2019

An Inclusive pension model for Kenya's informal sector with late entries and early exit rates.

Lagat, Cherono Asumpta
Strathmore Institute of Mathematical Sciences
Strathmore University

An Inclusive Pension Model for Kenya's Informal Sector With Late Entries and Early Exit Rates

Lagat, Cherono Asumpta

ubmitted in partial fulfillment of the requirements for the Masters of Science in Mathematical Finance at Strathmore University

Strathmore Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

June, 2019

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the dissertation contains no material previously published or written by any other person except where due reference is made in the dissertation itself.

Lagat, Cherono Asumpta

....

June 3^{rd}, 2019.

Approval

This dissertation of Lagat, Cherono Asumpta was reviewed and approved by the following:

Dr. Lucy Muthoni,
Lecturer, Strathmore Institute of Mathematical Sciences,
Strathmore University.

Mr. Ferdinand Othieno,
Dean, Strathmore Institute of Mathematical Sciences,
Strathmore University.

Professor Ruth Kiraka,
Dean, School of Graduate Studies,
Strathmore University.

Abstract

The purpose of this study is three-fold: first, we develop a pension model that uses pre-retirement mobile phone airtime expenditures to accumulate the pension fund. Secondly, we calculate the exit and entry rates into the comprehensive pension scheme. Finally, we determine the expenditure patterns experienced post-retirement and use these patterns to advise on the daily amount required to be charged per minute above the current rate in order to facilitate a comfortable post-retirement life.

The data utilized in this study was retrieved from various secondary sources. Inflation and interest rates data were retrieved from Kenya's Central Bank database. The entry and exit data into informal pension schemes was retrieved from Eagle Africa the administrators of Mbao Pension scheme the largest informal pension scheme in Kenya. The mortality rates were retrieved from the World Health Organization and the life expectancy from World Atlas, Lancet and World Life Expectancy. Pre-retirement data was retrieved from November 2013 from an integrated survey on land ownership and tenure, provision, access and control of basic services, asset ownership, financial resources, evictions and demolition of houses, as well as thirty-two key informant interviews with informal small-scale service providers facilitated by Strathmore University. The inflation and interest rates were forecasted using ARIMA (1,0,5)-GARCH (0,1) model while the backward entry and exit data points were simulated in R.

Our results show that an unemployed Kenyan spends approximately KShs. 2, 000.37 a month considering inflation this amount will translate KShs. 4, 025.45 to maintain the same life style post-retirement assuming the person joins the scheme at 18 and exits at the age of 55. Given the expenditure pre-retirement of this group of people, it will require them to be charged KShs. 3.41 per minute above the current rate in order to raise an amount sufficient to

sustain their lifestyle post-retirement.

Contents

List of Figures

List of Tables

List of Abbreviations

HDI	Human Development Index
ILO	International Labor Organization
KSH	Kenya Shilling
PRB	Population Reference Bureau
RBA	Retirement Benefit Authority
CIR	Cox Ingersoll Rox
CBK	Central Bank of Kenya
NSSF	National Security Social Fund

Chapter 1

Introduction

1.1 Background of Study

Pension funds/schemes are defined as pools of funds which are collected and invested majorly through contributions by beneficiaries to provide for the future in the case of retirement or any situation that may make the beneficiary not to go to work (Davis, 2005). Pension funds provide the medium for people to finance themselves after retirement or people not in a position to work through paying a lump sum or by periodical payments (referred to as annuities) while also supplying funds to the government and other corporations for investment and consumption.

Pension schemes enable the efficient use of taxpayers' funds through pooling and efficiently controlling assets thus achieving economies of scale. The pension one gets helps to boost the economy as they spend the money in the local economy to purchase products, which will then lead to a rise in profits for the businesses manufacturing and processing the products. The broader economy is also boosted which provides a win-win for employers, employees, taxpayers and local government. Pension plans promote economic growth as long as pension models have been formulated correctly. Pension also helps in reducing risks since the pension risks are shared on a wider audience such as the social fund. There are dierent ways in which one can contribute pension as highlighted below:

(i) **State pension** where a certain amount is deducted monthly from your salary and paid to the government and once you reach retirement age, the government starts paying you a fixed fee every month. In Kenya, state pension through the National Social Security Fund is compulsory to contribute to if one is employed, while private pension is not mandatory.

(ii) **Private pension** where one decides to personally save for their retirement without being prompted by the state or the employer. This is similar to the normal savings account except that it attracts tax exemption. As opposed to the state pension where a determined amount is deducted every month, private pension contributions can be made as a fixed monthly or annual amount or even in a lump sum format. This is very common with the self-employed and small to medium enterprises.

(iii) **Company pension** scheme is the third very common scheme, which is set up by the employer as one of the benefits of working for them. One makes a percentage contribution to the pension scheme regularly and in most cases, the employer also contributes a certain amount to the employee's pension account. In Kenya, employers offer this in addition to the state pension scheme.

