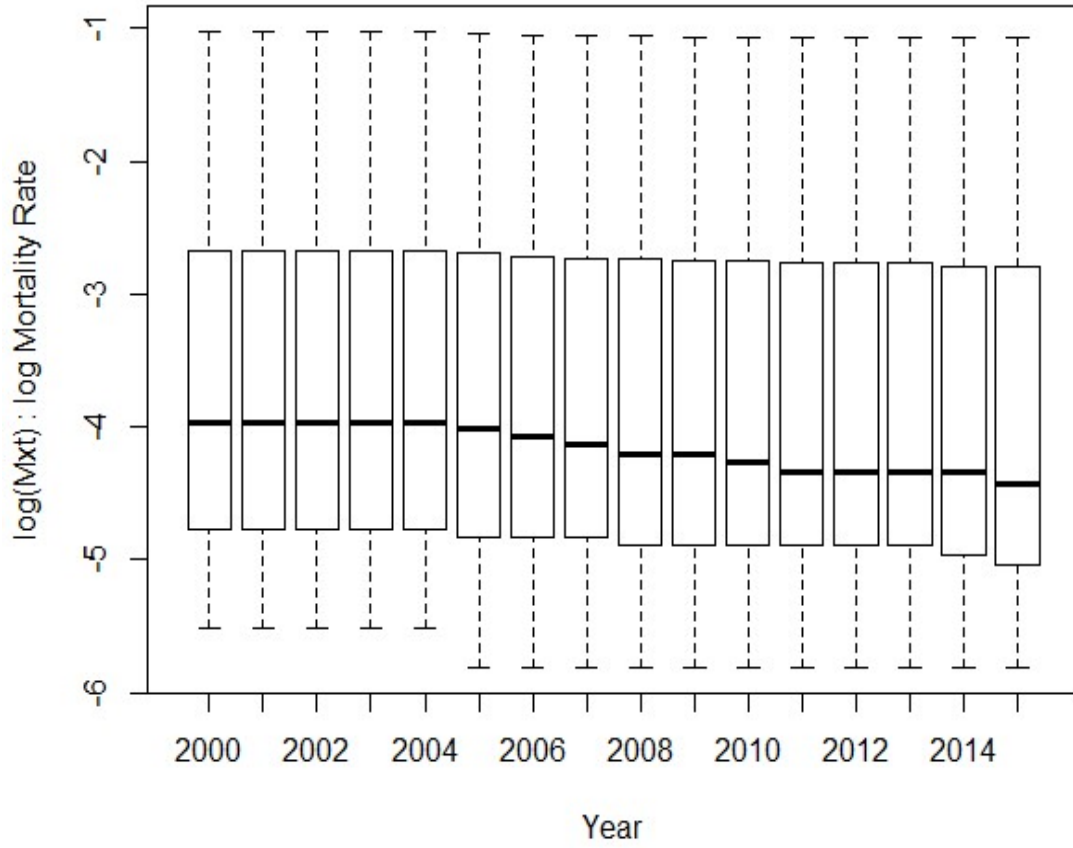


Figure 4.3: Boxplots for Male Log Mortality Rates across Time-period

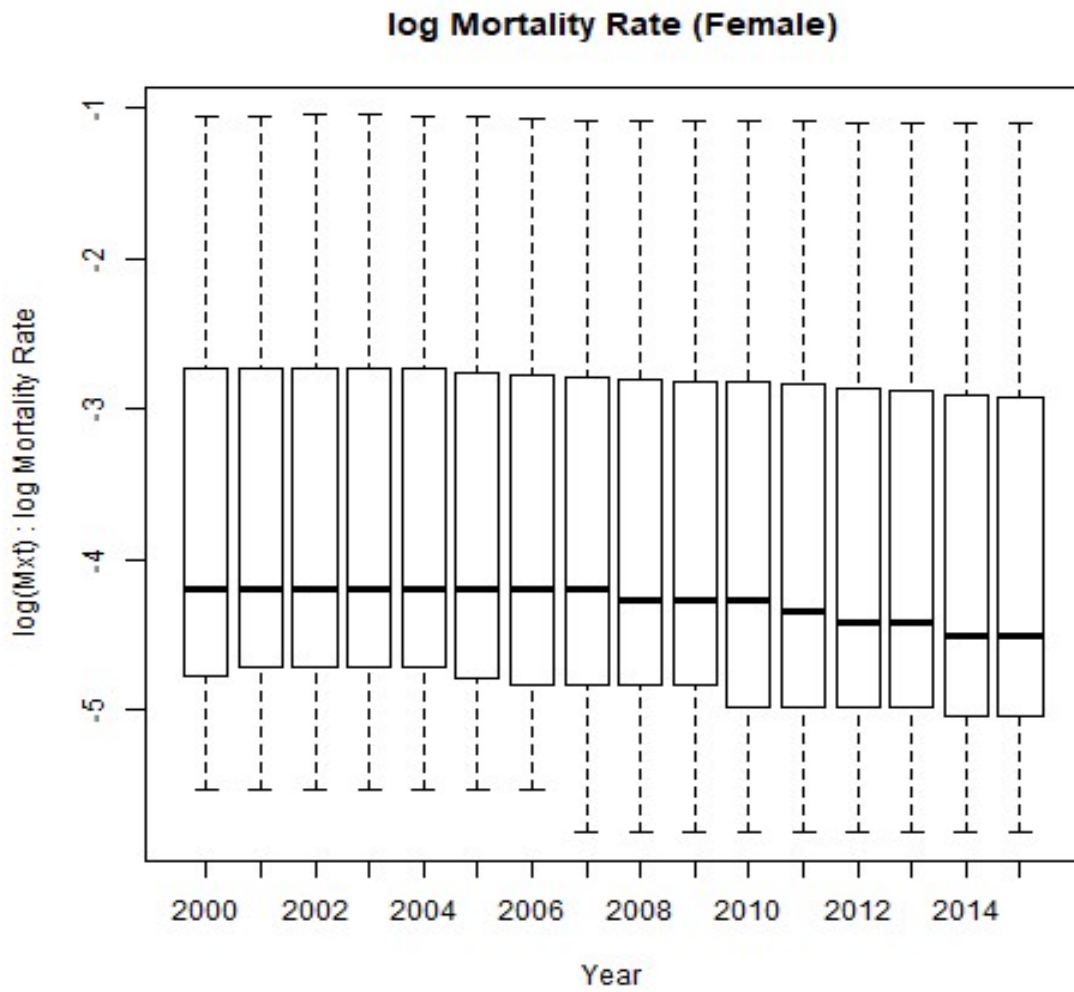


Figure 4.4: Boxplots for Female Log Mortality Rates

4.4.2 Boxplots for Mortality and Log Mortality Rates

The boxplots are displayed in Figures 4.1 to 4.4. The plots in Figures 4.1 and 4.2 reflect the presence of outliers in the male and female mortality data. However, a close look at the plots in Figures 4.3 and 4.4 reflect the absence of outliers in the log mortality rates. It implies that the outliers disappeared after the logarithm of the mortality rates was taken. This transformation shows the effect of the logarithm function on the mortality rates

Table 4.3: Kolmogorov Smirnov Normality Test Across Ages

Ages	Male		Female	
	D	P-value	D	P-value
<1	0.53188	9.413e-05	0.8875	2.005e-13
1-4	0.50439	0.0005827	0.52631	0.0001173
5-9	0.50199	0.0006294	0.50439	0.0005827
10-14	0.5012	0.0006457	0.50199	0.0006294
15-19	0.5016	0.0006375	0.5012	0.0006457
20-24	0.50239	0.0006213	0.5016	0.0006375
25-29	0.50239	0.0006213	0.5016	0.0006375
30-34	0.50279	0.0006134	0.50239	0.0006213
35-39	0.50359	0.0005979	0.50279	0.0006134
40-44	0.50439	0.0005827	0.50319	0.0006056
45-49	0.50479	0.0005752	0.50359	0.0005979
50-54	0.50638	0.0005463	0.50399	0.0005902
55-59	0.50838	0.000512	0.50479	0.0005752
60-64	0.51237	0.0004495	0.50678	0.0005392
65-69	0.51914	0.0003594	0.51077	0.0004736
70-74	0.53148	0.0002373	0.51755	0.0003789
75-79	0.55211	0.0001161	0.52989	0.0002505
80-84	0.58512	3.492e-05	0.55013	0.0001245
85+	0.63420	5.144e-06	0.58356	3.702e-05

Table 4.4: Kolmogorov Normality Test across Time Period

Year	Male		Female	
	D	P-value	D	P-value
2000	0.87713	4.019e-13	0.87831	3.717e-13
2001	0.87654	4.18e-13	0.87713	4.019e-13
2002	0.87595	4.348e-13	0.87654	4.18e-13
2003	0.87595	4.348e-13	0.87654	4.18e-13
2004	0.87713	4.019e-13	0.87713	4.019e-13
2005	0.87889	3.574e-13	0.87889	3.574e-13
2006	0.88064	3.18e-13	0.88122	3.059e-13
2007	0.88179	2.943e-13	0.88237	2.831e-13
2008	0.88295	2.724e-13	0.88352	2.621e-13
2009	0.88352	2.621e-13	0.88409	2.522e-13
2010	0.88352	2.621e-13	0.88467	2.427e-13
2011	0.88409	2.522e-13	0.88524	2.336e-13
2012	0.88467	2.427e-13	0.8858	2.248e-13
2013	0.88467	2.427e-13	0.88637	2.164e-13
2014	0.88524	2.336e-13	0.88694	2.083e-13
2015	0.88524	2.336e-13	0.8875	2.005e-13

Table 4.5: Shapiro Wilks Normality Test across Age-groups

	Male		Female	
	W	P-value	W	P-value
<1	0.95104	0.5064	0.95321	0.5422
1-4	0.93371	0.2789	0.93208	0.2629
5-9	0.81401	0.004205	0.75896	0.0008163
10-14	0.59088	1.33e-05	0.63827	0.000038
15-19	0.54643	5.272e-06	-	-
20-24	0.59088	1.33e-05	0.67592	0.0000927
25-29	0.79358	0.002241	0.79624	0.002429
30-34	0.7502	0.0006388	0.87207	0.02926
35-39	0.84934	0.01332	0.78364	0.001665
40-44	0.76031	0.0008479	0.7286	0.000355
45-49	0.83054	0.007138	0.84624	0.01200
50-54	0.83342	0.007842	0.75859	0.0008078
55-59	0.84462	0.01137	0.84462	0.01137
60-64	0.86872	0.02599	0.92434	0.19800
65-69	0.86736	0.02479	0.87618	0.03385
70-74	0.88328	0.04366	0.89819	0.07518
75-79	0.88345	0.04394	0.91133	0.12220
80-84	0.85928	0.01871	0.89424	0.06504
85+	0.83964	0.009627	0.87264	0.02985

Table 4.6: Shapiro Wilks Normality Test across Time Period

Year	Male		Female	
	W	P-value	W	P-value
2000	0.66547	2.217e-05	0.65229	1.582e-05
2001	0.66462	2.169e-05	0.64946	1.474e-05
2002	0.66269	2.064e-05	0.64817	1.426e-05
2003	0.66126	1.989e-05	0.64731	1.396e-05
2004	0.66017	1.934e-05	0.64464	1.306e-05
2005	0.65995	1.923e-05	0.64209	1.225e-05
2006	0.65663	1.767e-05	0.63896	1.134e-05
2007	0.65377	1.643e-05	0.63823	1.113e-05
2008	0.65036	1.507e-05	0.63532	1.036e-05
2009	0.64766	1.408e-05	0.63136	9.402e-06
2010	0.64528	1.327e-05	0.62978	9.046e-06
2011	0.64172	1.214e-05	0.6266	8.371e-06
2012	0.63959	1.151e-05	0.6252	8.09e-06
2013	0.63833	1.116e-05	0.6229	7.654e-06
2014	0.63502	1.028e-05	0.61865	6.909e-06
2015	0.63229	9.617e-06	0.61745	6.713e-06

4.4.3 Test for Normality Using Kolmogorov Smirnov and Shapiro-Wilks Procedures

Method 1: Kolmogorov Smirnov Normality Test

Tables 4.3 and 4.4 show the tabular presentations of the Kolmogorov Smirnov normality test.

Hypothesis test:

H_0 : The data set is normally distributed

H_1 : The data set is not normally distributed

As shown in Table 4.3, the P-values are lesser than or equal to 0.000645 across the 19 age-groups for both male and female data. From Table 4.4, the P-values are all lesser or equal to 4.348×10^{-13} across the time period for the male mortality data and 4.18×10^{-13} for the females. It is concluded that the Kolmogorov Smirnov test was significant at 1%, confirming non-normality for both gender across age-groups and time periods, respectively.

Method 2: Shapiro-Wilks Normality Test

The results for the Shapiro-Wilks normality are displayed in Tables 4.5 and 4.6.

Hypothesis test:

H_0 : The data set is normally distributed

H_1 : The data set is not normally distributed

As shown in Table 4.5, the P-values are all lesser than or equal to 0.04394 in about seventeen age-groups for the male mortality data with two age-groups (0-1 and 1-4) having P-values greater than or equal to 0.2789. For the female mortality data, the P-values are all lesser than or equal to 0.07518 in about fifteen age-groups with four age-groups (0-1, 1-4, 60-64, 75-79) having P-values greater than or equal to 0.1222. From Table 4.6, the P-values across the time period are all lesser than or equal to 2.217×10^{-5} for the male mortality data and 1.582×10^{-5} for the females. Hence, it is concluded that across the time periods, the Shapiro-Wilks test was significant at 1%, confirming non-

normality for both gender. The procedure also confirmed non-normality for males in 17 age-groups and females in 15 age-groups at 5% and 10% significance level respectively.

In summary, while the Kolmogorov-Smirnov procedure was significant at 1%, confirming non-normality with P-values lesser than or equal to 0.000645 for both gender across all the age-groups, the Shapiro-Wilks procedure confirmed non-normality with P-values lesser than or equal to 0.04394 in 17 age-groups for males and with P-values lesser than or equal to 0.07518 in 15 age-groups for females at 5% and 10% significance level, respectively. However, across the time periods, the Kolmogorov-Smirnov, Shapiro-Wilks procedures were significant at 1%, confirming non-normality with P-values lesser than or equal to $(4.35 \times 10^{-13}, 2.22 \times 10^{-5})$ and $(4.18 \times 10^{-13}, 1.58 \times 10^{-13})$ for males and females, respectively.

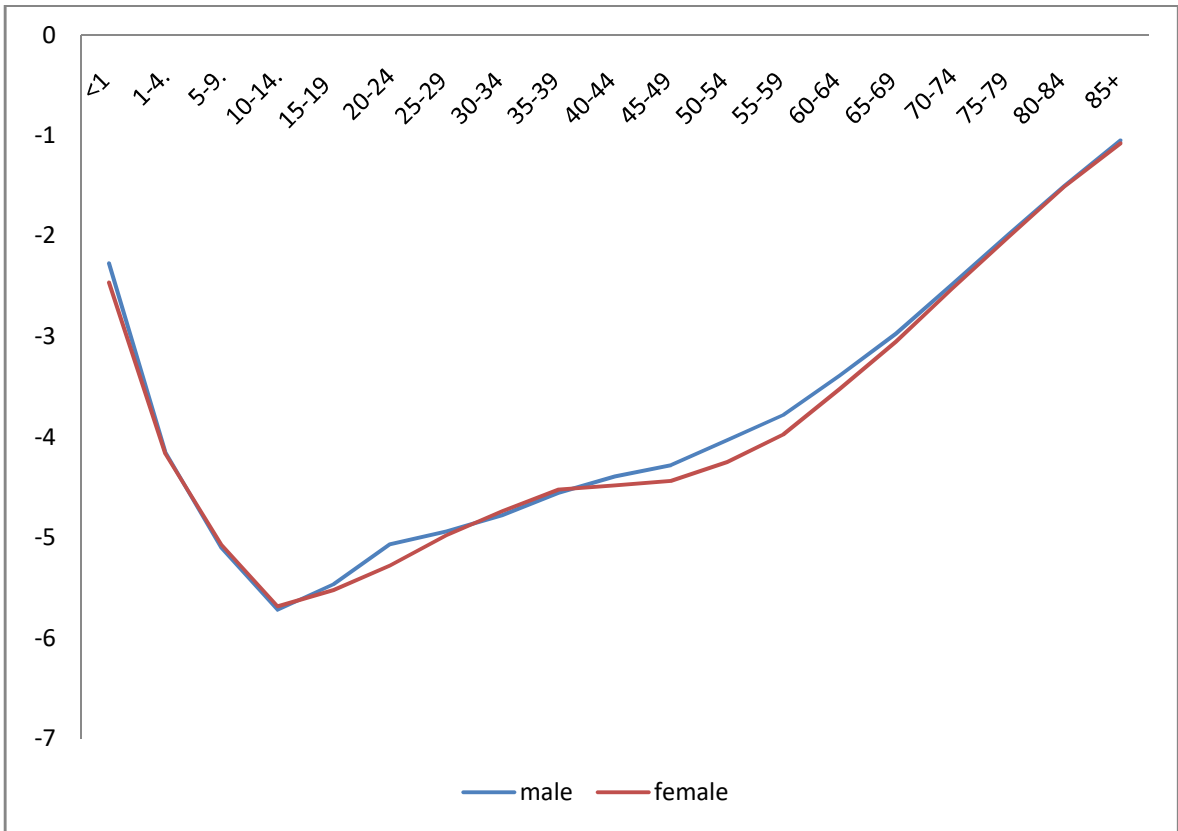


Figure 4.5: General Pattern of Mortality by Age (Lee-Carter)

4.4.4 General Pattern of Mortality by Age (\hat{a}_x) -Lee-Carter Model

a. Within each gender

As shown in Figure 4.5, \hat{a}_x values decreased with age for age-group 0-1 to age-group 10-14 with values -2.27346 to -5.71924 for the males and -2.27346 to -5.68328 for the females. Mortality pattern increased with age from age-groups 10-14 to 85⁺ within both gender. The mortality curve shows a downward trend from age-group less than 1 to 14 implying that mortality is decreasing with age for those age-groups and an upward trend from age-groups 10 upwards indicating an increasing mortality rate with age. In summary, infant mortality rate is higher than child mortality rate for both sexes while younger ages (from age-group 10-14) have a lower mortality rate than the older ages.

b. Between gender

The estimated \hat{a}_x are almost identical between males and females as shown in Figure 4.5 especially in the age-groups 1-14 and 65-85⁺. However, a closer inspection reveals that parameter \hat{a}_x is slightly higher for males than females in about 15 out of the 19 age-groups. Mortality seems to be higher for male infants and children in the age-groups 0-4, 15-29, 40-85⁺ than their female counterparts. Moreover, \hat{a}_x values are higher for females than males in the age-groups 5-9, 10-14, 30-34, 35-39 with \hat{a}_x values -5.06958, -5.68328, -4.73847 and -4.52107 respectively.

One of the reasons behind the afore-mentioned disparity is attributable to differences in exposure to risk for the two sexes. Females in the age-group 5-14 and 30-39 are prone to the risk of sexual violations, Abortion, maternal mortality and so on. Males in the age-group 15-29 in many developing countries are prone to social vices like crime, intensive labour and so on while males in the age-groups 40 and above are exposed to the risk of family responsibilities, stress, and several health issues like heart failure etc.

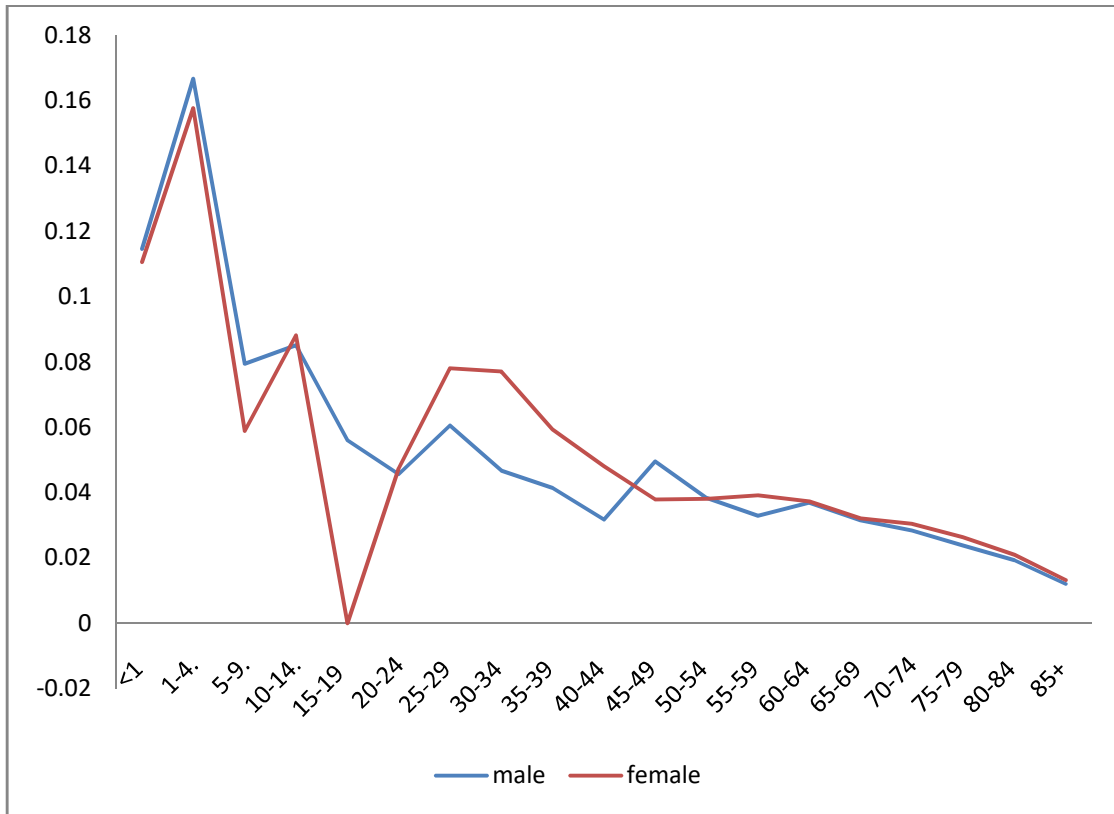


Figure 4.6: Relative pace of change in mortality by age (Lee-Carter)

4.4.5 Relative Pace of Change in Mortality by Age (\hat{b}_x) -Lee-Carter Model

a. Within each gender

Parameter \hat{b}_x is the relative pace of change in mortality by age. It describes the tendency of mortality at age x to change as the general level of mortality changes. The larger the values of \hat{b}_x at a particular age-group, the more fluctuant the mortality rate at that age-group as compared to the general level of mortality change. For developed Countries, parameter \hat{b}_x is expected to exhibit a downward trend. A look at Figure 4.6 shows that the \hat{b}_x curve exhibits a fluctuating increase and decrease across the age-groups for both males and females. This fluctuation is more noticeable in the middle age-group 25-49 for both sexes and more prominent in the male population than the females'. However for both gender, it is observed that despite the noticeable fluctuation, \hat{b}_x values are generally decreasing with age with longer interval of decrease in the age-groups 60-85⁺ for the males and 55-85⁺ for the females. For both sexes, age-group 1-4 has the largest \hat{b}_x values with 0.16664 for males and 0.15759 for females implying that mortality at age 1-4 has the tendency to change at a faster pace relative to the general level of mortality. Negative \hat{b}_x values is not present in the male estimates but observed in the females' at age-group 15-19. Lee and Carter (1992) states that \hat{b}_x values could be negative for some ages indicating that mortality at those ages tend to rise (especially when it comes in contact with negative \hat{k}_t) while falling at other ages.

b. Between gender

As shown in Figure 4.6, visible differences are observed between males and females across the age-groups. Very close ties exist at the very young (0-4) and older (60-85) age-groups. Larger discrepancies are observed in the age-groups 15-19 and 30-34. Out of the 19 age-groups studied, it is observed that \hat{b}_x values are greater for males than females in only 6 age-groups; 0-1,1-4,5-9,15-19,45-49 and 50-54. This means that the female mortality pattern exhibits a greater tendency to change relative to the general mortality pattern compared to the males in 13 age-groups which mostly comprises the child-bearing ages and adult.

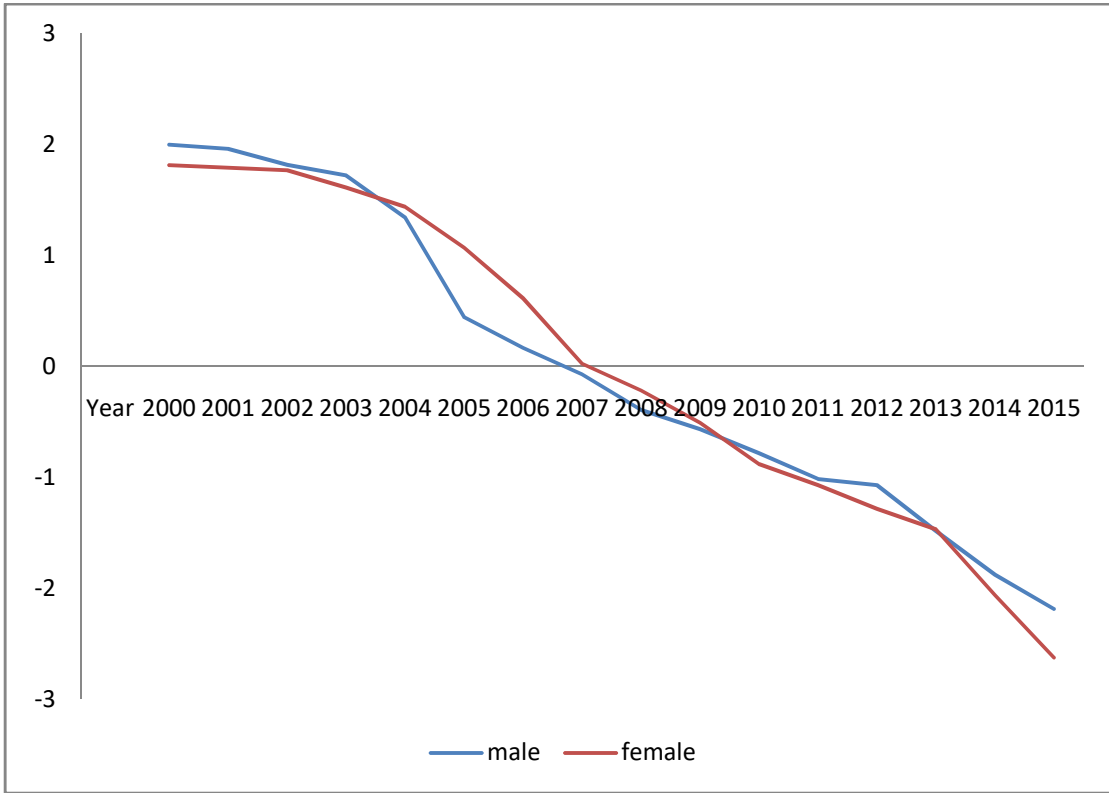


Figure 4.7: Mortality Time Trend (Lee-Carter)

4.4.6 Mortality Time Trend (\hat{k}_t) - Lee-Carter Model

a. Within each gender

The mortality time trend or index in Figure 4.7 shows a downward trend from 2000 to 2015 for both sexes. A closer look at the curve reveals that the pattern is not a straight linear trend but curve-linear in nature. Also, the rate of decline is not sporadic but gradual. Within the male population, \hat{k}_t values range from 1.9968 in year 2000 to -2.1860 in year 2015 and 1.8132 in year 2000 to -2.6268 in year 2015 for the female population. Moreover, negative \hat{k}_t values are observed for the males from 2007-2015 and 2008-2015 for the females.

The consequence of negative \hat{k}_t values for the afore-mentioned years is that mortality trend for age-groups with negative \hat{b}_x values will increase due to negative values for both \hat{b}_x and \hat{k}_t . However, age-groups with positive \hat{k}_t values and negative \hat{b}_x values will have a decreasing mortality pattern. Negative \hat{b}_x value is observed in age-group 15-19 for the females indicating that the estimated mortality for age-group 15-19 decreased from 2000-2007 and increased from 2008-2015.

b. Between gender

A comparison between the two sexes highlights the fact that \hat{k}_t values are greater for males than females in 9 out of 16 years; 2000-2003, 2010-2012, 2014-2015 indicating a higher mortality index for males than females for those years. The females have higher \hat{k}_t values in the years 2004-2009 and 2013. The two sexes have close ties in the years 2002, 2009 and 2013. However, larger discrepancies are observed in the year 2005, 2006 and 2015.

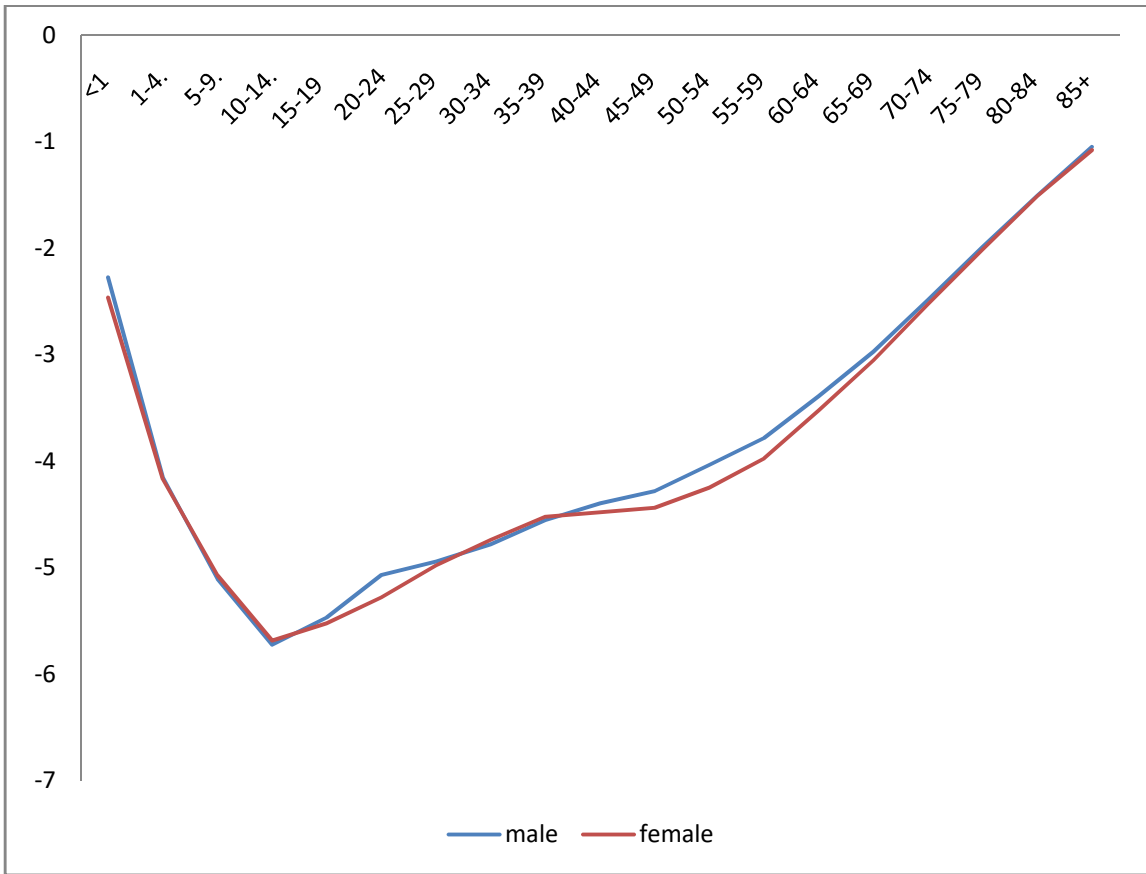


Figure 4.8: General pattern of mortality by age (Brouhns)

4.4.7 General Pattern of Mortality by Age (\hat{a}_x) -Brouhns Model

a. Within each gender

The graphical display of parameter \hat{a}_x under the Brouhns model is shown in Figure 4.8. Just as in the Lee and Carter (1992), within each gender, \hat{a}_x values decreased with age for age-group 0-1 to age-group 10-14 with values -2.27336 to -5.72147 for the males and -2.46189 to -5.68328 for the females. Mortality pattern increased with age from age-group 10-14 to 85⁺ with \hat{a}_x values moving from -5.72147 to -1.049931 for males and -5.68328 to -1.078266 for females. As further shown in Figure 4.8, the mortality curve shows a downward trend from age-groups 0⁺ to 14 implying that mortality is decreasing with age for those age-groups and an upward trend from age-groups 10 to 85⁺ indicating an increasing mortality rate with age. Infant mortality rate is higher than child mortality rate for both sexes while younger ages (from age-group 10-14) have a lower mortality rate than the older ages.

b. Between gender

The estimated \hat{a}_x are almost identical between males and females especially in the age-groups 1-14 and 65-85⁺. Mortality is higher for male infants and children in the age-groups 0-4, 15-29, 40-85⁺ than their female counterparts. Moreover, \hat{a}_x values are higher for females than males in the age-groups 5-9, 10-14, 30-34 and 35-39. In summary, the general pattern of mortality by age is slightly higher for males than females in about 15 out of the 19 age-groups.

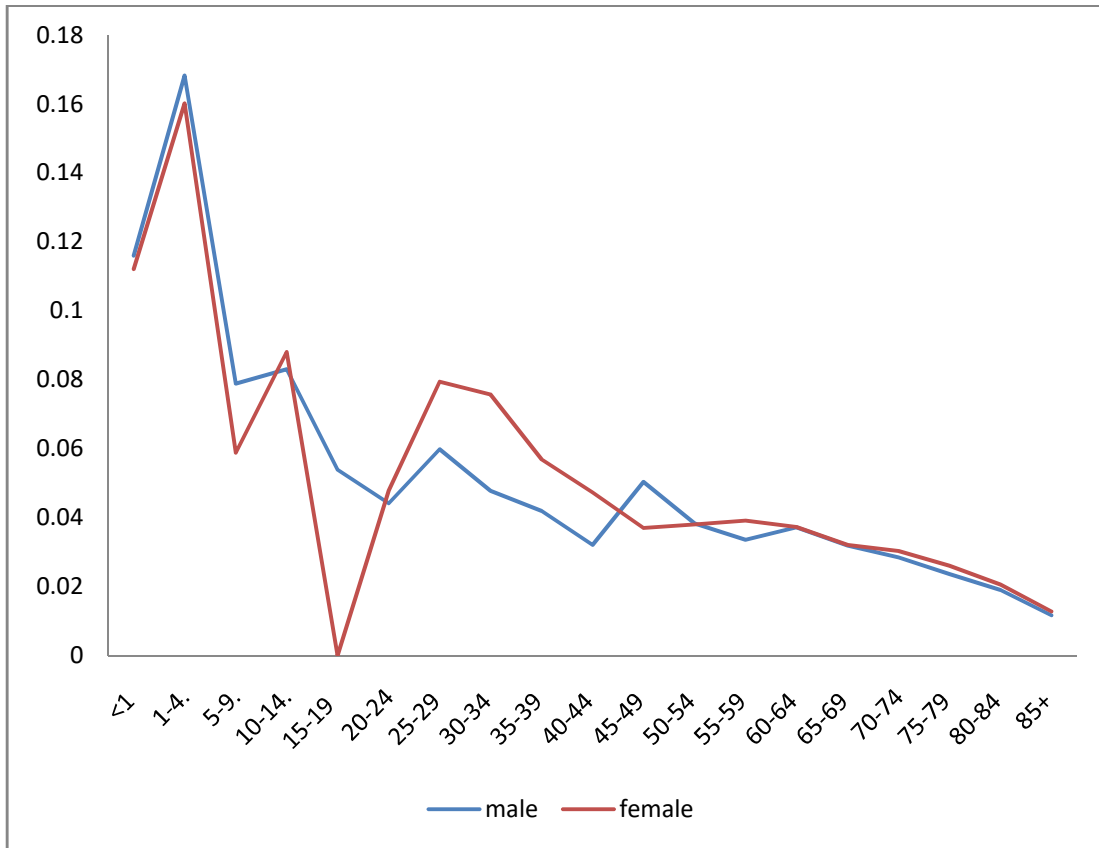


Figure 4.9: Relative pace of change in mortality by age (Brouhns)

4.4.8 Relative Pace of Change in Mortality by Age (\hat{b}_x) – Brouhns Model

a. Within each gender

The trend curve of \hat{b}_x seems not well-behaved as a fluctuating increase and decrease is observed in figure 4.9 within each gender. This fluctuation is more prominent in the males than the females'. A long steep slope downwards is observed between age-groups 1-4 and 5-9 and between 10-14 and 15-19 for the females. For the males, a steep slope downwards exists between age-groups 1-4 and 5-9. The highest peak for males and females is found at age-group 1-4 with 0.168243 and 0.160198 respectively. This implies that age-group 1-4 will have a more fluctuant mortality pattern than any other age-group for both gender. The deepest depression is observed at age-group 15-19 for the females and 85+ for the males meaning that the aforementioned age-groups will have the least fluctuant mortality pattern. Negative values are not present in both gender.

b. Between gender

Between the two sexes, closer ties exist at ages 0-4 and 60-85 years. Larger discrepancies are observed at age-groups 15-19 and 25-39 years. Longer intervals of decreasing trend are seen in the female curve compared to the males as shown in Figure 4.9. \hat{b}_x values are greater for females than males in 13 age-groups. This implies that with respect to the Brouhns model, females generally have a more fluctuant mortality than males.

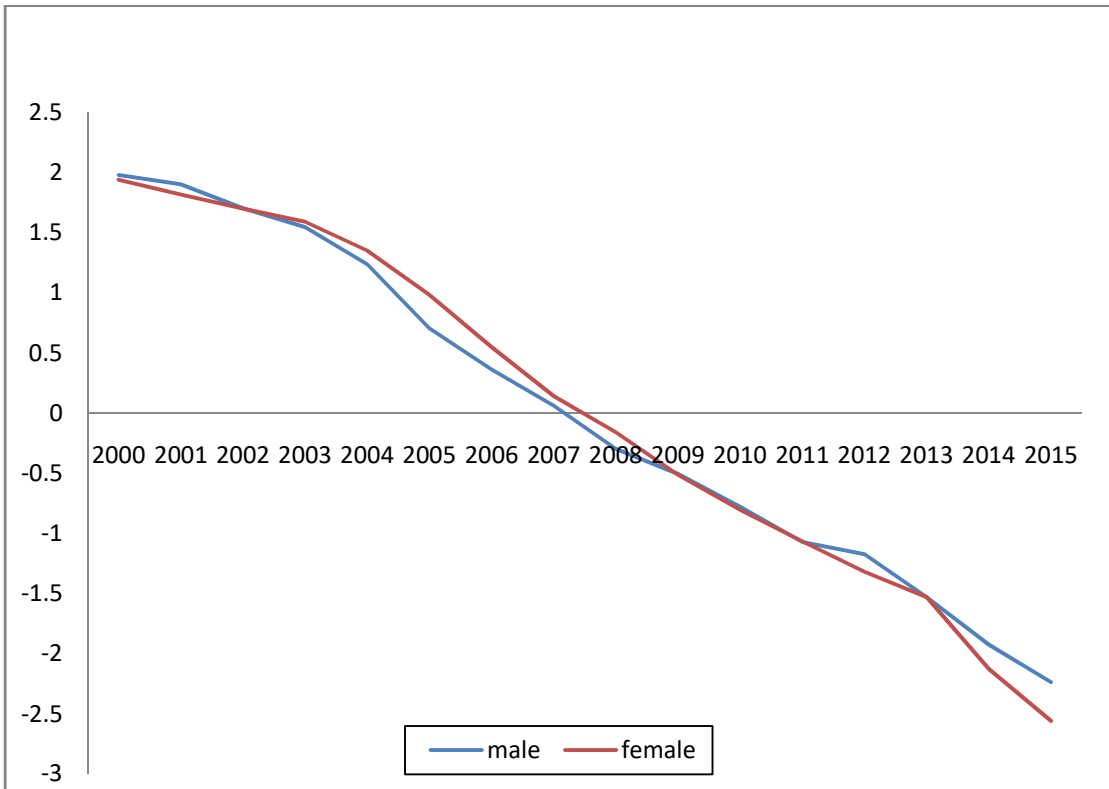


Figure 4.10: mortality time trend (Brouhns)

4.4.9 Mortality Time Trend (\hat{k}_t) – Brouhns Model

a. Within each gender

The mortality index in Figure 4.10 shows a downward trend for both sexes. \hat{k}_t values for the male population ranges from 1.981958 in year 2000 to -2.236734 in year 2015. For the females, it ranges from 1.940734 in year 2000 to -2.559394 in year 2015. Moreover, negative \hat{k}_t values are observed for both sexes from 2008 to 2015. The implication of negative \hat{k}_t values has been discussed earlier. This implies that the estimated age-specific mortality will have a decreasing trend from 2008 to 2015 since the Brouhns model does not have negative \hat{b}_x values for both gender.

b. Between gender

A comparison between the two sexes shows that \hat{k}_t values are greater for males than females in 9 out of 16 years. Females have larger \hat{k}_t values in the years 2003-2008 and 2011. The difference in \hat{k}_t values between year 2000 and 2015 is 4.218692 for males and 4.500128 for females. This implies that the female folk experienced a faster decline in mortality compared to the males during the time period 2000-2015.

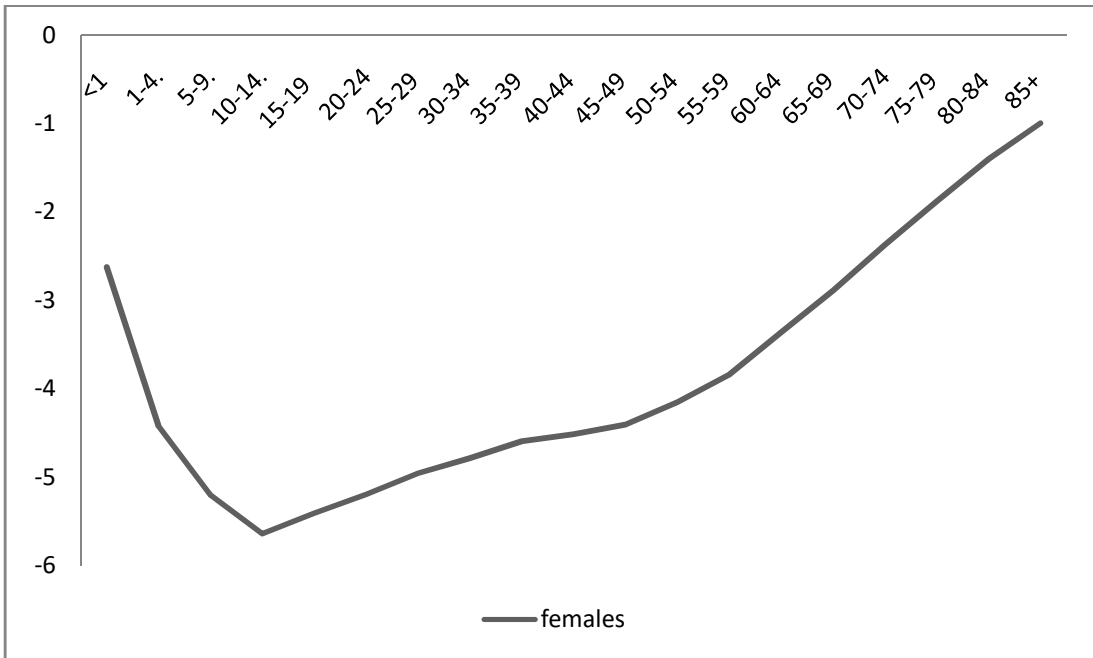


Figure 4.11: General pattern of mortality by age (Renshaw Haberman)

4.4.10 General Pattern of Mortality by Age (\hat{a}_x) - Renshaw Haberman Model

From Figure 4.11, the general pattern of mortality decreased from ages < 1 to 10-14, with \hat{a}_x values ranging from -2.62376 to -5.63828. A gradual increase is observed from ages 10-14 to 85+. The mortality curve shown in Figure 4.11 also illustrates the fact that infant mortality rate is higher than child mortality rate which is in turn higher than mortality at ages 5-14. Younger ages (from age-group 10-14) have a lower mortality rate than the older ages.

Furthermore, the Renshaw–Haberman model did not converge for the male mortality data. This result agrees with Currie (2016) who stated that the convergence for the Renshaw –Haberman model could be problematic. He however suggested the use of the Lee-Carter model parameters as starting values for the Renshaw–Haberman model. This approach was utilized in this study but it didn't work for the male mortality data.

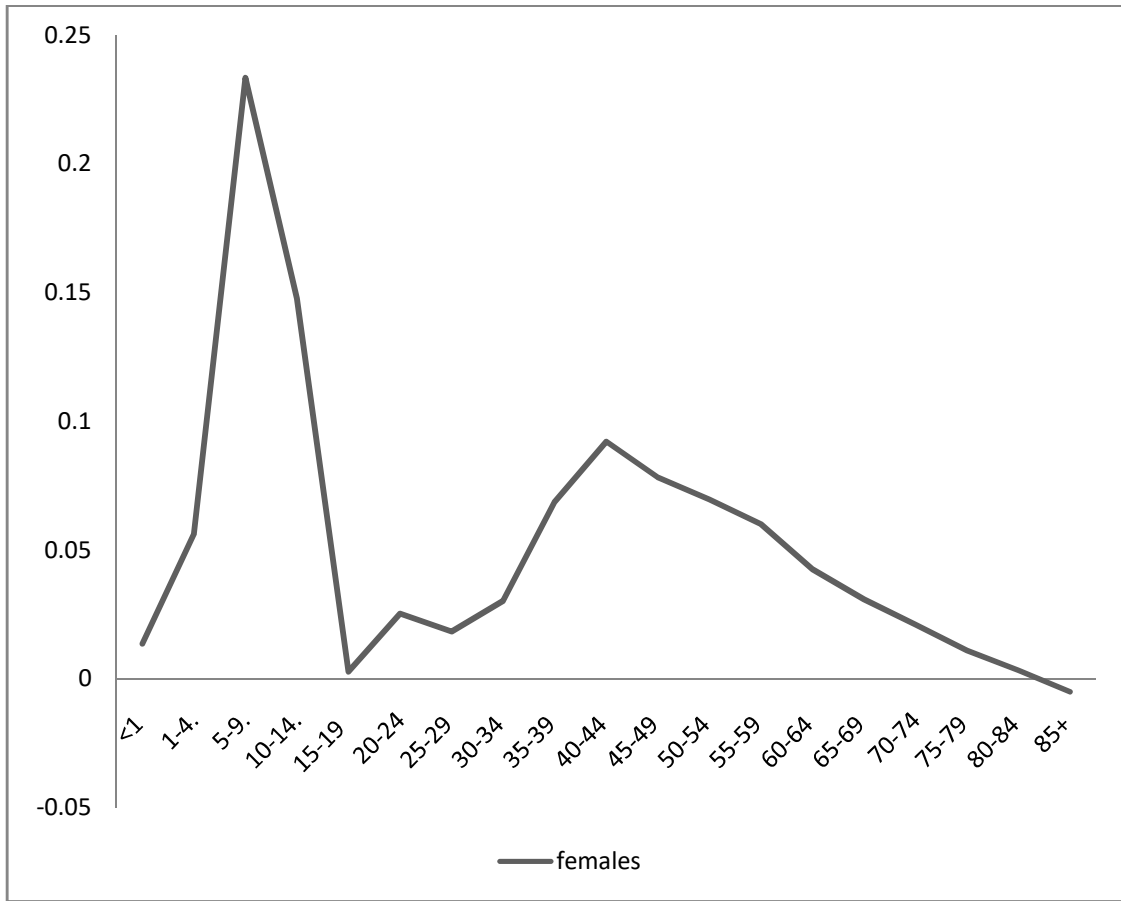


Figure 4.12: Relative pace of change in mortality by age (Renshaw Haberman)

4.4.11 Relative Pace of Change in Mortality by Age (\hat{b}_x) - Renshaw Haberman Model

The graphical illustration for parameter \hat{b}_x is displayed in Figure 4.12. A fluctuating increase and decrease is seen in the \hat{b}_x curve. For some age-groups, \hat{b}_x values are actually decreasing with longer intervals of decrease observed between age-group 40-85⁺. Age-group 5-9 has the highest fluctuant mortality while age-group 85⁺ has the least fluctuant mortality. Negative \hat{b}_x value is observed only at age-group 85⁺.

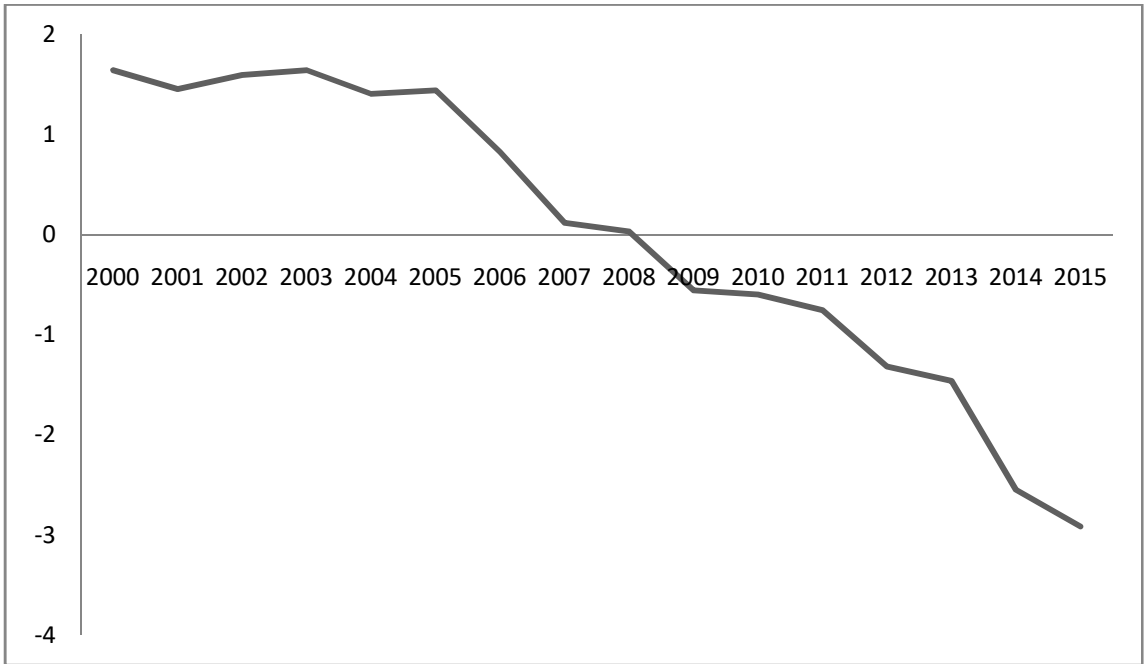


Figure 4.13: mortality time trend (Renshaw-Haberman)

4.4.12 Mortality Time Trend (\hat{k}_t) - Renshaw-Haberman Model

The mortality index for the Renshaw-Haberman model as presented in Figure 4.13 is generally decreasing with time. A fluctuating increase and decrease is seen between year 2000 and 2005. A close look at Figure 4.13 shows that the mortality trend is not exactly linear. Positive \hat{k}_t values range from 1.641282 in 2000 to 0.03022 in 2008. Negative values are observed from the year 2009-2015. Since negative \hat{b}_x value is observed only in age-group 85^+ for the Renshaw-Haberman model, it indicates that the estimated mortality for age-group 85^+ will decrease from 2000-2008 and increase from 2009-2015.

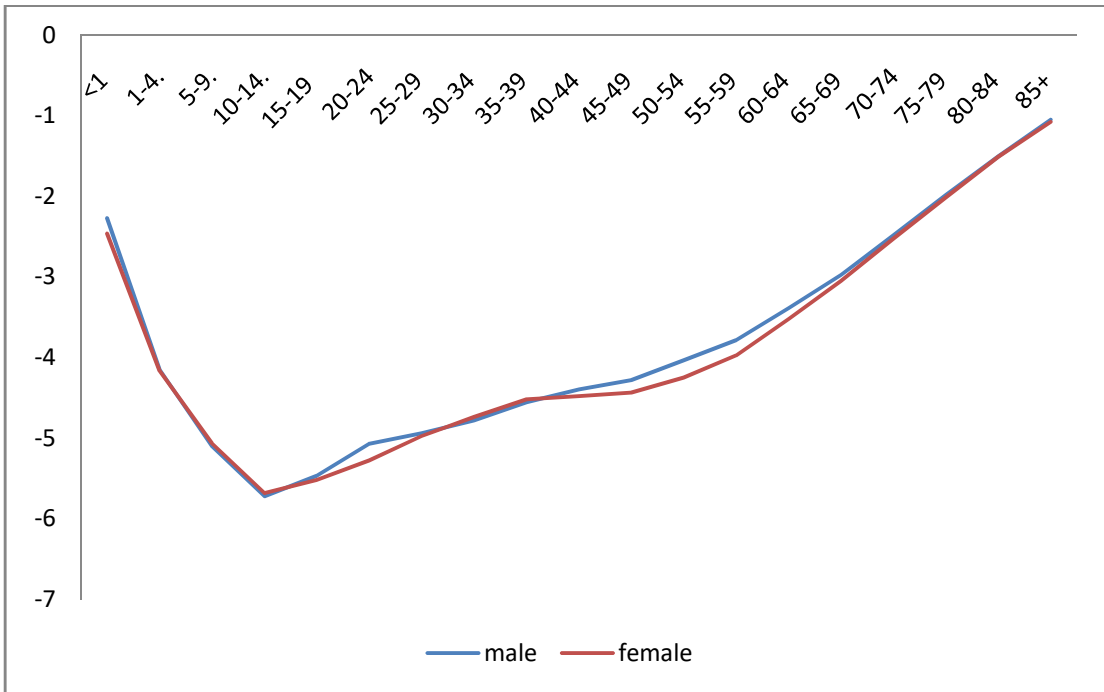


Figure 4.14: General pattern of mortality by age (Gamma-normal Lee-Carter model)

4.4.13 General Pattern of Mortality by Age (\hat{a}_x) - Gamma-Normal Lee-Carter Model

Within each gender

As shown in Tables 4.7 and 4.8, \hat{a}_x values decreased with age for age-group 0-1 to age-group 10-14 with values -2.27346 to -5.71924 for the males and -2.27346 to -5.68328 for the females. Mortality pattern increased with age from age-groups 10-14 to 85⁺ within both gender. The mortality curve in Figure 4.14 shows a downward trend from age-group < 1 to 14 implying that mortality is decreasing with age for those age-groups and an upward trend from age-groups 10 upwards indicating an increasing mortality rate with age. Infant mortality rate is higher than child mortality rate for both sexes while younger ages (from age-group 10-14) have a lower mortality rate than the older ages.

Between gender

From Figure 4.14, the estimated \hat{a}_x are almost identical between males and females especially in the age-groups 1-14 and 65-85⁺. However, taking a closer look reveals the fact that parameter \hat{a}_x is slightly higher for males than females in about 15 out of the 19 age-groups. Mortality is higher for male infants and children in the age-groups 0-4, 15-29, 40-85⁺ than their female counterparts. Moreover, \hat{a}_x values are higher for females than males in the age-groups 5-9, 10-14, 30-34, 35-39. Possible reasons that could be responsible for the afore-mentioned disparity has been discussed earlier.

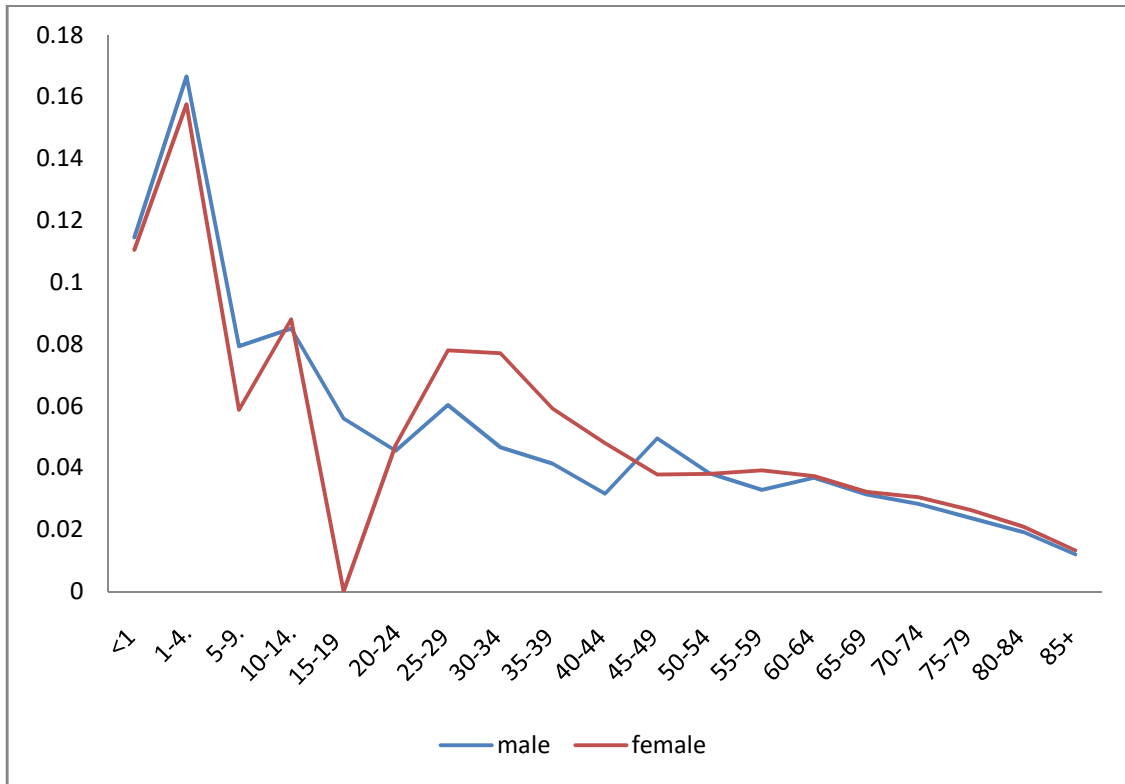


Figure 4.15: Relative pace of change in mortality by age (Gamma-normal Lee-Carter model)

4.4.14 Relative Pace of Change in Mortality by Age (\hat{b}_x)- Gamma-Normal Lee-Carter Model

Within each gender

As stated earlier, Parameter \hat{b}_x describes the tendency of mortality at age x to change as the general level of mortality changes. From Figure 4.15, the \hat{b}_x curve exhibits a fluctuating increase and decrease across the age-groups for both male and female population. This fluctuation is more noticeable in the middle age-group 25-49 for both sexes and more prominent in the male population than the females'. From tables 4.9 and 4.10, age-group 1-4 has the largest \hat{b}_x values for both gender, with 0.16664658 for males and 0.15761933 for females respectively. This means that mortality at age-group 1-4 has the most fluctuant mortality pattern relative to the general level of mortality. Age-group 85⁺ has the least fluctuant mortality pattern for the males while for the females; it is age-group 15-19. Negative \hat{b}_x value is not present in both gender.

Between gender

As shown in Figure 4.15, visible differences are observed between males and females across the age-groups. Close ties are seen at age-group 0-4 and 60-85 years. Larger discrepancies are observed in the age-groups 15-19 and 30-34. Out of the 19 age-groups, it is observed that \hat{b}_x values are greater for females than males in 13 age-groups which comprises the child-bearing ages and adult females. This implies that the female mortality pattern exhibits a greater tendency to fluctuate compared to the males.

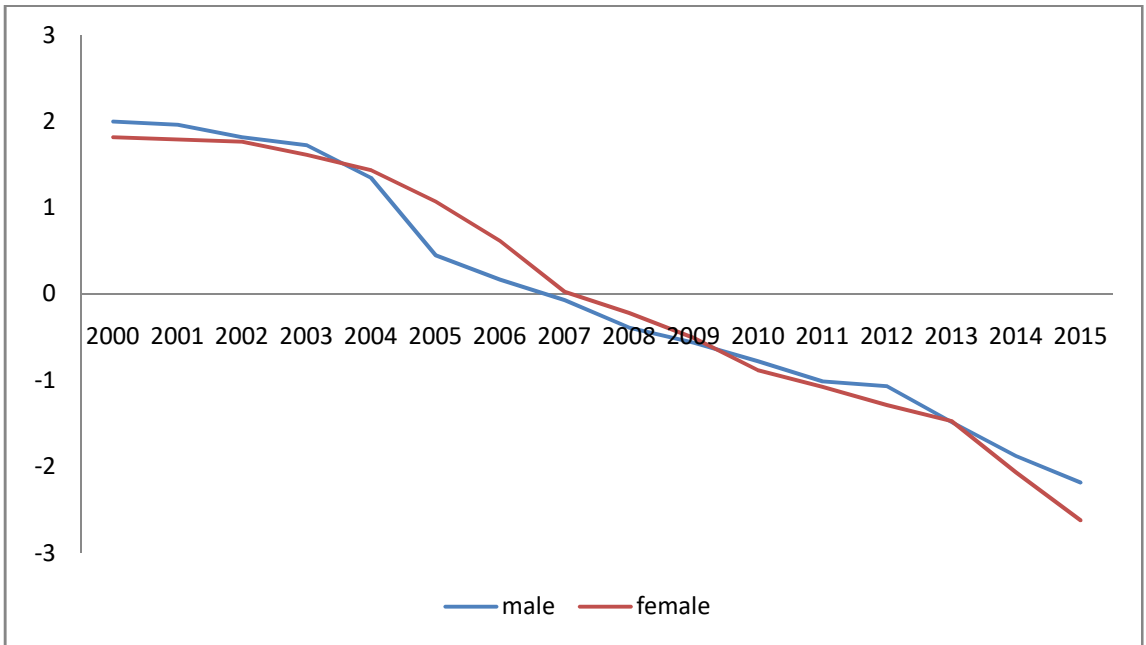


Figure 4.16: mortality time trend (Gamma-normal Lee-Carter model)

4.4.15 Mortality Time Trend(\hat{k}_t)- Gamma-Normal Lee-Carter Model

Within each gender

\hat{k}_t values for females and males are displayed in Tables 4.11 and 4.12. The mortality index in Figure 4.16 shows a downward trend from 2000 to 2015 for both sexes. Within the male population, \hat{k}_t values range from 1.9962 in year 2000 to -2.18536 in year 2015 and 1.81570 in year 2000 to -2.6202 in year 2015 for the female population. Moreover, negative \hat{k}_t values are observed from 2007-2015 for the males and from 2008-2015 for the females.

The implication of negative \hat{k}_t values as discussed by Hanna (2007) is that mortality pattern for age-groups with negative \hat{b}_x values will increase due to negative values for both \hat{b}_x and \hat{k}_t . However, age-groups with positive \hat{k}_t values and negative \hat{b}_x values will have a decreasing mortality pattern will increase due to negative values for both \hat{b}_x and \hat{k}_t while the other age-groups are experiencing a decreasing mortality pattern. The consequence is having mortality curves crossing each other. A merit of the proposed model is that negative \hat{b}_x value is not present in both the male and female data.

Between gender

A comparison between the two sexes shows that \hat{k}_t values are greater for males than females in 9 out of 16 years; 2000-2003, 2010-2012 and 2014-2015 indicating a higher mortality index for males than females for those years. The two sexes have close ties in the years 2002, 2009 and 2013. However, larger discrepancies are observed in the year 2005, 2006 and 2015.

Table 4.7: Comparison of the MLEs of Parameter α_x under the Four Models for Nigeria Female Mortality Data

	LC	BR	RH	GNLC
<1	-2.461892	-2.461851	-2.6237619	-2.461892
1-4	-4.164983	-4.165188	-4.4210464	-4.164983
5-9	-5.069584	-5.070319	-5.2007739	-5.069584
10-14	-5.683282	-5.685167	-5.6382799	-5.683282
15-19	-5.521461	-5.521461	-5.4084323	-5.521461
20-24	-5.278079	-5.280118	-5.1915108	-5.278079
25-29	-4.977923	-4.979045	-4.9542377	-4.977923
30-34	-4.738466	-4.738359	-4.7842194	-4.738466
35-39	-4.521071	-4.519965	-4.5915937	-4.521071
40-44	-4.479415	-4.478644	-4.5137096	-4.479415
45-49	-4.435550	-4.435540	-4.4027834	-4.435550
50-54	-4.248468	-4.248570	-4.1516312	-4.248468
55-59	-3.974881	-3.974918	-3.8380744	-3.974881
60-64	-3.522756	-3.522780	-3.3607945	-3.522756
65-69	-3.052064	-3.052062	-2.8875846	-3.052064
70-74	-2.525927	-2.525934	-2.3698408	-2.525927
75-79	-2.013919	-2.013953	-1.8768102	-2.013919
80-84	-1.512042	-1.512117	-1.4006350	-1.512042
85+	-1.078266	-1.078415	-0.9954274	-1.078266

Table 4.8: Comparison of the MLEs of Parameter α_x under the Four Models for Nigeria Male Mortality Data

	LC	BR	RH	GNLC
<1	-2.273458	-2.273360	-	-2.273458
1-4	-4.149711	-4.149656	-	-4.149711
5-9	-5.102009	-5.102974	-	-5.102009
10-14	-5.719242	-5.721468	-	-5.719242
15-19	-5.465675	-5.467189	-	-5.465675
20-24	-5.067824	-5.068268	-	-5.067824
25-29	-4.939385	-4.940139	-	-4.939385
30-34	-4.777768	-4.777951	-	-4.777768
35-39	-4.552705	-4.552664	-	-4.552705
40-44	-4.393268	-4.393211	-	-4.393268
45-49	-4.280251	-4.280244	-	-4.280251
50-54	-4.032856	-4.032906	-	-4.032856
55-59	-3.781564	-3.781608	-	-3.781564
60-64	-3.388248	-3.388273	-	-3.388248
65-69	-2.972004	-2.972015	-	-2.972004
70-74	-2.484435	-2.484459	-	-2.484435
75-79	-1.987872	-1.987935	-	-1.987872
80-84	-1.504879	-1.504973	-	-1.504879
85+	-1.049787	-1.049931	-	-1.049787

4.4.16 Comparison of MLEs of Parameter α_x across all the Models

Tables 4.7 and 4.8, respectively show the maximum likelihood estimates of Parameter $\hat{\alpha}_x$ for female and male mortality data across the models. Parameter $\hat{\alpha}_x$ represents the general pattern of mortality by age. A similarity is noticed in the general pattern of mortality by age for all the models under consideration. Generally, the results show that $\hat{\alpha}_x$ values decreased with age for age-groups 0-1 to age-groups 10-14, but increased with age for age-groups 10-14 to 85⁺. The younger ages (from age-group 10-14) have a lower mortality rate than the older ages. As expected of a developing Country, there is a clear indication that Infant mortality rate is higher than child mortality rate which is in turn greater than mortality rate at ages 5-14 years for both gender. Across the models, mortality is greater for males than females except for the child-bearing age-groups.

Table 4.9: Comparison of the MLEs of Parameter b_x under the Four Models for Nigeria Female Mortality Data

	LC	BR	RH	GNLC
<1	1.104649e-01	1.121297e-01	0.013588631	0.1105305631
1-4	1.575868e-01	1.601981e-01	0.056164504	0.1576193323
5-9	5.882967e-02	5.888257e-02	0.233382645	0.0587414376
10-14	8.805851e-02	8.803642e-02	0.147707367	0.0880098908
15-19	-1.394130e-17	1.638890e-16	0.002829039	0.0000267991
20-24	4.742123e-02	4.790192e-02	0.025341864	0.0474709506
25-29	7.796972e-02	7.943062e-02	0.018282374	0.0779606398
30-34	7.702937e-02	7.574334e-02	0.030265294	0.0770945235
35-39	5.927341e-02	5.686279e-02	0.068614830	0.0591113584
40-44	4.805914e-02	4.730043e-02	0.092077470	0.0479492269
45-49	3.783256e-02	3.706698e-02	0.078183912	0.0378279402
50-54	3.807446e-02	3.802905e-02	0.069666548	0.0380193337
55-59	3.910521e-02	3.916809e-02	0.060073108	0.0391321893
60-64	3.730551e-02	3.732741e-02	0.042542319	0.0373319652
65-69	3.213565e-02	3.211966e-02	0.030853607	0.0321785613
70-74	3.041920e-02	3.031044e-02	0.021081115	0.0304437756
75-79	2.631621e-02	2.608956e-02	0.011018519	0.0263550986
80-84	2.089943e-02	2.060288e-02	0.003354608	0.0209355434
85+	1.321902e-02	1.279999e-02	-0.005027751	0.0132608706

Table 4.10: Comparison of the MLEs of Parameter b_x under the Four Models for Male Mortality Data

	LC	BR	RH	GNLC
<1	0.11453640	0.11596226	-	0.11455018
1-4	0.16664015	0.16824386	-	0.16664658
5-9	0.07942992	0.07884561	-	0.07935779
10-14	0.08506221	0.08300317	-	0.08505607
15-19	0.05595923	0.05385829	-	0.05605301
20-24	0.04557947	0.04417464	-	0.04557819
25-29	0.06045002	0.05983338	-	0.06038905
30-34	0.04671098	0.04777042	-	0.04668092
35-39	0.04136255	0.04196989	-	0.04137113
40-44	0.03166568	0.03211921	-	0.03165884
45-49	0.04955134	0.05037755	-	0.04955335
50-54	0.03834928	0.03833411	-	0.03834905
55-59	0.03290210	0.03356607	-	0.03290987
60-64	0.03687730	0.03720575	-	0.03688395
65-69	0.03149564	0.03189516	-	0.03150348
70-74	0.02833183	0.02845730	-	0.02833892
75-79	0.02381239	0.02368113	-	0.02382046
80-84	0.01926668	0.01903534	-	0.01927410
85+	0.01201684	0.01166686	-	0.01202506

4.4.17 Comparison of MLEs of Parameter b_x across all the Models

The MLEs of the Parameter \hat{b}_x for the models are presented in Tables 4.9 and 4.10. Parameter \hat{b}_x is the relative pace of change in mortality by age. It describes the tendency of mortality at age x to change as the general level of mortality changes. \hat{b}_x values exhibit a fluctuating increase and decrease across all the models for both male and female population. The Renshaw-Haberman model differs a bit from other models. However, closer ties exist between the Lee-Carter and the proposed model due to the similarity in their distribution framework. Across the models, age-group 1-4 has the most fluctuant mortality pattern since it has the highest \hat{b}_x value while females have a more fluctuant mortality than males. Negative \hat{b}_x value is observed in both the Lee-Carter and Renshaw-Haberman models but absent in the Brouhns and Gamma-normal Lee-Carter models. Lee and Carter (1992) states that \hat{b}_x values could be negative for some ages indicating that mortality at those ages tend to rise (especially when it comes in contact with negative \hat{k}_t) while falling at other ages.

Table 4.11: Comparison of the MLEs of Parameter k_t under the Four Models for Nigeria Female Mortality Data

Year	LC	BR	RH	GNLC	RH (l_{t-x})
2000	1.81326912	1.940734	1.641282	1.81569534	0.40313861
2001	1.78974088	1.817955	1.452249	1.78923244	0.37291575
2002	1.76584276	1.702361	1.592320	1.76397174	0.32659933
2003	1.61194797	1.590959	1.639793	1.61028575	0.29941709
2004	1.43708036	1.353757	1.404944	1.43580639	0.27456479
2005	1.07020184	0.9865951	1.441419	1.07014721	0.22543857
2006	0.61382262	0.5478901	0.823625	0.61380454	0.19659732
2007	0.02544335	0.1428722	0.117060	0.02643208	0.18240987
2008	-0.21730417	-0.1596062	0.030222	-0.21710828	0.13673352
2009	-0.50929413	-0.5121536	-0.555158	-0.50941185	0.10077095
2010	-0.88255690	-0.8050801	-0.597509	-0.88339855	0.08062507
2011	-1.07274024	-1.067653	-0.754415	-1.07429290	0.04754275
2012	-1.28517343	-1.319157	-1.315529	-1.28576824	0.04237538
2013	-1.47015979	-1.530603	-1.461180	-1.47183060	-0.0156427
2014	-2.06331801	-2.129477	-2.545847	-2.06338316	-0.0298911
2015	-2.62680223	-2.559394	-2.913276	-2.62018189	-0.0547512

Table 4.12: Comparison of the MLEs of Parameter k_t under the Four Models for Nigeria Male Mortality Data

	LC	BR	RH	GNLC
2000	1.99682611	1.981958	-	1.99617121
2001	1.95916968	1.904966	-	1.95844299
2002	1.81575738	1.704728	-	1.81515433
2003	1.72101574	1.547027	-	1.72115585
2004	1.34166655	1.239463	-	1.34233038
2005	0.44377420	0.7081821	-	0.44616172
2006	0.16556124	0.3634139	-	0.16667418
2007	-0.07282974	0.0623292	-	-0.07259202
2008	-0.39214449	-0.2969154	-	-0.39259812
2009	-0.56588780	-0.5029136	-	-0.56643923
2010	-0.78093251	-0.7776134	-	-0.78198529
2011	-1.01470182	-1.070649	-	-1.01560960
2012	-1.06997913	-1.173066	-	-1.07076482
2013	-1.48331496	-1.529730	-	-1.48336704
2014	-1.87795029	-1.924445	-	-1.87737528
2015	-2.18603016	-2.236734	-	-2.18535928

4.4.18 Comparison of MLEs of Parameter k_t across all the Models

Parameter \hat{k}_t

The MLEs of the mortality trend of all the models are shown in Tables 4.11 and 4.12. Parameter \hat{k}_t is known as the mortality index. According to Lee and Carter (1992), when \hat{k}_t is linear in time, mortality at each age changes at its own constant exponential rate. For the male population, closer ties are noticed between the Lee-Carter and the proposed. However for the females, all the models are closely knitted except for the Renshaw-Haberman. Across all the models, the females have a lower mortality index compared to the males.

Parameter \hat{l}_{t-x}

The Parameter \hat{l}_{t-x} is only present in the Renshaw-Haberman model. It represents the impact of the cohort history on mortality trend. The \hat{l}_{t-x} trend is smoother than the \hat{k}_t trend. From Table 4.11, the highest peak is observed at year 2000 with 0.40313861 while its lowest value is observed in year 2015. Negative \hat{l}_{t-x} are observed in from 2013-2015.

Table 4.13: Estimates of Parameters α under the Gamma-Normal Lee-Carter Model

	α
Male	0.9501383
Female	0.9722425

4.4.19 Estimates of Parameters α under the Gamma-Normal Lee-Carter Model

Parameter $\hat{\alpha}$ in the proposed Gamma-Normal Lee-Carter model denotes the shape parameter. It gives us an idea of the degree of asymmetry of the distribution. Its demographic representation implies additional effects that might exist and could act constantly across age and time in human mortality experience especially in developing Countries. $\hat{\alpha}$ is 0.9501383 and 0.9722425 for the male and female population respectively. This means that heterogeneity is higher than and more important for the females than the males. This parameter is absent in the Lee-Carter model and the variants considered in this study.

Table 4.14: Measures of Goodness of Fit for Male and Female when $h = 16$

Male					
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	HQC
GNLC	624.9016	-1147.8032	-958.23479	-1126.75558	-1071.9714
LC	625.9756	-1147.9512	-954.66576	-1125.99104	-1070.6325
BR	-1934.92	3973.84	4167.12544	3995.80016	4051.1587
RH	-20964355	41928812	41929001.6	41928941.9	41929285.3
Female					
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	HQC
GNLC	579.7472	-1057.4944	-867.92599	-1036.44678	-981.6626
LC	580.9978	-1057.9956	-864.71016	-1036.03544	-980.6769
BR	-2142.02	4388.04	4581.32544	4410.00016	4465.3587
RH	-19407.37	39144.74	39758.0496	39274.653	39390.0782

Table 4.15: Measures of Goodness of Fit results for the mortality model when $h = 16$

MALE

	MSE	RMSE
GNLC	0.0009590635	0.03096875
LC	0.0009590584	0.03096867
BR	0.001001066	0.03163963
RH	0.0002772892	0.01665200

FEMALE

	MSE	RMSE
GNLC	0.001289327	0.03590719
LC	0.001289305	0.03590689
BR	0.001281768	0.03580179
RH	0.0004074057	0.02018429

4.4.20 Measures of Goodness of Fit for Male and Female when $h = 16$

An evaluation of model performance is based on the values of the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Corrected Akaike Information Criterion (CAIC) and the Hannan Quinine information criterion.

The results of the BIC, CAIC and HQC for both sexes as indicated in table 4.14 confirm that the proposed model outperformed the Lee-Carter model while the Renshaw-Haberman model did better than the Brouhns with respect to information loss. Comparison could not be made using the information Criterion across the models because of the different error structures as buttressed by Burnham and Anderson (2002). Furthermore, Table 4.15 points out that for both sexes, the Renshaw-Haberman model has the lowest root mean square error compared to other models with very little disparity between the Lee-Carter model and the Gamma-normal lee-Carter model.

Table 4.16: Forecast Estimates for the Lee-Carter Model (Male)

	Mean	Lower		Upper	
		80%	95%	80%	95%
2016	-2.4649	-2.73163	-2.87283	-2.19816	-2.05695
2017	-2.7437	-3.12097	-3.32066	-2.36653	-2.16684
2018	-3.0226	-3.48461	-3.72917	-2.56061	-2.31604
2019	-3.3015	-3.83493	-4.11734	-2.76799	-2.48559
2020	-3.5803	-4.17676	-4.49249	-2.98388	-2.66815
2021	-3.8592	-4.51254	-4.85841	-3.20581	-2.85994
2022	-4.138	-4.84375	-5.21733	-3.43232	-3.05874
2023	-4.4169	-5.17133	-5.57071	-3.66245	-3.26307
2024	-4.6957	-5.49596	-5.91956	-3.89554	-3.47194
2025	-4.9746	-5.8181	-6.26462	-4.13111	-3.6846
2026	-5.2535	-6.13813	-6.60644	-4.3688	-3.90049
2027	-5.5323	-6.45632	-6.94546	-4.60832	-4.11919
2028	-5.8112	-6.77291	-7.28202	-4.8494	-4.34034
2029	-6.0900	-7.08807	-7.61639	-5.092	-4.56368
2030	-6.3689	-7.40195	-7.94882	-5.3358	-4.78896
2031	-6.6477	-7.71469	-8.2795	-5.5808	-5.0160
2032	-6.9266	-8.02639	-8.60857	-5.8268	-5.24464
2033	-7.2055	-8.33713	-8.93619	-6.0738	-5.47474
2034	-7.4843	-8.64699	-9.26248	-6.3216	-5.70617
2035	-7.7632	-8.95606	-9.58753	-6.5703	-5.93883

Table 4.17: Forecast Estimates for the Lee-Carter Model (Female)

	Mean	Lower		Upper	
		80%	95%	80%	95%
2016	-3.19019	-3.4096	-3.52576	-2.97077	-2.85462
2017	-3.75363	-4.24426	-4.50399	-3.263	-3.00327
2018	-4.31707	-5.13806	-5.57266	-3.49609	-3.06149
2019	-4.88052	-6.08231	-6.7185	-3.67872	-3.04253
2020	-5.44396	-7.0712	-7.93261	-3.81672	-2.95531
2021	-6.0074	-8.1005	-9.20853	-3.9143	-2.80628
2022	-6.57084	-9.16702	-10.5414	-3.97467	-2.60034
2023	-7.13429	-10.2682	-11.9272	-4.00039	-2.341401
2024	-7.69773	-11.4019	-13.3628	-3.99355	-2.03267
2025	-8.26117	-12.5664	-14.8455	-3.9559	-1.67683
2026	-8.82462	-13.7603	-16.3731	-3.88896	-1.27618
2027	-9.38806	-14.9821	-17.9434	-3.794	-0.83269
2028	-9.9515	-16.2308	-19.5549	-3.6722	-0.34813
2029	-10.5149	-17.5054	-21.2059	-3.5245	0.175974
2030	-11.0784	-18.8049	-22.895	-3.3519	0.738221
2031	-11.6418	-20.1285	-24.621	-3.1552	1.337366
2032	-12.2053	-21.4755	-26.3828	-2.9351	1.972274
2033	-12.7687	-22.8452	-28.1793	-2.6922	2.64191
2034	-13.3322	-24.237	-30.0096	-2.4273	3.345323
2035	-13.8956	-25.6503	-31.8728	-2.1409	4.081635

4.4.21 Forecast Estimates for the Lee-Carter Model

In the family of Lee-Carter models, the dynamics of mortality are driven by the period and cohort indexes. Lee and Carter (1992) believes that error is minimized when forecast is done with a scalar quantity rather than with a vector quantity. In this study, for all the models including the proposed, what went into forecast is the mortality index or time trend (parameter k_t) of each model within 80-95% prediction interval for a 20-year period.

The forecasts estimates for the Lee-Carter model are presented in Tables 4.16 and 4.17. Generally, the mean estimates are higher for males than females. For the males, the mean forecast ranges from -2.464 in 2016 to -7.763 in 2035. However, for the females, the mean forecast ranges from -3.1902 in 2016 to -13.896 in 2035. For the prediction intervals, lower and upper 80 represent the lower and upper bounds of 80% prediction interval while lower and upper 95 represent the lower and upper bounds of 95% prediction interval. The 80% prediction interval implies that there is an 80% probability that the predicted mortality index will fall somewhere between the lower and upper bounds from 2016 to 2035. This is the same implication for a 95% prediction interval. Generally, it is observed that the forecast trend is decreasing for both gender under the Lee-Carter model.

Table 4.18: Forecast Estimates for the Brouhns Model (Male)

	Mean	Lower		Upper	
		80%	95%	80%	95%
2016	-2.518	-2.66908	-2.74907	-2.36688	-2.28689
2017	-2.7992	-3.01291	-3.12603	-2.58554	-2.47242
2018	-3.0805	-3.34219	-3.48073	-2.81876	-2.68022
2019	-3.3617	-3.66392	-3.82389	-3.05952	-2.89954
2020	-3.643	-3.98083	-4.15969	-3.30509	-3.12624
2021	-3.9242	-4.29433	-4.49026	-3.55409	-3.35816
2022	-4.2055	-4.60523	-4.81686	-3.80568	-3.59406
2023	-4.4867	-4.91408	-5.14032	-4.05933	-3.83309
2024	-4.7679	-5.22125	-5.46121	-4.31465	-4.07469
2025	-5.0492	-5.52701	-5.77996	-4.57138	-4.31843
2026	-5.3304	-5.83158	-6.09687	-4.8293	-4.56401
2027	-5.6117	-6.13511	-6.4122	-5.08826	-4.81118
2028	-5.8929	-6.43773	-6.72613	-5.3481	-5.05974
2029	-6.1742	-6.73954	-7.03883	-5.6088	-5.30953
2030	-6.4554	-7.04063	-7.35042	-5.8702	-5.56043
2031	-6.7367	-7.34107	-7.66102	-6.1323	-5.81232
2032	-7.0179	-7.64092	-7.97072	-6.3949	-6.06512
2033	-7.2992	-7.94023	-8.27958	-6.6581	-6.31874
2034	-7.5804	-8.23904	-8.5877	-6.9218	-6.57312
2035	-7.8617	-8.5374	-8.89511	-7.1859	-6.8282

Table 4.19: Forecast Estimates for the Brouhns Model (Female)

	Mean	Lower		Upper	
		80%	95%	80%	95%
2016	-2.9893	-3.156	-3.2443	-2.8226	-2.73434
2017	-3.4192	-3.792	-3.9894	-3.0464	-2.8491
2018	-3.8491	-4.4729	-4.8031	-3.2254	-2.8951
2019	-4.2791	-5.1922	-5.6756	-3.3659	-2.8826
2020	-4.709	-5.9454	-6.5999	-3.4726	-2.8181
2021	-5.1389	-6.7292	-7.5711	-3.5485	-2.7067
2022	-5.5688	-7.5414	-8.5856	-3.5962	-2.552
2023	-5.9987	-8.3799	-9.6404	-3.6176	-2.3571
2024	-6.4286	-9.2431	-10.733	-3.6142	-2.1243
2025	-6.8586	-10.13	-11.861	-3.5874	-1.8558
2026	-7.2885	-11.039	-13.024	-3.5383	-1.5531
2027	-7.7184	-11.969	-14.219	-3.468	-1.218
2028	-8.1483	-12.919	-15.445	-3.3773	-0.8516
2029	-8.5782	-13.89	-16.701	-3.2669	-0.4552
2030	-9.0082	-14.879	-17.986	-3.1375	-0.0298
2031	-9.4381	-15.886	-19.3	-2.9899	0.42358
2032	-9.868	-16.912	-20.64	-2.8244	0.90418
2033	-10.298	-17.954	-22.007	-2.6418	1.41116
2034	-10.728	-19.013	-23.399	-2.4423	1.94381
2035	-11.158	-20.089	-24.817	-2.2265	2.50145

4.4.22 Forecast Estimates for the Brouhns Model

The forecasts estimates for the Brouhns model are presented in tables 4.18 and 4.19. Just as in the Lee-Carter model, it is observed that the mean estimates are higher for males than females. For the males, the mean forecast ranges from -2.518 in 2016 to -7.862 in 2035. However, for the females, the mean forecast ranges from -2.989 in 2016 to -11.158 in 2035. For the prediction intervals, lower and upper 80 represent the lower and upper bounds of 80% prediction interval while lower and upper 95 represent the lower and upper bounds of 95% prediction interval. The 80% prediction interval means that there is an 80% probability that the predicted mortality index will fall somewhere between the lower and upper bounds from 2016 to 2035. This is the same meaning for a 95% prediction interval. The trend estimate exhibits a decreasing trend for both gender.

Table 4.20: Forecast Estimates for the Gamma-Normal Lee-Carter model (male)

	Mean	Lower		Upper	
		80	95	80	95
2016	-0.629	-0.6973	-0.7335	-0.5606	-0.5244
2017	-0.7005	-0.7972	-0.8484	-0.6039	-0.5527
2018	-0.7721	-0.8905	-0.9531	-0.6537	-0.5911
2019	-0.8437	-0.9804	-1.0527	-0.707	-0.6346
2020	-0.9152	-1.0681	-1.149	-0.7624	-0.6815
2021	-0.9868	-1.1542	-1.2428	-0.8194	-0.7308
2022	-1.0584	-1.2392	-1.3349	-0.8775	-0.7818
2023	-1.1299	-1.3232	-1.4256	-0.9366	-0.8343
2024	-1.2015	-1.4065	-1.5151	-0.9965	-0.8879
2025	-1.2731	-1.4892	-1.6036	-1.0569	-0.9425
2026	-1.3446	-1.5713	-1.6913	-1.118	-0.998
2027	-1.4162	-1.653	-1.7783	-1.1794	-1.0541
2028	-1.4878	-1.7342	-1.8646	-1.2413	-1.1109
2029	-1.5593	-1.8151	-1.9504	-1.3036	-1.1682
2030	-1.6309	-1.8956	-2.0357	-1.3662	-1.2261
2031	-1.7025	-1.9759	-2.1206	-1.4291	-1.2844
2032	-1.774	-2.0558	-2.205	-1.4922	-1.3431
2033	-1.8456	-2.1356	-2.2891	-1.5556	-1.4021
2034	-1.9172	-2.2151	-2.3728	-1.6193	-1.4615
2035	-1.9887	-2.2944	-2.4562	-1.6831	-1.5213

Table 4.21: Forecast Estimates for the Gamma-Normal Lee-Carter model (female)

	Mean	Lower		upper	
		80%	95%	80%	95%
2016	-0.5358	-0.5730	-0.5927	-0.4987	-0.479
2017	-0.6306	-0.7138	-0.7578	-0.5475	-0.5035
2018	-0.7254	-0.8645	-0.9382	-0.5863	-0.5127
2019	-0.8202	-1.0238	-1.1316	-0.6166	-0.5088
2020	-0.9150	-1.1907	-1.3367	-0.6393	-0.4934
2021	-1.0098	-1.3644	-1.5522	-0.6552	-0.4675
2022	-1.1046	-1.5445	-1.7773	-0.6648	-0.432
2023	-1.1994	-1.7304	-2.0114	-0.6685	-0.3874
2024	-1.2942	-1.9218	-2.2540	-0.6667	-0.3345
2025	-1.389	-2.1184	-2.5045	-0.6596	-0.2735
2026	-1.4838	-2.3200	-2.7626	-0.6476	-0.2050
2027	-1.5786	-2.5263	-3.028	-0.6309	-0.1292
2028	-1.6734	-2.7372	-3.3004	-0.6096	-0.0464
2029	-1.7682	-2.9525	-3.5794	-0.5839	0.04302
2030	-1.863	-3.172	-3.8649	-0.554	0.13894
2031	-1.9578	-3.3956	-4.1567	-0.52	0.2411
2032	-2.0526	-3.6231	-4.4545	-0.4821	0.34933
2033	-2.1474	-3.8545	-4.7582	-0.4403	0.46343
2034	-2.2422	-4.0896	-5.0676	-0.3947	0.58326
2035	-2.337	-4.3284	-5.3826	-0.3455	0.70867

4.4.23 Forecast Estimates for the Gamma-Normal Lee-Carter Model

The forecasts estimates for the Gamma-Normal Lee-Carter model are presented in tables 4.20 and 4.21. Unlike the results obtained from the Lee-Carter and the Brouhns' models, estimates of the mean forecast for the Gamma-Normal Lee-Carter model were higher for females than males during the first five years of forecast (year 2016-2020) and then higher for males than females from 2021 to 2035. For the males, the mean forecast ranges from -0.629 in 2016 to -1.989 in 2035. However, for the females, the mean forecast ranges from -0.536 in 2016 to -2.337 in 2035. For the prediction intervals, lower and upper 80 represent the lower and upper bounds of 80% prediction interval while lower and upper 95 represent the lower and upper bounds of 95% prediction interval. As stated earlier, the 80% prediction interval indicates that there is an 80% probability that the predicted mortality index will fall somewhere between the lower and upper bounds from 2016 to 2035. This also holds for a 95% prediction interval. Generally it is observed that the forecast trend is decreasing for both gender with an intersection almost occurring in the year 2020.

4.4.24 Comparison of Mean Forecast Estimates across the Models

The results for all the models are presented in tables 4.16 to 4.21. Across the models, for both male and female, the proposed Gamma-Normal Lee-Carter had higher estimates compared to the other two models. The Brouhns model had higher estimates than the Lee-Carter for the female forecast, while a very close tie existed between the two for the male forecast.