Under the National Social Security Act of 2013, both employed and the self-employed and their dependents qualify as contributing members of the National Social Security Fund (NSSF) that was established in 1965 through an Act of Parliament Cap 258 of the Laws of Kenya. The contributing members are expected to make their contributions on monthly basis, with the lowest amount payable being KShs. 200 per month for the voluntary members (self-employed) and KShs. 360 per month for the employed.

The Mbao pension, established in 2009 by RBA, targeted people in the informal sector who had seasonal income, lack of access to institutional banking and running Small Market Enterprise. This was the first pension scheme in Kenya to cater for the informal sector who currently are the majority in the labor force in Kenya. The Mbao pension has a minimum amount of KShs. 20 per day and an individual can save daily, weekly, monthly or yearly. It also has no penalty when one withdraws from the scheme. This has become largest successful scheme in the world (Kabare K., 2018). In the 2018 economic survey report from the Kenya National Bureau of Statistics (KNBS), the informal sector contributed 83.4% of the total employment and 787.8 thousand jobs were created in 2017 when compared to 110 thousand created in the modern sector. This shows that the informal sector are the majority and more people join the informal sector every year and there is a need to address the issues that affect the informal sector who do not have consistent monthly sources of income. One of the ways to address it is to come up with a pension that is suitable to the informal sector in order to mitigate old age poverty.

1.2 Pension in Kenya

Kenya's population at the last census in 2009, as reported by the Kenya National Bureau of Statistics (KNBS) was at 38.6 Million. The number of those recorded to be formally employed was at 10.5 Million with only 1.4 Million registered under a retirement benefit scheme. The main reason why people do not take up pension is the mentality that their children will take care of them when they grow old. According to the same statistics, Kenya's birth rate had dropped from 4.6 children per household in 2010 to 3.9 children per household in 2014. The Kenya National Bureau of Statistics also states that the percentage of the age groups as recorded at January 2018 are as follows: 0 to 14 years was 40.02, 19.15% from 15 to 24 years, 33.91% from 25 to 54; 3.92% from 55 to 64 and 3% for those above 65 years.

According to Sundeep (2008), the population of Kenya is still young but is projected to age. It is also estimated that by the time today's new workers retire, the percentage of people above 55 years is predicted to increase to almost triple. Sundeep (2008) studied the Pension System in Kenya and in his research, he found that majority of workers in Africa, Kenya included, are informally employed either in the urban sector or the agricultural sector. They also discovered that the rate of the formal sector workforce has been decreasing rapidly over the past two decades. However, the female gender constitutes 50.1% of the total population but only 29.4% of the female population are formally employed. These proportion of females that are formally employed earn 33% less than the male gender. Kenya's current retirement benefits scheme is made up of the following pension plans:

Scheme Type	National Social Security Fund	Public Service Pension Schemes	Occupational Schemes	Individual Schemes
Legal Structure	Act of Parliament	Act of Parliament	Established under Trust	Established under Trust
Membership	Employees in formal sector establishments with 5+employees excluding public service employees	All public service employees, including civil servants, teachers and disciplined forces. Separate scheme- for armed forces	Formal sector workers in companies that operate retirement schemes	Open to all on voluntary basis
Funding	Funded	Not Funded	Funded	Funded
Regulation	RBA	Act of Parliament	RBA	RBA

Table 1.1: Kenya's Current Retirement Benefit Scheme

There are four criteria that are used in order to evaluate the components of the current retirement benefits pension system in Kenya. These criteria include:

(i) **Adequacy** – The returns cover the full extent of the population making them sufficient to mitigate poverty in the old age and to establish reliable methods that ensure no poverty for the larger population.

(ii) **Affordability** – The returns are made affordable to people, organizations and the society at large.

(iii) **Sustainability** – The benefits are financially sound through the time horizon of the beneficiary.

(iv) **Robustness** – The benefits scheme has the capacity to withstand both good and bad news (major shocks) such as demographic trends.

(Caroline, 2007) studied pensions in developing countries and discovered that older people in Africa need money and special care but very few of them are actually willing to take up a pension scheme at an early age. According to Caroline van Dullemen (Caroline, 2007), 20% of the world's people living on less than a dollar a day are over 60 years (estimated 100 million people). (Caroline, 2007) seeks to establish an intergenerational solidarity which would strengthen economic and social ties in order to alleviate poverty for older people themselves and intergenerational effects as well, which will stimulate more people to enroll in schools and those already in schools to continue with their studies and thus improve the nutrition and basic health of the young people.