Table 4.22: Measures of Goodness of fit for the forecast when $h = 20$

Male						
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	σ^2	S.E($\hat{\theta}$)
GNLC	23.20	-42.40	-40.99	-41.40	0.002844	0.0133
LC	2.78	-1.55	-0.14	-0.55	0.043320	0.0519
BR	11.30	-18.60	-17.19	-17.60	0.013900	0.0294
Female						
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	σ^2	-
GNLC	29.70	-57.40	-56.76	-57.06	0.000841	-
LC	4.84	-7.69	-7.05	-7.39	0.02931	-
BR	8.69	-15.38	-14.74	-15.04	0.01692	-

4.4.25 Measures of Goodness of Fit for the Forecast when $h = 20$

Across the models, the male data set fits best under a random walk with drift model, Arima (0 1 0) while the female data fit best under an Arima (0 2 0). As displayed in table 4.22, for the male data, the standard error of the drift parameter ($\hat{\theta}$) is lowest for the proposed model. For both the male and female dataset, the Gamma-normal Lee-Carter has the minimum variance with 0.002844 and 0.000841 respectively. The proposed model also has the lowest AIC, BIC and CAIC values which represents the distribution with minimum information loss. The Renshaw Haberman model did not converge.

Table 4.23: Measures of Goodness of fit results for the simulated data

years = 5, ages = 15					
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	HQC
GNLC	-71.5810	207.162	281.3216	257.4477	236.7731
LC	-71.5817	209.1634	285.6405	263.8951	239.6998
BR	-59782.7	119631.3	119707.8	119653.3	119746.6
RH	-37822.3	75746.5	75864.69	75876.41	76219.84
years = 20, ages = 20					
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	HQC
GNLC	-198.322	510.6444	738.1579	529.9777	600.7427
LC	-198.338	512.6758	744.1807	532.7462	604.3548
BR	-208805	417725.8	417957.3	417745.9	417817.5
RH	-174534	349259	349642.2	349320.5	349410.7
years = 30, ages = 20					
	$\log [L(\hat{\theta})]$	AIC	BIC	CAIC	HQC
GNLC	-281.065	696.1302	990.7245	715.8727	810.8098
LC	-281.100	698.2004	997.1916	726.5509	814.5917
BR	-296181	592498	592797	592515.7	592614.4
RH	-266975	534181.4	534691.4	534237.6	534379.9
years = 40, ages = 25					
	$\log [L(\hat{\theta})]$	AIC	BIC		
GNLC	-556.957	1287.91	1714.889		
LC	-540.4348	1256.87	1688.752		
BR	-564036	1128248	1128680		
RH	-530638.5	1061579	1062320		

4.4.26 Measures of Goodness of Fit Results for the Simulated Data

For the simulated data, the AIC, BIC and CAIC were lowest for the gamma-normal Lee-Carter than other models. The Brouhns and the Renshaw Haberman models did not converge for both gender. Furthermore, the results presented in Table 4.23 show that the proposed Gamma-Normal Lee-Carter model could not work beyond 30 years. When the time period extended to 40 years, the Lee-Carter model outperformed the Gamma-Normal Lee-Carter model. This brings to mind the central limit theorem whose attributes has been discussed in the previous chapter.

CHAPTER FIVE

Summary, Conclusion, Contribution to Knowledge and Suggestion for Further Study

5.1 Summary

The Lee-Carter model has become a reference point for mortality modelling in most developed Countries. Attempts in using the model for describing the pattern of mortality in developing countries with defective data situation, have resulted in the violation of the normality assumption. This is due to its inability to accommodate variability across all age-groups. Therefore, in this study, a Gamma-Normal Lee-Carter model was developed and compared to some existing variants of the Lee-Carter model which include: Lee and Carter (1992), Brouhns (2002) and Renshaw and Haberman (2006) model. The proposed model was able to accommodate variability at different ages better than the other existing variants used.

5.2 Conclusion

When it comes to modelling heterogeneous mortality data, which is usually the situation in many developing countries, the Gamma-Normal Lee-Carter model should be considered as a better alternative to the referenced classical Lee-Carter. Hence, the modified Lee-Carter model is recommended for modelling mortality data from developing countries.

5.3 Contribution to Knowledge

The limitation of the classical Lee-Carter model with respect to mortality data from developing countries has been successfully overcome as the proposed Gamma-Normal Lee-Carter model can be effectively used when the dynamics of mortality change are heterogeneous within a given population.

5.4 Suggestion for further study

This method can also be applied to other mortality models that assume a Poisson or Normal distribution framework such as the Brouhns and the Renshaw Haberman model.

REFERENCES

- Akaike, H., 1973. Information theory and an extension of the maximum likelihood Principle. *Proceeding of the Second International Symposium on Information Theory*. Eds. B. N. Petrov and F. Csaki. Budapest: Akademiai Kiado. 267-281.
- Alexander, C., Cordeiro, G. M., Ortega, E. M. M., & Sarabia, J. M. 2012. Generalized beta-generated distributions. *Computational Statistics & Data Analysis*. 56: 1880–1897.
- Alho, J. M. 1990. Stochastic methods in population forecasting. *International Journal of forecasting* 6.4: 521-530.
- Alho, J. M. & Spencer, B. D. 1985. Uncertain population forecasting. *Journal of the American Statistical Society* 80.390: 306-314.
- Bamidele, M. A. & Adejumo, A. O. 2014. Stochastic Modelling of Mortality Risks in Nigeria Using Lee-Carter and Renshaw-Haberman Model. *Mathematical Theory and Modelling* 4.6: 74-86.
- Bedrick, E. J. & Tsai, C. L. 1994. Model selection for multivariate regression in small Samples. *Biometrics* 50.1: 226-231.
- Bolker, B. 2017. Maximum likelihood estimation and analysis with the bbmle package. Retrieved June 7, 2018, from <https://cran.rproject.org/web/packages/bbmle/vignettes/mle2.pdf>
- Bongaarts, J. 2004. Long-Range Trends in adult mortality: Models and projection methods *The population council*. 2004.192: 1-14.
- Booth, H., Maindonald, J. & Smith, L. 2002. Applying Lee-Cater under Conditions of Variable Mortality Decline. *Population Studies* 56.3: 325-336.
- Booth, H., Tickle, L. & Smith, L. 2005. Evaluation of the variants of the Lee-Carter method of forecasting mortality: A multi-Country comparison. *New Zealand Population Review* 31.1: 13-34.
- Box, G. E. P. & Jenkins, G. M. 1976. *Time Series Analysis: Forecasting and Control*. Rev. ed. San Francisco: Holden Day
- Bradley, D. J. 1991. Malaria. *Disease and Mortality in Sub-Saharan Africa*. Eds. R. G.

- Feachem and D. T. Jamison. Oxford, England: Oxford University Press for the World Bank.
- Brouhns, N., Denuit, M. & Vermunt J. 2002. A Poisson log-bilinear regression approach to the construction of projected life-tables. *Insurance: Mathematics and Economics* 31.3: 373–393.
- Bowman, K. O., & Shenton, B. R. 1975. Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and $\sqrt{b_2}$. *Biometrika* 62.2: 243-250
- Brown, J. R. 2003. Redistribution and insurance: mandatory annuitization with mortality heterogeneity. *Journal of Risk and Insurance*.70.1: 17-41
- Burnham, K. P. & Anderson, D. R. 2002. *Model selection and multimodel inference: A practical information theoretic-approach*. 2nded. New York: Springer-Verlag.
- Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. 1983. *Graphical Methods for Data Analysis*. Duxbury Press: Boston
- Chukwu, A. U. & Oladipupo, O. E. 2012. Modelling adult mortality in Nigeria: An Analysis based on the Lee-carter model. *Studies in Mathematical Sciences* 5.2: 1-11.
- Chukwu, A. U. & Adegbilero-Iwari, O. E. 2015. The performance of the Lee-Carter model on heterogeneous adult mortality data. *Mathematical Theory and Modeling* 5.4: 198-207
- Currie, I. D. 2016. On Fitting Generalized Linear and non-linear Models of Mortality. *Scandinavian Actuarial Journal* 4: 356-383.
- D'Agostino, R. B. 1986. Tests for normal distribution. *Goodness-of-fit Techniques*. Eds. R. B. D'Agostino & M. A. Stephens. New York: Marcel Dekker. 367-420
- D'Amato, V., Haberman, S. & Russolillo, M. 2009. Efficient bootstrap applied to the Poisson log-bilinear Lee Carter model. In *ASMDA. Proceedings of the International Conference Applied Stochastic Models and Data Analysis* 13: 374. Vilnius Gediminas Technical University, Department of Construction Economics & Property.
- Delwarde, A., Denuit, M. & Partrat, C. 2007. Negative binomial version of the Lee –

- Carter model for mortality forecasting. *Applied Stochastic Models in Business and Industry* 23.5: 385-401
- Doray, L. G. 2008. Inference for Logistic-type Models for the Force of Mortality. *International Symposium on Living to 100 and Beyond, Society of Actuaries, SOA Monograph M-LI08-1*: 18.
- Doray, L. G. & Tang, K. O. 2011. Projection of Mortality rates at advanced ages in Canada with a New Lee-Carter type model. *Presented at the Living to 100 Society of Actuaries Symposium Orlando, Fla. January 5-7, 2011.*
- Eugene, N., Lee, C., Famoye, F. 2002. Beta-normal distribution and its applications. *Communication in Statistics- Theory and Methods* 31.4: 497-512.
- Ewbank, D. C. & Gribble, J. N. 1993. *Effects of Health Programs on Child Mortality in Sub-Saharan Africa*. Washington: National Academy Press.
- Girosi, F. & King, G. 2007. Understanding the Lee-Carter mortality forecasting method. Retrieved July 13, 2011 from <http://gking.harvard.edu/files/lc.pdf>.
- Haberman, S. & Butt, Z., 2010. Ilc: A Collection of R Functions for Fitting a Class of Lee-Carter Mortality Models using Iterative Fitting Algorithm. *Cass Actuarial Research Paper* 190.
- Hanna, S. 2007. Applying the Lee-Carter model to countries in Eastern Europe and the former Soviet Union. *Journal of Applied Statistics*, 33.2: 255-272
- Hannan, E. J. & Quinn, B. G. 1979. The determination of the order of an auto-Regression. *Journal of the Royal Statistical Society: Series B (Methodological)* 41.2: 190-195.
- Hurvich, C. M., Shumway, R. H. & Tsai, C. L. 1990. Improved estimators of Kullback-Leibler information for autoregressive model selection in small samples. *Biometrika* 77.4: 709-719.
- Hurvich, C. M. & Tsai, C. L. 1989. Regression and Time series model selection in small samples. *Biometrika* 76.2: 297-307.
- Hurvich, C. M. & Tsai, C. L. 1993. A corrected Akaike information criterion for vector auto-regressive model selection. *Journal of Time series Analysis* 14.3: 271-279.

- Hyndman, R. J. & Ullah, M. S. 2007. Robust forecasting of mortality and fertility Rates: A functional data approach. *Computational Statistics and Data Analysis* 51.10: 4942-4956.
- Igawa T. 2013. Analysis of the Residual Structure of the Lee–Carter Model: The Case of Japanese Mortality. *Asia-Pacific Journal of Risk and Insurance* 7.2: 53-80
- Kannisto, V. 1992. Frailty and survival. *Genus* 47: 101-118
- Keyfitz, N. 1991. Population and development within the ecosphere: one view of the literature. *Population index* 5-22.
- Keyfitz, N. & Caswell, H. 2005. *Applied mathematical demography (Vol. 47)*. New York: Springer.
- Koissi, M., & Shapiro, A. 2008. The Lee-carter model under the condition of variables age-specific parameters. *Presented at the 43rd Actuarial Research Conference, Regina, Canada August 2008*.
- Koissi M., Shapiro A. & Högnäs G. 2004. Fitting and Forecasting Mortality Rates for Nordic Countries Using the Lee-Carter method. *Actuarial Research Clearing House* 1:21
- Kolmogorov, A. 1933. Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari* 4: 83–91.
- Kumaraswamy, P. 1980. Generalized probability density-function for double-bounded Random processes. *Journal of Hydrology* 46.1-2: 79–88.
- Lee, R. & Carter, L. 1992. Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association* 87.419: 659-671.
- Lee, R. & Miller, T. 2001. Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography* 38.4: 537-549
- Li, S. & Chan, W. 2007. The Lee-Carter model for forecasting mortality, revisited. *North American Actuarial Journal* 11.1: 68-87
- Li, N., Lee, R. & Tuljapurkar, S. 2002. Using the Lee-Carter method to forecast mortality in populations with limited data. *International Statistical Review* 72.1: 19-36
- Lima, M., Cordeiro, G., & Ortega, E. 2015. A new extension of the normal

- distribution. *Journal of Data Science* 13.2: 385-408
- Makeham, W.M. 1860. On the law of mortality and the construction of actuary tables. *Journal of the Institute Actuaries* 8.6: 301-310
- Manton, K. G., Akushevich, I. & Kravchenko, J. 2009. *Cancer Mortality and Morbidity Patterns in the US Population*. New York, NY.: Springer. 37-87.
- Nadarajah, S., Cordeiro, G., & Ortega, E. 2015. The Zografos–Balakrishnan G family of distributions: mathematical properties and application. *Communication in Statistics-Theory and Methods* 44.1: 186–215
- Olshansky, S. J. 1988. On forecasting mortality. *The Milbank Quarterly* 482-530.
- Pescim, R. R., Cordeiro, G. M., Demétrio, C. G., Ortega, E. M. & Nadarajah, S. 2012. The new class of Kummer beta generalized distributions. *SORT-Statistics and Operations Research Transactions* 36.2: 153-180.
- Pedroza, C. 2006. A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics* 7.4: 530–550. doi:10.1093/biostatistics/kxj024
- Pollard, J. H. 1987. Projection of age-specific mortality rates. *Population Bulletin of the United Nations* 21/22: 55-69.
- Renshaw, A. E. & Haberman, S. 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38.3: 556 –570
- Renshaw, A. E. & Haberman, S. 2008. On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee–Carter modelling. *Insurance: Mathematics and Economics* 42.2: 797-816.
- Reese, S. 2015. Asymptotic inference in the Lee-Carter model for modelling mortality rates. *Working Paper/Department of Economics, School of Economics and Management, Lund University* 16.
- Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics*. 6.2: 461-464.
- Shapiro, S. S. & Wilk, M. B. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52.3/4: 591-611.
- Smirnov, N. 1948. Table for estimating the goodness of fit of empirical distributions.

- Annals of Mathematical Statistics* 19.2: 279–281.
doi :10.1214/aoms/1177730256.
- Stacy, E. W. 1962. A generalization of the gamma distribution. *The Annals of Mathematical Statistics* 33.3: 1187-1192.
- Stephens, M. A. 1974. EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association* 69.347: 730-737.
- Taruvinga, R., Gachira, W., Chiwanza, W. & Nkomo, D. 2017. Comparison of the Lee-Carter and arch in modelling and forecasting mortality in Zimbabwe. *Asian Journal of Economic Modelling* 5.1: 11-22
- Timaeus, I. M. 1999. Mortality in Sub-Saharan Africa. *Health and Mortality: Issues of Global Concern*. Eds. J. Chamie & R. L. Cliquet. Brussels: Population Division, United Nations and Population and Family Study Centre (CBGS). 110-131.
- Villegas, A., Kaishev, V. K. & Millossovich, P. 2015. StMoMo: An R package for Stochastic mortality modelling. In *7th Australasian Actuarial Education and Research Symposium*. Retrieved June 7, 2018, from <http://www.papers.ssrn.com>.
- Wang, H. & Preston, S. H., 2009. Forecasting United States mortality using cohort smoking histories. *Proceedings of the National Academy of Sciences*, 106.2: 393-398.
- Wang J. 2007. “Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model. *Mathematical Statistics, Stockholm University*. Retrieved Dec 23, 2009, from <http://www.math.su.se/matstat>
- Wilmoth, J. 1993. Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. *Technical report, Department of Demography, University of California, Berkeley U.S.A.*
- World Health Organisation. 2017. Global Health observatory. Retrieved Feb 23, 2018, from <http://apps.who.int/gho/data/view.main.61200?lang=en>
- Yaffee, R. & McGee, M. 2000. *Introduction to time-series analysis and forecasting With application of SAS and SPSS*. Orlando: Academic Press, INC.

- Zhao, B. B. 2012. A modified Lee-Carter model for analysing short-base period data. *Population Studies* 66.1: 39-52.
- Zografos, K. & Balakrishnan, N. 2009. On families of Beta- and generalized Gamma-Generated distributions and associated inference. *Statistical Methodology* 6.4: 344-362.

APPENDIX A

Table 1.1: Age-specific Mortality rates Nigeria (Males)

	2000	2001	2002	2003	2004	2005	2006	2007
<1	0.132	0.128	0.124	0.121	0.117	0.113	0.109	0.105
1-4	0.022	0.022	0.021	0.02	0.019	0.018	0.017	0.016
5-9	0.007	0.007	0.007	0.007	0.007	0.006	0.006	0.006
10-14	0.004	0.004	0.004	0.004	0.004	0.003	0.003	0.003
15-19	0.005	0.005	0.005	0.005	0.004	0.004	0.004	0.004
20-24	0.007	0.007	0.007	0.007	0.007	0.006	0.006	0.006
25-29	0.008	0.008	0.008	0.008	0.008	0.007	0.007	0.007
30-34	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
35-39	0.011	0.011	0.011	0.012	0.011	0.011	0.011	0.011
40-44	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.012
45-49	0.015	0.015	0.015	0.015	0.015	0.015	0.014	0.014
50-54	0.019	0.019	0.019	0.019	0.019	0.018	0.018	0.018
55-59	0.024	0.024	0.024	0.024	0.024	0.024	0.023	0.023
60-64	0.036	0.036	0.036	0.036	0.036	0.035	0.034	0.034
65-69	0.054	0.054	0.054	0.054	0.054	0.053	0.052	0.051
70-74	0.087	0.088	0.088	0.088	0.087	0.086	0.084	0.083
75-79	0.142	0.143	0.144	0.144	0.142	0.14	0.138	0.136
80-84	0.229	0.23	0.231	0.231	0.229	0.226	0.223	0.221
85+	0.357	0.358	0.359	0.359	0.357	0.354	0.350	0.348

Table 1.1.1: Age-specific Mortality rates Nigeria (Males)

	2008	2009	2010	2011	2012	2013	2014	2015
<1	0.101	0.098	0.094	0.091	0.088	0.085	0.082	0.08
1-4	0.015	0.015	0.014	0.013	0.013	0.012	0.011	0.011
5-9	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005
10-14	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
15-19	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
20-24	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
25-29	0.007	0.007	0.007	0.007	0.007	0.007	0.006	0.006
30-34	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.007
35-39	0.011	0.01	0.01	0.01	0.01	0.01	0.01	0.009
40-44	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.011
45-49	0.014	0.013	0.013	0.013	0.013	0.013	0.013	0.012
50-54	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.016
55-59	0.023	0.022	0.022	0.022	0.022	0.022	0.021	0.021
60-64	0.033	0.033	0.033	0.032	0.032	0.032	0.032	0.031
65-69	0.051	0.05	0.05	0.049	0.049	0.049	0.048	0.048
70-74	0.082	0.082	0.081	0.081	0.08	0.08	0.079	0.079
75-79	0.135	0.134	0.134	0.133	0.133	0.132	0.132	0.131
80-84	0.219	0.218	0.218	0.217	0.216	0.216	0.215	0.215
85+	0.347	0.346	0.345	0.345	0.345	0.344	0.344	0.343

Table 1.2: Age-specific Mortality rates Nigeria (Females)

	2000	2001	2002	2003	2004	2005	2006
<1	0.110	0.107	0.103	0.100	0.097	0.093	0.090
1-4	0.022	0.021	0.020	0.020	0.019	0.018	0.017
5-9	0.007	0.007	0.007	0.007	0.007	0.007	0.006
10-14	0.004	0.004	0.004	0.004	0.004	0.004	0.004
15-19	0.004	0.004	0.004	0.004	0.004	0.004	0.004
20-24	0.006	0.006	0.006	0.005	0.005	0.005	0.005
25-29	0.008	0.008	0.008	0.008	0.008	0.007	0.007
30-34	0.009	0.01	0.01	0.01	0.01	0.01	0.009
35-39	0.011	0.011	0.012	0.012	0.012	0.012	0.012
40-44	0.012	0.012	0.012	0.012	0.012	0.012	0.012
45-49	0.012	0.012	0.013	0.013	0.013	0.012	0.012
50-54	0.015	0.015	0.015	0.015	0.015	0.015	0.015
55-59	0.02	0.02	0.02	0.02	0.02	0.02	0.019
60-64	0.031	0.031	0.032	0.032	0.031	0.031	0.03
65-69	0.05	0.05	0.05	0.05	0.05	0.049	0.048
70-74	0.084	0.084	0.085	0.085	0.084	0.083	0.081
75-79	0.139	0.14	0.141	0.141	0.139	0.137	0.135
80-84	0.227	0.229	0.23	0.23	0.229	0.226	0.222
85+	0.347	0.349	0.35	0.35	0.348	0.345	0.342

Table 1.2.1: Age-specific Mortality rates Nigeria (Females)

	2007	2008	2009	2010	2011	2012	2013	2014	2015
<1	0.087	0.084	0.081	0.078	0.075	0.073	0.070	0.068	0.066
1-4	0.016	0.015	0.014	0.014	0.013	0.012	0.012	0.011	0.011
5-9	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.005
10-14	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
15-19	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
20-24	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.004
25-29	0.007	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006
30-34	0.009	0.009	0.009	0.008	0.008	0.008	0.008	0.007	0.007
35-39	0.011	0.011	0.011	0.011	0.011	0.011	0.01	0.009	0.008
40-44	0.012	0.012	0.011	0.011	0.011	0.011	0.011	0.01	0.009
45-49	0.012	0.012	0.012	0.011	0.012	0.012	0.011	0.011	0.01
50-54	0.015	0.014	0.014	0.014	0.014	0.014	0.014	0.013	0.012
55-59	0.019	0.019	0.018	0.018	0.018	0.018	0.018	0.017	0.017
60-64	0.03	0.029	0.029	0.029	0.028	0.028	0.028	0.027	0.027
65-69	0.047	0.047	0.046	0.046	0.046	0.045	0.045	0.044	0.044
70-74	0.08	0.079	0.078	0.078	0.077	0.077	0.076	0.075	0.075
75-79	0.133	0.132	0.131	0.13	0.129	0.129	0.128	0.127	0.126
80-84	0.22	0.218	0.217	0.216	0.215	0.214	0.213	0.212	0.211
85+	0.339	0.337	0.336	0.335	0.335	0.334	0.333	0.332	0.332

Table 1.3: Number of Deaths for Nigeria (Males)

	2000	2001	2002	2003	2004	2005	2006	2007
<1	12090	11780	11460	11160	10840	10500	10160	9810
1-4	7460	7200	6940	6650	6370	6100	5830	5580
5-9	2816	2801	2775	2740	2680	2605	2534	2478
10-14	1457	1453	1441	1428	1407	1374	1345	1323
15-19	1775	1774	1761	1751	1738	1705	1681	1663
20-24	2518	2517	2493	2477	2465	2411	2375	2345
25-29	2803	2822	2814	2811	2799	2739	2701	2653
30-34	3128	3174	3199	3226	3232	3189	3159	3094
35-39	3562	3657	3733	3796	3812	3762	3748	3679
40-44	3860	3937	4002	4059	4078	4057	4072	4025
45-49	4145	4209	4267	4315	4331	4315	4314	4272
50-54	4821	4877	4932	4980	4999	4995	4996	4971
55-59	5673	5722	5775	5821	5842	5853	5852	5850
60-64	7217	7272	7343	7412	7453	7493	7501	7530
65-69	8758	8828	8915	8991	9056	9148	9198	9285
70-74	10020	10092	10187	10271	10379	10550	10677	10854
75-79	9346	9385	9455	9534	9688	9946	10169	10440
80-84	6072	6061	6082	6135	6286	6545	6794	7069
85+	2479	2440	2425	2444	2542	2714	2895	3080

Table 1.3.1: Number of Deaths for Nigeria (Males)

	2008	2009	2010	2011	2012	2013	2014	2015
<1	9490	9170	8870	8570	8300	8040	7800	7570
1-4	5330	5090	4860	4660	4450	4260	4090	3920
5-9	2446	2428	2416	2404	2383	2349	2310	2243
10-14	1313	1310	1308	1308	1304	1292	1285	1250
15-19	1658	1656	1658	1667	1669	1670	1672	1642
20-24	2332	2328	2324	2329	2324	2323	2332	2301
25-29	2639	2597	2584	2592	2579	2570	2567	2486
30-34	3090	2999	2979	3008	3000	2990	2976	2808
35-39	3705	3583	3588	3650	3651	3647	3628	3352
40-44	4061	3931	3918	3991	4010	4024	4022	3757
45-49	4294	4216	4218	4297	4336	4355	4349	4137
50-54	4986	4939	4936	4982	5011	5031	5040	4927
55-59	5856	5862	5865	5880	5898	5911	5921	5919
60-64	7540	7603	7623	7618	7635	7653	7673	7788
65-69	9320	9454	9509	9504	9538	9574	9617	9849
70-74	10941	11163	11272	11290	11363	11442	11527	11888
75-79	10589	10875	11035	11092	11208	11333	11462	11898
80-84	7226	7475	7628	7700	7815	7937	8061	8418
85+	3183	3321	3408	3458	3527	3600	3669	3849

Table 1.4: Number of Deaths for Nigeria (Females)

	2000	2001	2002	2003	2004	2005	2006	2007
<1	10260	9970	9690	9380	9080	8780	8490	8210
1-4	7510	7270	7000	6740	6470	6200	5920	5650
5-9	2961	2936	2903	2862	2797	2715	2637	2576
10-14	1584	1575	1561	1546	1521	1484	1452	1428
15-19	1725	1718	1704	1690	1669	1633	1603	1579
20-24	2113	2110	2098	2085	2062	2022	1987	1956
25-29	2777	2815	2840	2854	2837	2787	2748	2682
30-34	3288	3364	3432	3493	3510	3484	3461	3372
35-39	3683	3811	3927	4025	4057	4025	4031	3959
40-44	3647	3736	3820	3897	3934	3934	3970	3932
45-49	3579	3645	3707	3763	3789	3784	3796	3762
50-54	4126	4183	4236	4283	4302	4294	4297	4268
55-59	4997	5044	5092	5136	5153	5149	5140	5125
60-64	6928	6989	7050	7103	7127	7136	7121	7131
65-69	9080	9150	9217	9269	9317	9374	9392	9462
70-74	11017	11078	11142	11197	11290	11434	11531	11699
75-79	10644	10656	10689	10740	10885	11122	11325	11599
80-84	7150	7098	7082	7116	7272	7532	7786	8084
85+	2931	2852	2810	2822	2929	3113	3312	3524

Table 1.4.1: Number of Deaths for Nigeria (Females)

	2008	2009	2010	2011	2012	2013	2014	2015
<1	7930	7650	7400	7140	6909	6690	6490	6290
1-4	5400	5160	4910	4700	4490	4300	4110	3940
5-9	2541	2520	2503	2489	2471	2448	2433	2397
10-14	1416	1410	1407	1405	1401	1393	1395	1375
15-19	1570	1563	1561	1564	1563	1556	1554	1537
20-24	1940	1925	1915	1909	1900	1887	1878	1856
25-29	2665	2578	2553	2553	2523	2440	2377	2250
30-34	3376	3219	3189	3220	3190	3031	2905	2650
35-39	4012	3847	3866	3942	3926	3711	3537	3118
40-44	3993	3846	3847	3933	3945	3782	3654	3290
45-49	3801	3715	3732	3822	3855	3751	3661	3429
50-54	4292	4233	4238	4292	4318	4260	4216	4100
55-59	5129	5123	5126	5146	5163	5161	5160	5153
60-64	7123	7182	7191	7182	7197	7259	7308	7423
65-69	9464	9608	9639	9613	9639	9782	9894	10149
70-74	11746	11998	12079	12065	12129	12375	12574	12988
75-79	11717	12051	12197	12233	12353	12675	12946	13456
80-84	8239	8547	8710	8786	8924	9219	9475	9918
85+	3645	3826	3936	4006	4104	4279	4434	4679

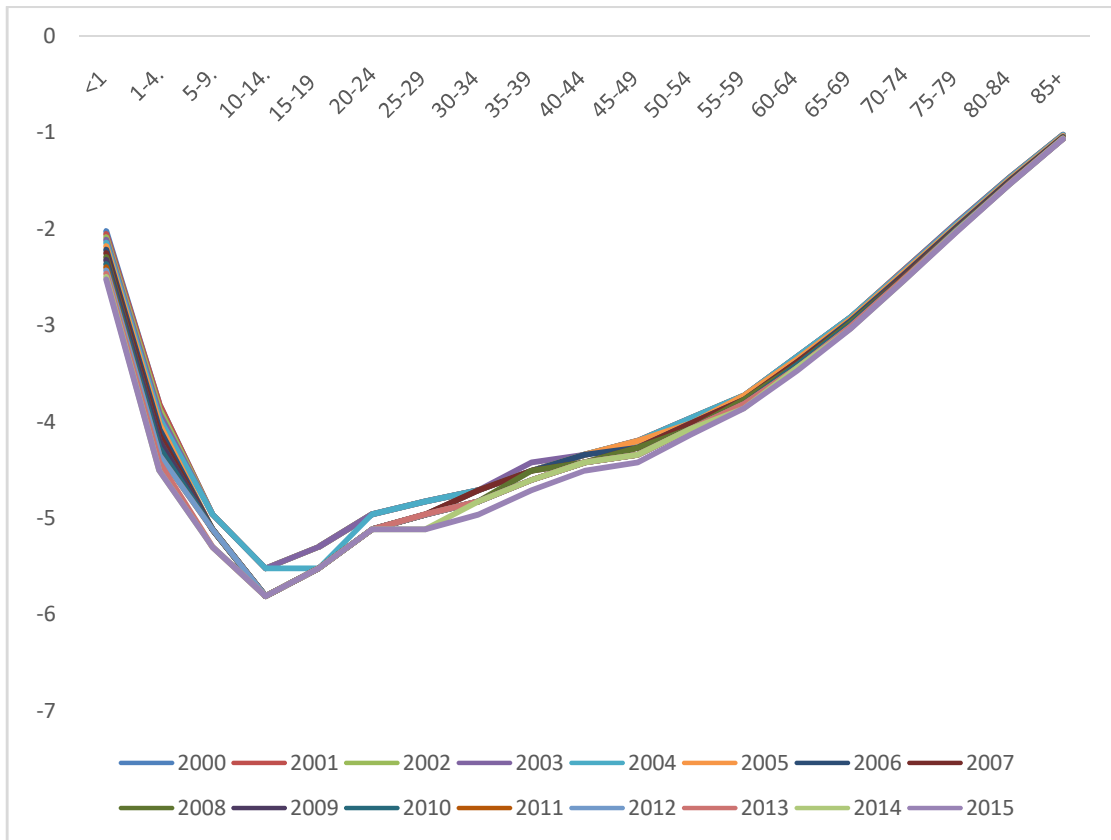


Figure 1: Annotated plots of log male mortality rates

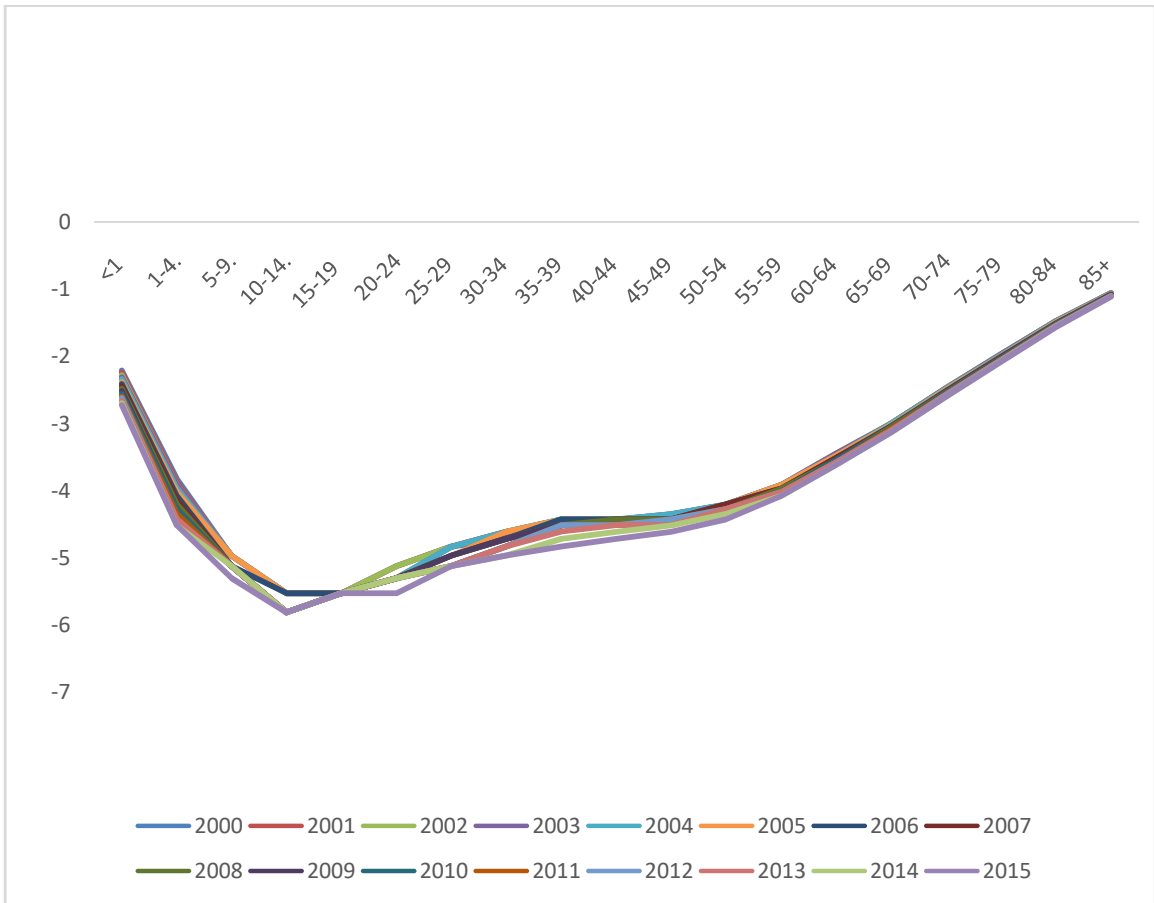


Figure 2: Annotated plots of log female mortality rates

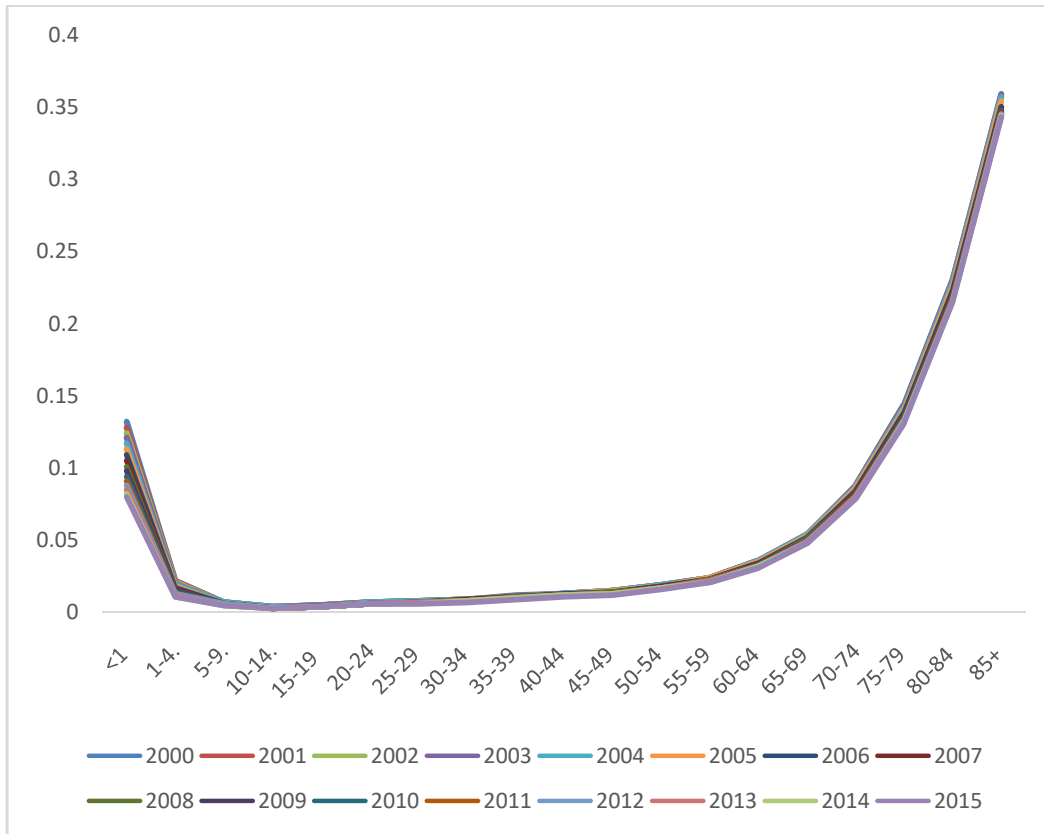


Figure 3: Annotated plots of untransformed male mortality rates (2000-2015)

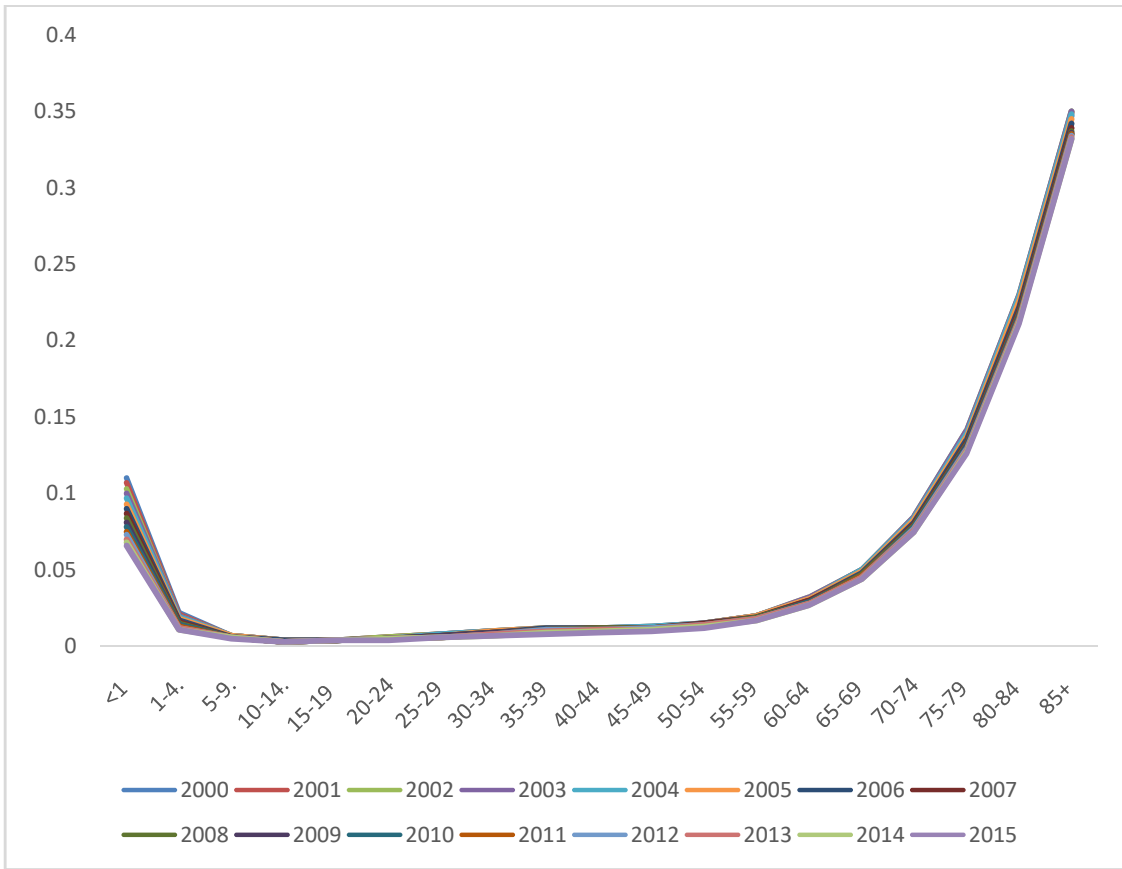


Figure 4: Annotated plots of untransformed female mortality rates (2000-2015)

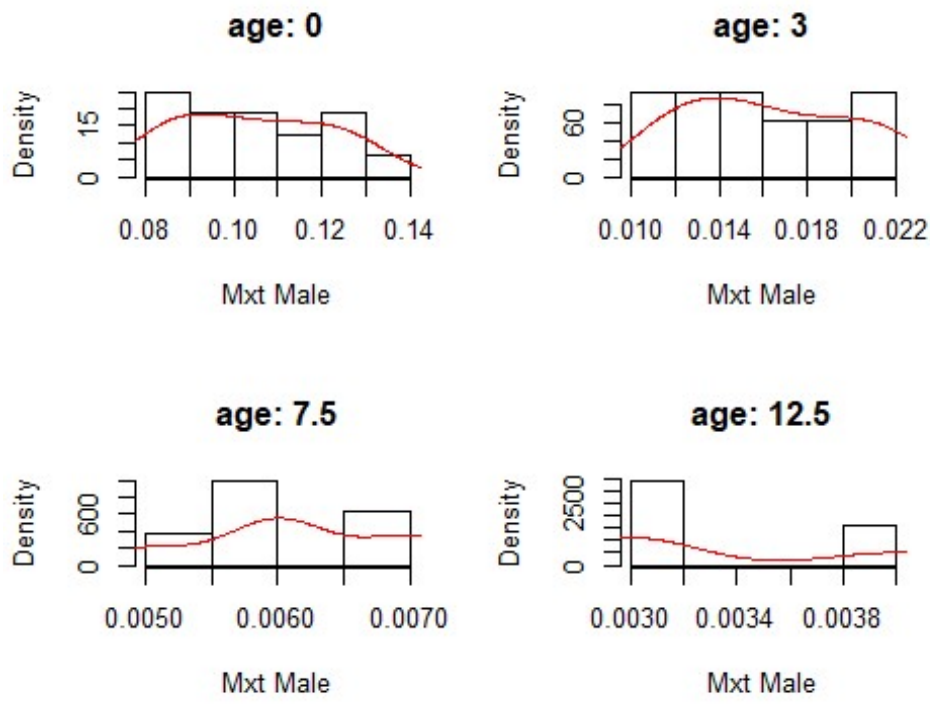


Figure 5: Histogram across age-groups 0-12.5 (male mortality rates)

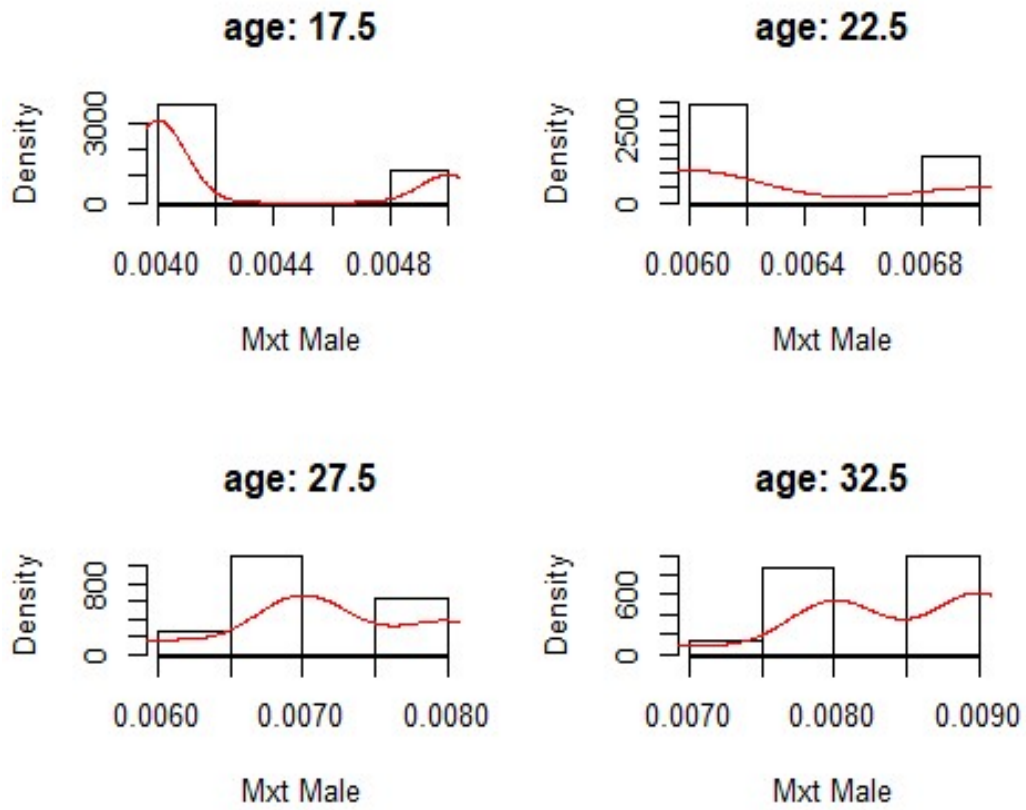


Figure 6: Histogram across age-groups 17.5-32.5 (male mortality rates)

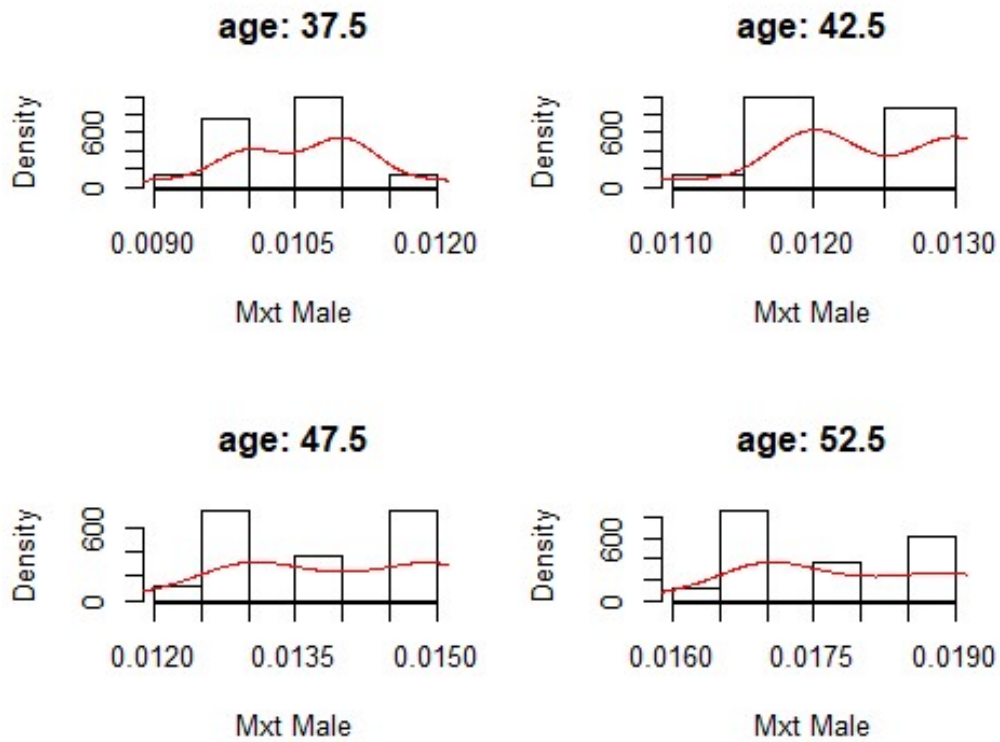


Figure 7: Histogram across age-groups 37.5-52.5 (male mortality rates)

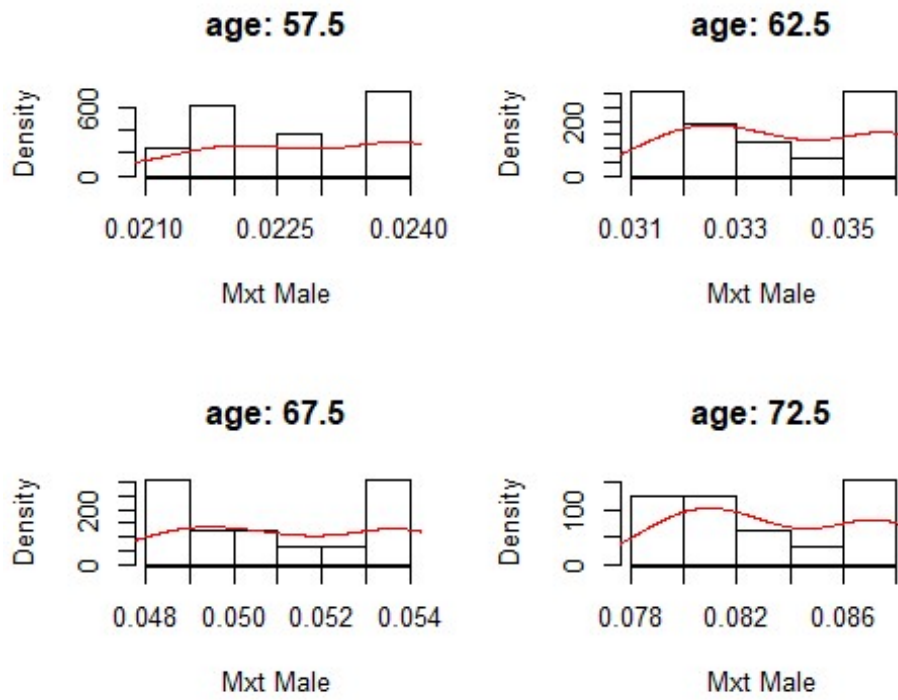


Figure 8: Histogram across age-groups 57.5-72.5 (male mortality rates)

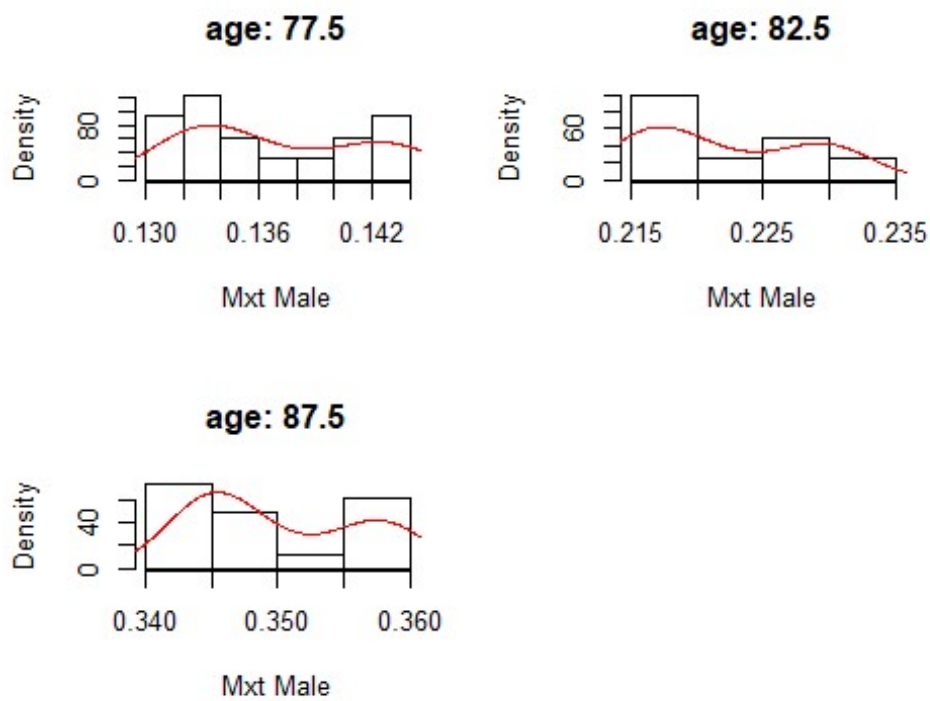


Figure 9: Histogram across age-groups 77.5-87.5 (male mortality rates)

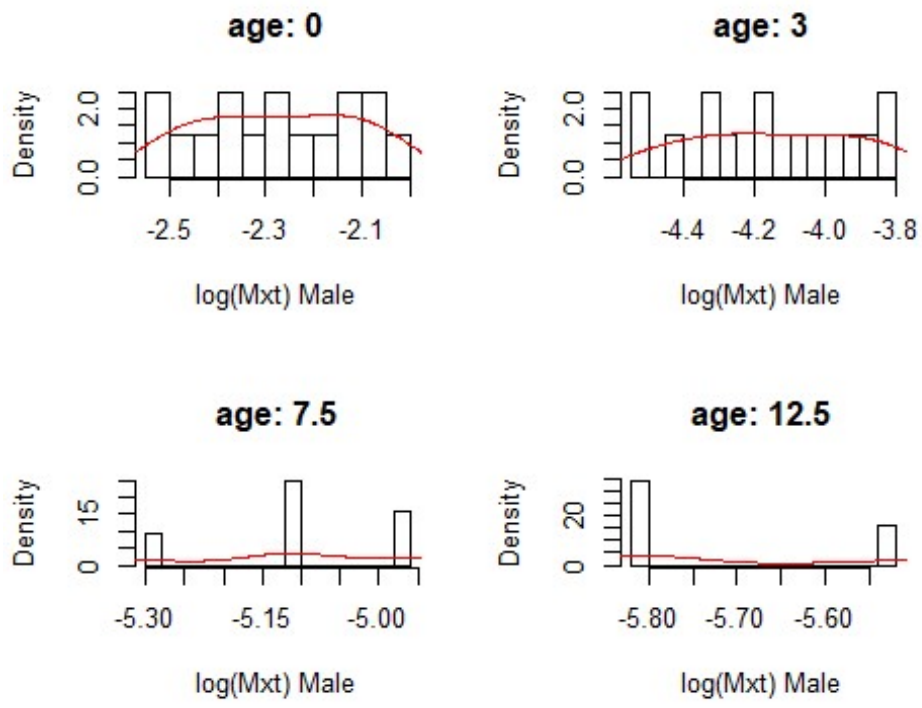


Figure 10: Histogram across age-groups 0-12.5 (male log mortality rates)

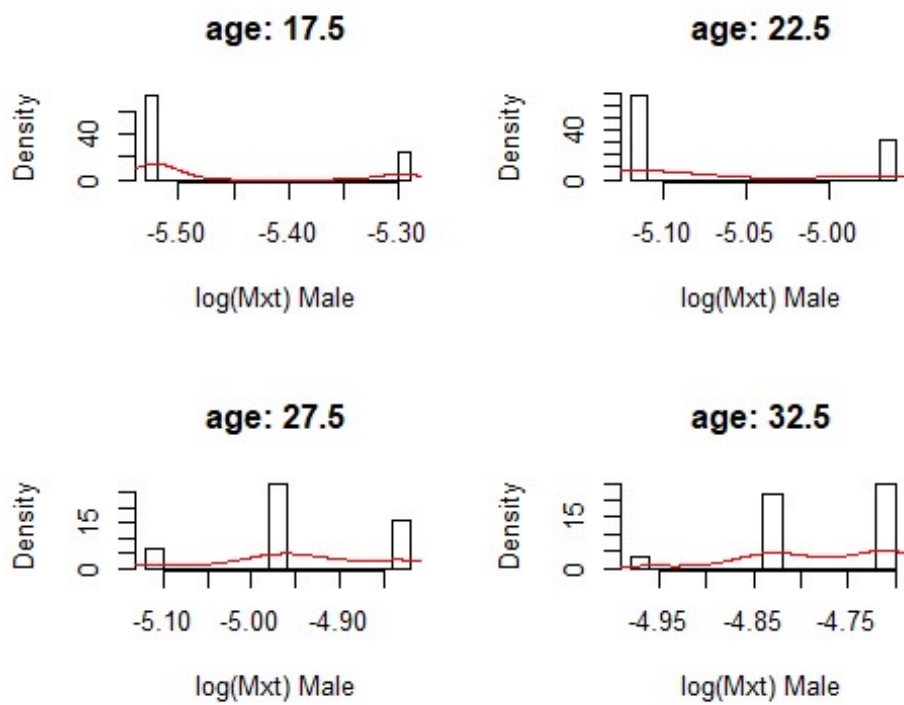


Figure 11: Histogram across age-groups 17.5-32.5 (male log mortality rates)

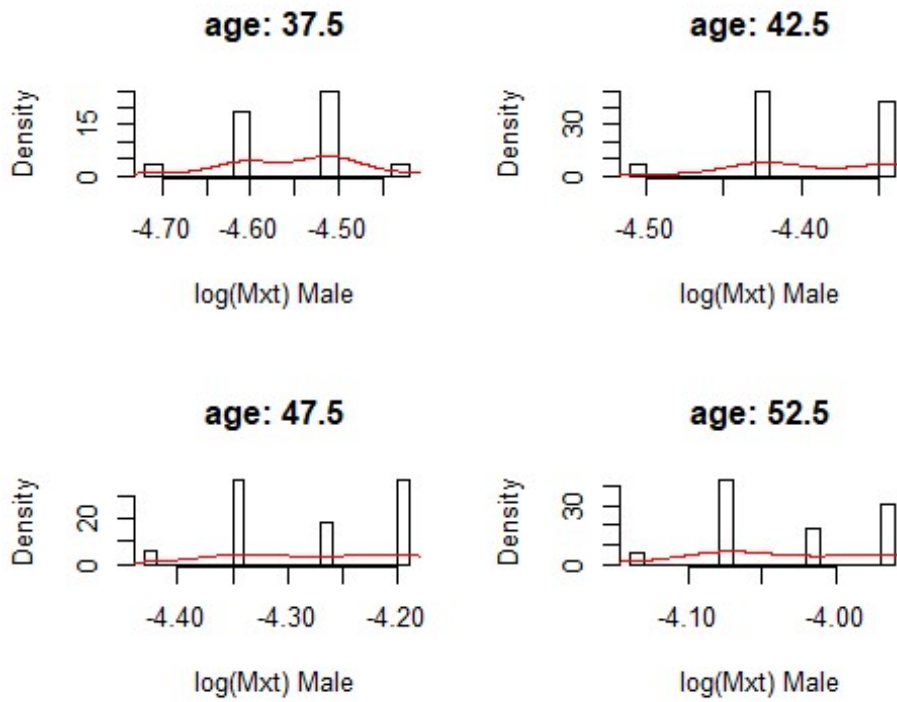


Figure 12: Histogram across age-groups 37.5-52.5 (male log mortality rates)

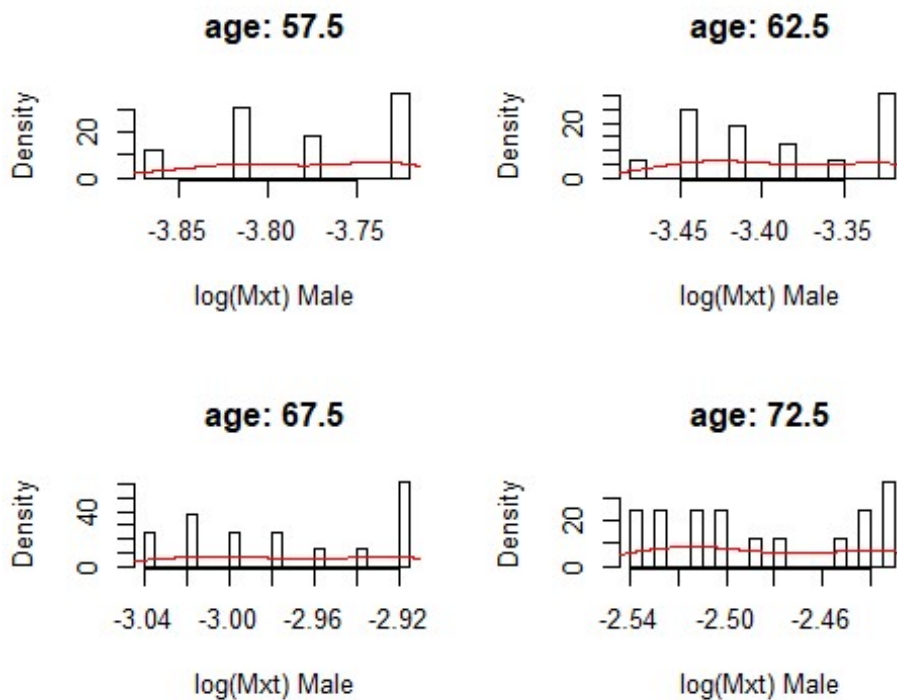


Figure 13: Histogram across age-groups 57.5-72.5 (male log mortality rates)

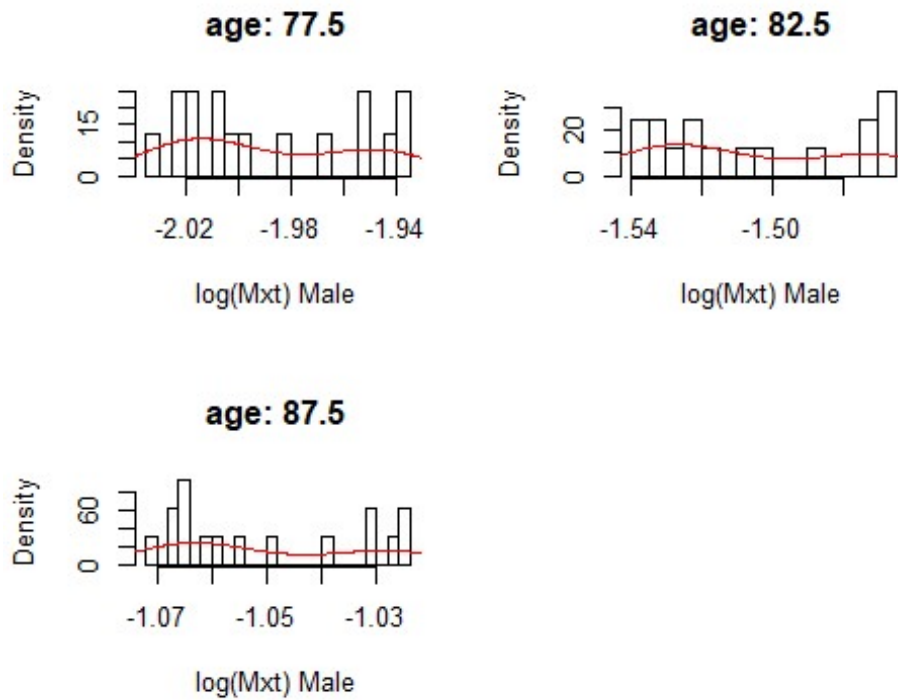


Figure 14: Histogram across age-groups 77.5-87.5 (male log mortality rates)

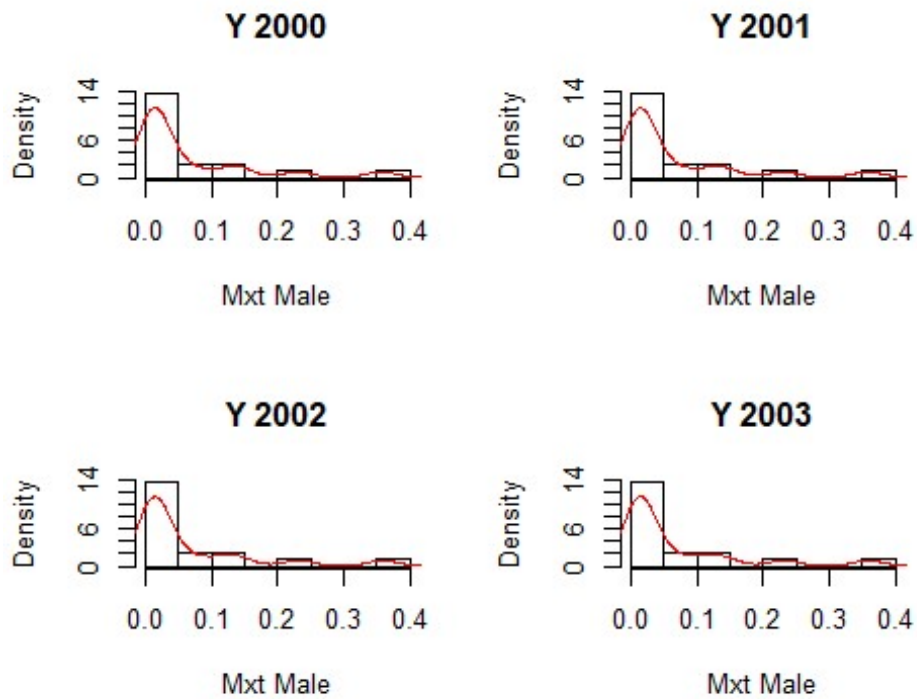


Figure 15: Histogram across period 2000-2003 (male mortality rates)

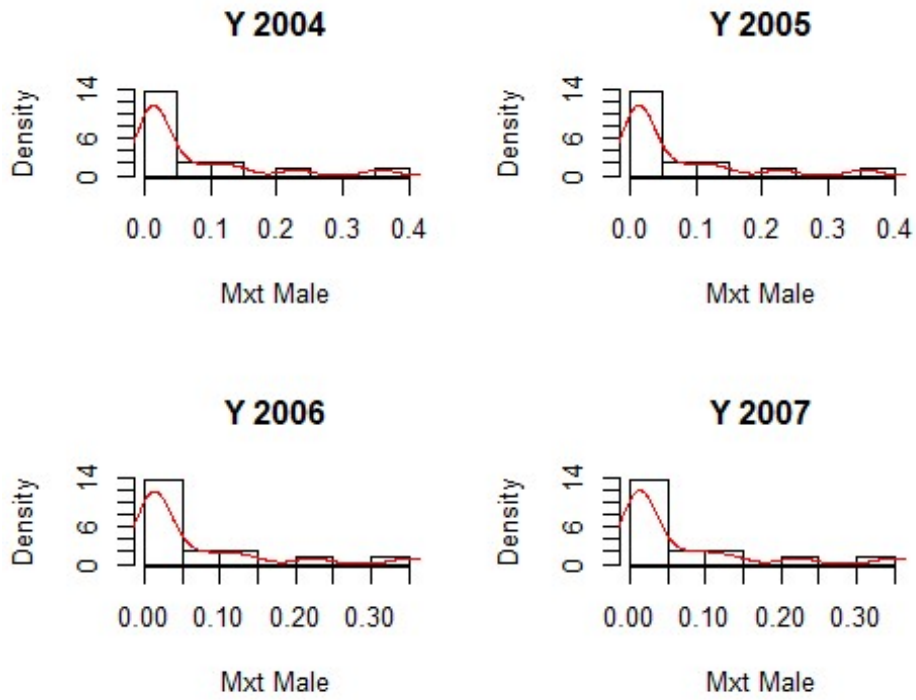


Figure 16: Histogram across period 2004-2007 (male mortality rates)

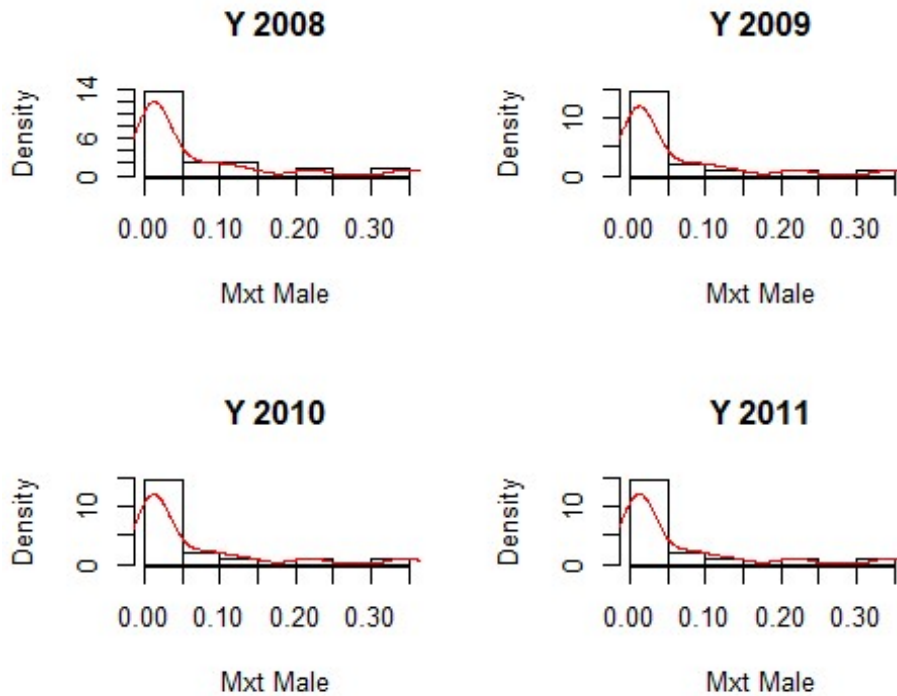


Figure 17: Histogram across period 2008-2011 (male mortality rates)

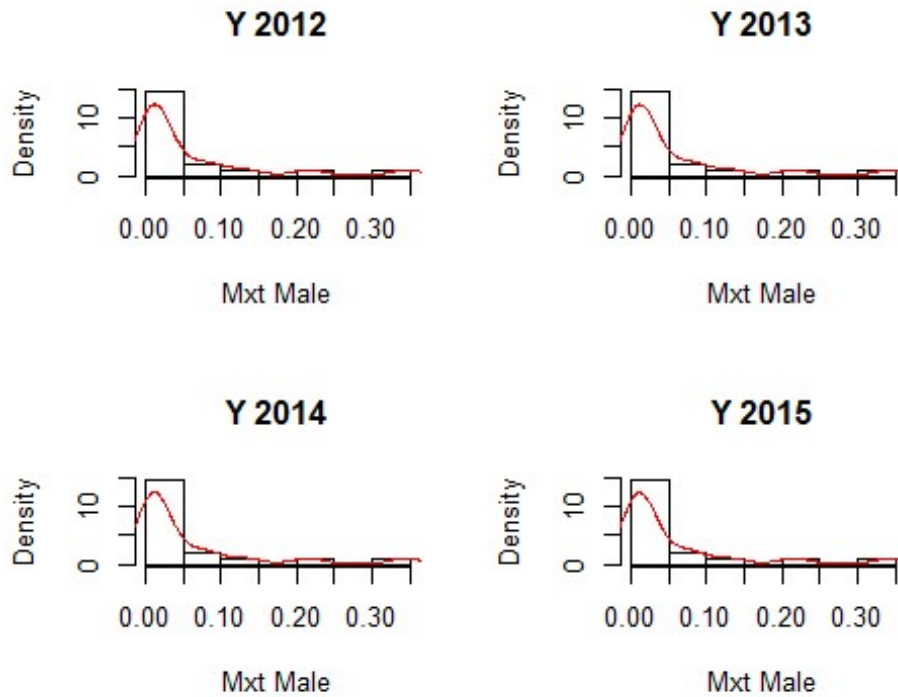


Figure 18: Histogram across period 2012-2015 (male mortality rates)

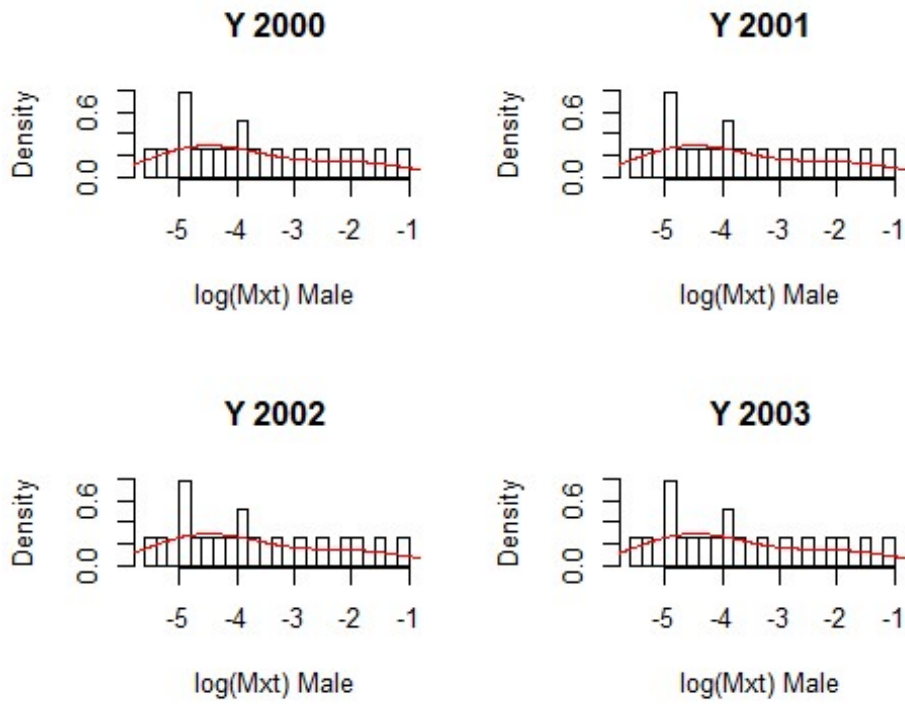


Figure 19: Histogram across period 2000-2003 (male log mortality rates)

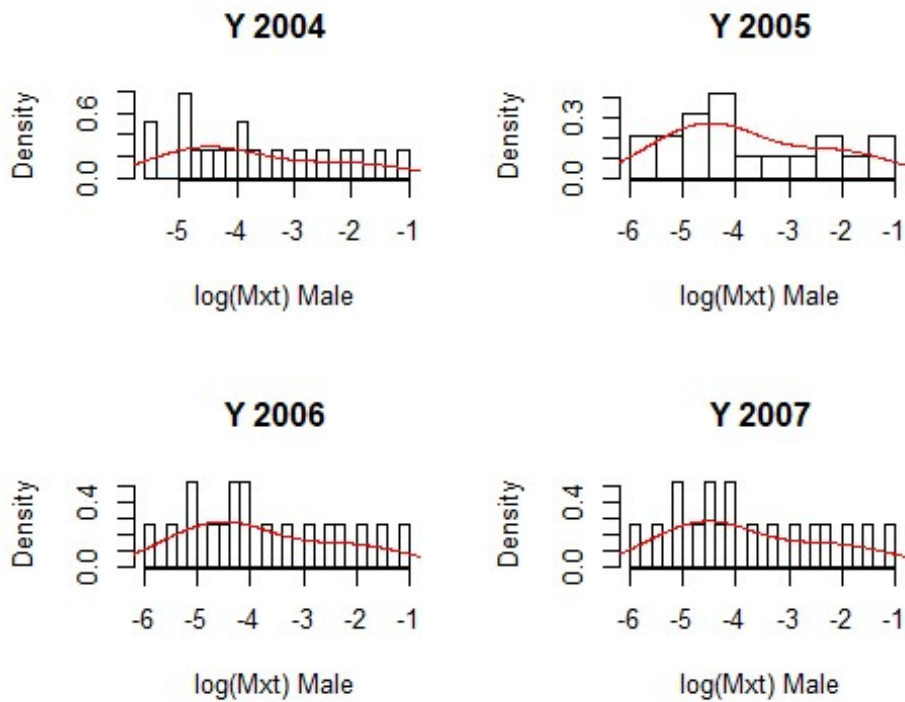


Figure 20: Histogram across period 2004-2007 (male log mortality rates)

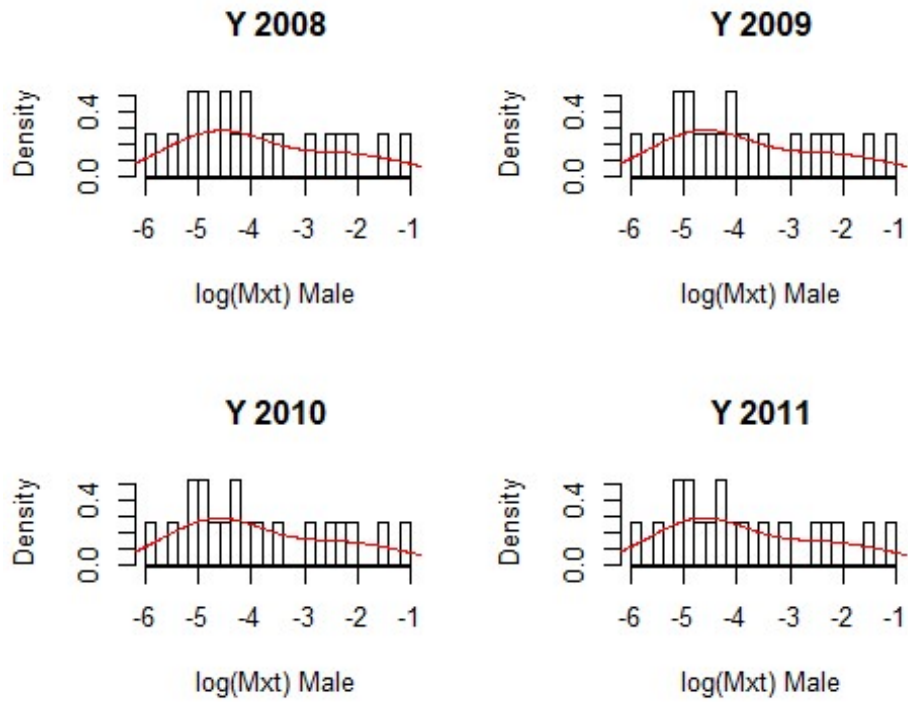


Figure 21: Histogram across period 2008-2011 (male log mortality rates)

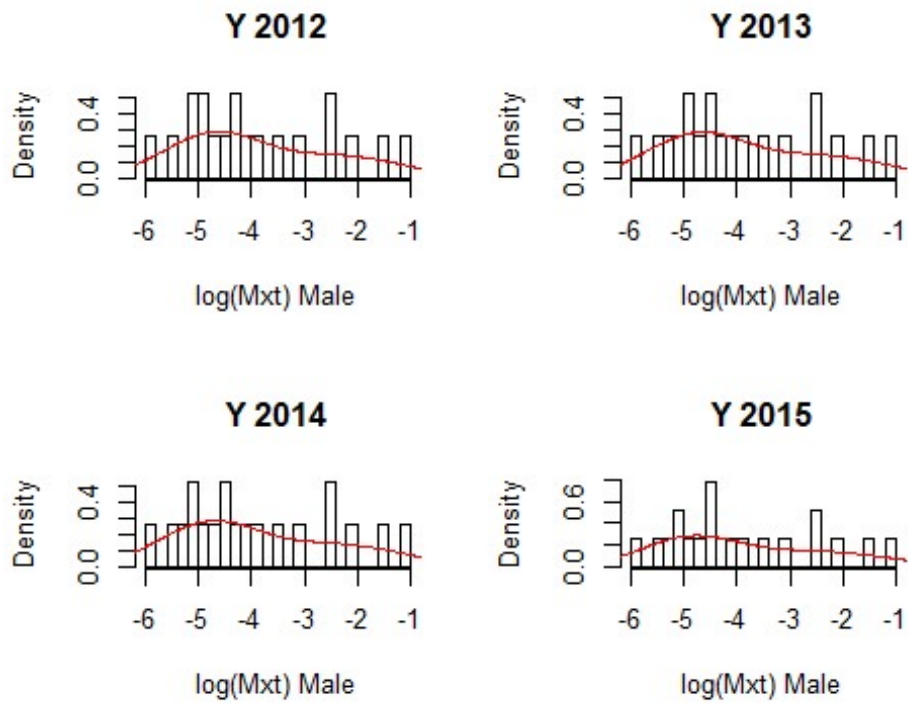


Figure 22: Histogram across period 2012-2015 (male log mortality rates)

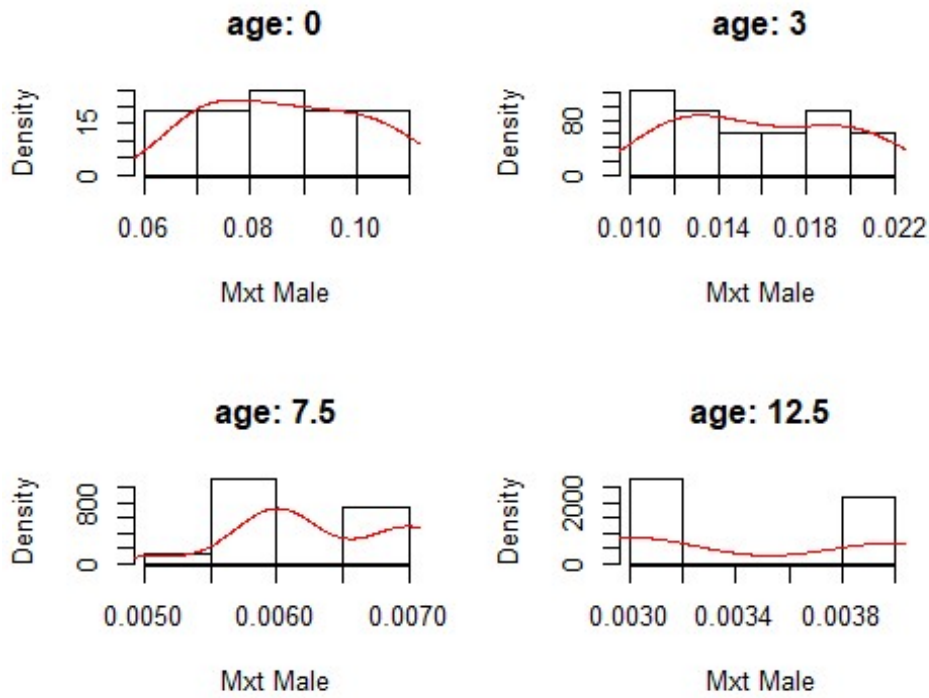


Figure 23: Histogram across age-groups 0-12.5 (female mortality rates)

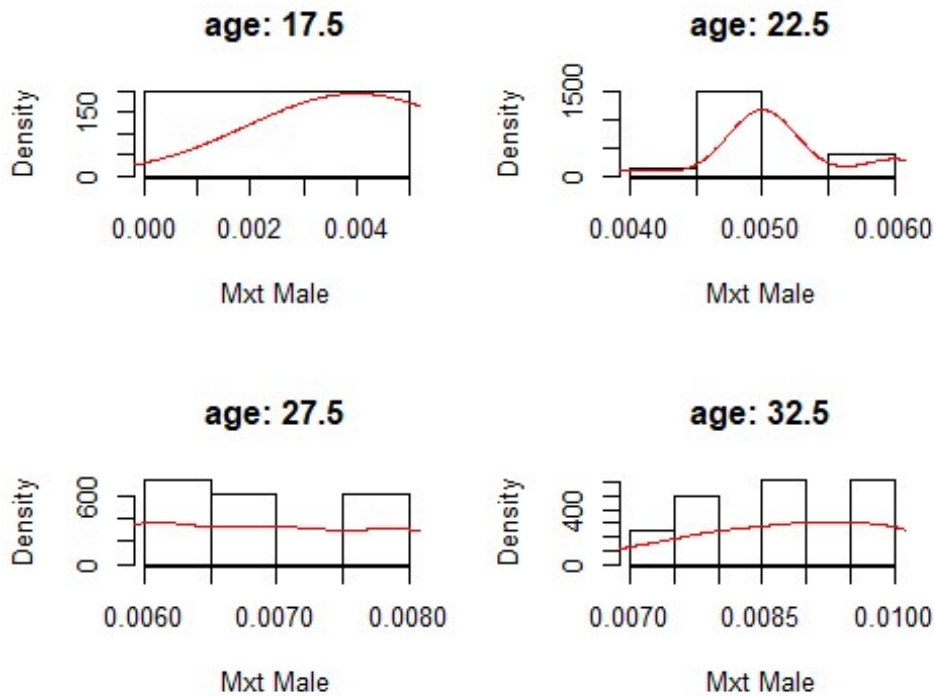


Figure 24: Histogram across age-groups 17.5-32.5 (female mortality rates)

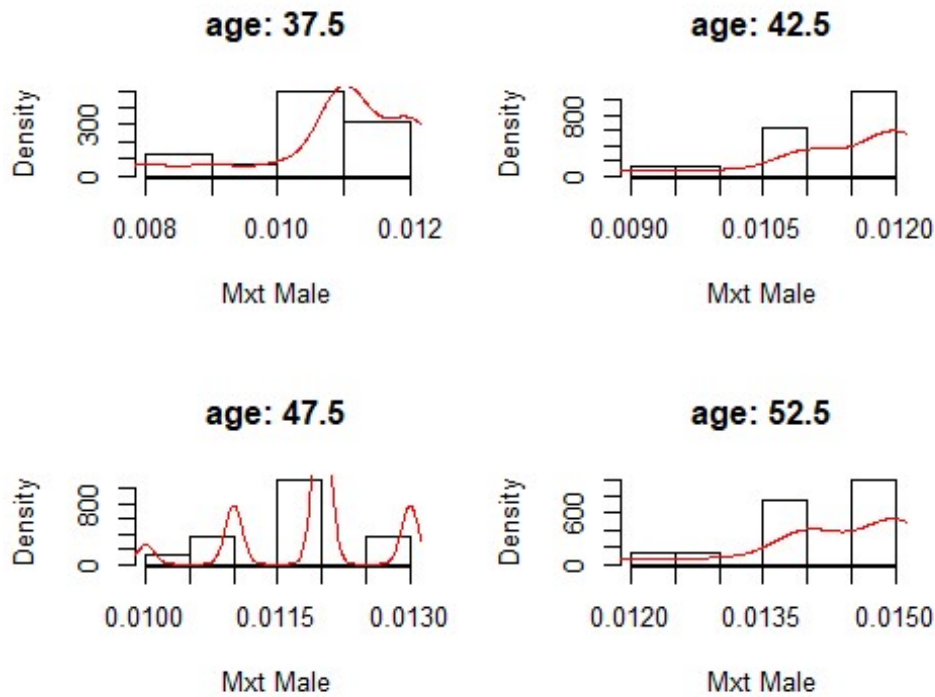


Figure 25: Histogram across age-groups 37.5-52.5 (female mortality rates)

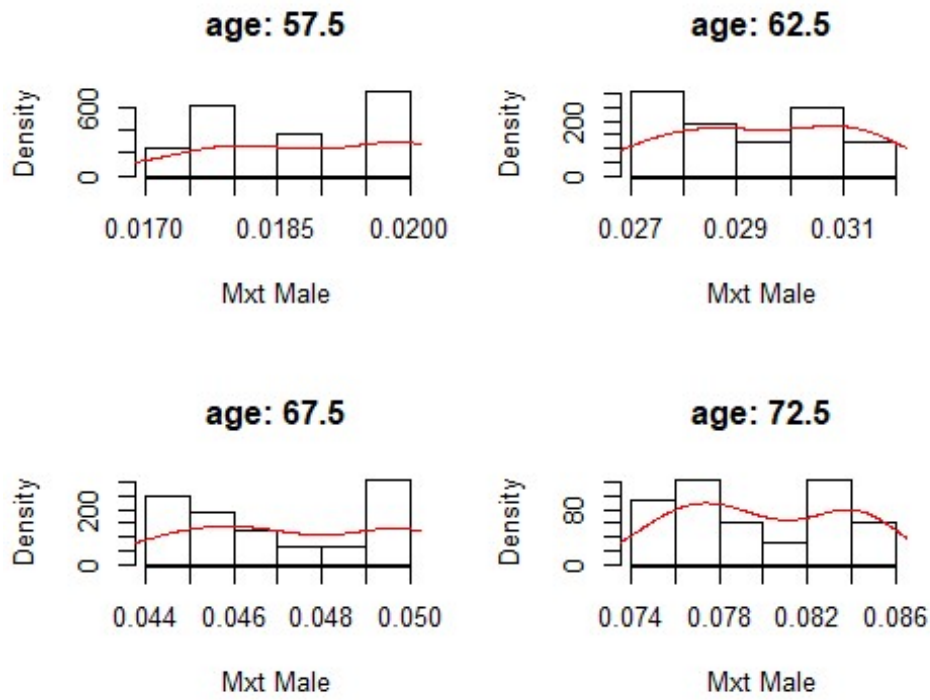


Figure 26: Histogram across age-groups 57.5-72.5 (female mortality rates)

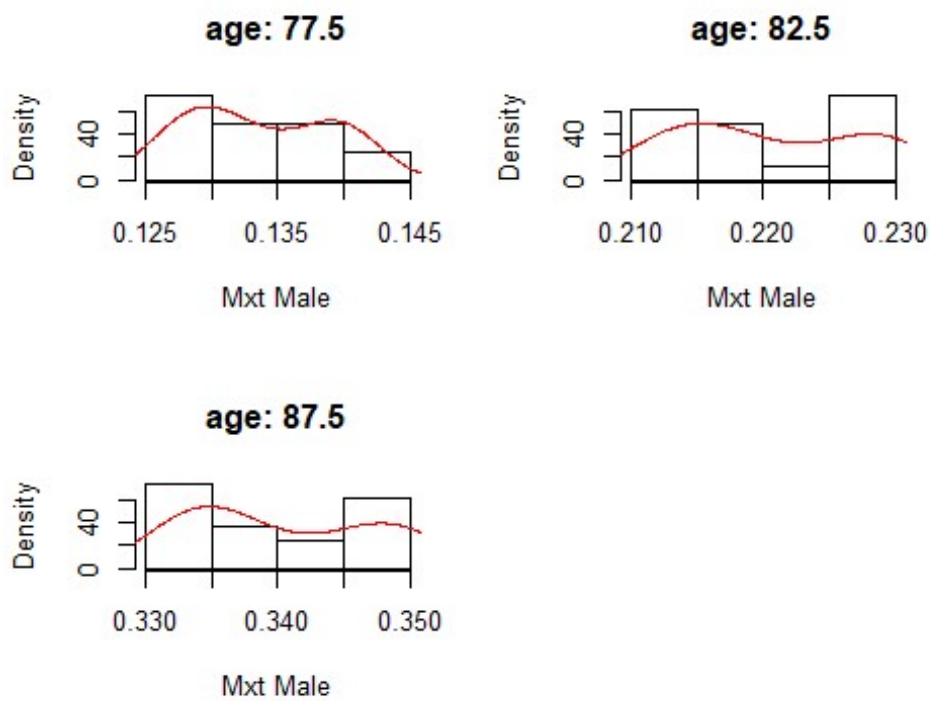


Figure 27: Histogram across age-groups 77.5-87.5 (female mortality rates)