1.3 The Informal Sector in Kenya and Pension Funds

(Cynthia Onyango et al, 2016) studied the attitude of the informal sector labor towards saving for retirement pensions. The research study aimed at distinguishing the perception of the workers towards the existing pensions with a particular concern to Mbao pension plan because the pension scheme is specifically designed for low-income earners. The authors studied six groups that were composed of three males and three female groups with three distinct age groups comprising of ages 18-29, 30-40 and 41-55years. Each group had an approximate number of six to eight members. The results of this study showed that participants were more focused on their immediate lives and did not give much thinking into their future. (Richard H. et al., 1981) studied an economic theory of self-control where the two authors established that the concept of self-control is incorporated in a theory of individual's preferences by modeling the person as an organization. According to the results derived, both individuals and organizations utilize the same techniques to curb obstacles created by the conflict between them.

(Richard H. et al, 2004) studied the use of behavioral economics to increase the employees' savings. The two authors explained that employees bear more responsibility in terms of deciding whether to save in defined-contribution plans as opposed to defined-benefit plans.

They also found out that employees either failed to join the plan or those in the defined-contribution plan, participated in a very low level lower than the estimated savings rates life cycle. (Richard H. et al, 2004) propose a prescriptive savings program known as Save More Tomorrow (SMarT). This program ensures that people commit in advance to allocating a fraction of their future salary to a retirement savings plan. (Maloney et al, 2007) study the informal sector, more so, how to care and measure it. In their research, they seek to define the term informality, why we may care about the term, and how to measure it.

1.4 Problem Statement

A vast majority of the world's 1.5 billion youth live in poor countries, with nearly 1.3 billion living in developing countries and one in five living on less than $1 a day. (World Development report, 2007). Low-income youth tend to start working early in life and enroll into complex financial transactions and engagements early in life. The youth who are out of school, are over represented even in places where the unemployment rates are low. (Emerging perspectives on youth Savings, 2011).

According to the United Nations, the median age in Kenya stands at 19.2 as at January 2018. Despite their huge population, the Kenyan youth are yet to fully adopt the culture of saving as reported by the FSD Kenya (Financial Sector Deepening Kenya, 2015). The report indicates that 52% of the non-savers are between the ages of 18 to 24 years. This can be explained by high levels of unemployment.

The United Nations Human Development Index (HDI) report 2017 puts the rate of unemployment in Kenya at 39.1%, the highest in the East African Region. The UN findings further indicate that youth unemployment in Kenya is the highest in East Africa at 17.6% when compared to its neighbors with Uganda at 6%, Tanzania at 6.3%, Ethiopia at 7.6%, and Rwanda having the lowest rate at 3%. The unemployment thus makes it hard for the youth to save.

Given the youthful nature of its population, Kenya is currently on the High Child Dependency ratio but as the years go by with the report that the life expectancy will grow to 73.9 years there is a likelihood for Kenya to be High Old age Dependency.

High Old age dependency can be reduced by enrolling in a good pension scheme. Currently, more than 12 million working Kenyans are not enrolled in any formal pension plan and about 10 million un-employed Kenyans may not be saving for post-retirement. Of interest to note is that majority of them have access to mobile phones, which can be used as a tool for saving towards retirement.

According to the Communication Authority of Kenya, mobile penetration in Kenya stands at 88.1% with 37.8 million subscribers as at 2015. The use of mobile among the youth has been increasing steadily with their consumption rise to 23.6 billion on mobile expense while spending KShs.64 billion annually on clothes and other accessories. This number indicate that the youth are capable of saving if they are encouraged to reduce their consumption and increase savings if there is adequate incentives or mechanisms. This incentive comes in terms of pension or an income for them which will generate after they are not able to work.

Recently, NSSF has tried to introduce the cashless platform through the MPESA to be able to attract the informal sector who are currently estimated to be around 14 million in Kenya to encourage them to save. Unfortunately, not many of them registered from the informal sector posing the above mentioned risk of high old age dependency during retirement. With this regard, the adults who are currently in both formal and informal sector need to start saving early for them not to depend on the younger ones for assistance during their retirement.

1.5 Research Objectives

The objective of this study is to:

(i) To calculate the post retirement amount considering various assumptions;

(ii) To calculate the exist and entry points;

(iii) To determine spending patterns post retirement and advise on the amount needed to be saved pre-retirement.

1.6 Research Significance

(i) The findings of this research will inform the designing of a pension fund that effectively and specifically provides for the needs of the informal sector and the unemployed.

(ii) It will also facilitate easier saving towards retirement for the informal sector due to the introduction of automatic savings from mobile airtime, reducing the need for one to make a deliberate effort to save. The high penetration of mobile phones in the country makes this achievable.

(iii) Given the dynamics considered in this study, the variation between the amounts needed post retirement and the actual savings is minimized, giving a platform that is reflective of savings and expenditure patterns of Kenyan youth and those in informal sector.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides a discussion on the relevant literature that has been reviewed in regards to early theories of consumption and saving. Among the issues featured in this chapter are the various pension schemes in Africa and the challenges they have and we relate this challenge to the Kenyan pension and come up with ideal pension plan for the informal sector in Kenya who are the majority. We also discuss factors that affect the performance of pension schemes in Kenya as we review the relevant literature and theories that elaborate pension schemes and Kenya's informal sector.