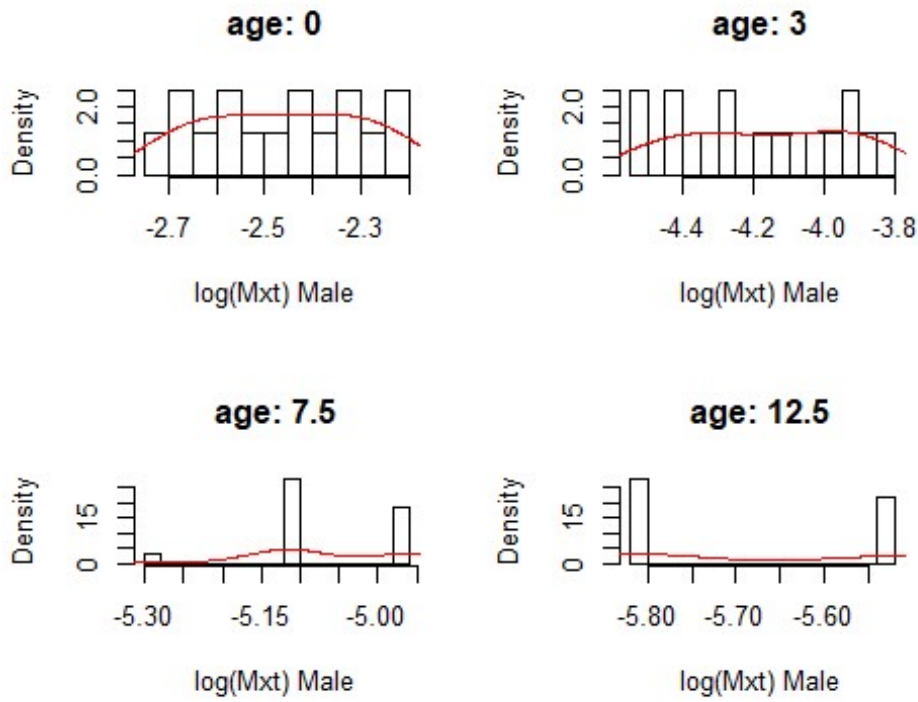


Figure 28: Histogram across age-groups 0-12.5 (female log mortality rates)

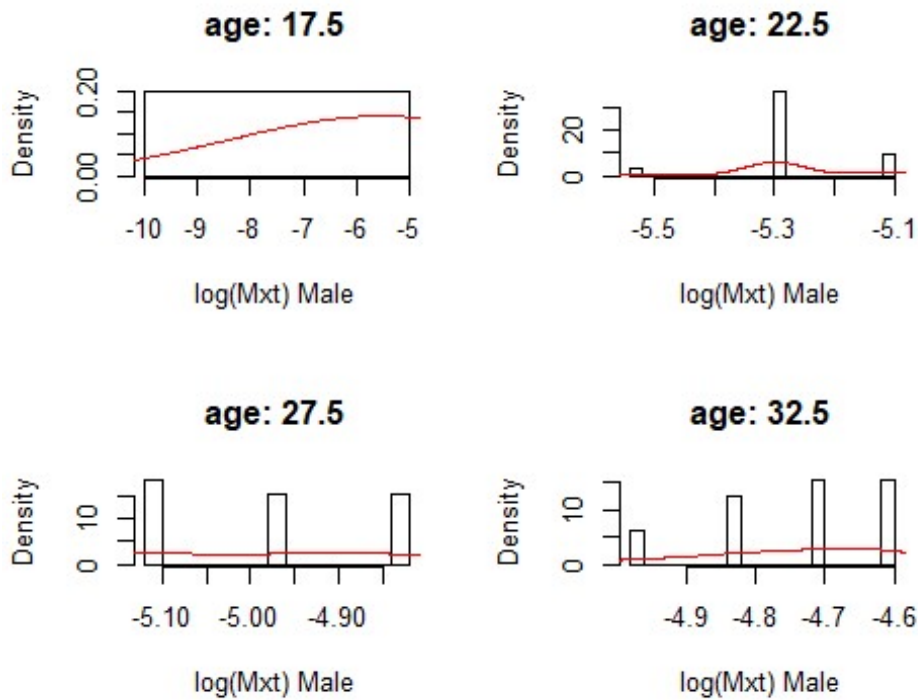


Figure 29: Histogram across age-groups 17.5-32.5 (female log mortality rates)

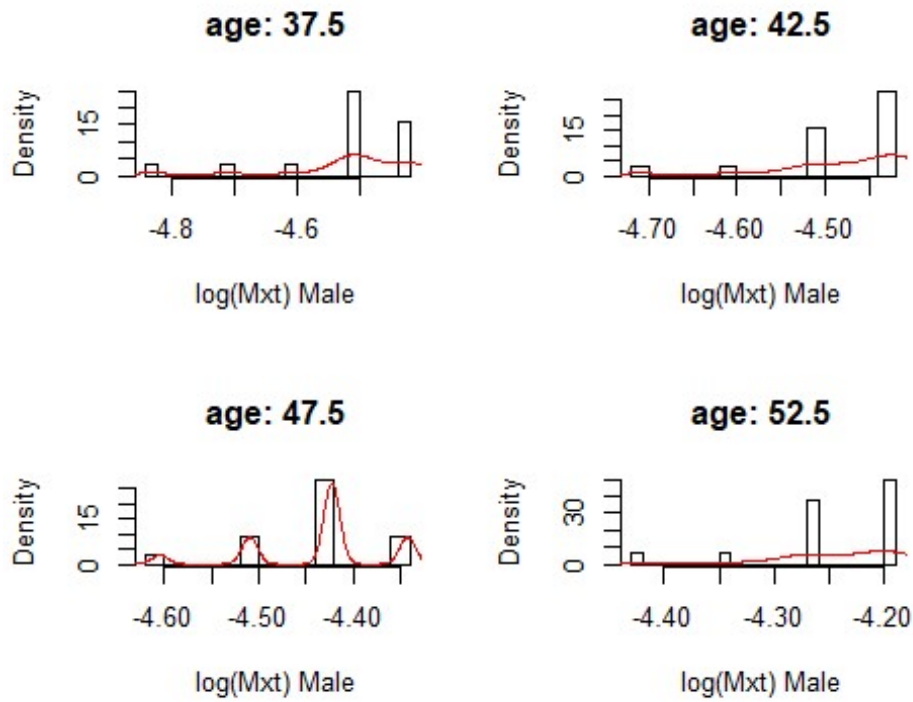


Figure 30: Histogram across age-groups 37.5-52.5 (female log mortality rates)

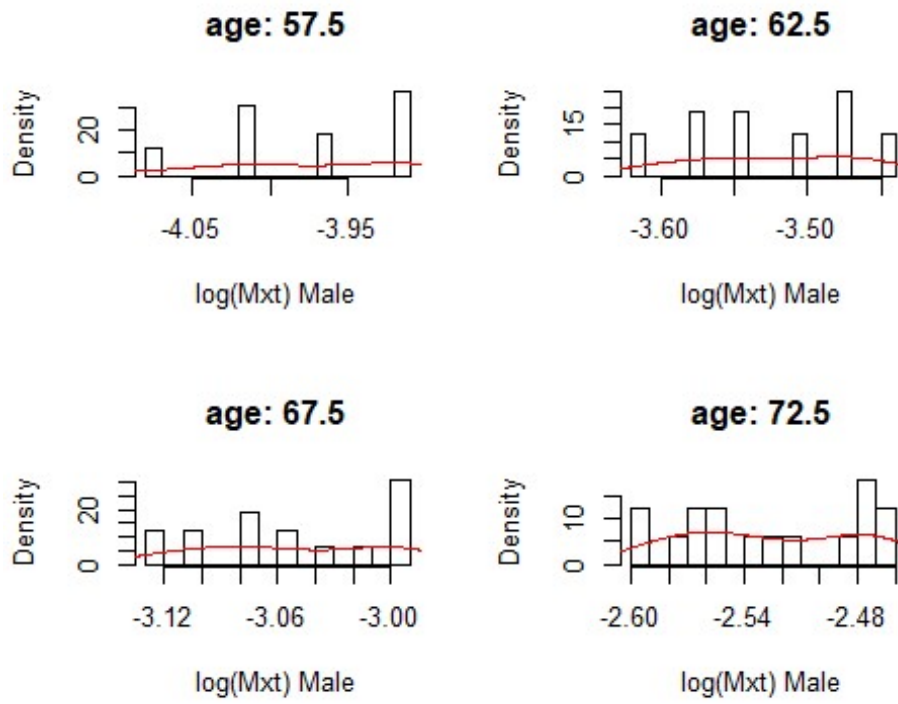


Figure 31: Histogram across age-groups 57.5-72.5 (female log mortality rates)

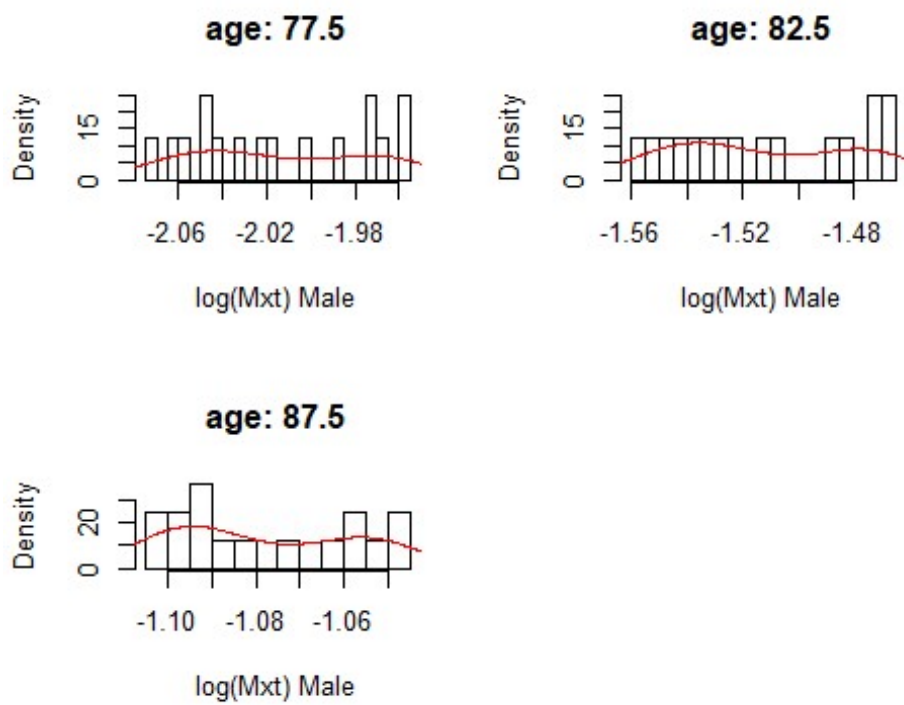


Figure 32: Histogram across age-groups 77.5-87.5 (female log mortality rates)

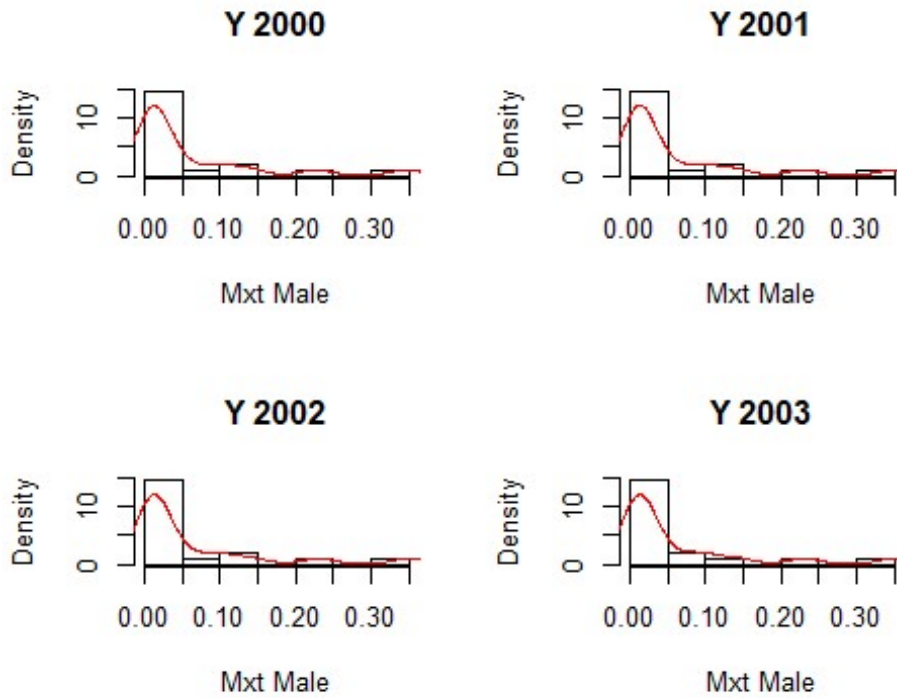


Figure 33: Histogram across period 2000-2003 (female mortality rates)

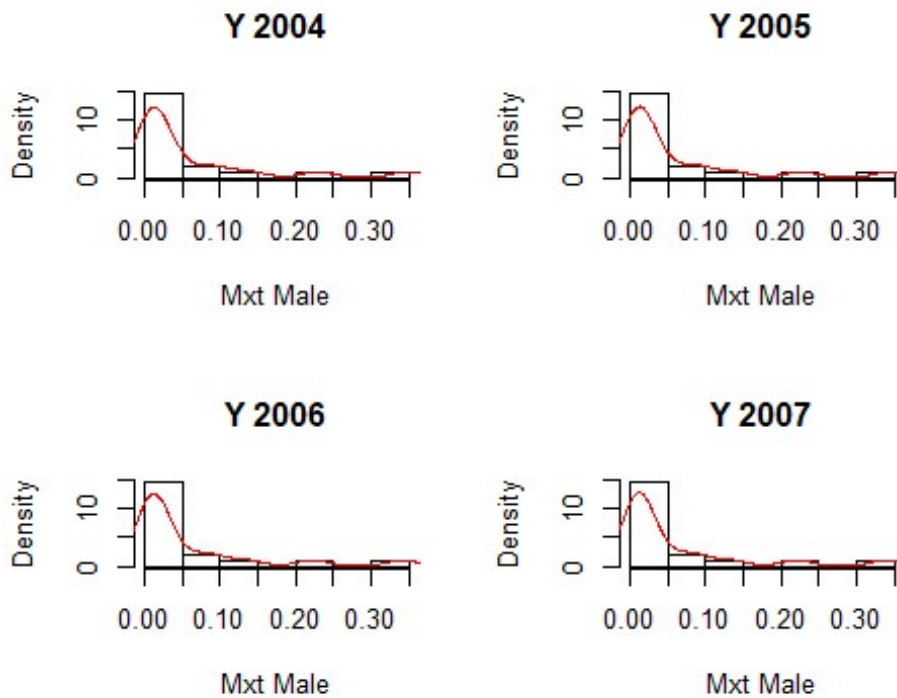


Figure 34: Histogram across period 2004-2007 (female mortality rates)

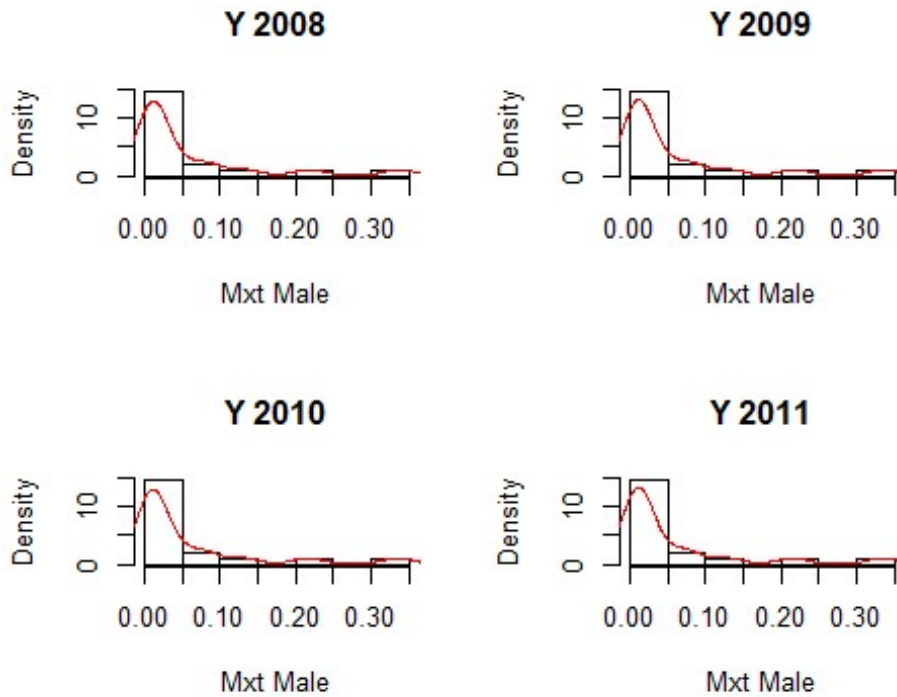


Figure 35: Histogram across period 2008-2011 (female mortality rates)

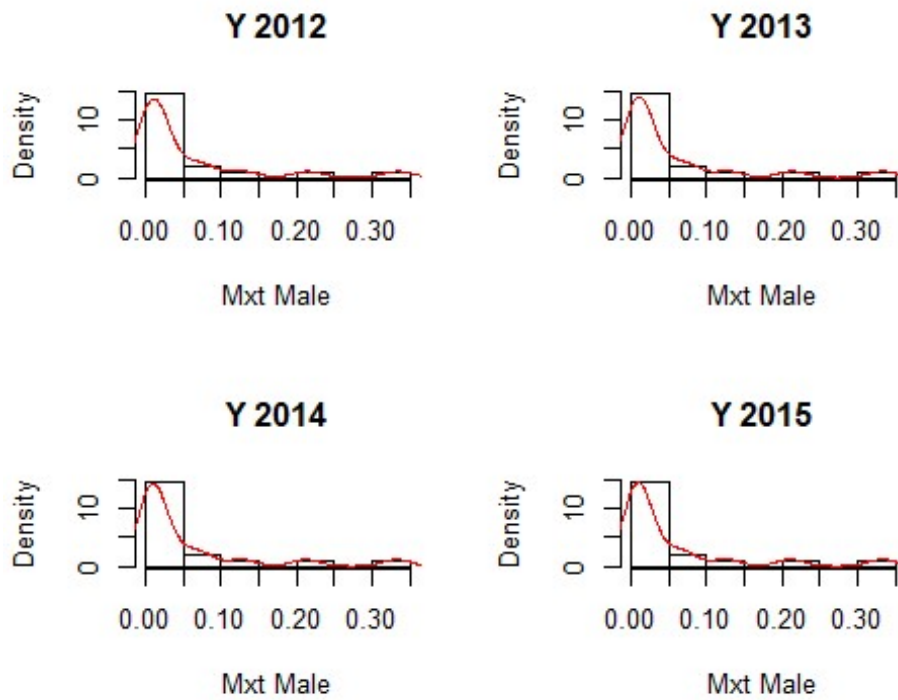


Figure 36: Histogram across period 2012-2015 (female mortality rates)

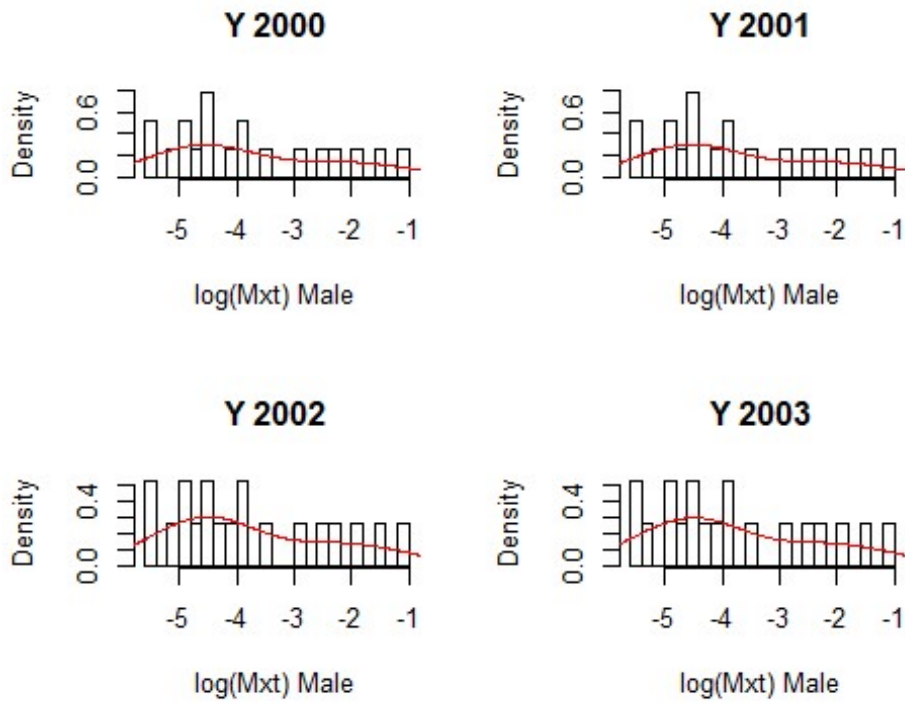


Figure 37: Histogram across period 2000-2003 (female log mortality rates)

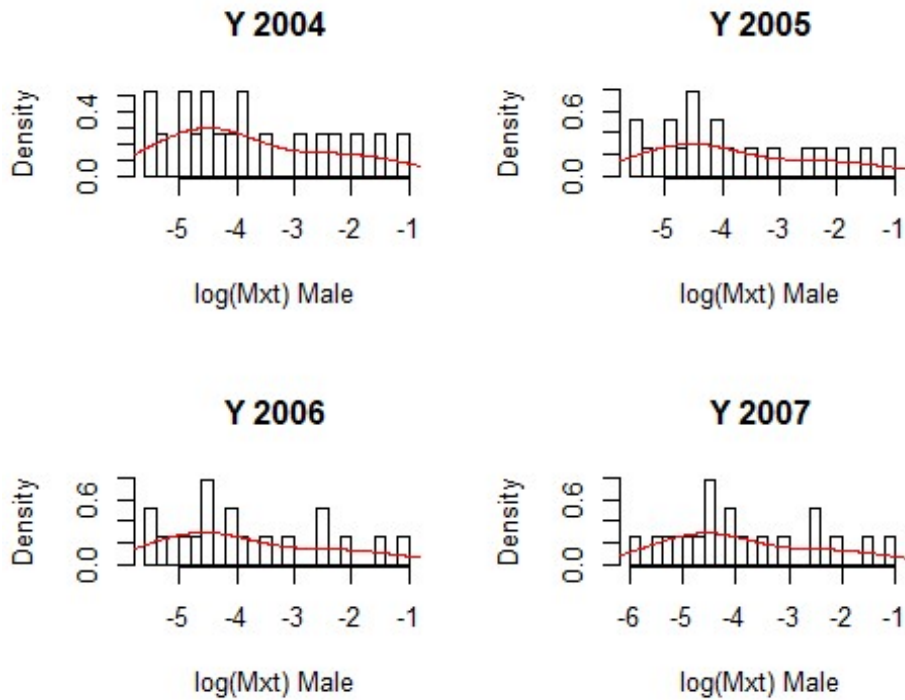


Figure 38: Histogram across period 2004-2007 (female log mortality rates)

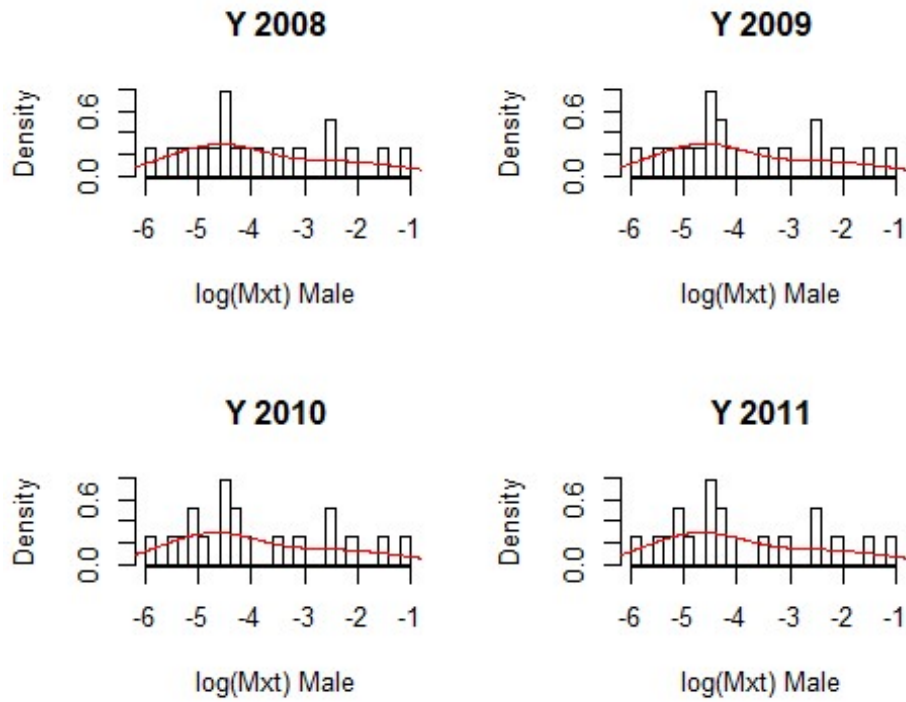


Figure 39: Histogram across period 2008-2011 (female log mortality rates)

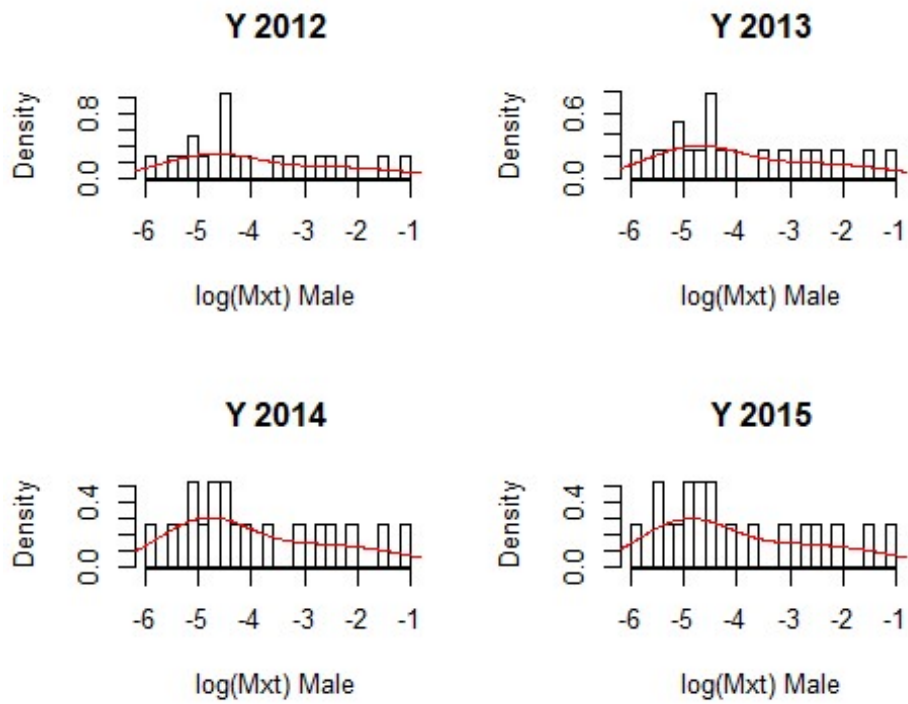


Figure 40: Histogram across period 2012-2015 (female log mortality rates)

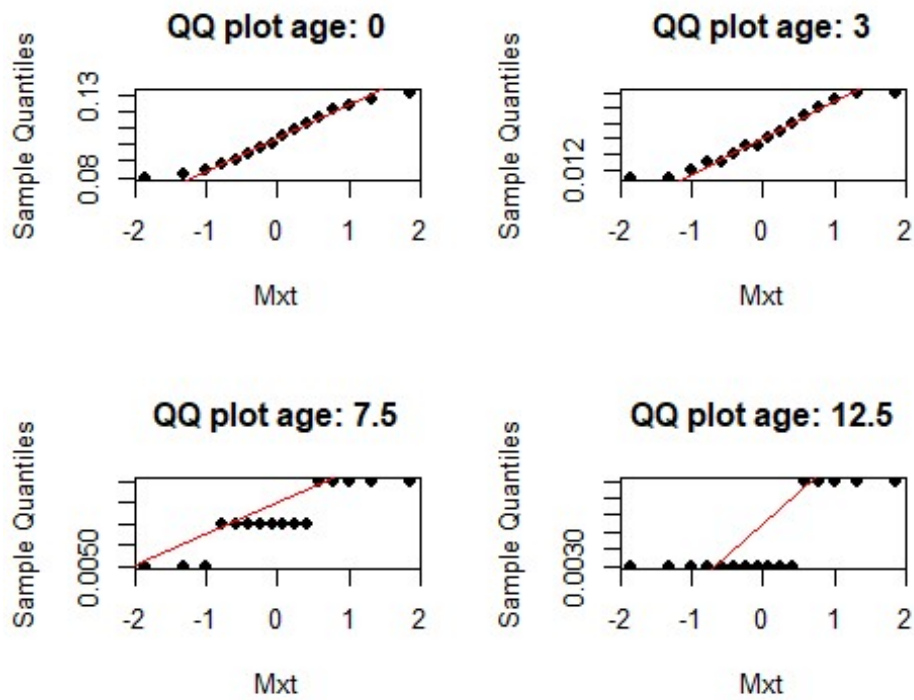


Figure 41: QQ plots across ages 0-12.5 (male mortality rates)



Figure 42: QQ plots across ages 17.5-32.5 (male mortality rates)

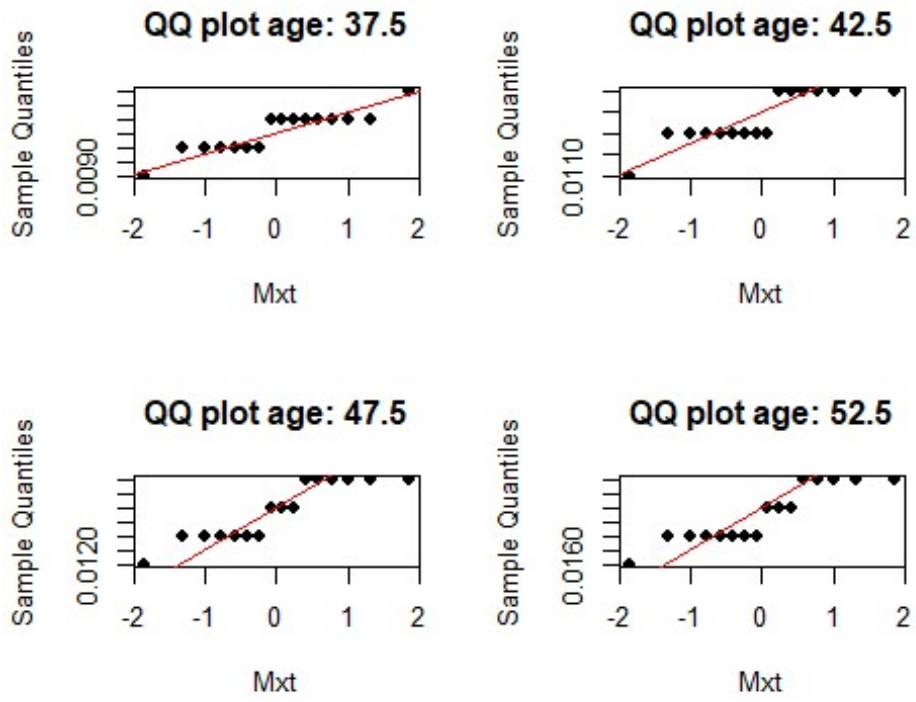


Figure 43: QQ plots across ages 37.5-52.5 (male mortality rates)

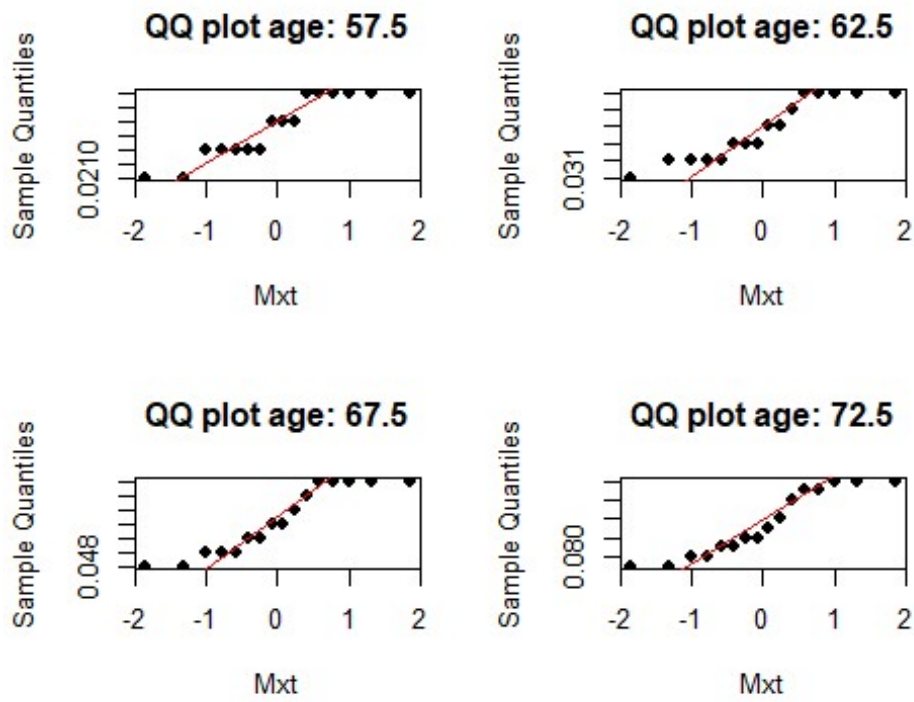


Figure 44: QQ plots across ages 57.5-72.5 (male mortality rates)

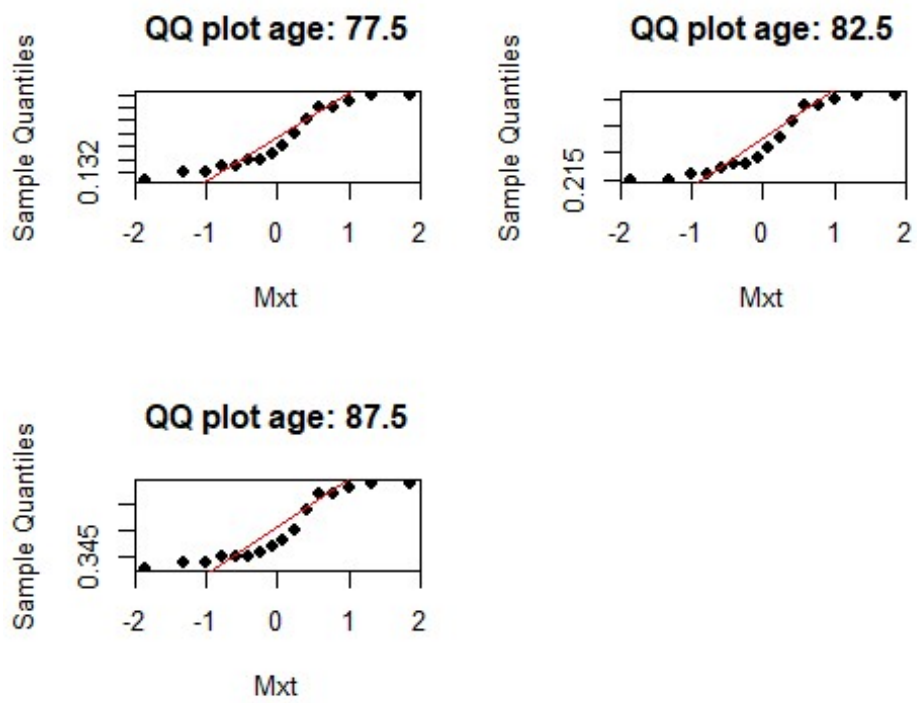


Figure 45: QQ plots across ages 77.5-87.5 (male mortality rates)

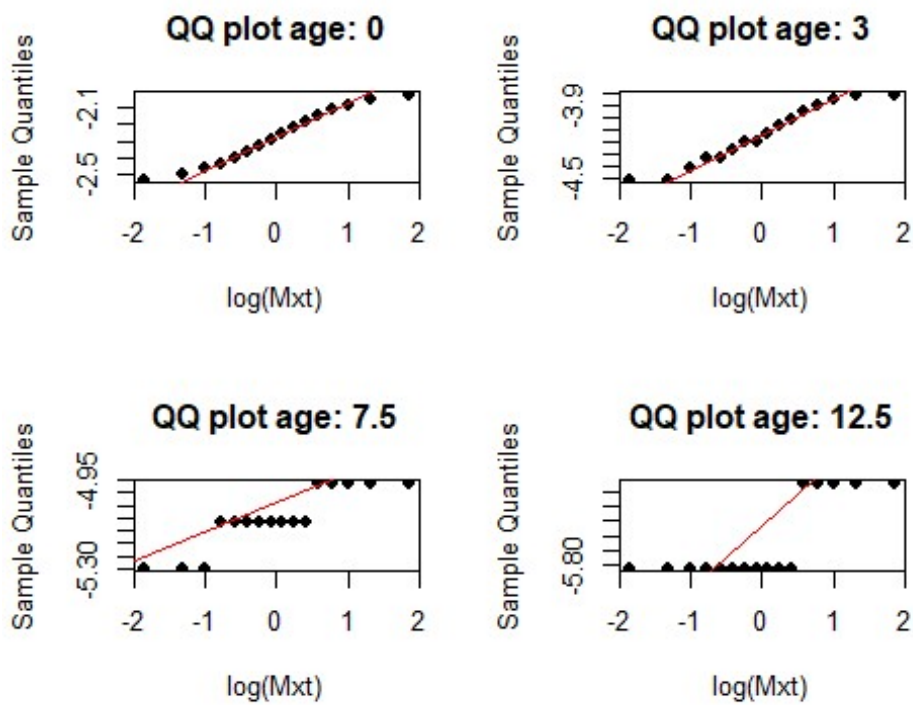


Figure 46: QQ plots across ages 0-12.5 (male log mortality rates)

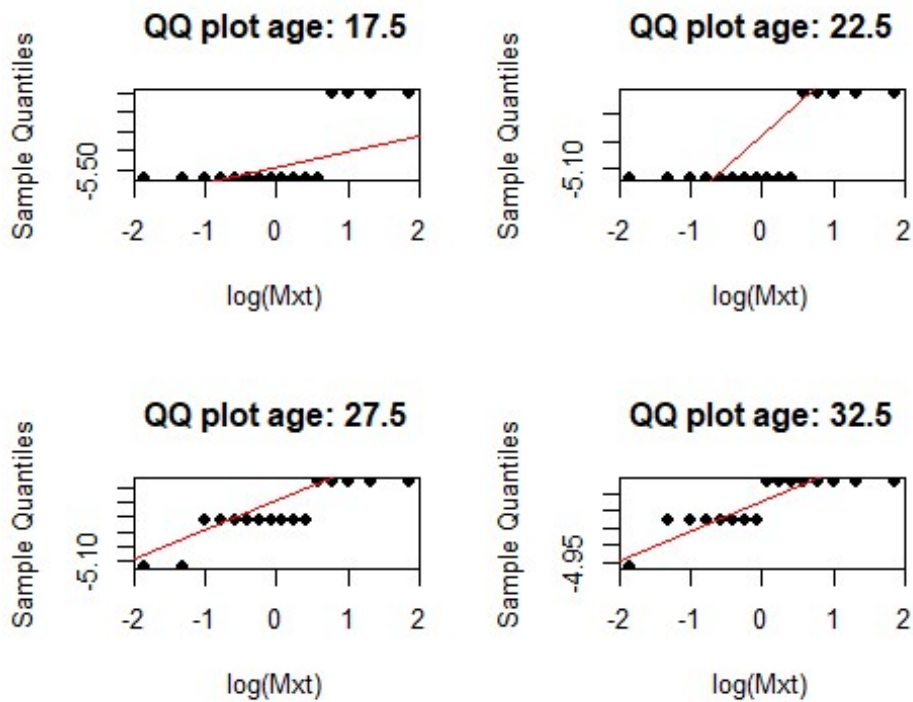


Figure 47: QQ plots across ages 17.5-32.5 (male log mortality rates)

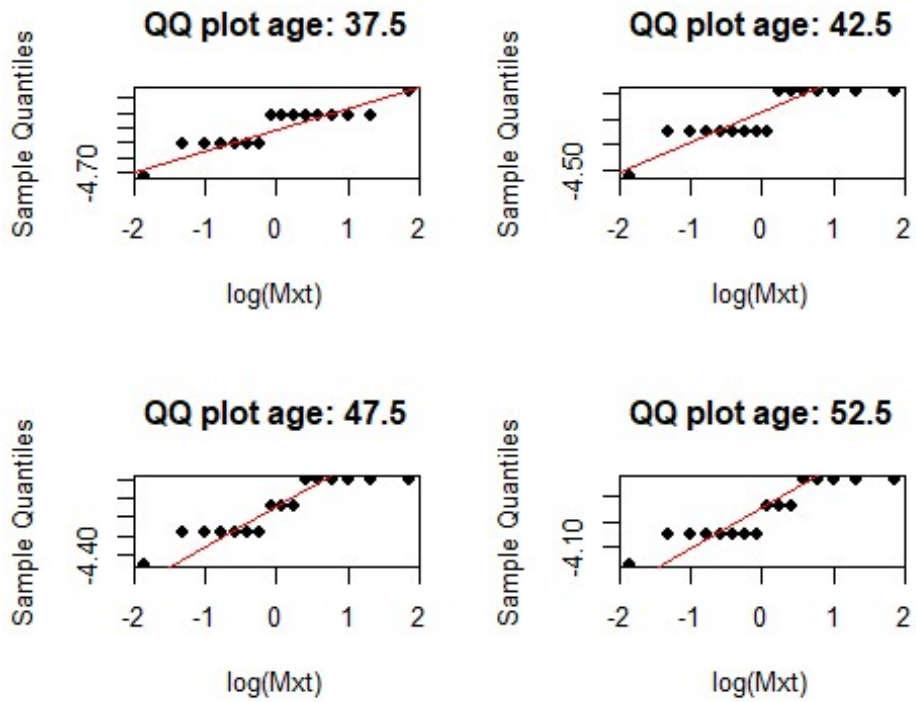


Figure 48: QQ plots across ages 37.5-52.5 (male mortality rates)

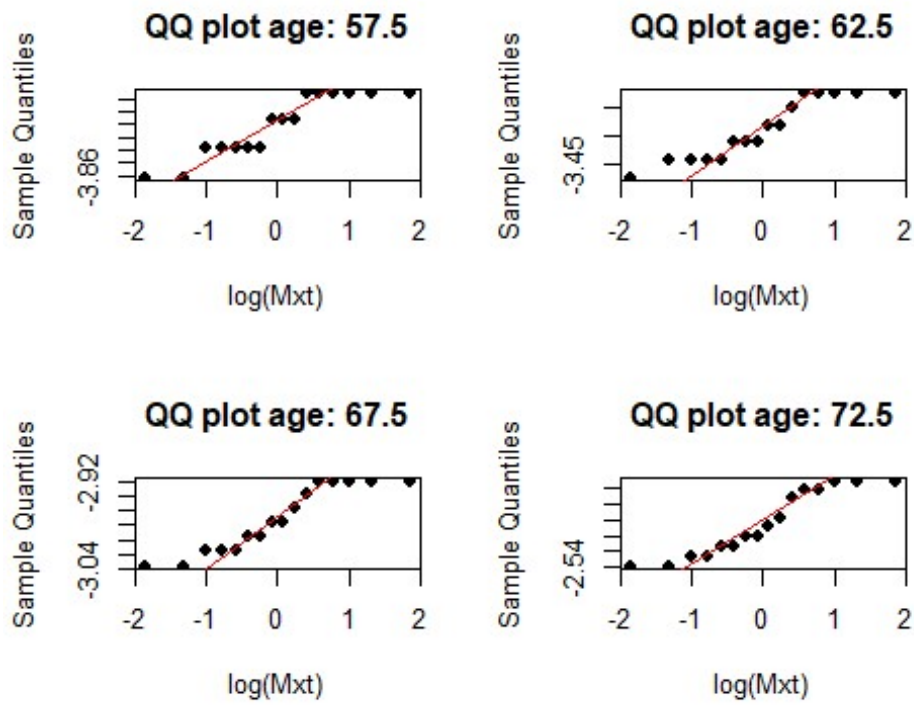


Figure 49: QQ plots across ages 57.5-72.5 (male log mortality rates)

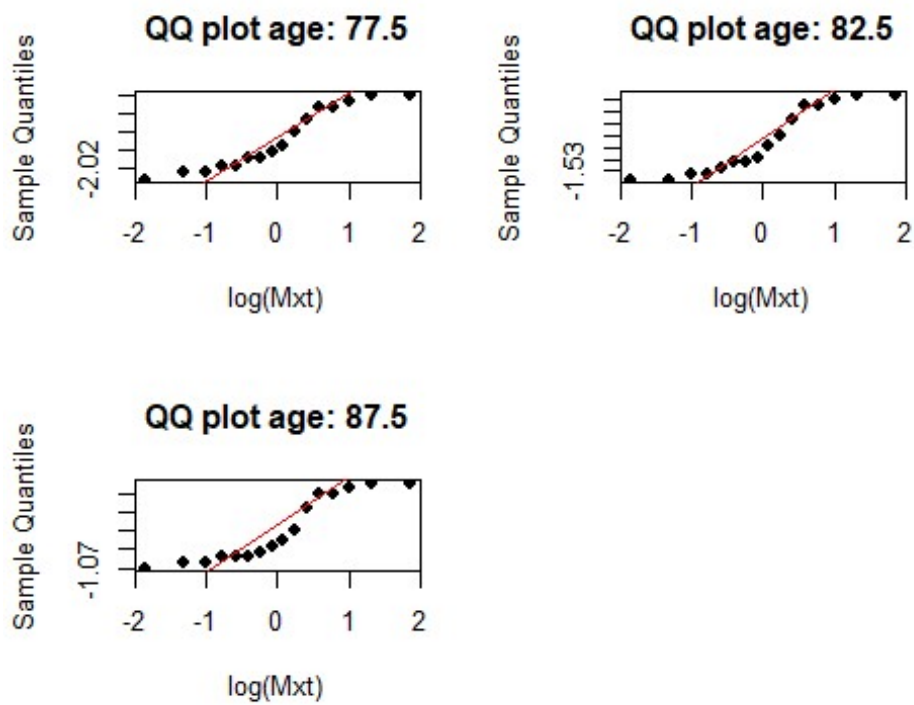


Figure 50: QQ Plots across age-groups 77.5-87.5 (male log mortality rates)

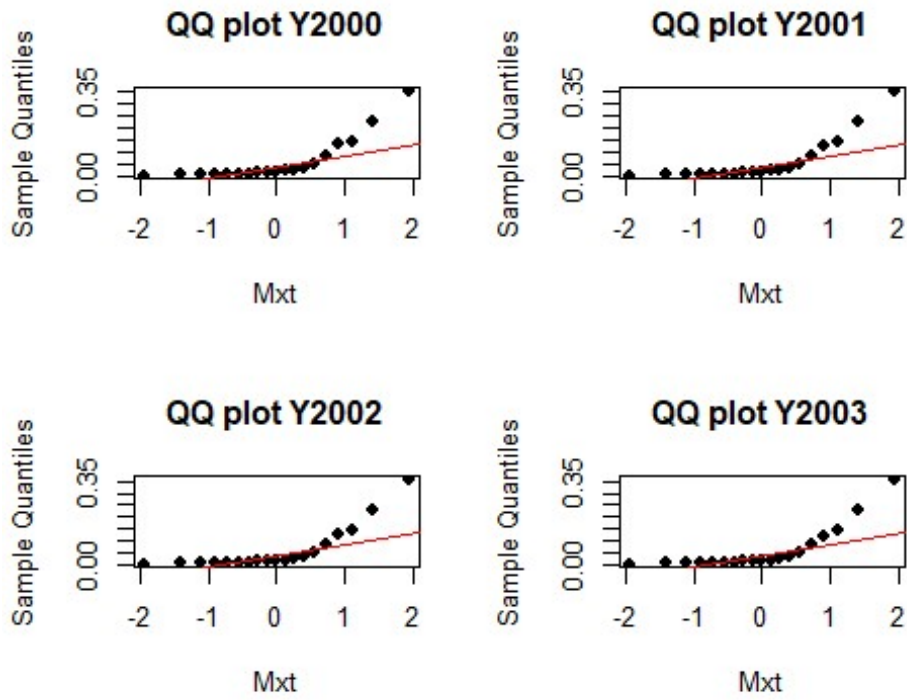


Figure 51: QQ Plots across period 2000-2003 (male mortality rates)

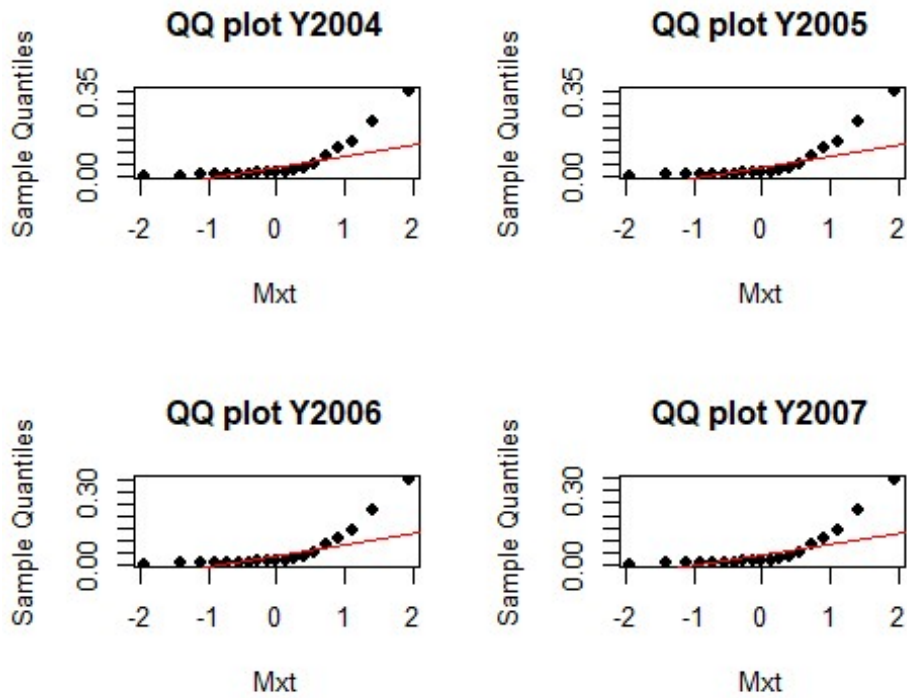


Figure 52: QQ Plots across period 2004-2007 (male mortality rates)

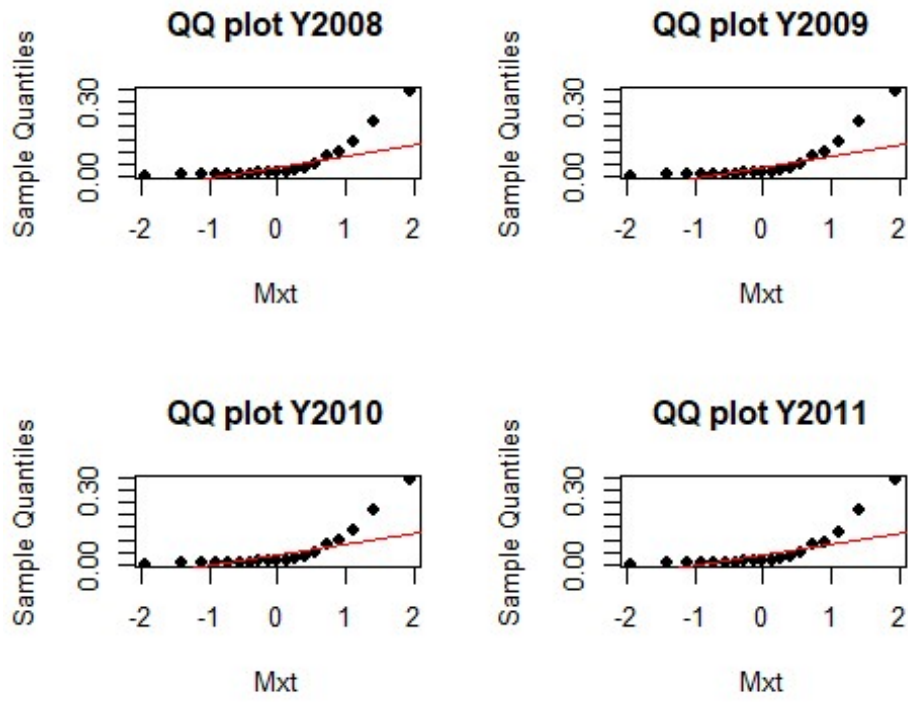


Figure 53: QQ Plots across period 2008-2011 (male mortality rates)

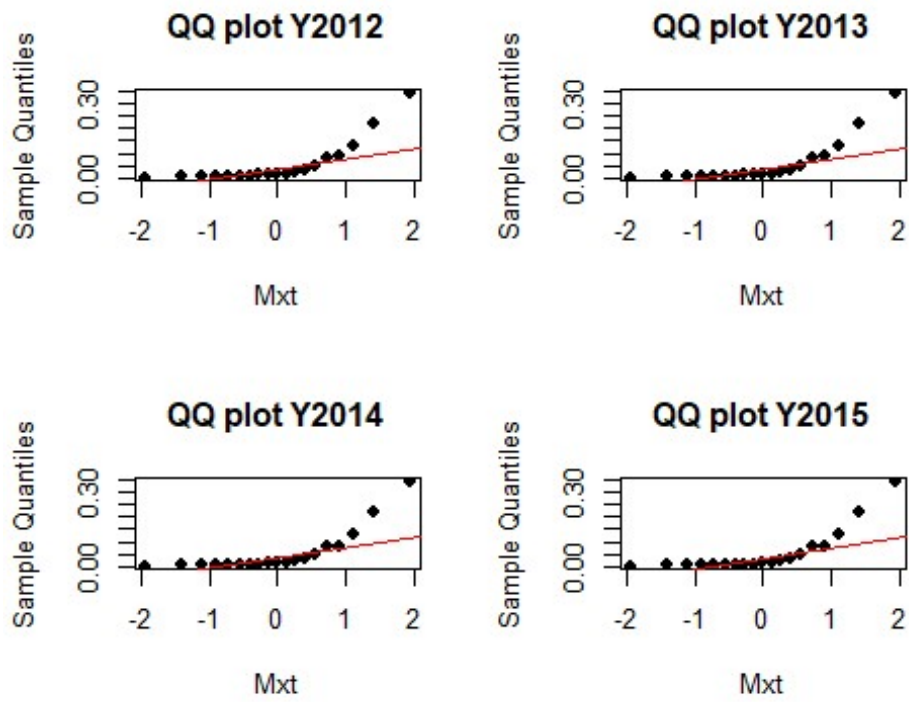


Figure 54: QQ Plots across period 2012-2015 (male mortality rates)

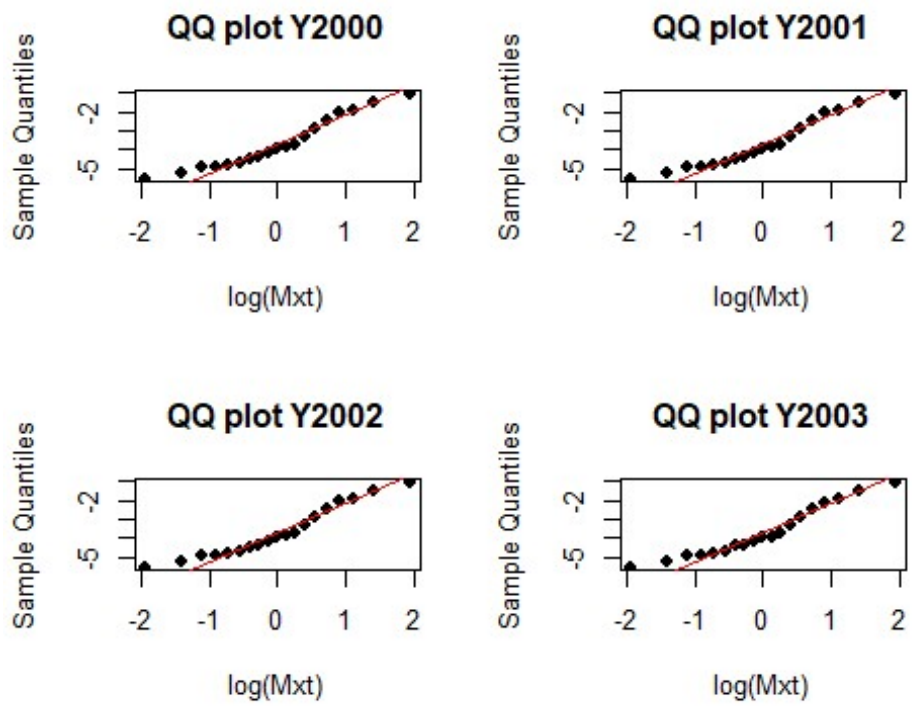


Figure 55: QQ Plots across period 2000-2003 (male log mortality rates)

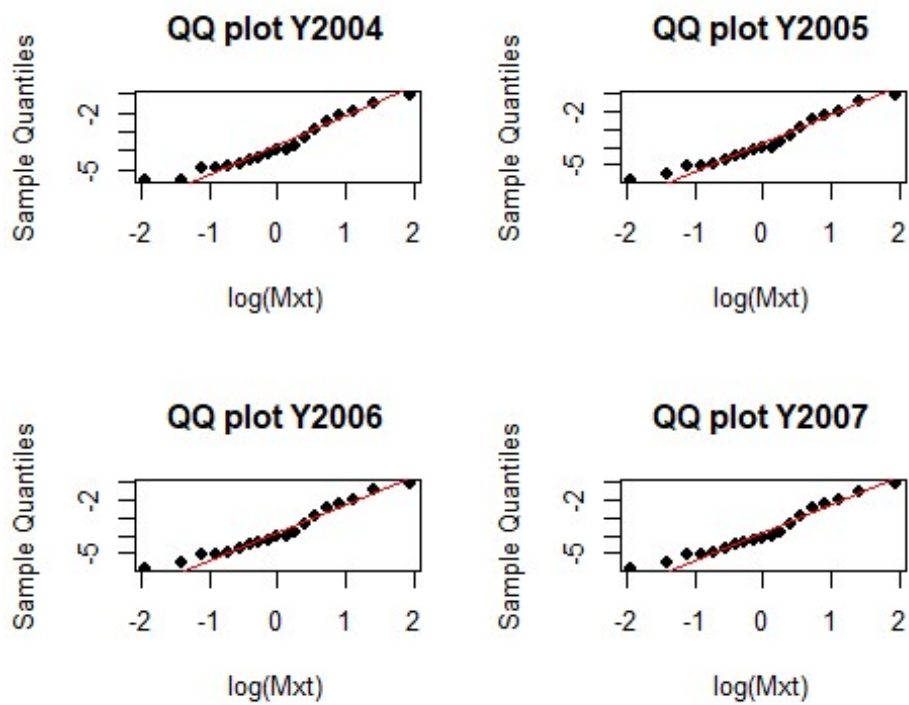


Figure 56: QQ Plots across period 2004-2007 (male log mortality rates)

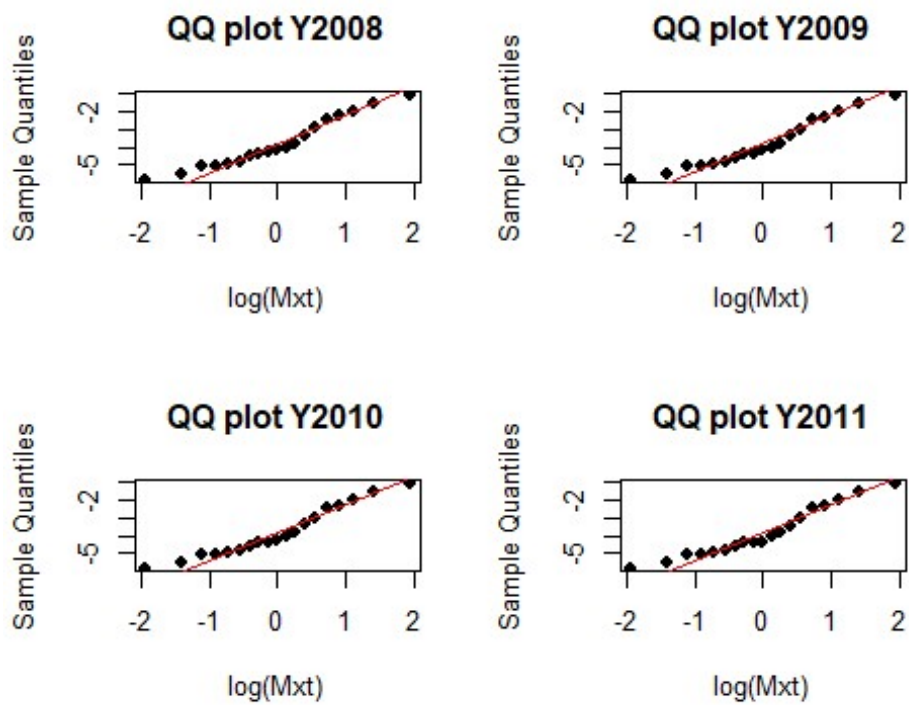


Figure 57: QQ Plots across period 2008-2011 (male log mortality rates)

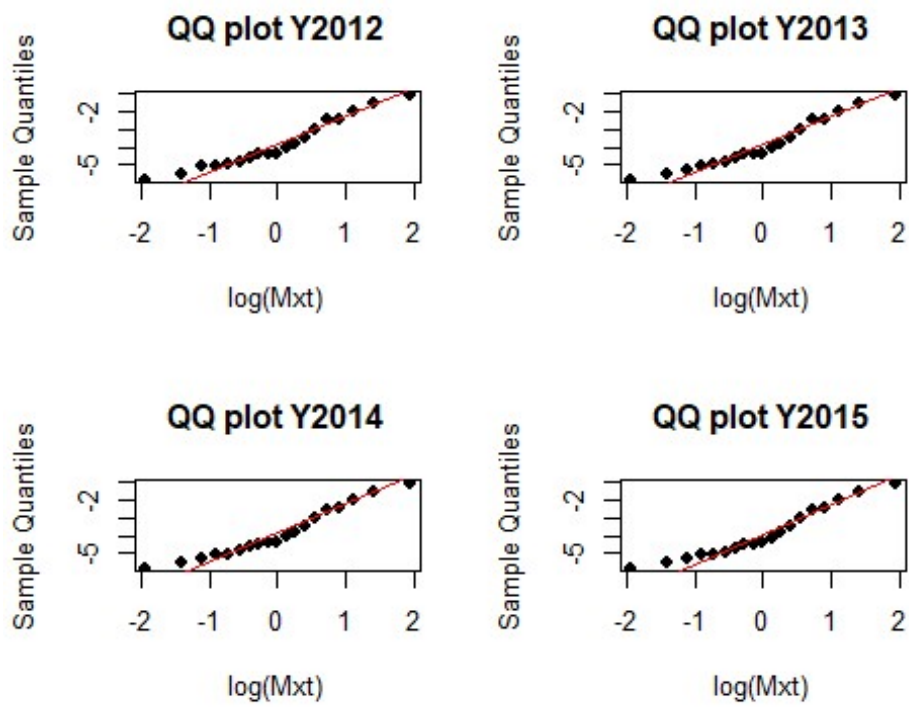


Figure 58: QQ Plots across period 2012-2015 (male log mortality rates)

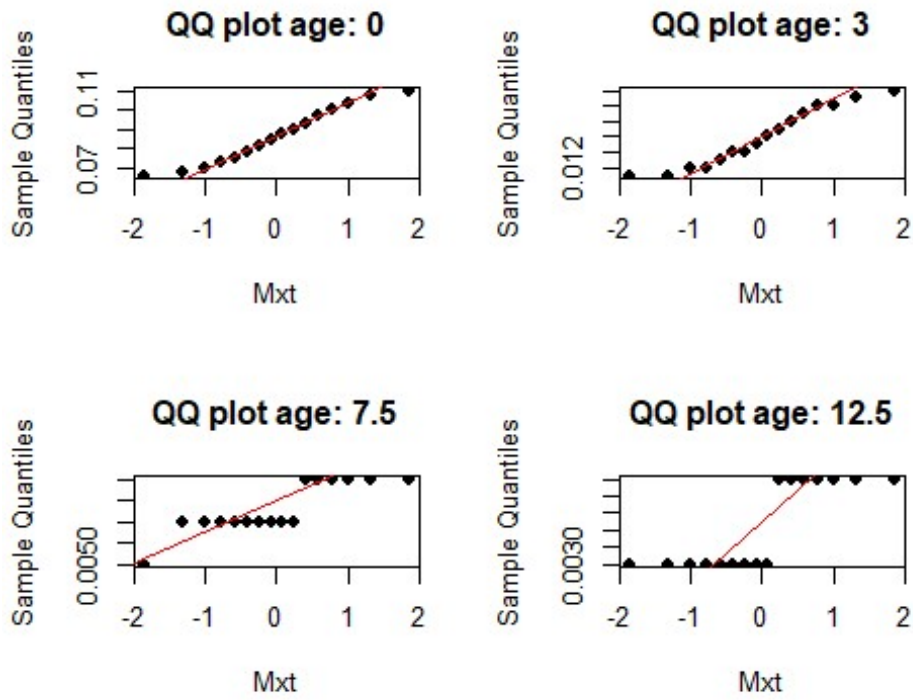


Figure 59: QQ Plots across age-groups 0-12.5 (female mortality rates)

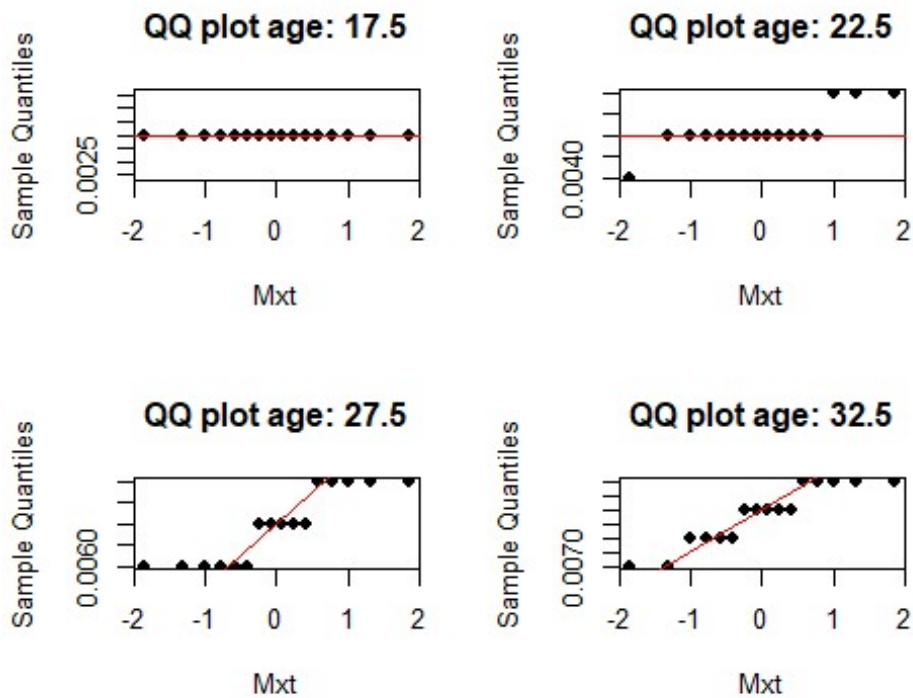


Figure 60: QQ Plots across age-groups 17.5-32.5 (female mortality rates)

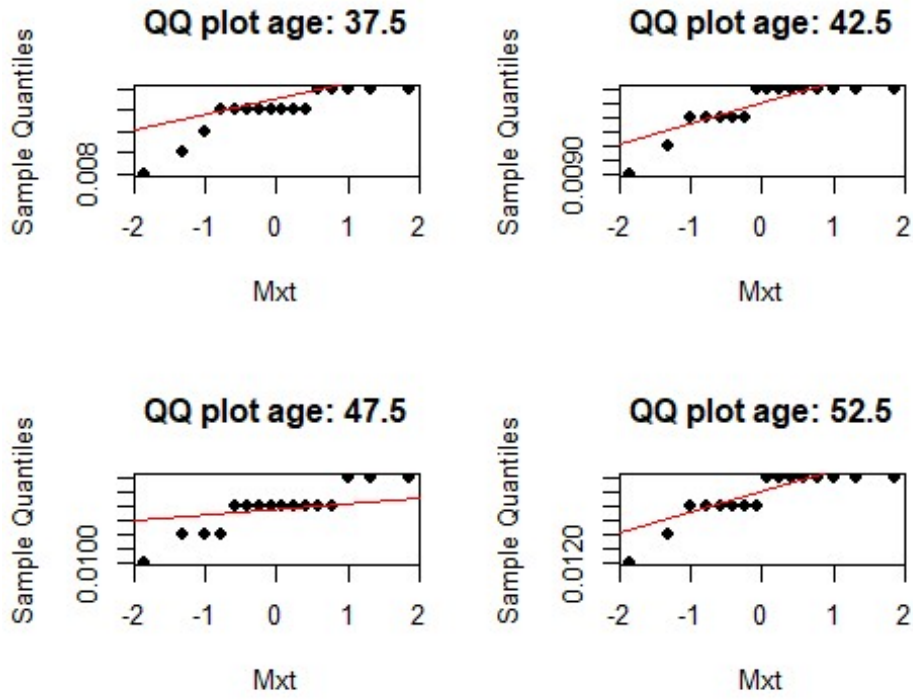


Figure 61: QQ Plots across age-groups 37.5-52.5 (female mortality rates)

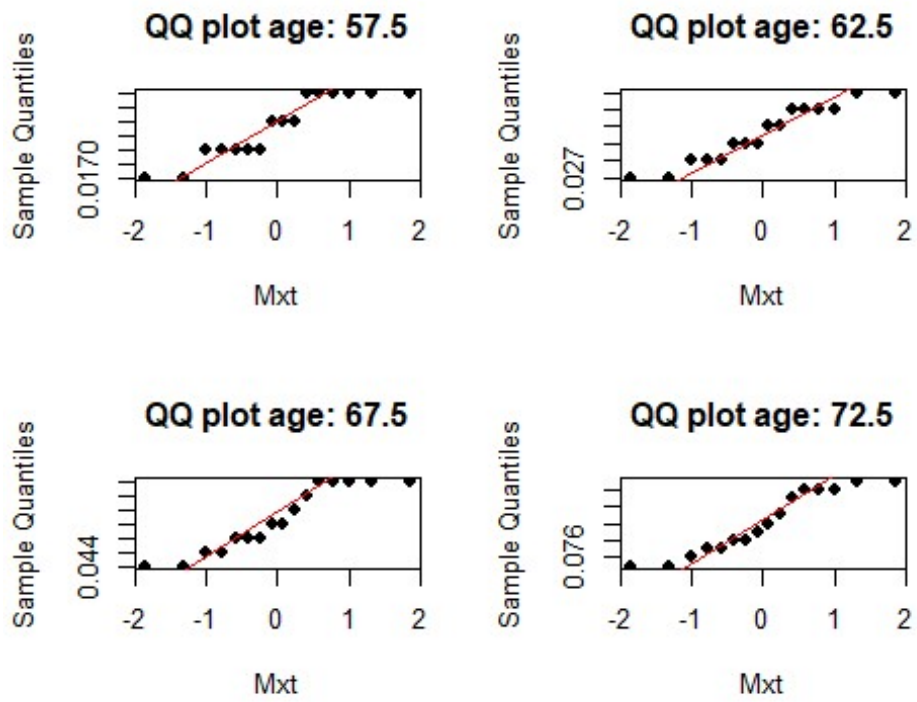


Figure 62: QQ Plots across age-groups 57.5-72.5 (female mortality rates)



Figure 63: QQ Plots across age-groups 77.5-87.5 (female mortality rates)

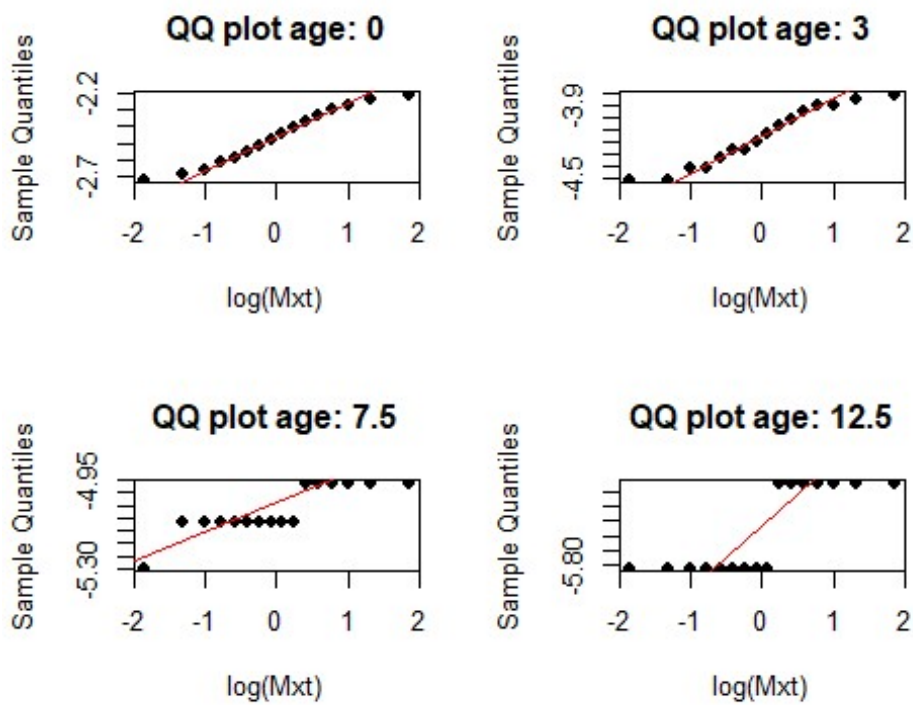


Figure 64: QQ Plots across age-groups 0-12.5 (female log mortality rates)

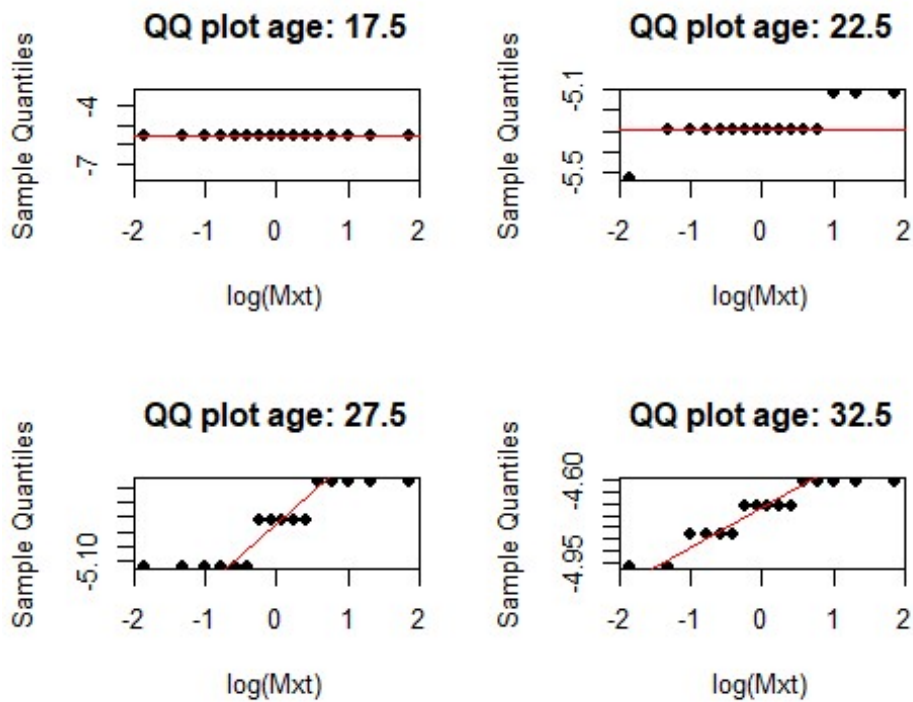


Figure 65: QQ Plots across age-groups 17.5-32.5 (female log mortality rates)

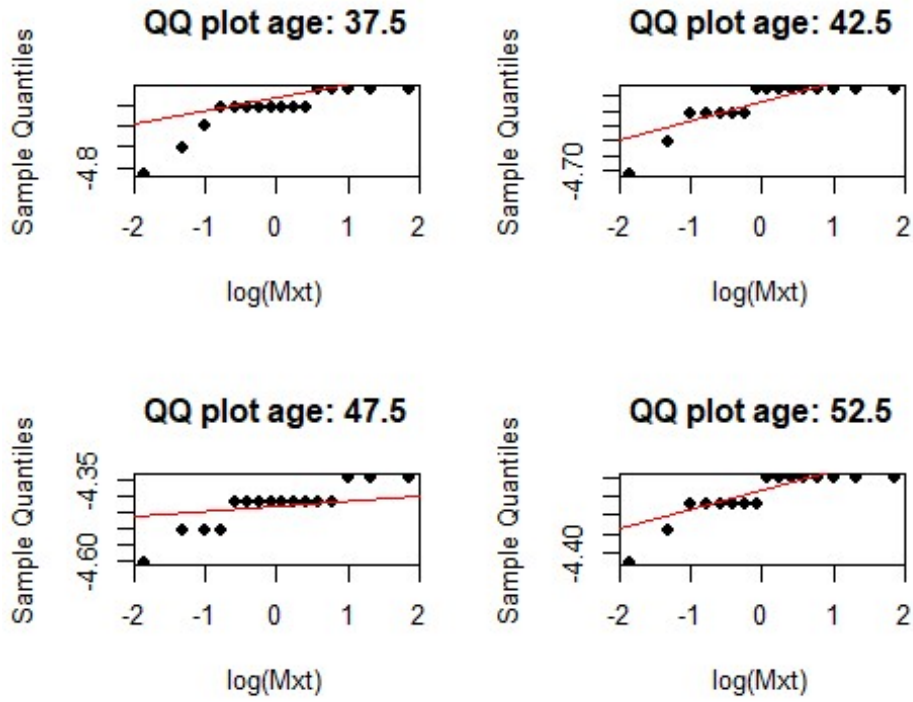


Figure 66: QQ Plots across age-groups 37.5-52.5 (female log mortality rates)

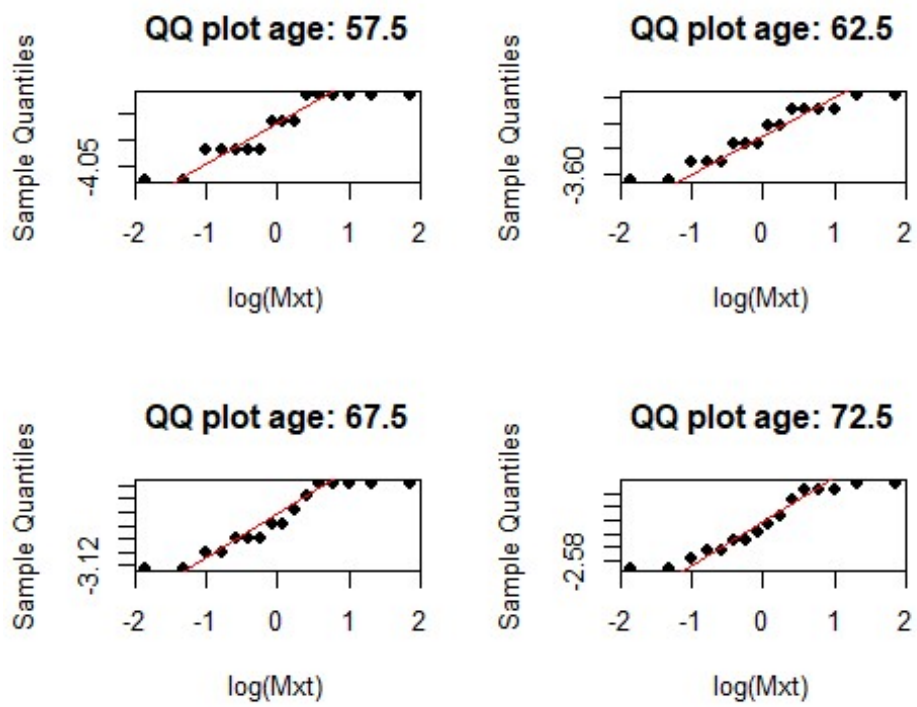


Figure 67: QQ Plots across age-groups 57.5-72.5 (female log mortality rates)

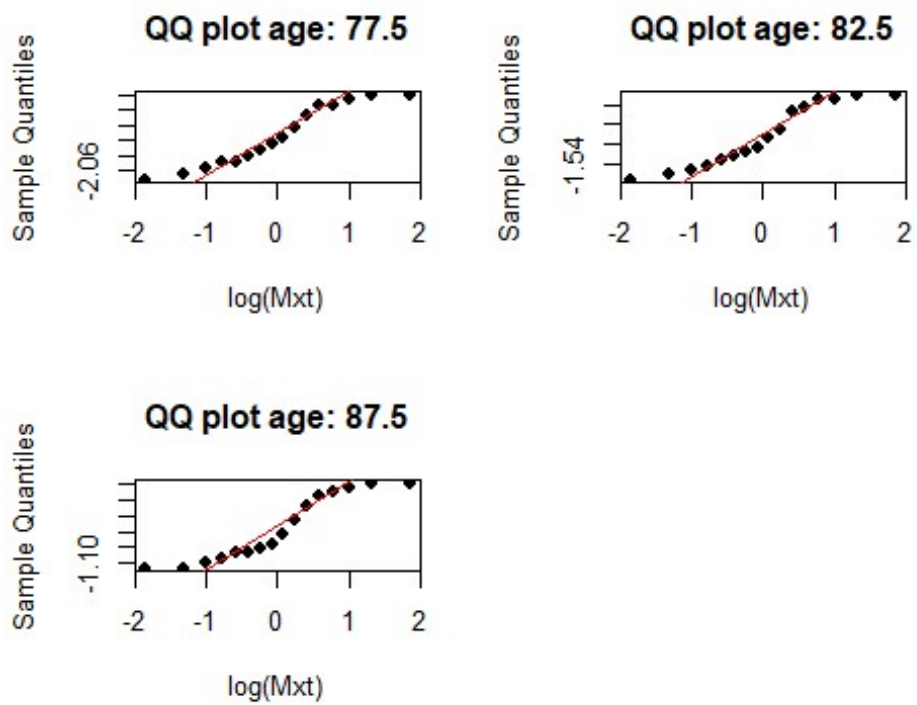


Figure 68: QQ Plots across age-groups 77.5-87.5 (female log mortality rates)

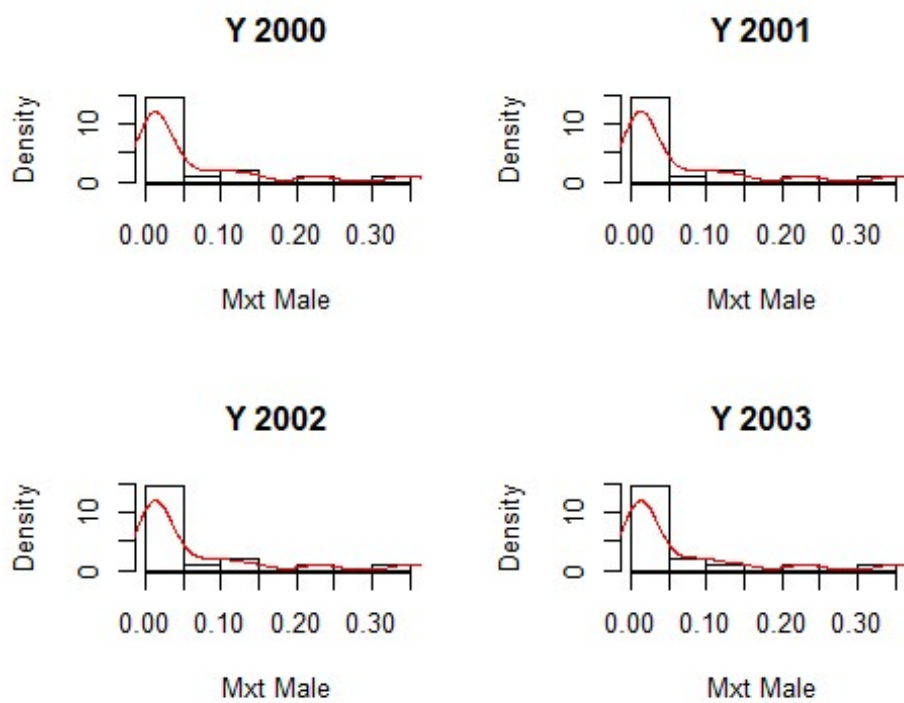


Figure 69: QQ Plots across period 2000-2003 (female mortality rates)

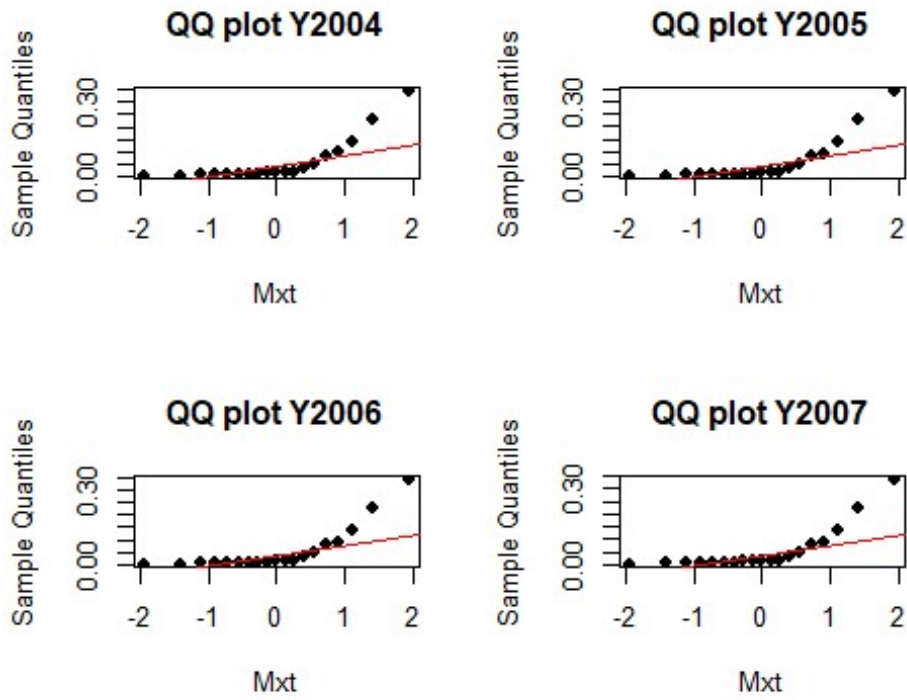


Figure 70: QQ Plots across period 2004-2007 (female mortality rates)

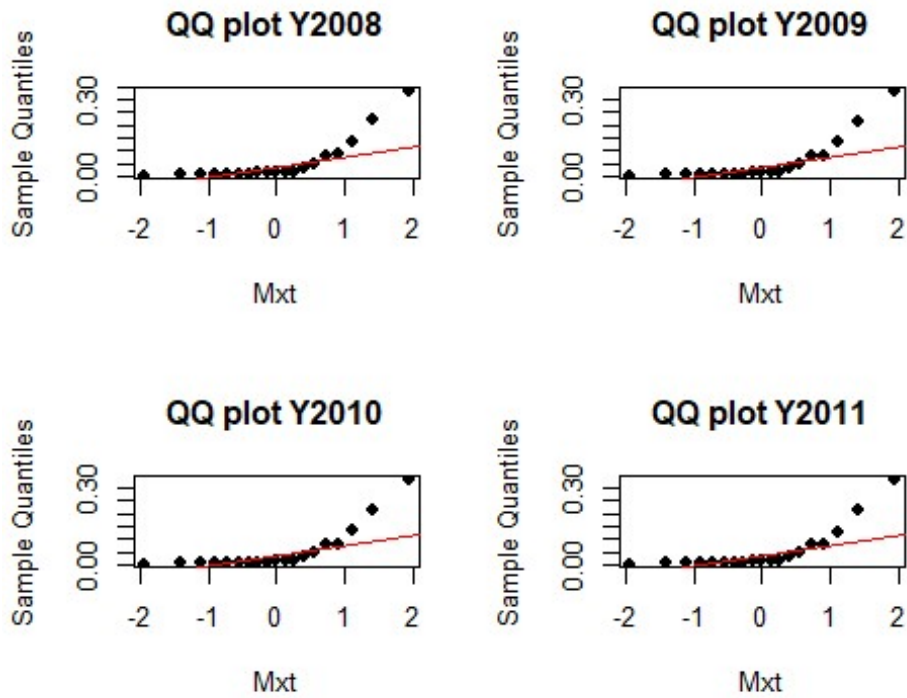


Figure 71: QQ Plots across period 2008-2011 (female mortality rates)

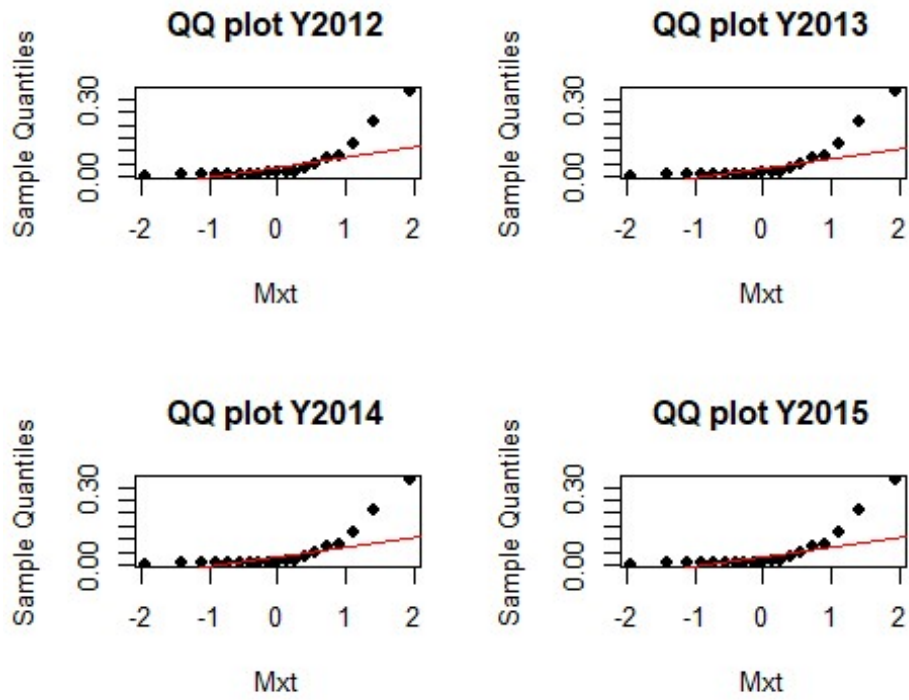


Figure 72: QQ Plots across period 2012-2015 (female mortality rates)