2.2 Theoretical Review

There is a lot of literature on pension funds and more so focused on the informal sector. However, recent developments indicate that the current pension funds, that is, the self-pension are not sufficient and effective to the informal sector.

2.2.1 The General Theory of Employment, Interest and Money

John Maynard Keynes came up with the first theory of saving which was named after him, the General Theory of Employment, Interest and Money in 1936. Maynard observed saving as basically another type of product that consumers could "purchase." As with other goods, Maynard reasoned that expenses on saving would increase with revenue. This created a potential challenge because when people devote income toward saving instead of consumption. Maynard and other economists agonized that there was a possibility that people would focus their income on savings rather than consumption as national incomes grew in the post-war era. (Sablik, 2016).

2.2.2 The Relative Income Theory

The Maynard' saving theory was ruled out after more research was done by Duesenberry. In the Duesenberry study he hypothesized that the level of consumption and saving by individual is not dependent on his income but also the habit of an individual. According to the relative income theory, for any given relative income distribution, the proportion of income saved by a household will tend to be a distinctive, and invariant, and its function will be increasing. The percentage saved will be independent of the absolute level of income. The relative theory also states that the aggregate saving proportion will be independent of the absolute level of income (Duesenberry, 1949).

2.2.3 The Life-Cycle Hypothesis Theory

The relative income theory was replaced by the Modigliani and Brumberg's (Modigliani et al., 1954) "life cycle theory of consumption" and permanent income hypothesis, (Friedman, 1957). In the permanent hypothesis theory states that the consumer will spend based on their expected future income. The consumer choices will change based on what they anticipate to receive. In the life-cycle hypothesis theory (Ando Modigliani, 1963), individuals do plan their saving and consumption in a way that allows them to maintain the same lifestyle throughout their life. They save when they have an income and dissave when they have retired. The consumption level remains the same throughout their lives.

2.3 Empirical Review

Thaler Shefrin (Thaler, 1981) study ruled out that the people are able to control the amount that one needs to save for future consumption. In their study they designed a model that captures an individual's behavior by using the concept self-control into the inter-temporal consumption. In the model a person was modeled as an organization. The author concluded that from the model, pension plan produces saving at no other ledger fees and penalty fees occur only when saving is withheld involuntarily. The members with mandatory pension plan had total savings higher than those without.

Time-discounting is an important factor that aects wealth accumulation. The general rule of impatient people is that they tend to spend their income abruptly without planning for it thus not saving for the future. People often put high value on instantaneous consumption as compared to consumption in the future thus self-control problems either knowingly or not. Another characteristic of impatient people with money is that they procrastinate important decisions in the view that such decisions will be made when the need arises. These kind of people need the aid of appropriate public policies and governance that will ease their self-control problems (Tanaka Murrok, 2012).

(Thaler Bernartzi, 2004) did a study using behavioral economics to increase employee saving through a program Save More Tomorrow (SMarT). In this program, employees commit to contribute a certain portion of their future salary increase towards retirement saving. This is supposed to help solve "default" behavior by use of standard automatic enrolment. The program was to help employees who are predicated to be saving less than the predicated life cycle saving rates. This plan has helped to avoid procrastination and also deal with the self-control on saving. The employee can opt out if they are pleased with the plan. They concluded that the program was successful and the automatic enrollment was a good example of libertarian paternalism. Employees who did nothing were presumed not to want to join the plan.

The (Child and Youth Finance, 2012) explain more about why the youth are a growing percentage of the Sub-African population, and why many are economically at risk. Financial inclusion for the young people is caused by a number of positive socio-economic results such as the particularly the promotion of a savings culture. Possessing savings has been linked to more income, better health and academic excellence (Child and Youth Finance International Research Working Group 2012; Sharma et al. 2015). The available data shows that the most of the young people in the Sub-Saharan Africa are not saving or using banking services. According to the World Bank, about only 9.3% of the youth in the Sub-Saharan Africa had saved in a formal financial institution in the previous year while only 14% of the youth had saved in the previous year in a savings club (Demirguc-Kunt et al, 2012). Recent national surveys of young people in Kenya estimated that 51% of the young people outlined that they had no access to any banking services while 84% of the young people outlined that they had not used any banking services in the past year. Recent national surveys of young people in Ghana estimated that 28% of the young people outlined that they had no access to any banking services while 68% of the young people outlined that they had not used any banking services in the past year.

The theories of asset accumulation often pay attention to personal ideologies such as economic resources, knowledge, and family support. However, organizational constructs also pay attention to the additional role of policies, and the products and services that are provided by financial institutions. These kinds of organizational constructs include restrictions, access to financial products and services, and incentives, and they may influence a savings culture when personal constructs are also enforced.