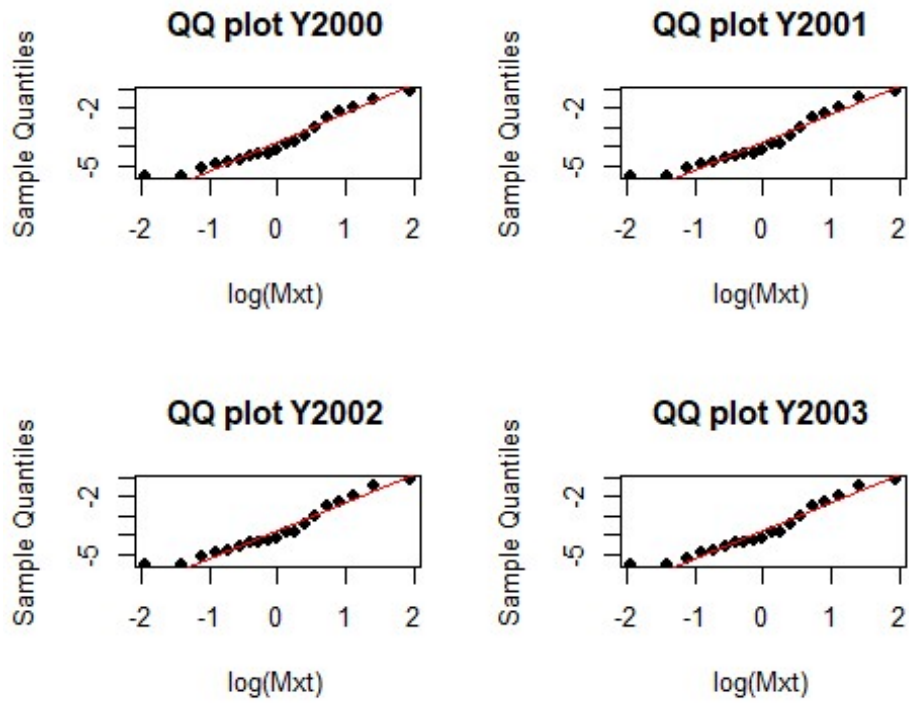


Figure 73: QQ Plots across period 2001-2003 (female log mortality rates)

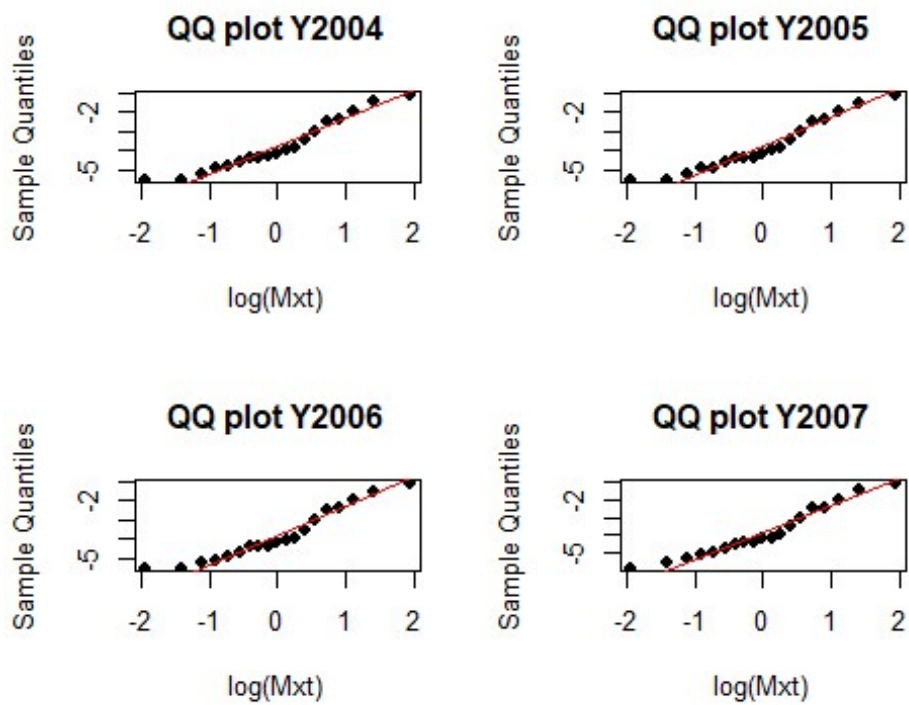


Figure 74: QQ Plots across period 2004-2007 (female log mortality rates)

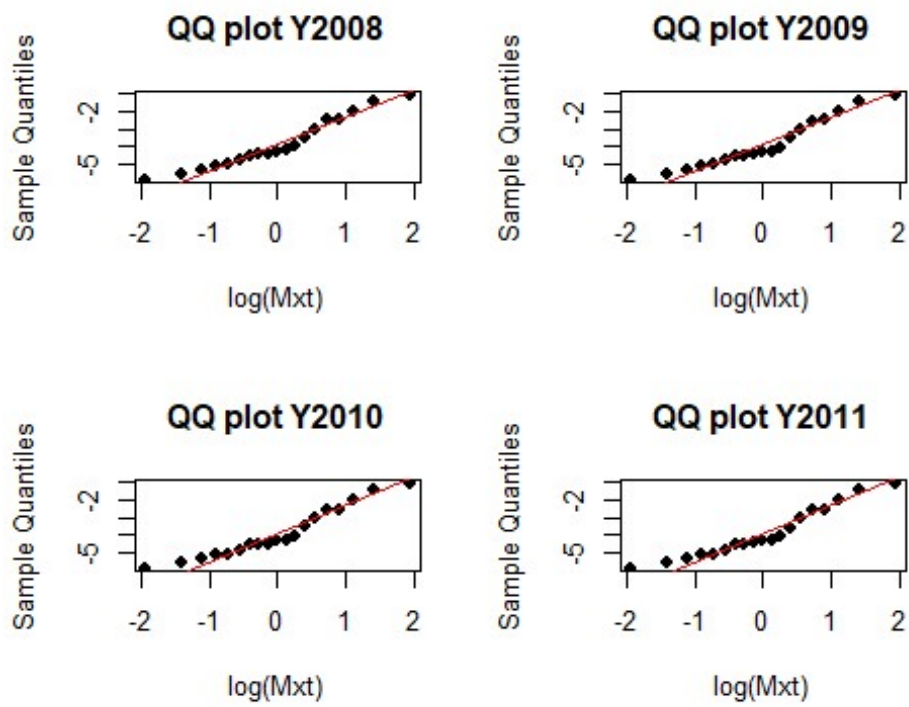


Figure 75: QQ Plots across period 2008-2011 (female log mortality rates)

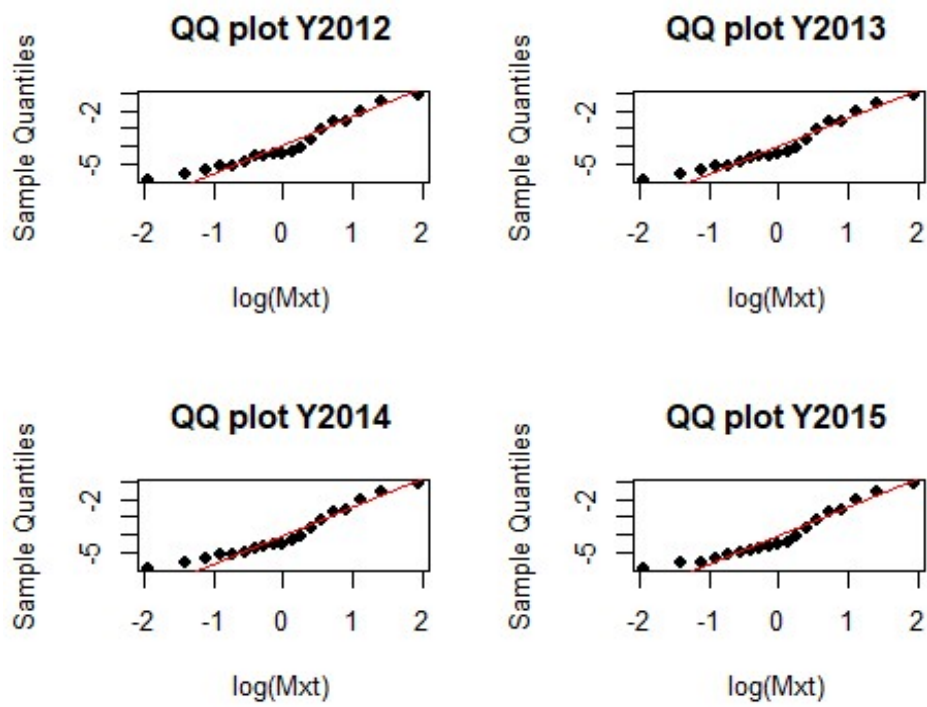


Figure 76: QQ Plots across period 2012-2015 (female log mortality rates)

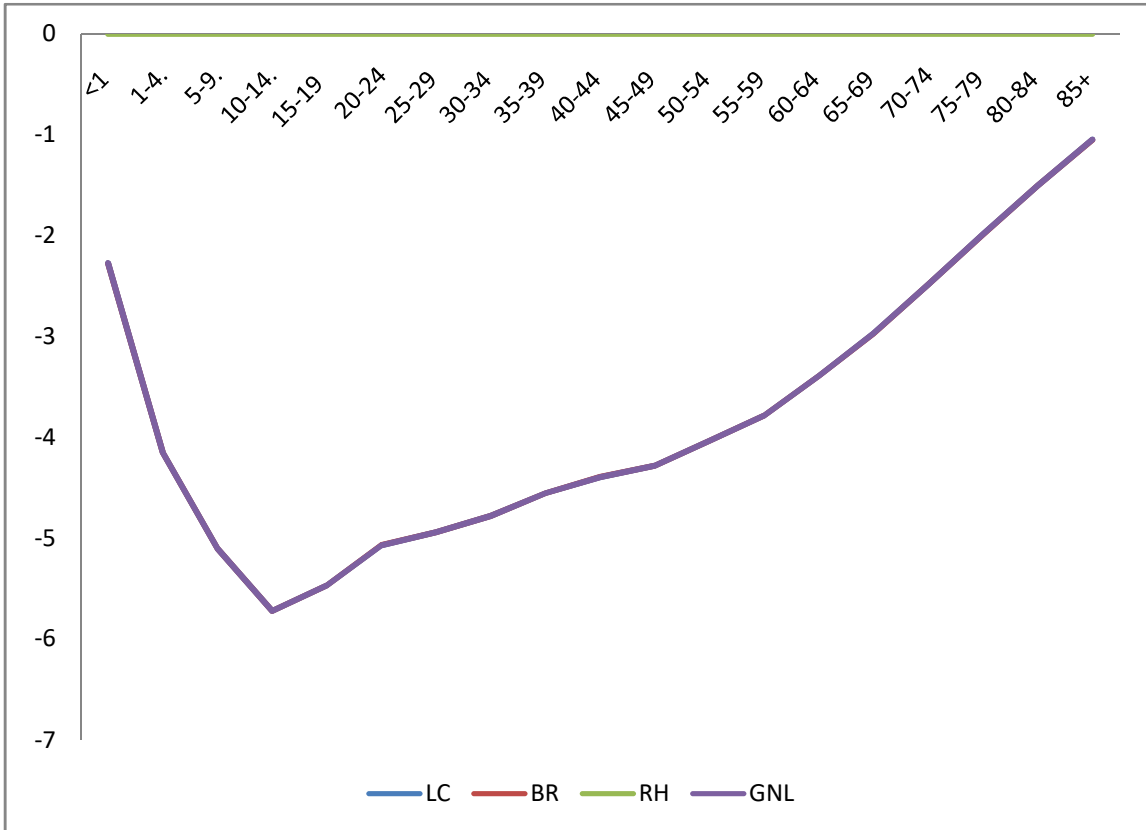


Figure 77: Parameter α_x for male

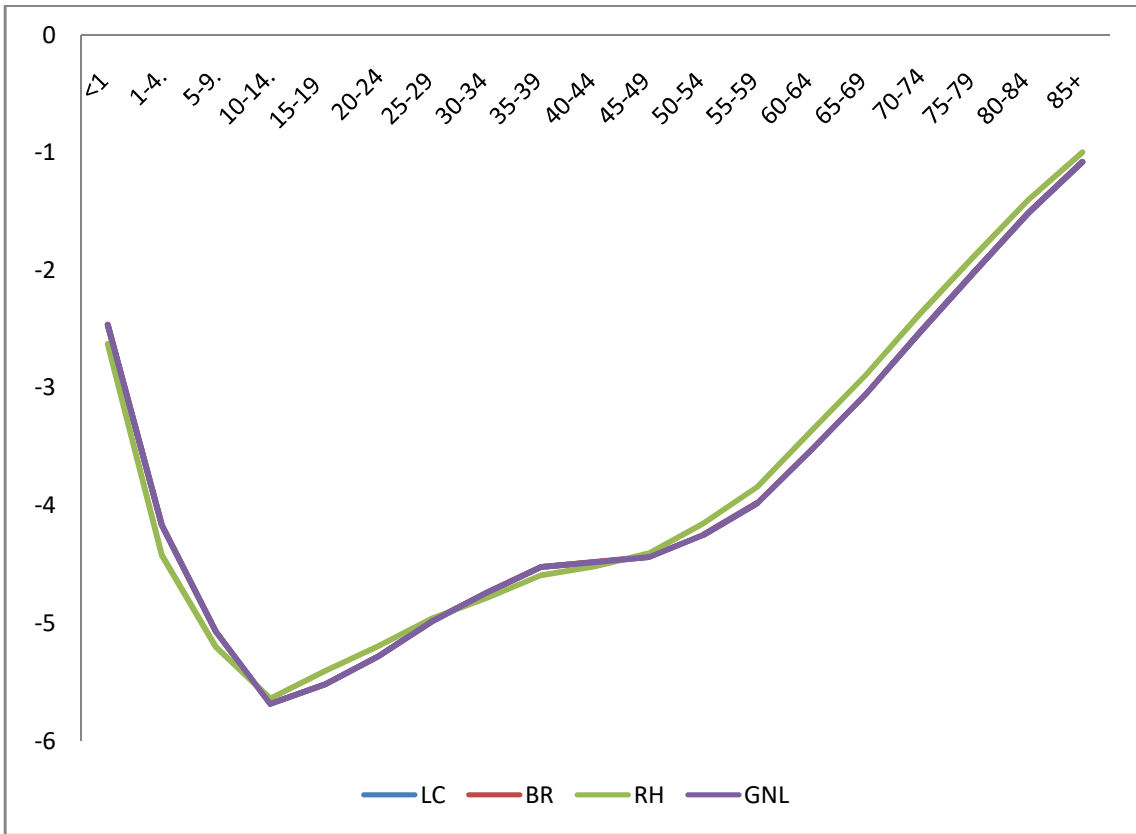


Figure 78: Parameter α_x for female

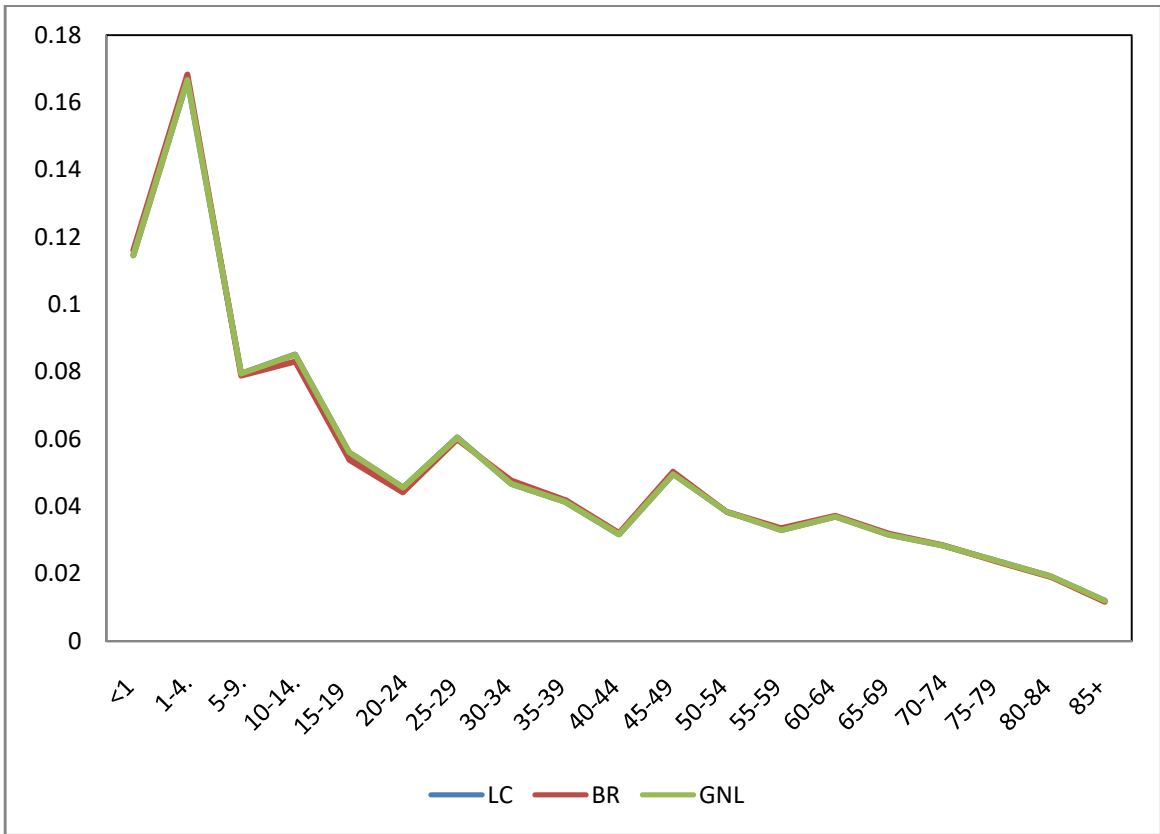


Figure 79: Parameter b_x for male

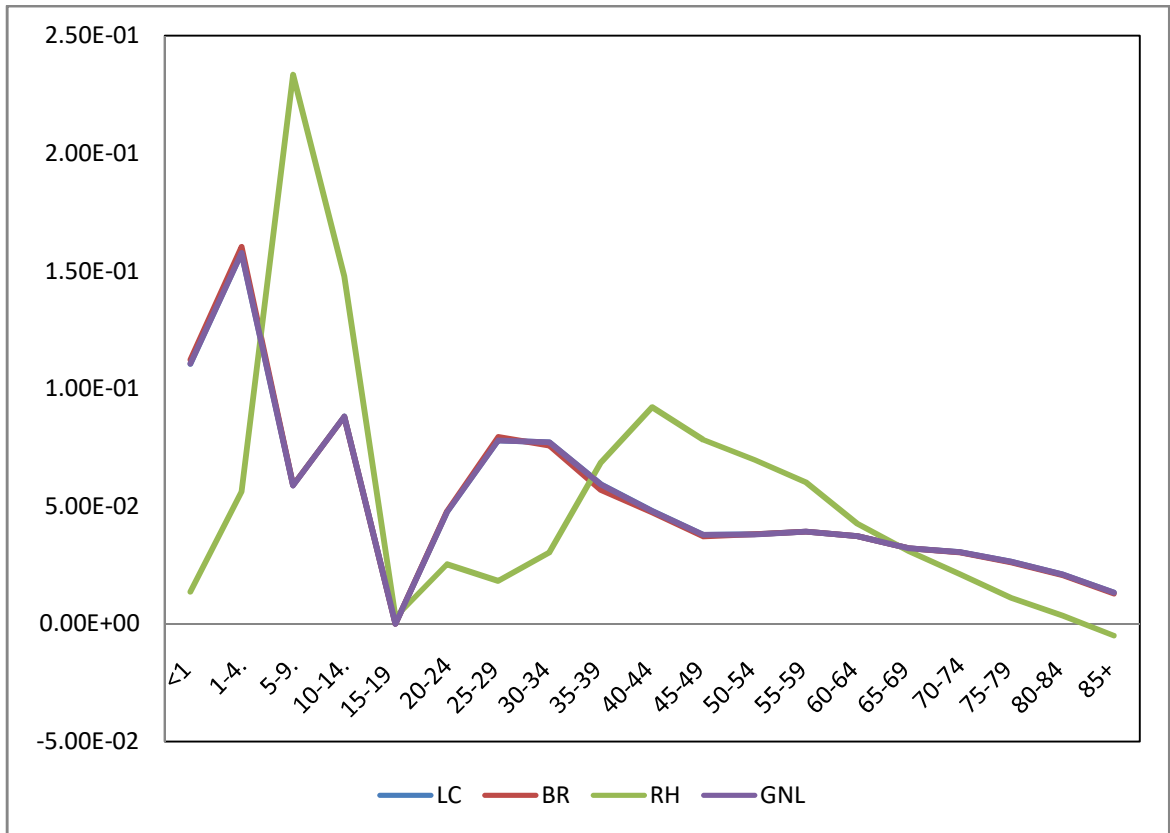


Figure 80: Parameter b_x for female

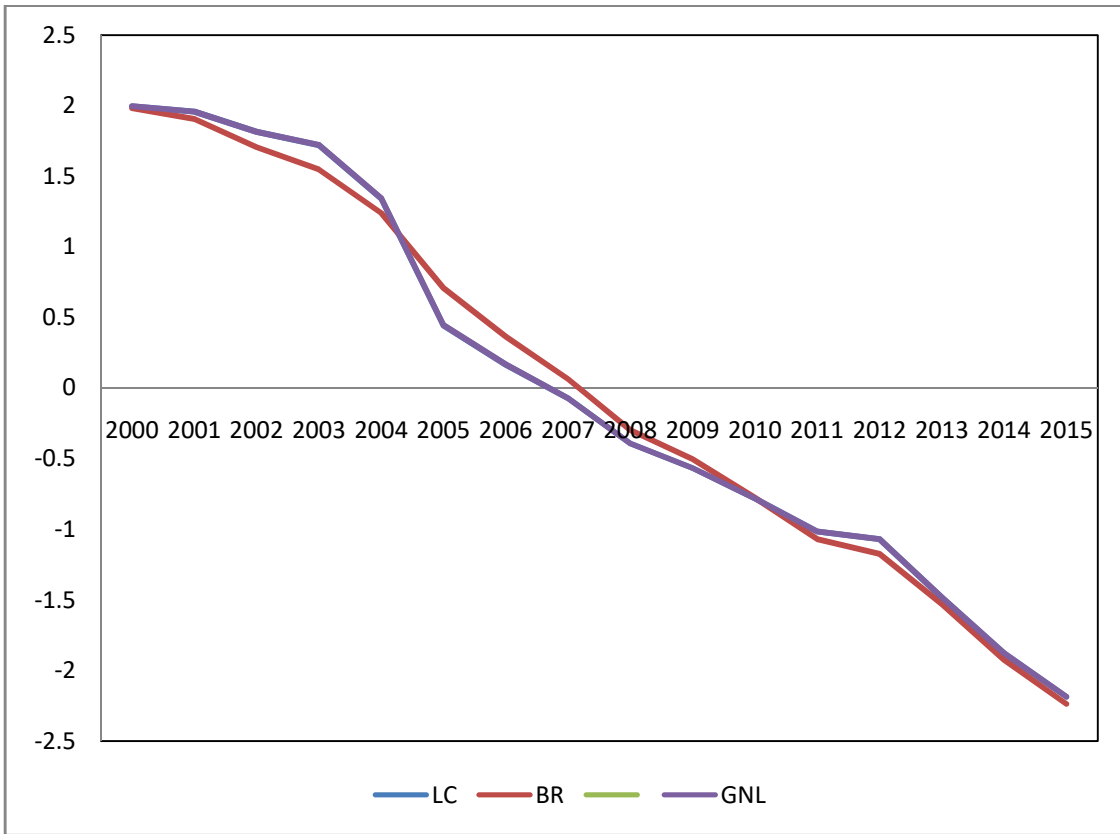


Figure 81: Parameter k_t for male

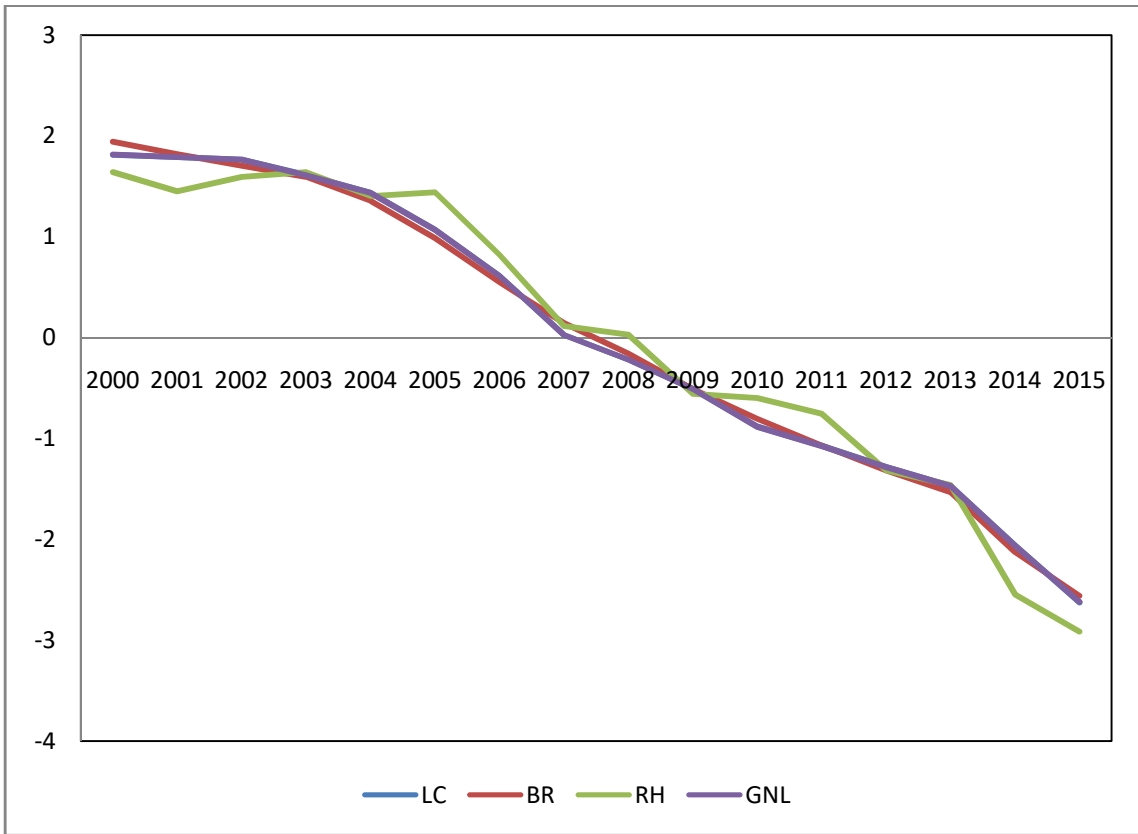


Figure 82: Parameter k_t for female

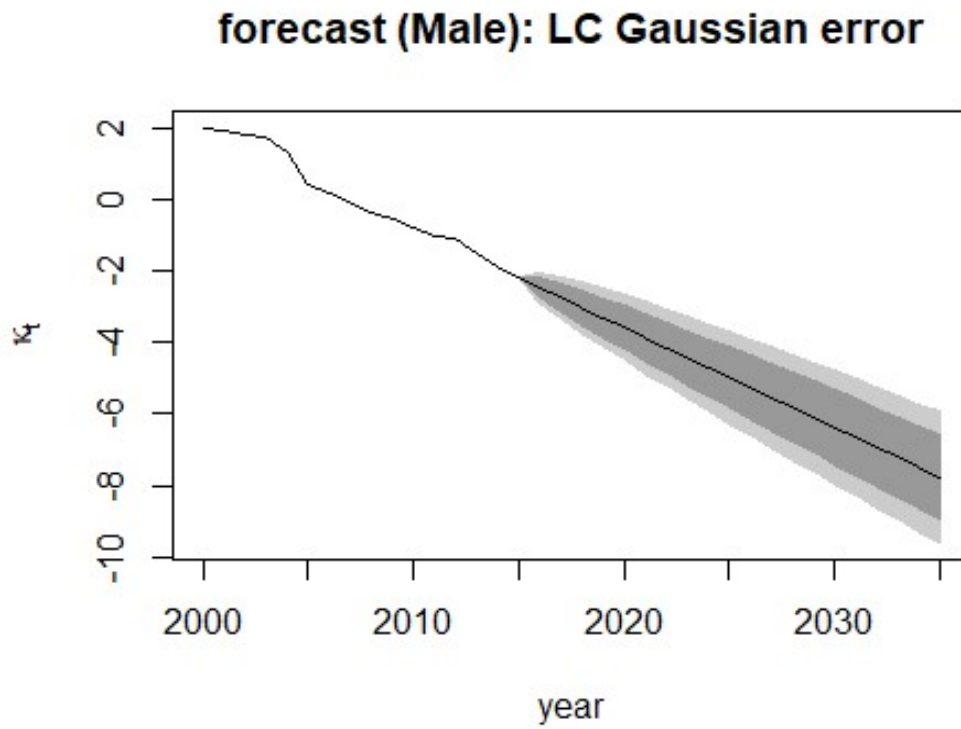


Figure 83: Fanplots using Lee-Carter model for males

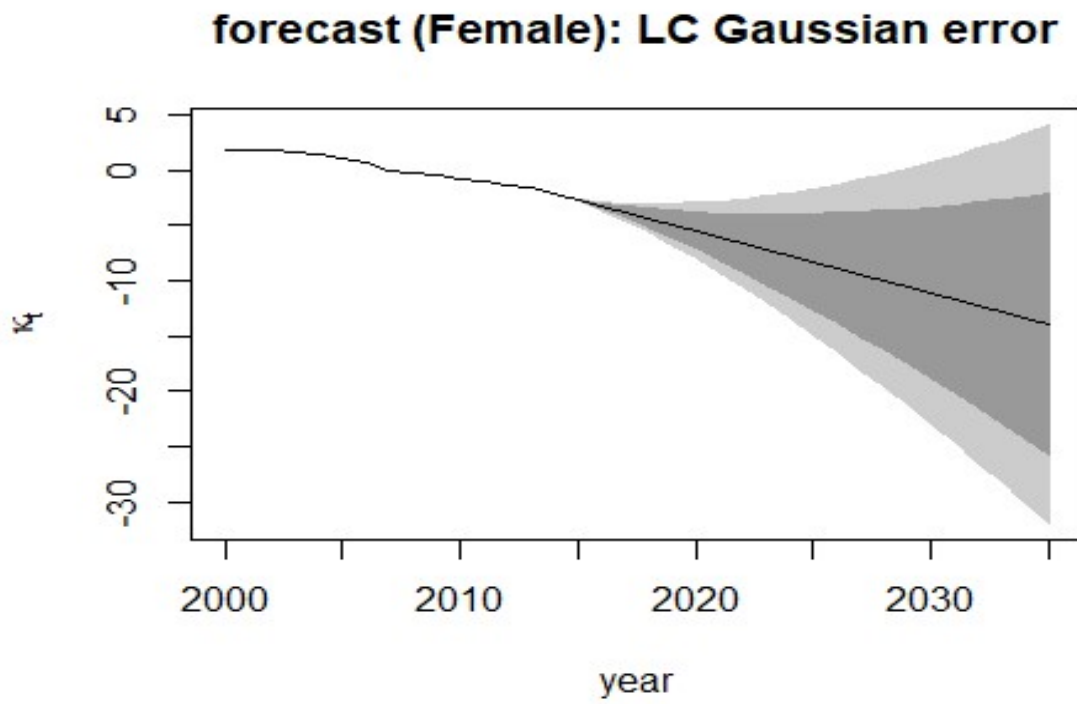


Figure 84: Fanplots using Lee-Carter model for females

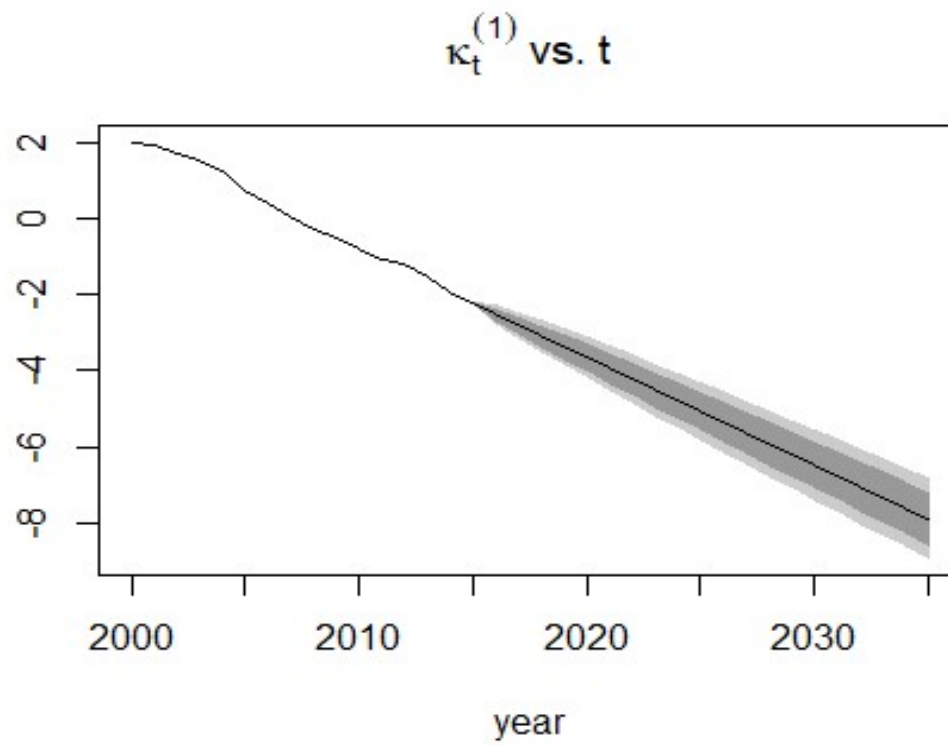


Figure 85: Fanplots using Brouhns model for male

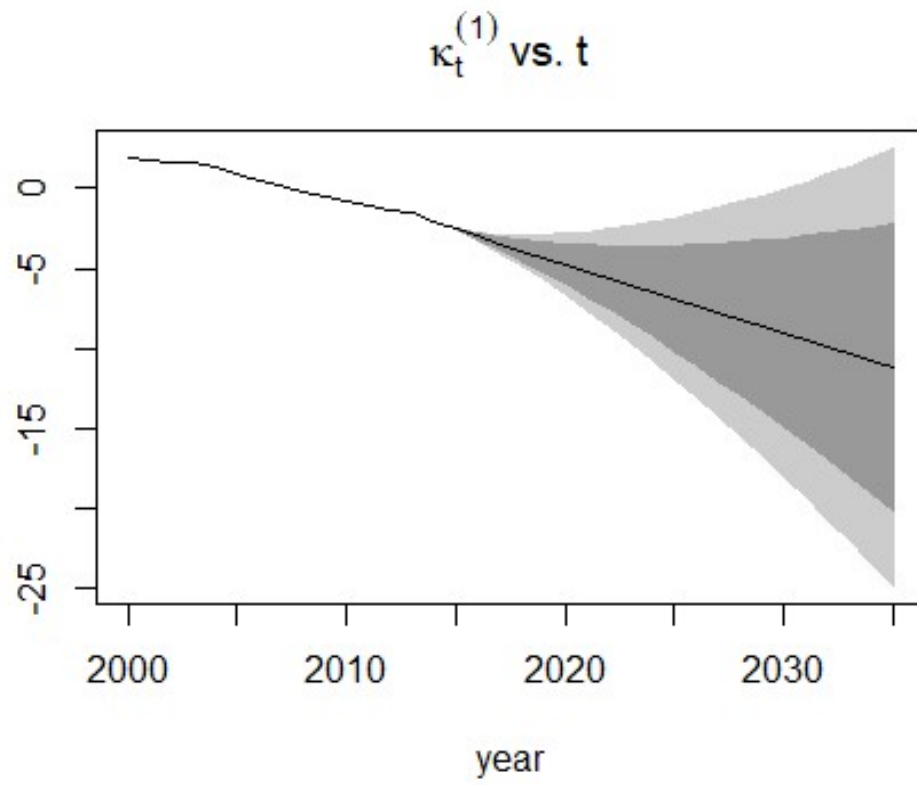


Figure 86: Fanplots using Brouhns model for female

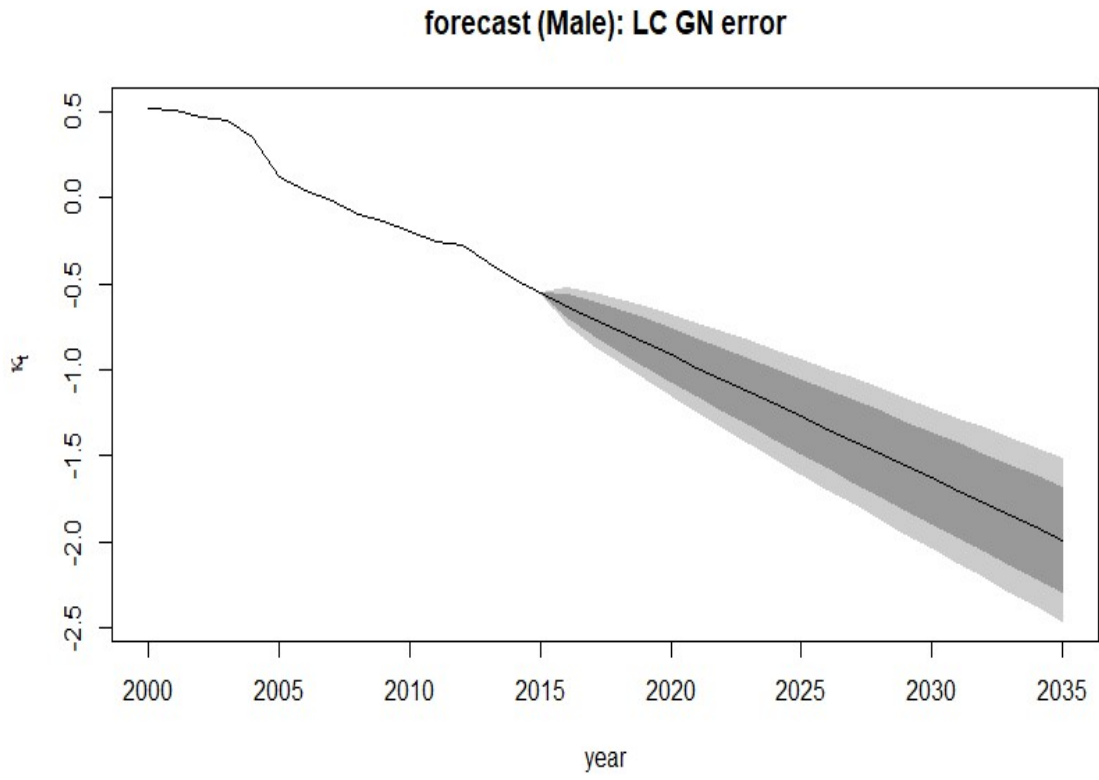


Figure 87: Fanplots using Gamma-normal model for males

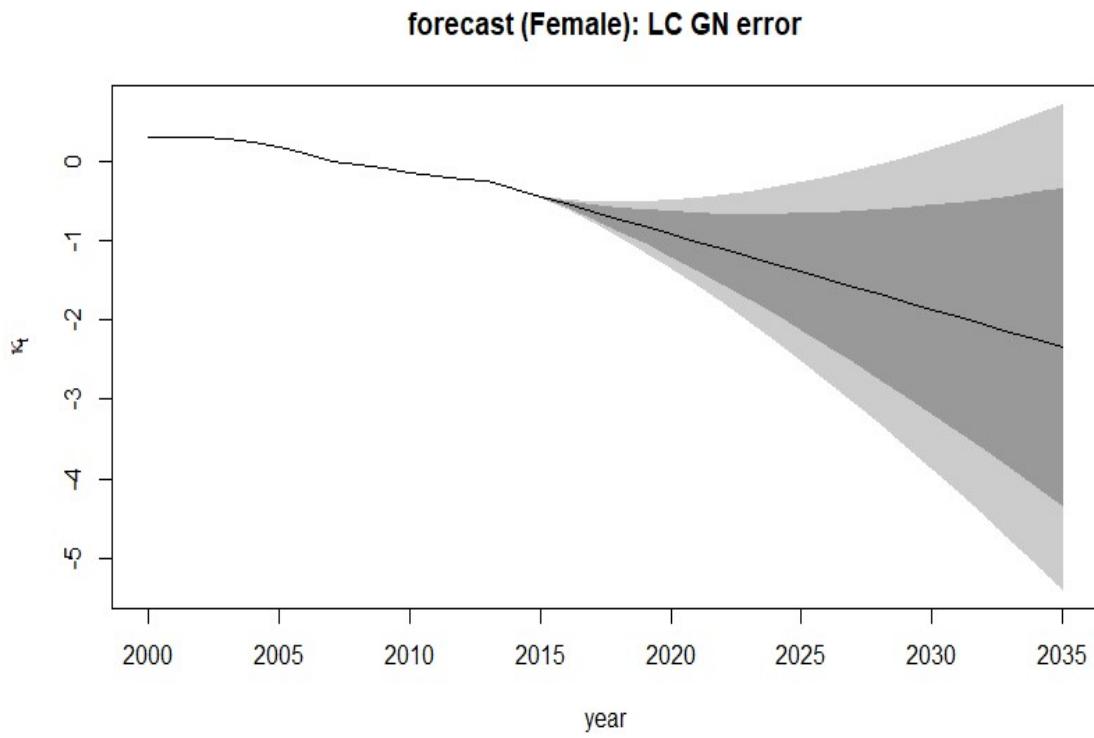


Figure 88: Forecast using Gamma-Normal model for female

APPENDIX B

R CODES USED IN THE STUDY

1. NORMALITY TESTS

Tables 4.3 and 4.4 -Kolmogorov-Smirnov test

```
# Kolmogorov-Smirnov Tests
```

```
# Across Ages
```

```
apply(Mxt_male,1,function(x) ks.test(x,'pnorm'))
```

```
# Across the years
```

```
apply(Mxt_male,2,function(x) ks.test(x,'pnorm'))
```

```
# Across Ages
```

```
apply(Mxt_female,1,function(x) ks.test(x,'pnorm'))
```

```
# Across the years
```

```
apply(Mxt_female,2,function(x) ks.test(x,'pnorm'))
```

Tables 4.5 and 4.6 -Shapiro-Wilk test

```
# Shapiro-Wilk test
```

```
# Across Ages
```

```
apply(Mxt_male,1,shapiro.test)
```

```
# Across the years
```

```
apply(Mxt_male,2,shapiro.test)
```

```
# Across Ages
```

```
apply(Mxt_female,2,1,shapiro.test)
```

```
# Across the years
```

```
apply(Mxt_female,2,shapiro.test)
```

2. MAXIMUM LIKELIHOOD ESTIMATE OF PARAMETERS, GOODNESS OF FIT AND PLOTS - Table 4.7 – Table 4.14

Older Variants

```
# loading Libraries
```

```
a. library(StMoMo)
```

```
b. library(demography)
```

```
c. library(readxl)
```

```
d. library(ilc)
```

Male Dataset

```
nig.dd_male
```

Model 1: Lee-Carter with Gaussian Error

```
LC_gausErr_Male <- lca.rh(nig.dd_male, model = 'lc',
```

```
error = 'gaussian',
```

```
verbose=F)
```

```
# Estimated Parameters
```

```
coef(LC_gausErr_Male)
```

```
# Model summary
```

```
LC_gausErr_Male
```

```
Plot
```

```

plot(LC_gausErr_Male)

###Model 2: Brouhns with Poisson Error

nig_male = StMoMoData(nig.dd_male,series = 'male')

nig_female = StMoMoData(nig.dd_female,series = 'female')

nig_male

nig_female

### Brouhns (Using log link i.e Poisson)

LC = lc(link = 'log')

LCfit_male <- fit(LC, data = nig_male)

LCfit_male

# Estimated Parameters

coef(LCfit_male)

# Goodness of fit

AIC(LCfit_male)

BIC(LCfit_male)

# Plot

plot(LCfit_male)

###Model 3: Renshaw-Haberman model (Using log Link)

# The model fitting didn't converge

```

```

RH <- rh(link = "log", cohortAgeFun = '1')

RHfit_male <- fit(RH, data = nig_male, start.ax = LCfit_male$ax,
start.bx = LCfit_male$bx, start.kt = LCfit_male$kt)

# Estimated Parameters

coef(RHfit_male)

# Goodness of fit

AIC(RHfit_male)

BIC(RHfit_male)

# Plot

plot(RHfit_male)

#### Female Dataset

nig.dd_female

### Model 1b: Lee-Carter with Gaussian Error

LC_gausErr_female <- lca.rh(nig.dd_female, model = 'lc',
error = 'gaussian',
verbose=F)

# Estimated Parameters

coef(LC_gausErr_female)

```



```

# Plot

plot(LC_gausErr_female)

# Model summary

LC_gausErr_female

####Model 2b: Brouhns –LC with Poisson Error

nig_female = StMoMoData(nig.dd_female,series = 'female')

nig_female

#### Brouhns- LC (Using log link i.e Poisson)

LC = lc(link = 'log')

LCfit_female <- fit(LC, data = nig_female)

LCfit_female

# Estimated Parameters

coef(LCfit_female)

# Goodness of fit

AIC(LCfit_female)

BIC(LCfit_female)

# Plot

plot(LCfit_female)

```

Model 3b: Renshaw-Haberman model

```
RH <- rh(link = "log", cohortAgeFun = '1')

RHfit_female <- fit(RH, data = nig_female, start.ax = LCfit_female$ax,
start.bx = LCfit_female$bx, start.kt = LCfit_female$kt)

# Estimated Parameters

Coef(RHfit_female)

# Goodness of fit

AIC(RHfit_female)

BIC(RHfit_female)

# Plot

plot(RHfit_female)

# Compute Log-Likelihood Gaussian- (Lee-Carter)

computeLogLikGaussian2 <- function(obs, fit) {

nobs = length(obs)

sigma2 = sum((obs-fit)^2)/(nobs);sigma2

res = (-0.5*log(2*pi)) - (0.5*log(sigma2)) - (((obs-fit)^2)/(2*sigma2))

sum(res,rm.na=T)

}

# Compute Log-Likelihood Poisson- (Brouhns and Renshaw-Haberman)
```

```

computeLogLikPoisson<- function(obs, fit) {

res<- (obs * log(fit) - fit - lfactorial(obs))

sum(res, na.rm=TRUE)

}

### Model 4: Gamma-Normal Lee-Carter (GNLC)

# loading Library

library (bbmle)

## Utility function

# GNLC Density Function

dgammaNorm<- function(x,a,mu,sd,log=FALSE){

d = dnorm(x,mu,sd) *

((1/gamma(a)) *

(-log(1 - pnorm(x,mu,sd)))^(a-1))

if (log) log(d) else d

}

# Convert an un-named vector parameter to list

toListGN<- function(pars,rw,cl){

nms =c(paste('a',1:rw,sep = ""),

paste('b',1:rw,sep = ""),

```

```

paste('k',1:cl,sep = ""),

      'alpha',

      'sigma')

structure(as.list(pars),names= nms)

}

#Compute Y predicted

computePredicted = function(ax,bx,kt){

ind = gl(cl,rw)

mat<- matrix(ax + bx*kt[ind], nrow=rw, ncol=cl)

  Ypred = array(mat)

  Ypred

}

# Compute Log-Likelihood

computeLogLikGN = function(Yobs,Ypred,alpha,sigma){

names(Yobs) = NULL

names(Ypred) = NULL

names(alpha) = NULL

names(sigma) = NULL

nobs = length(Yobs)

```

```

LL = -nobs*(lgamma(alpha)) +
sum(log(dnorm(Yobs,Ypred,sigma))) +
      (alpha -1)*sum(log(-log(1- pnorm(Yobs,Ypred,sigma))))

LL
}

```

Imposing Constraint

```

imposeConstraint=function(CoeffList){

tempBx = CoeffList$bx

tempKt = CoeffList$kt

c1 = mean(tempKt)

c2 = sum(tempBx)

# bx

bx = tempBx/c2

# kt

kt = c2*(tempKt - c1)

list(bx=bx,kt=kt)

}

```

Compute MSE

```

computeMSE = function(Yobs,Ypred){

```

```

nobs = length(Yobs)

rss = (Yobs - Ypred)^2

mse = sum(rss)/nobs

mse

}

### a.Male Dataset

maleDataList = getSpecDataOnly('male')

ages<- c(0,3,7.5,12.5,17.5,22.5,

        27.5,32.5,37.5,42.5,47.5,

        52.5,57.5,62.5,67.5,72.5,

        77.5,82.5,87.5)

years = 2000:2015

Mxt_m = maleDataList$Mxt_m

rownames(Mxt_m) = ages

colnames(Mxt_m) = years

nAges = nrow(Mxt_m)

nYears = ncol(Mxt_m)

# Data Setup

```

```

log_Mxt_m = log(Mxt_m)

Data_male = log_Mxt_m

# Parameters

rw = nAges;rw

cl = nYears;cl

cat('Number of parameter needed for GN',eval(rw+rw+cl+2))

##### GNLC Log-likelihood #####

LLGN <- function(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11,a12,a13,a14,a15,a16,a17,a18,a19,

                b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,b12,b13,b14,b15,b16,b17,b18,b19,

                k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,

                alpha,

                sigma){

  AX = c(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,

        a11,a12,a13,a14,a15,a16,a17,a18,a19)

  BX = c(b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,

        b12,b13,b14,b15,b16,b17,b18,b19)

  KT= c(k1,k2,k3,k4,k5,k6,k7,k8,k9,

        k10,k11,k12,k13,k14,k15,k16)

  ff = gl(cl,rw) #

```

```

mt<- matrix(AX + BX*KT[ff], nrow=rw, ncol=cl)

Y.pred = array(mt)

Y = array(Data_male)

R = suppressWarnings(

dgammaNorm(Y,a =alpha, mu = Y.pred, sd = sigma, log = TRUE))

-sum(R)

}

# StartUp Parameters

fixedAs = apply(Data_male,1,mean,na.rm = T)

startPars <- c(fixedAs,rep(1/rw,rw),rep(0,cl),0.2,1)

length(startPars)

##### Fitting MLE 1

mleFit_male =mle2(minuslogl = LLGN,

fixed = toListFixed(fixedAs,rw),

start= toListGN(startPars,rw,cl),

control=list(maxit=3000)

)

mleFit_male

## Extracting Coefficients

```



```

coeffs_male = extractCoefficients(mleFit_male)

coeffs_male

## Imposing Constraints

cstCoef_male = imposeConstraint(coeffs_male)

sum(cstCoef_male$bx)

sum(cstCoef_male$kt)

# Final Model Parameter

## Extracting Coefficients

FinalPars = list(ax =

fixedAs, bx=cstCoef_male$bx,

kt=cstCoef_male$kt,alpha=coeffs_male$alpha,

sigma=coeffs_male$sigma)

# Final Parameters

Yobs = array(log_Mxt_m)

# Recompute Likelihood

Ypred = computePredicted(FinalPars$ax,FinalPars$bx,FinalPars$kt)

# Log- Likelihood

#a

llik = computeLogLikGN(Yobs = Yobs,Ypred = Ypred,

```

```

alpha = FinalPars$alpha,

sigma = FinalPars$sigma)

# b

llik2 = computeLogLikGN2(Yobs = Yobs,

                        Ypred = computePredicted(coeffs_male$sax,

                                                coeffs_male$bx,

                                                coeffs_male$kt),

alpha = coeffs_male$alpha,

sigma = coeffs_male$sigma)

llik

llik2

sum(dgammaNorm(x = array(Data_male),

                    a = FinalPars$alpha,

                    mu = Ypred,

                    sd = FinalPars$sigma,

                    log = T))

# Compute MSE

MSE_m = computeMSE(Yobs, Ypred); MSE_m

# RMSE

```

```

RMSE_m = sqrt(MSE_m);RMSE_m

# Output File

# output.m = list(model = mleFit_male, unconstrCoeff = coeffs_male,

#       ConstCoef = FinalPars,

#       Yobs = Yobs, Ypred = Ypred,

#       llik1 = llik, llik2 = llik2

# )

# save(output.m, file = 'mleFitGN_male.rda')

```

b. Female Dataset

```

# loading Library

library(bbmle)

FemaleDataList = getSpecDataOnly('female')

ages<- c(0,3,7.5,12.5,17.5,22.5,

         27.5,32.5,37.5,42.5,47.5,

         52.5,57.5,62.5,67.5,72.5,

         77.5,82.5,87.5)

years = 2000:2015

Mxt_f = FemaleDataList$Mxt_f

```

```

rownames(Mxt_f) = ages

colnames(Mxt_f) = years

nAges = nrow(Mxt_f)

nYears = ncol(Mxt_f)

# Data Setup

log_Mxt_f = log(Mxt_f)

Data_Female = log_Mxt_f

# Parameters

rw = nAges;rw

cl = nYears;cl

cat('Number of parameter needed for Norm',eval(rw+rw+cl+1))

cat('Number of parameter needed for GN',eval(rw+rw+cl+2))

##### GNLC Log-likelihood#####

LLGN <- function(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11,a12,a13,a14,a15,a16,a17,a18,a19,

                b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,b12,b13,b14,b15,b16,b17,b18,b19,

                k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,

                alpha,

                sigma){

  AX = c(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,

```

```

a11,a12,a13,a14,a15,a16,a17,a18,a19)

BX = c(b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,

b12,b13,b14,b15,b16,b17,b18,b19)

KT= c(k1,k2,k3,k4,k5,k6,k7,k8,k9,

k10,k11,k12,k13,k14,k15,k16)

ff = gl(cl,rw) #

mt<- matrix(AX + BX*KT[ff], nrow=rw, ncol=cl)

Y.pred = array(mt)

# Y = array(Data_Female)

Y = array(log(Mxt_f))

R = suppressWarnings(

dgammaNorm(Y,a =alpha, mu = Y.pred, sd = sigma, log = TRUE))

-sum(R)

}

# StartUp Parameters

fixedAs = apply(Data_Female,1,mean,na.rm = T)

startPars <- c(fixedAs,rep(1/rw,rw),rep(0,cl),0.2,1)

length(startPars)

#### Fitting MLE ####

```

```

mleFit_female =mle2(minuslogl = LLGN,

fixed = toListFixed(fixedAs,rw),

start= toListGN(startPars,rw,cl),

control=list(maxit=3000)

)

mleFit_female

## Extracting Coefficients

coeffs_female = extractCoefficients(mleFit_female)

coeffs_female

## Imposing Constraints

cstCoef_female = imposeConstraint(coeffs_female)

sum(cstCoef_female$bx)

sum(cstCoef_female$kt)

# Final Model Parameter

## Extracting Coefficients

FinalPars = list(ax = fixedAs, bx=cstCoef_female$bx,

kt=cstCoef_female$kt,alpha=coeffs_female$alpha,

sigma=coeffs_female$sigma)

# FinalParameters

```

```

Yobs = array(log_Mxt_f)

# Recompute Likelihood

Ypred = computePredicted(FinalPars$ax,FinalPars$bx,FinalPars$kt)

# Log Lik

#a

llik = computeLogLikGN(Yobs = Yobs,Ypred = Ypred,

alpha = FinalPars$alpha,

sigma = FinalPars$sigma)

# b

llik2 = computeLogLikGN2(Yobs = Yobs,

                        Ypred = computePredicted(coeffs_female$ax,

                                                  coeffs_female$bx,

                                                  coeffs_female$kt),

alpha = coeffs_female$alpha,

sigma = coeffs_female$sigma)

llik

llik2

sum(dgammaNorm(x = array(Data_Female),

                  a = FinalPars$alpha,

```

```

mu = Ypred,

sd = FinalPars$sigma,

log = T))

# Compute MSE

MSE_f = computeMSE(Yobs,Ypred); MSE_f

# RMSE

RMSE_f = sqrt(MSE_f);RMSE_f

# Output File

# output.f = list(model = mleFit_female, unconstrCoeff = coeffs_female,

#   ConstCoef = FinalPars,

#   Yobs = Yobs,Ypred = Ypred,

#   llik1 = llik,llik2 = llik2

#   )

# save(output.f,file = 'mleFitGN_female.rda')

```

2b. ROOT MEAN SQUARE ERROR AND MEAN SQUARE ERROR-TABLE 4.15

```

##### Male dataset

## Lee-Carter – Gaussian

res_gm = LC_gaus_male$residuals$y

```



```

nobs = length(as.array(LC_gaus_male$fitted$y))

mse_gm = sum((res_gm)^2)/nobs; mse_gm

rmse_gm = sqrt(mse_gm); rmse_gm

## Brouhns- LC poisson

Ypred_pm <- LC_pois_male$fittingModel$fitted.values

Yobs <- LC_pois_male$fittingModel$y

# a = Ypred_pm - Yobs

# a [1]

res_pm <- LC_pois_male$fittingModel$residuals

mse_pm = sum(res_pm^2)/nobs;mse_pm

rmse_pm = sqrt(mse_pm);rmse_pm

## Renshaw-Haberman

res_rhm <- RHfit_male$fittingModel$residuals

mse_rhm <- sum(res_rhm^2)/nobs; mse_rhm

rmse_rhm <- sqrt(mse_rhm); rmse_rhm

#### Female dataset

## Lee-Carter - Gaussian

res_gf = LC_gaus_female$residuals$y

```

```

nobs = length(as.array(LC_gaus_female$fitted$y))

mse_gf = sum((res_gf)^2)/nobs; mse_gf

rmse_gf = sqrt(mse_gf); rmse_gf

## Brouhns - LC poisson

# Ypred_pf <- LC_pois_female$fittingModel$fitted.values

# Yobs <- LC_pois_female$fittingModel$y

# a = Ypred_pf - Yobs

# a [1]

res_pf <- LC_pois_female$fittingModel$residuals

mse_pf = sum(res_pf^2)/nobs; mse_pf

rmse_pf = sqrt(mse_pf); rmse_pf

## Renshaw-Haberman

res_rhf <- RHfit_female$fittingModel$residuals

mse_rhf <- sum(res_rhf^2)/nobs; mse_rhf

rmse_rhf <- sqrt(mse_rhf); rmse_rhf

```

3. SIMULATION STUDY- TABLE 4.23

```

# loading Libraries
a) library(knitr)
b) library(StMoMo)
c) library(bbmle)

```

```

##### Simulation Study for t = 5 #####
## Specification
ages = 0:14
nAges = length(ages)
years = 2000:2004
nYears = length(years)

## Simulate Data
set.seed(452)
n5_Dxt = matrix(runif(n = nAges*nYears, min = 1000, max = 10000), ncol = nYears)
#n5_Dxt = matrix(rpois(n = nAges*nYears, lambda = 12000), ncol = nYears)
set.seed(912)
n5_Ext = matrix(runif(n = nAges*nYears, min = 20000, max = 300000), ncol = nYears)
colnames(n5_Dxt) = colnames(n5_Ext) = years
rownames(n5_Dxt) = rownames(n5_Ext) = ages
n5_Mxt = n5_Dxt/n5_Ext
n5_lnMxt = log(n5_Mxt)

## Setup data for Modelling
# Demogdata
dim(n5_Dxt)
head(n5_lnMxt)
# Modelling
## Gamma Normal- Lee-Carter
#### Data Setup
Yobs_GN_n5 <- array(n5_lnMxt)
#### Log Likelihood
LLGN      <-      function(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11,a12,a13,a14,a15,
                           b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,b12,b13,b14,b15,
                           k1,k2,k3,k4,k5,

```

```

alpha, sigma){
  AX = c(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,
        a11,a12,a13,a14,a15)
  BX = c(b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,
        b12,b13,b14,b15)
  KT= c(k1,k2,k3,k4,k5)
ff = gl(nYears,nAges) #
mt<- matrix(AX + BX*KT[ff], nrow=rw, ncol=cl)
  Y.pred = array(mt)
  Y = array(n5_1nMxt)
  R = suppressWarnings(
dgammaNorm(Y,a =alpha, mu = Y.pred, sd = sigma, log = TRUE))
  -sum(R)
}
##### Setup Starting Parameters
rw = nAges;rw
cl = nYears;cl
cat('Number of parameter needed for GN',eval(rw+rw+cl+2))
fixedAs = apply(n5_1nMxt,1,mean,na.rm = T)
# startPars <- c(rep(0,rw),rep(0,rw),rep(0,cl),0.2,1)
startPars <- c(fixedAs,rep(1/rw,rw),rep(0,cl),0.2,1)
length(startPars)
# startPars
##### Fitting MLE
n5_GNfit =mle2(minuslogl = LLGN,
fixed = toListFixed(fixedAs,rw),
start= toListGN(startPars,rw,cl),
control=list(maxit=99999)
      # method = 'Nelder-Mead'
)

```

```

n5_GNfit
##### Make GN model Output
n5_GNfitOutput = MakeGNoutput2(model = n5_GNfit)

# Get Predictions (y_hats)
Yhat_GN <- computeFittedY_GN(n5_GNfitOutput)
head(Yhat_GN)

##### Recompute LogLikelihood
alp_m = n5_GNfitOutput$unconstrCoeff$alpha
sig_m = n5_GNfitOutput$unconstrCoeff$sigma
logLik_GNm <- computeLLGN(Yobs = Yobs_GN_n5,
                          Ypred = Yhat_GN,
                          alpha = alp_m, sigma = sig_m)
logLik_GNm

## Lee-Carter Gaussian
##### Sources internal functions
loadGaussianReqSource()

##### Data Setup
# Convert to demogdata
#Required library(StMoMo)
S1_demogdata <- demography::demogdata(data = n5_Mxt ,pop = n5_Ext,
ages = ages, years = years,
type = 'mortality',
label = 'S1',
name = 'S1')
S1_StMoMoData = StMoMo::StMoMoData(data = S1_demogdata)
S1_StMoMoData

```

```

##### Fit Lee-Carter - Gaussian
{r include=FALSE}

##### Model P1: LC with Gaussian Error stmomo #####
LC.g = lc2(link = 'gaussian')
fitLCgaus<- fit.StMoMo2(LC.g, data = S1_StMoMoData)
fitLCgaus
Yhat_Gaus_N5 <- computeFittedY(fitLCgaus)
head(Yhat_Gaus_N5)
computeLogLikGaussian2(obs = Yobs_GN_n5,fit = Yhat_Gaus_N5)

## Inference
mm<- matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S1)')

## Fit Brouhns (LC Poisson)
##### Model P1: LC poisson Error stmomo #####
LC.p = lc(link = 'log')
fitLCpois<- fit(LC.p, data = S1_StMoMoData)
fitLCpois

## Inference
mm<- matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S1)')

## Fit Renshaw -Haberman (RH)
##### Model P1: LC with Gaussian Error
RH = rh(link = "log", cohortAgeFun = '1')

```

```

fitRH<- fit(RH, data = S1_StMoMoData)
fitRH

## Inference
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'RH (S1)')

## Performance of the models
gmno = matrix(c(logLik_GNm, length(startPars)-(2+3)),nrow = 1)
colnames(gmno) <- c('logLikelihood','Number of Parameters')
kable(gmno,format = 'pandoc', caption = 'GN (S1)')

#Lee-Carter - Gaussian
mm<-
matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus),fitLCgaus$npar),nrow =
1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of Parameters')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S1)')

# Brouhns (LC – Poisson)
mm<-
matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois),fitLCpois$npar),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC', 'Number of Parameters')
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S1)')

# Renshaw-Haberman
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH),fitRH$npar),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of parameters')
kable(mm,format = 'pandoc',caption = 'RH (S1)')

```

```

#####
##### Simulation Study for t = 20 #####
## Specification
ages = 0:19
nAges = length(ages)
years = 1997:2016
nYears = length(years)

## Simulate Data
set.seed(812)
n20_Dxt = matrix(rpois(n = nAges*nYears, lambda = 7000), ncol = nYears)
set.seed(912)
n20_Ext = matrix(runif(n = nAges*nYears, min = 20000, max = 100000), ncol = nYears)
n20_Mxt = n20_Dxt/n20_Ext
n20_lnMxt = log(n20_Mxt)

## Setup data for Moodelling
# Demogdata
dim(n20_Dxt)

# Modelling
## Gamma Normal Lee-Carter (GNLC)
#### Data Setup
Yobs_GN_n20 <- array(n20_lnMxt)

#### Log Likelihood
LLGN <- function(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11,a12,a13,a14,a15,
                 a16,a17,a18,a19,a20,
                 b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,b12,b13,b14,b15,

```



```

        b16,b17,b18,b19,b20,
        k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,k17,k18,k19,k20,
alpha,
sigma){
  AX = c(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,
        a11,a12,a13,a14,a15
,a16,a17,a18,a19,a20
  )

  BX = c(b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,
        b12,b13,b14,b15
,b16,b17,b18,b19,b20
  )

  KT= c(k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,k17,k18,k19,k20)
ff = gl(nYears,nAges) #
mt<- matrix(AX + BX*KT[ff], nrow=rw, ncol=cl)
  Y.pred = array(mt)
  Y = array(n20_ lnMxt)
  R = suppressWarnings(
dgammaNorm(Y,a =alpha, mu = Y.pred, sd = sigma, log = TRUE))
  -sum(R)
}
##### Setup Starting Parameters
rw = nAges;rw
cl = nYears;cl
cat('Number of parameter needed for GN',eval(rw+rw+cl+2))
fixedAs = apply(n20_ lnMxt,1,mean,na.rm = T)
# startPars <- c(rep(0,rw),rep(0,rw),rep(0,cl),0.2,1)
startPars <- c(fixedAs,rep(1/rw,rw),rep(0,cl),0.2,1)

```

```

length(startPars)
startPars

##### Fitting GNLC MLE
n20_GNfit = mle2(minuslogl = LLGN,
fixed = toListFixed(fixedAs,rw),
start= toListGN(startPars,rw,cl),
control=list(maxit=9999999)
)
n20_GNfit
##### Make GNLC model Output
n20_GNfitOutput = MakeGNoutput2(model = n20_GNfit)
# Get Predictions (y_hats)
Yhat_GN <- computeFittedY_GN(n20_GNfitOutput)

##### Recompute LogLikelihood
alp_m = n20_GNfitOutput$unconstrCoeff$alpha
sig_m = n20_GNfitOutput$unconstrCoeff$sigma
logLik_GNm <- computeLLGN(Yobs = Yobs_GN_n20,
                          Ypred = Yhat_GN,
alpha = alp_m,sigma = sig_m)
logLik_GNm

## Lee-Carter – (LC Gaussian)
##### Data Setup
# Convert to demogdata
library(StMoMo)
S2_demogdata <- demography::demogdata(data = n20_Mxt ,pop = n20_Ext,
ages = ages,years = years,
type = 'mortality',

```

```

label = 'S2',
name = 'S2')
S2_StMoMoData = StMoMoData(data = S2_demogdata)
S2_StMoMoData
## Fit Lee-Carter - Gaussian
##### Model P1: LC with Gaussian Error stmomo #####
LC.g = lc2(link = 'gaussian')
fitLCgaus<- fit.StMoMo2(LC.g, data = S2_StMoMoData)
fitLCgaus
Yhat_Gaus_N20 <- computeFittedY(fitLCgaus)
head(Yhat_Gaus_N20)
LLgaus_N20 = computeLogLikGaussian2(obs = Yobs_GN_n20,fit = Yhat_Gaus_N20)
LLgaus_N20

## Inference
mm<- matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S2)')

## Fit Brouhns (LC – Poisson)
##### Model P1: LC with Gaussian Error stmomo #####
LC.p = lc(link = 'log')
fitLCpois<- fit(LC.p, data = S2_StMoMoData)
fitLCpois

## Inference
mm<- matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S2)')

```

Fit Renshaw-Haberman

Model P1: LC with Gaussian Error stmomo

```
RH = rh(link = "log", cohortAgeFun = '1')
```

```
fitRH<- fit(RH, data = S2_StMoMoData)
```

```
fitRH
```

Inference

```
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH)),nrow = 1)
```

```
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
```

```
kable(mm,format = 'pandoc',caption = 'RH (S2)')
```

Performance

Inference

```
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH)),nrow = 1)
```

```
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
```

```
kable(mm,format = 'pandoc',caption = 'RH (S2)')
```

Performance

Gamma Normal Lee-Carter

```
gmno = matrix(c(logLik_GNm, length(startPars)-5),nrow = 1)
```

```
colnames(gmno) <- c('logLikelihood','Number of Parameters')
```

```
kable(gmno,format = 'pandoc', caption = 'GN (S2)')
```

Lee-Carter – (LC Gaussian)

```
mm<-
```

```
matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus),fitLCgaus$npar),nrow =
```

```
1)
```

```
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of Parameters')
```

```
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S2)')
```

Brouhns – (LC Poisson)

```
mm<-  
matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois),fitLCpois$npar),nrow = 1)  
colnames(mm) <- c('Deviance','Loglikelihood','AIC', 'Number of Parameters')  
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S2)')
```

Renshaw-Haberman –(RH)

```
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH),fitRH$npar),nrow = 1)  
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of parameters')  
kable(mm,format = 'pandoc',caption = 'RH (S2)')
```

```
#####
```

Simulation Study 3: t = 30

Specifications

```
ages = 0:19  
nAges = length(ages)  
years = 1988:2017  
nYears = length(years)
```

Simulate Data

```
set.seed(930)  
n30_Dxt = matrix(rpois(n = nAges*nYears, lambda = 7000), ncol = nYears)  
  
set.seed(9233)  
n30_Ext = matrix(runif(n = nAges*nYears, min = 20000, max = 100000), ncol = nYears)  
  
n30_Mxt = n30_Dxt/n30_Ext  
n30_lnMxt = log(n30_Mxt)
```

```

## Setup data for Moodelling
# Demogdata
dim(n30_Dxt)

# Modelling
## Gamma -Normal Lee-Carter
#### Data Setup
Yobs_GN_n30 <- array(n30_InMxt)

## Log Likelihood

LLGN <-
function(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11,a12,a13,a14,a15,a16,a17,a18,a19,a20,
         b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,b12,b13,b14,b15,b16,b17,b18,b19,b20,
         k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,k17,k18,k19,k20,
         k21,k22,k23,k24,k25,k26,k27,k28,k29,k30,
         alpha,
         sigma){
  AX = c(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,
         a11,a12,a13,a14,a15,a16,a17,a18,a19,a20)

  BX = c(b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,
         b12,b13,b14,b15,b16,b17,b18,b19,b20)

  KT= c(k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,k17,k18,k19,k20,
        k21,k22,k23,k24,k25,k26,k27,k28,k29,k30)

  ff = gl(nYears,nAges) #
  mt<- matrix(AX + BX*KT[ff], nrow=rw, ncol=cl)
  Y.pred = array(mt)

```

```

Y = array(n30_InMxt)
R = suppressWarnings(
dgammaNorm(Y,a =alpha, mu = Y.pred, sd = sigma, log = TRUE))
-sum(R)
}
##### Setup Starting Parameters
rw = nAges;rw
cl = nYears;cl
cat('Number of parameter needed for GN',eval(rw+rw+cl+2))
fixedAs = apply(n30_InMxt,1,mean,na.rm = T)
startPars <- c(fixedAs,rep(1/rw,rw),rep(0,cl),0.2,1)
length(startPars)
startPars

##### Fitting MLE
n30_GNfit =mle2(minuslogl = LLGN,
fixed = toListFixed(fixedAs,rw),
start= toListGN(startPars,rw,cl),
control=list(maxit=69999)
          #method = 'Nelder-Mead'
)
n30_GNfit
##### Make GN model Output

n30_GNfitOutput = MakeGNoutput2(model = n30_GNfit)

# Get Predictions (y_hats)
Yhat_GN <- computeFittedY_GN(n30_GNfitOutput)

##### Recompute LogLikelihood

```

```

alp_m = n30_GNfitOutput$unconstrCoeff$alpha
sig_m = n30_GNfitOutput$unconstrCoeff$sigma
logLik_GNm <- computeLLGN(Yobs = Yobs_GN_n30,
                          Ypred = Yhat_GN,
                          alpha = alp_m,sigma = sig_m)
logLik_GNm

## Lee-Carter – ( LC Gaussian)
#### Data Setup
# Convert to demogdata
library(StMoMo)
S3_demogdata <- demography::demogdata(data = n30_Mxt ,pop = n30_Ext,
ages = ages,years = years,
type = 'mortality',
label = 'S3',
name = 'S3')
S3_StMoMoData = StMoMoData(data = S3_demogdata)
S3_StMoMoData
## Fit Lee-Carter -LC Gaussian
#### Model P1: LC with Gaussian Error stmomo ####
LC.g = lc2(link = 'gaussian')
fitLCgaus<- fit.StMoMo2(LC.g, data = S3_StMoMoData)
fitLCgaus
## Inference
mm<- matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S3)')

## Fit Brouhns -LC Poisson
#### Model P1: LC with Gaussian Error stmomo ####

```



```

LC.p = lc(link = 'log')
fitLCpois<- fit(LC.p, data = S3_StMoMoData)
fitLCpois
## Inference
mm<- matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S3)')

## Fit Renshaw-Haberman -RH
##### Model P1: LC with Gaussian Error stmomo #####
RH = rh(link = "log", cohortAgeFun = '1')
fitRH<- fit(RH, data = S3_StMoMoData)
fitRH

## Inference
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'RH (S3)')

## Performance
# Gamma -Normal Lee-Carter
gmno = matrix(c(logLik_GNm, length(startPars)),nrow = 1)
colnames(gmno) <- c('logLikelihood','Number of Parameters')
kable(gmno,format = 'pandoc', caption = 'GN (S3)')
#Lee-Carter – (LC Gaussian)
mm<-
matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus),fitLCgaus$npar),nrow =
1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of Parameters')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S3)')

```

Brouhns – (LC Poisson)

```
mm<-  
matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois),fitLCpois$npar),nrow = 1)  
colnames(mm) <- c('Deviance','Loglikelihood','AIC', 'Number of Parameters')  
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S3)')
```

Renshaw-Haberman (RH)

```
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH),fitRH$npar),nrow = 1)  
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of parameters')  
kable(mm,format = 'pandoc',caption = 'RH (S3)')
```

```
#####
```

```
##### Simulation Study 3: N = 40 #####
```

Specification

Number of Age-set : 25n

Number of Years: 40n

ages = 0:24

nAges = length(ages)

years = 1978:2017

nYears = length(years)

Simulate Data

```
set.seed(812)
```

```
# n40_Dxt = matrix(runif(n = nAges*nYears, min = 1000, max = 20000), ncol = nYears)
```

```
n40_Dxt = matrix(rpois(n = nAges*nYears, lambda = 7000), ncol = nYears)
```

```
set.seed(912)
```

```
n40_Ext = matrix(runif(n = nAges*nYears, min = 20000, max = 100000), ncol = nYears)
```

```

n40_Mxt = n40_Dxt/n40_Ext
n40_lnMxt = log(n40_Mxt)
## Setup data for Moodelling
# Demogdata
dim(n40_Dxt)

# Modelling
## Gamma Normal Lee-Carter
##### Data Setup
Yobs_GN_n40 <- array(n40_lnMxt)
##### Log Likelihood
LLGN <-
function(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11,a12,a13,a14,a15,a16,a17,a18,a19,a20,
         a21,a22,a23,a24,a25,
         b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,b12,b13,b14,b15,b16,b17,b18,b19,b20,
         b21,b22,b23,b24,b25,
         k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,k17,k18,k19,k20,
         k21,k22,k23,k24,k25,k26,k27,k28,k29,k30,k31,k32,k33,k34,k35,k36,k37,k38,k39,k40,
         alpha,
         sigma){
  AX = c(a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,
         a11,a12,a13,a14,a15,a16,a17,a18,a19,a20,
         a21,a22,a23,a24,a25)
  BX = c(b1,b2,b3,b4,b5,b6,b7,b8,b9,b10,b11,
         b12,b13,b14,b15,b16,b17,b18,b19,b20,
         b21,b22,b23,b24,b25)
  KT= c(k1,k2,k3,k4,k5,k6,k7,k8,k9,k10,k11,k12,k13,k14,k15,k16,k17,k18,k19,k20,

```

k21,k22,k23,k24,k25,k26,k27,k28,k29,k30,k31,k32,k33,k34,k35,k36,k37,k38,k39,k40)

```
ff = gl(nYears,nAges) #
mt<- matrix(AX + BX*KT[ff], nrow=rw, ncol=cl)
Y.pred = array(mt)
Y = array(n40_lnMxt)
R = suppressWarnings(
dgammaNorm(Y,a =alpha, mu = Y.pred, sd = sigma, log = TRUE))
-sum(R)
}
##### Setup Starting Parameters
rw = nAges;rw
cl = nYears;cl
cat('Number of parameter needed for GN',eval(rw+rw+cl+2))
fixedAs = apply(n40_lnMxt,1,mean,na.rm = T)
startPars <- c(fixedAs,rep(1/rw,rw),rep(0,cl),0.2,1)
length(startPars)
startPars

##### Fitting MLE
n40_GNfit =mle2(minuslogl = LLGN,
fixed = toListFixed(fixedAs,rw),
start= toListGN(startPars,rw,cl),
control=list(maxit=9999999),
method = 'Nelder-Mead'
)
n40_GNfit
##### Make GN model Output
n40_GNfitOutput = MakeGNoutput2(model = n40_GNfit)
```

```

# Get Predictions (y_hats)
Yhat_GN <- computeFittedY_GN(n40_GNfitOutput)

##### Recompute LogLikelihood
alp_m = n40_GNfitOutput$unconstrCoeff$alpha
sig_m = n40_GNfitOutput$unconstrCoeff$sigma
logLik_GNm <- computeLLGN(Yobs = Yobs_GN_n40,
                          Ypred = Yhat_GN,
                          alpha = alp_m, sigma = sig_m)
logLik_GNm

## GN - Gaussian
##Data Setup
# Convert to demogdata
library(StMoMo)
S3_demogdata <- demography::demogdata(data = n40_Mxt ,pop = n40_Ext,
ages = ages, years = years,
type = 'mortality',
label = 'S3',
name = 'S3')

S3_StMoMoData = StMoMoData(data = S3_demogdata)
S3_StMoMoData
## Fit Lee-Carter- (LC-Gaussian)
{r include=FALSE}
##### Model P1: LC with Gaussian Error stmomo #####
LC.g = lc2(link = 'gaussian')
fitLCgaus<- fit.StMoMo2(LC.g, data = S3_StMoMoData)
fitLCgaus

```

```

## Inference
mm<- matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S3)')

```

Fit **Brouhns (LC – Poisson)**

```

{r include=FALSE}
##### Model P1: LC with Gaussian Error stmomo #####
LC.p = lc(link = 'log')
fitLCpois<- fit(LC.p, data = S3_StMoMoData)

```

fitLCpois

```

## Inference
mm<- matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S3)')

```

Fit **Renshaw-Haberman (RH)**

```

{r include=FALSE}
##### Model P1: LC with Gaussian Error stmomo #####
RH = rh(link = "log", cohortAgeFun = '1')
fitRH<- fit(RH, data = S3_StMoMoData)

```

fitRH

```

## Inference
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH)),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC')
kable(mm,format = 'pandoc',caption = 'RH (S3)')

```

Performance

```
# Gamma-Normal Lee-Carter
gmno = matrix(c(logLik_GNm, length(startPars)),nrow = 1)
colnames(gmno) <- c('logLikelihood','Number of Parameters')
kable(gmno,format = 'pandoc', caption = 'GN (S3)')
```

#Lee-Carter – (LC Gaussian)

```
mm<-
matrix(c(fitLCgaus$deviance,fitLCgaus$loglik,AIC(fitLCgaus),fitLCgaus$npar),nrow =
1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of Parameters')
kable(mm,format = 'pandoc',caption = 'LC - Gaussian (S3)')
```

Brouhns – (LC Poisson)

```
mm<-
matrix(c(fitLCpois$deviance,fitLCpois$loglik,AIC(fitLCpois),fitLCpois$npar),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC', 'Number of Parameters')
kable(mm,format = 'pandoc',caption = 'LC - Poisson (S3)')
```

Renshaw-Haberman (RH)

```
mm<- matrix(c(fitRH$deviance,fitRH$loglik,AIC(fitRH),fitRH$npar),nrow = 1)
colnames(mm) <- c('Deviance','Loglikelihood','AIC','Number of parameters')
kable(mm,format = 'pandoc',caption = 'RH (S3)')
```

4. CODES FOR FORECAST AND FANPLOTS

```
# library (StMoMo)
```

```
## Lee-Carter model
```

Male – Table 4.16

```
forecast_LC_gausErr_m_kt.t <- forecast.LCGN(  
  
object = LCgaussianStMoMo.m,  
  
h = 20,  
  
level = c(80, 95),  
  
oxt = NULL,  
  
jumpchoice = "fit",  
  
kt.method = "iarima",  
  
kt.order = NULL,  
  
kt.include.constant = TRUE,  
  
kt.lookback = NULL,  
  
N=1)  
  
forecast_LC_gausErr_m_kt.t
```

Fan plot – (Appendix)

```
object_male =makeForecastPlotObj(LCgaussianStMoMo.m,forecast_LC_gausErr_m_kt.t)  
  
plot.GNfor(x= object_male,colour = "grey60",main='forecast (Male): LC Gaussian error')
```

Female – Table 4.17

```
forecast_LC_gausErr_f_kt.t <- forecast.LCGN(  
  
object = LCgaussianStMoMo.f,
```



```

h = 20,

level = c(80, 95),

oxt = NULL,

jumpchoice = "fit",

kt.method = "iarima",

kt.order = NULL,

kt.include.constant = TRUE,

kt.lookback = NULL,

N=1)

forecast_LC_gausErr_f_kt.t

```

Fan plot – (Appendix)

```

object_female = makeForecastPlotObj(LCgaussianStMoMo.f,forecast_LC_gausErr_f_kt.t)

plot.GNfor(x= object_female,colour = "grey60",main='forecast (Female): LC Gaussian
error')

```

Brouhns model

Male – Table 4.18

```
#ARIMA
```

```
forecastArima_LC_pErr_m <- forecast(LC_pois_male,
```

```
h=20,kt.method = "iarima",
```

```

kt.order = NULL)

forecastArima_LC_pErr_m$kt.f

# Forecast kt

par(mfrow = c(1,1))

# Fan plot – (Appendix)

plot(forecastRW_LC_pErr_m,only.kt = TRUE)

# Female – Table 4.19
#ARIMA
forecastArima_LC_pErr_f<- forecast(LC_pois_female,
                                   h=20,kt.method = "iarima",
                                   kt.order = NULL)

forecastArima_LC_pErr_f$kt.f

# Forecast kt

par(mfrow = c(1,1))

# Fan plot – (Appendix)

plot(forecastArima_LC_pErr_f,only.kt = TRUE)

## Gamma Normal Lee-Carter model

# Male – Table 4.20

forecastGN_m_kt.t <- forecast.LCGN(

object = fitGN_m_Object,

h = 20,

level = c(80, 95),

```

```

oxt = NULL,

jumpchoice = "fit",

kt.method = "iarima",

kt.order = NULL,

kt.include.constant = TRUE,

kt.lookback = NULL,

N=1)

forecastGN_m_kt.t

# Fan Plot – (Appendix)

object_male = makeForcastPlotObj(fitGN_m_Object,forecastGN_m_kt.t)

plot.GNfor(x= object_male,colour = "grey60",main='forecast (Male): LC GN error')

# Female – Table 4.21

forecastGN_f_kt.t <- forecast.LCGN(

object = fitGN_f_Object,

h = 20,

oxt = NULL,

jumpchoice = "fit",

level = c(80, 95),

kt.method = "iarima",

```

```

kt.order = NULL,

kt.include.constant = TRUE,

kt.lookback = NULL,

N=1)

forecastGN_f_kt.t

# Fan plot- (Appendix)

object_female = makeForecastPlotObj(fitGN_f_Object,forecastGN_f_kt.t)

plot.GNfor(x= object_female,colour = "grey60",main='forecast (Female): LC GN error')

```

CODES FOR APPENDIX A

Annotated plots of log mortality and untransformed mortality rates -Figures 1 and 2

```

par(mfrow = c(1,2))

plot_dd(nig.dd_male, xlim=c(0, 90), lpar=list(x.int=-0.2, y.int=0.9, cex=0.85))

plot_dd(nig.dd_female, xlim=c(0, 90), lpar=list(x.int=-0.2, y.int=0.9, cex=0.85))

par(mfrow = c(1,2))

plot_dd(nig.dd_male,transf=F, xlim=c(0, 90), lpar=list(x.int=-0.2, y.int=0.9, cex=0.85))

plot_dd(nig.dd_female,transf=F, xlim=c(0, 90), lpar=list(x.int=-0.2, y.int=0.9, cex=0.85))

```

HISTOGRAMS AND QQ PLOTS

Histograms

Male data

Histograms for mortality rates (age-groups) – Figure 3

```
ages<- rownames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_male)){

hist(Mxt_male[,i],probability=T,

main=paste('A',ages[i]),xlab = 'Mxt Male')

lines(density(Mxt_male[,i]),col=2)

}
```

Histograms for log-mortality rate (age-groups) – Figure 4

```
ages<- rownames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_male)){

hist(log(Mxt_male[,i]),breaks = 17,probability=T,

main=paste('A',ages[i]),xlab = 'log(Mxt) Male')

lines(density(log(Mxt_male[,i])),col=2)

}
```

Histograms for mortality rates (time-period) –Figure 5

```

years<- colnames(Mxt_male)

# Mxt - mortality rates

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_male)){

hist(Mxt_male[,i],probability=T,

main=paste('Y',years[i]),xlab = 'Mxt Male')

lines(density(Mxt_male[,i]),col=2)

}

# Histograms for log-mortality rate (time-period) – Figure 6

years<- colnames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_male)){

hist(log(Mxt_male[,i]),breaks = 17,probability=T,

main=paste('Y',years[i]),xlab = 'log(Mxt) Male')

lines(density(log(Mxt_male[,i])),col=2)

}

## Female data

# Histograms for mortality rates (age-groups) – Figure 7

ages<- rownames(Mxt_female)

```

```

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_female)){

hist(Mxt_female[,i],probability=T,

main=paste('A',ages[i]),xlab = 'Mxt female')

lines(density(Mxt_female[,i]),col=2)

}

```

Histograms for log-mortality rates (age-groups) – Figure 8

```

ages<- rownames(Mxt_female)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_female)){

hist(log(Mxt_female[,i]),breaks = 17,probability=T,

main=paste('A',ages[i]),xlab = 'log(Mxt) female')

lines(density(log(Mxt_female[,i])),col=2)

}

```

Histograms for mortality rates (time-period) – Figure 9

```

years<- colnames(Mxt_female)

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_female)){

```

```

hist(Mxt_female[,i],probability=T,
main=paste('Y',years[i]),xlab = 'Mxt female')

lines(density(Mxt_female[,i]),col=2)

}

# Histograms for log-mortality rates (time-period) – Figure 10

years<- colnames(Mxt_female)

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_female)){

hist(log(Mxt_female[,i]),breaks = 17,probability=T,
main=paste('Y',years[i]),xlab = 'log(Mxt) female')

lines(density(log(Mxt_female[,i])),col=2)

}

### Quartile - Quartile Plot

## Male data

# QQ Plots for mortality rates (time-period)-Appendix A

years<- colnames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_male)){

qqnorm(Mxt_male[,i],main=paste("QQ plot Y",years[i],sep = ")),

```



```

xlab="Mxt",pch=19)

qqline(Mxt_male[,i],col=2)

}

# QQ Plots for log-mortality rates (time-period) – Appendix A

years<- colnames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_male)){

qqnorm(log(Mxt_male[,i]),main=paste("QQ plot Y",years[i],sep = "),

xlab="log(Mxt)",pch=19)

qqline(log(Mxt_male[,i]),col=2)

}

# QQ Plots for mortality rates (age-groups) – Appendix A

ages<- rownames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_male)){

qqnorm(Mxt_male[,i],main=paste("QQ plot A",ages[i],sep = "),

xlab="Mxt",pch=19)

qqline(Mxt_male[,i],col=2)

}

```

QQ Plots for mortality rates (age-groups) – Appendix A

```
ages<- rownames(Mxt_male)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_male)){

qqnorm(log(Mxt_male[,i]),main=paste("QQ plot A",ages[i],sep = " ),

xlab="log(Mxt)",pch=19)

qqline(log(Mxt_male[,i]),col=2)

}
```

Female data

QQ Plots for mortality rates (time-period) – Appendix A

```
years<- colnames(Mxt_female)

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_female)){

qqnorm(Mxt_male[,i],main=paste("QQ plot Y",years[i],sep = " ),

xlab="Mxt",pch=19)

qqline(Mxt_female[,i],col=2)

}
```

QQ Plots for log-mortality rates (time-period) – Appendix A

```
years<- colnames(Mxt_female)
```

```

par(mfrow=c(4,4))

for(i in 1:ncol(Mxt_female)){

qqnorm(log(Mxt_female[,i]),main=paste("QQ plot Y",years[i],sep = " ),
xlab="log(Mxt)",pch=19)

qqline(log(Mxt_female[,i]),col=2)

}

```

QQ Plots for mortality rates (age-groups) - Appendix A

```

ages<- rownames(Mxt_female)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_female)){

qqnorm(Mxt_female[,i],main=paste("QQ plot A",ages[i],sep = " ),
xlab="Mxt",pch=19)

qqline(Mxt_female[,i],col=2)

}

```

QQ Plots for log-mortality rates (age-groups)- Appendix A

```

ages<- rownames(Mxt_female)

par(mfrow=c(4,4))

for(i in 1:nrow(Mxt_female)){

qqnorm(log(Mxt_female[,i]),main=paste("QQ plot A",ages[i],sep = " ),

```

```
xlab="log(Mxt)",pch=19)  
  
qqline(log(Mxt_female[,i]),col=2)  
  
}
```