In the study they carried out they interviewed the youth who had participated in the youth project to find out what aects their savings behavior's youth were from Kenya and Ghana. They concluded the study by mentioning that the youth did not save due to lack of financial resources, conflicting demands on youth income and lack of encouragement.

A qualitative research was done to examine the attitudes of informal sector workers towards saving for retirement in Kenya. They divided the informal workers into six groups, three male groups and three female groups. Each group consistent of between eight to twelve participants. The findings from the research was that informal sector have a positive attitude towards saving. The participants were also asked if they had thought about retirement majority between the age of 18-29 age group admitted that they had a very little thought of the kind of life they would live when they are old. Majority of them are concentrating to make ends meet and addressing immediate needs than to save for retirement. The perception in all groups is that they did not trust the state to run their pension contribution.

The innovations of mobile money are critical in empowering the young people and thus enhance youth financial inclusion. Youth financial inclusion, on the other hand, has impelled support all over the world thus gaining the attention of the financial sector, professionals, policymakers, and scholars. Mobile technology has been the leading factor to the increase in access to information in the world and among rural and remote areas more so, financial products and services such as Kenya's MPESA.

A research was done which showed that the young adults across the developing countries want access to low-cost, flexible and transparent savings products. Mobile money has enabled low-income earners to access flexible and low-cost financial products and services. Mobile money can provide a flexible and accessible saving for the youth and this can be transformed into a pension fund for them since they can access it easily and be able to check how much they have contributed on a monthly basis. This will encourage them to save at the same time it will not be costly and this can reduce the number of trips one can make when going to the bank to check how much money they have saved. Mobile money will encourage the youth to plan and save for their future without any hustle since it will be automatically deducted from the services they are using and this will reduce the youth from spending so much money on clothes.

According to a study done by the RBA majority of the retirees use their pension money to pay medical bills since most insurance companies do not offer medical covers to the aged because of the high risk.
In Kenya due to the urbanization most of the families do not live within the same area and care for the aged is not guaranteed as it was before and the aged are left to fend for themselves.

The retirement Consumption Puzzle is about individuals not being able to plan for expected, large and salient reduction in income. Observing that individuals decrease their debt and in-crease their savings is the is the opposite of what a rational agent would do and thus puzzling from (Modigliani,1954)'s life cycle model perspective. (Olaf son et al, 2017).

According to a study done there is a drop of consumption after retirement and this could be either because of a drop of expenditure or a fall of spending in order to correct the use of over consumption. Majorly the reason was because of the lack of early planning of retirement. In Kenya, according the RBA, this could be a result of reducing spending in order to correct the over consumption after retirement and this is evident by the report done in 2016 where most of the retirees spent their pension saving either to finish house, to help their loved ones by funding their education or to use their pension to get medical checkups.

Any rational plan is to increase saving now in order to be able to live a comfortable considering the fact that life after retirement changes. This is the first paper to document the retirement consumption enigma that has affected the informal sector, more so, in Kenya. The RBA did a survey between the month of February 2017 to April 2017.The sample was drawn from a list of retirees who had retired in the last five years. In the survey they did majority of the retirees had dependents who they had to support financially.52.46% of the retirees depended on farming for their economic activities.

The survey also helped to identify the number of years they had saved and majority of them had saved for 25 years yet they still said that the savings were not enough for them after they had retired and wished they had saved more. 62.72% of the retirees mentioned that disposable income was a deterrent for them to make enough contribution and 26.52% indicated they had no information about making enough contribution. The was also an issue in the survey where by 41.5% changed jobs before retirement they withdrew their accumulated savings and this was going to aect them after retirement.

The retirees also indicated that it was worthwhile to save since they were able to receive income to cater for their every day needs. They also indicated that they wished they had saved enough before retirement.

2.4 Attitudes of Informal Sector Workers towards Saving for Retirement

Every Kenyan citizen, regardless of their financial status or their employment status, is entitled to the right to social security as stated in the country's constitution. The right to social security is also one of the core objectives of Kenya's Vision 2030 in order to revamp the standard of lives of the citizens. There are numerous social protection funding plans in Kenya which are either personal, tax-financed such as the Inua Jamii Senior Citizens' Grant, government-run, health insurance, or contributory schemes such as NSSF and Mbao Pension Plan. Currently, the pension schemes coverage in Kenya is at 23% (which is very low) and comprises of majorly formal workforce registered under the NSSF.

(Kabare, 2018) prepared a report on the Mbao Pension Plan more so to inform on the savings culture for the non-formal sector. The Mbao Pension Plan is a voluntary private retirement benefit scheme which was established in 2009 with the main of reaching out to the informal sector in order for them to pool funds and invest their savings thus mitigating the old-age poverty. The informal sector is one of the largest growing labor forces in Kenya comprising of 83% of adults in Kenya and is expected to grow even rapidly. Working in the informal sector has its own ups and downs with the downside being job insecurity, social insecurity, uncouth working conditions and merger salaries. Majority of the informal sector lack social insecurity and no set infrastructure of the older people leading to a crisis after retirement where most old people depend on their children. Without the ability to put aside money for after retirement, most old people are left with little or less than average income security and find themselves continuing to work even after they have hit the retirement age. This is one of the major causes of illness and disability for the old-aged people. The Mbao Pension Plan is more like a provident fund that is well-suited for the informal sector as compared to a pension scheme. The main objectives of the Mbao Pension Plan are as shown in the figure 2.1 below:

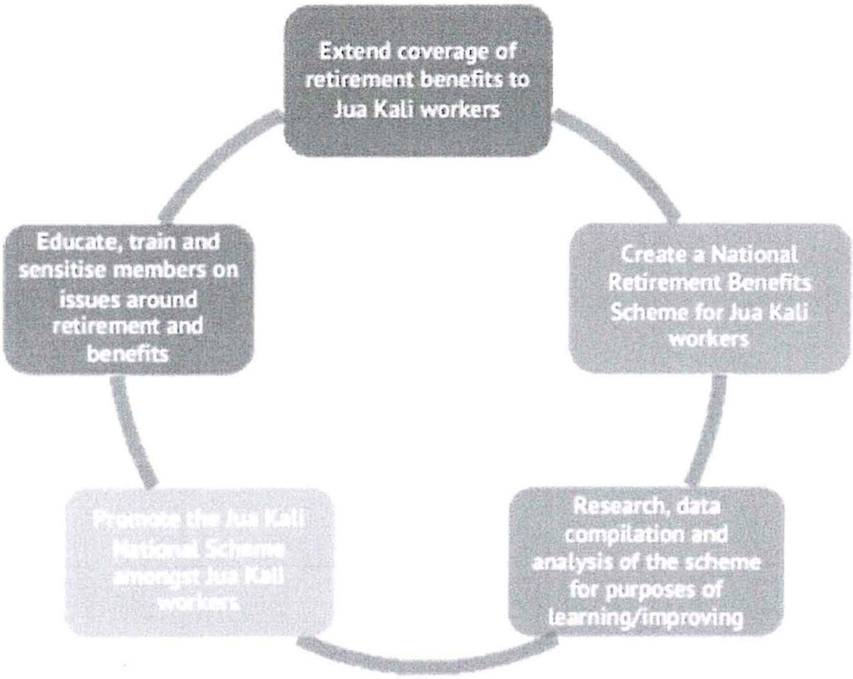


Figure 2.1: Mbao Pension Scheme

Another pension plan scheme that was established in 2018 to address this challenge is called the 'Inua Jamii Senior Citizens' Grant'. The Inua Jamii Senior Citizens' Grant provides USD20 per month to all citizens that are above 70 years. However, the major challenge of this pension plan is that it is majorly tax-financed and as such it does not allow the working population to set aside their own income for retirement. This shows that there is a clear need to tackle the various challenges and gaps in the pension schemes in Kenya.

2.5 Success Factors that Facilitate a Sustainable Micro Pension Scheme

Micro-pension refers to the long-term saving by low-income earners more so the informal sector laborers in order to secure income after retirement or on old age. Micro-pension schemes were typically launched as defined contribution pension schemes in order to provide contributions that are collected at a member's own convenience. The micro-pension scheme model requires there to be an equilibrium between economic viability and creation of sufficient returns to the policyholders. In order to establish the critical success aspects for a micro-pension scheme to operate effectively, the realistic models for its contraption, regulatory framework, problems that would affect its implementation and the strategies that could address these problems need to be clearly defined.

(Hu et al, 2009) studied the pension coverage and the informal sector workers and the authors discover that most pension reforms in the world have mostly focused on the formal sector leaving out the workforce in the informal sector out of the pension dispositions. This is more so in developing countries especially low-income earners who are prone to economic volatility and change. There is potential for micro-pensions to play a critical role in the country's social security system if integrated well by ensuring risk is shared appropriately.

2.6 Barriers to Retirement Benefits Schemes Participation by Low-Income Earners

One of the core challenges facing social security immunity in the world today is that more than 50% of the world's total population are excluded from social security protection.
According to (Lazarus, 2006), social security non-coverage is dominant in the Sub-Saharan part of Africa and also the South Asia with more than 90% of the population lacking cover. One of the main factors that has enhanced low social security coverage is the increase in informalization or the flexibility in the workforce market conditions leaving out most of the workers from social security. (Lazarus, 2006) attempted to explain this challenge of poor social security coverage by utilizing secondary information gathered from different works of scholars and concluded that the results may be controversial thereby leading to existing gaps in social security coverage.

The youth are growing rapidly and majority of them are economically vulnerable. One of the priorities internationally is to ensure financial and social inclusion to all people. However, this is a major concern as financial inclusion for the youth is inhibited by the lack of a savings behavior by the youth. Majority of the youth in Kenya are not however saving due to a number of reasons such as inadequate monthly income.

(Zou et al, 2015) studied the facilitators and obstacles in promoting a savings culture for the youth in particular concern for Ghana and Kenya. In their study, the authors sought to understand the factors that influence the youth to save money and the factors that hinder the youth from saving money. The authors did so by conducting interviews of four groups that were made up of the young adults, a parent, a guardian, and a school staff in both countries. The results in the study indicated that one of the major factors that encouraged the youth to save is the support they got from either their parents, their guardians, the school they were in or from financial institutions. The core hindrance of a savings culture among the youth is that majority of the youth struggle with limited financial planning and resources. (Zou et al, 2015) then went ahead to understand the individual and institutional reasons why the young adults in Kenya, Ghana and other Sub-Saharan African states were not saving given that promoting financial inclusion is one of the major functions of banking institutions. One of the questions that arose in the study was how much income people needed in order for them to maintain their standard of living before retirement (Zou et al, 2015). The authors answered this question by studying the expectations of people upon retirement and their acuity of their quality of life after retirement.

2.7 Research Gap

With a perfect pension system all members would be covered regardless of their employment type. Kenya, in a bid to move towards the ideal scenario, undertook the following: All employers were required to contribute towards employees' pension regardless of number of employees, as long as the said employees earned above a given value, Mwaura K., (2009). However, this did not cater for those under the value stipulated in the Act. This led to the development of Kulegalega Campaign Retirement Benefit Authority (2016) which required that every employed person had to be pensionable, regardless of the amount of their salaries. This ensured that all salaried people became pensionable. However, the unemployed, and those in informal sector had no platform to save towards their pension.

To address this , Mbao Pension Scheme was introduced to cater for the unemployed and those in informal sector, Kwena (2018). It required a daily saving of KShs. 20 per day for the members in the scheme. The Mbao pension plan, though for those in the informal sector, had two major weaknesses:

(i) As much as 30% of the subscribers did not save consistently as required and eventually exited the scheme. (Wangari and Muthoni 2019) addressed this by creating a model where contributions are made relative to airtime consumption. The contributions would be deducted automatically saving the member the deliberate effort to save.

(ii) The members have to pay a flat rate of Ksh. 20 daily. This amount could be prohibitively high for some, and thus discourage them from saving. (Wangari and Muthoni 2019) addressed this by creating a model where a person would save according to what 'they could afford', given their airtime consumption patterns.

Though (Wangari and Muthoni 2019) addressed most of the valid concerns that made savings towards retirement a hurdle, their assumptions were flawed. They assumed that:

(i) The members join when they are 21 years old and leave at 55 years.

(ii) There will be no lapses and withdrawals

(iii) Expenditure patterns remain the same pre and post retirement

These assumptions are not accurate since in Kenya, a person is considered at adult at the age of 18. In this study, we adjust the age of entry from 21 to age 18.

The second assumption is also inaccurate since given the Mbao Pension Scheme, we see the lapse and exit rates do exist. In this study, we consider these rates and adjust the model as necessary.

Finally, the assumption that the expenditure patters remain the same pre and post retirement is inaccurate, and has been shown to be so by several studies. One of the objectives of this study is to determine the expenditure patterns post retirement and indicate this as a variable in our model.

Chapter 3

Methodology

3.1 Introduction

This chapter elaborates the methodology that was used to accomplish the already established research objective. The research design, target population and sampling, sampling design, sample size, data collection and analysis, are briefly illustrated. The study develops a model that examines the practicality of calling charges as a means of saving. It will also determine the spending patterns for those in retirement and will advise on extra cost to be charged for the calls given the spending patterns post retirement.

3.2 Research Design

This study is exploratory where data was deduced from a census of the pension funds registered by the Retirement Benefits Authority (RBA) in Kenya. The use of USSD as a tool for pension schemes is still yet to be explored and thus the study utilizes that perspective and comes up with a suitable and sustainable model geared towards the informal sector that recalculates the amount for daily saving considering the discounting assumptions and determines spending patterns for those in retirement as adjusted for lapse rates.

3.3 Population and Sampling

The total population refers to the entire universe of a system of interest. The total population is the spectrum of people to which the study can be generalized (Johnston et al, 2009). According to the (Retirement Benefits Authority, 2013) there are one thousand two hundred and sixteen registered pension schemes in Kenya. The total number of University students that were approached to participate in the study were 250 but only responsive to this study were 100. The reason why not every student in the study as responsive was because having the students download the Bandicoot application was a challenge because it required data or Wi-Fi to install and some did not have enough memory space on their phones as well. The major reason why the university students were chosen is because they clearly represent the informal sector since they have similar features such as not being in formal employment, and having either an unpredictable income, a low income or both.

3.4 Data Collection and Analysis

All the data used in this study is secondary data. The pensions data was collected from Eagle Africa while the data on the airtime expenses was collected using Bandicoot. Bandicoot is a toolbox in Python, distributed under MIT license, that is used to extract more than 1442 features from a mobile phone's metadata in order to analyze and visualize them. Other data include the expenditure data, the interest rates and inflation rates data, and the life expectancy data. The expenditure data was collected from residence of Mukuru wa Reuben, Mukuru wa Njenga, Viwandani, Kiandutu and Kibera to aid in modelling the consumption patterns of low-income earners that work in the informal sector. The interest rates and inflation rates data were gotten from the country's Central Bank of Kenya website from the year 2010 to 2018 where as the life table (life expectancy data) of Kenyans was retrieved from the World Health Organization (WHO). Analysis of the data is done in R programming software and MS Excel.

3.5 The Model

3.5.1 The Assumptions

There are various assumptions that we will make before coming up with the pension model. These include:

(i) It is the assumption that the individuals will begin saving for pension from the age of 18 and will retire at 55 years.

(ii) The individual will live to age of 68 years.

(iii) The model assumes late entrants and early withdrawals.

(iv) Finally, the expenditure patterns are not similar for pre and post retirement.

3.5.2 Estimation of model parameters

(i) **Interest Rates**

The interest rates are merely the monthly weighted aggregate savings rates on commercial banks. This data was deduced from the country's Central Bank database from 2010 to 2018 which was forecasted using the ARIMA-GARCH model in order to be used to generate the pension model. The rates show a continuous time mackov process which is stochastic in nature. The assumption is that future interest rates will always be positive.

The ARIMA model is a combination of the Autoregressive (AR) model and the Moving Average (MA) model which are univariate time series models. The ARIMA model utilizes past information of a time series to forecast future values for the time series in the order of $ARIMA\ (p, d, q)$ where p, d, and q are integers that are either greater than or equal to zero. The first parameter p gives the number of autoregressive lags, the second parameter d refers to the order of integration that makes the data stationary, and the third parameter q refers to the number of moving average lags. (Pankratz, 1983; Pfaff, 2008).

A process $\{y_t\}$ is said to be $ARIMA\ (p, d, q)$ if $\Delta^d y_t = (1-L)^d y_t$ is $ARMA\ (p,\ q)$.

In general,

$$\phi(L)(1-L)^d y_t = \theta\varepsilon_t; \{\varepsilon_t\} \sim WN\left(0, \sigma^2\right) \quad (3.5.2.1)$$

Where ε_t follows a white noise (WN) and Δ is the difference operator.

We then define the Lag operator by $L^k y_t = y_{t-k}$ and the autoregressive operator and moving average operator are defined as follows:

$$\phi(L) = 1 - \phi_1(L) - \phi_2\left(L^2\right) - \cdots - \phi_p\left(L^p\right) \quad (3.5.2.2)$$

$$\theta(L) = 1 + \theta(L) + \theta_2\left(L^2\right) + \cdots + \theta_p\left(L^p\right) \quad (3.5.2.3)$$

Where ϕ and θ are the standard autoregressive (AR) and moving average (MA) polynomials of order p and q in variable L, $\phi(L) \neq 0\ for\ |\phi| < 1$, the process $\{y_t\}$ is stationary if and only if , in which case it reduces to an $ARMA(p, q)$ process.

The final time series model is usually selected using a penalty function statistic such as Akaike Information Criterion (AIC or AICc) or Bayesian Information Criterion (BIC) (Akaike, 1974; Schwarz, 1978). The Akaike Information Criterion, constant Akaike Information Criterion, and Bayesian Information Criterion are measures of the goodness of fit for a forecasted statistical model. We tend to choose models that the lowest AIC, BIC or AICc. In the general case:

$$AIC = 2k - 2\log(L)\ or AIC = 2k - 2nlog\left(\frac{RSS}{n}\right) \quad (3.5.2.4)$$

Note that, (Burnham et al, 1998) insist that since AICc converges to AIC as n gets large, AICc should be employed regardless of the sample size.

After a model has passed all the diagnostic checks, it becomes sufficient for predicting. Forecasting or predicting is the procedure of estimating events whose actual outcomes have not yet happened. For a stationary time, series Y_h an *ARMA* $(p,\ q)$ model is expressed as:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_p \varepsilon_{t-p} \varepsilon_t \sim i.i.d.(0, 2\sigma^2) \quad (3.5.2.5)$$

Where ε_t , is a white noise.

The *ARMA* $(p,\ q)$ model can be expressed as weighted sum of white noises ε_t , as:

$$y_t = \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \cdots + \Psi_k \varepsilon_{t-k} \quad (3.5.2.6)$$

Where Ψ weights are functions of the modal parameters Φ 's and θ 's.

The accuracy of the models can be compared using some statistics such as the mean error (ME), the mean absolute error (MAE), the root mean square error. (RMSE), the mean absolute percentage error (MAPE), and the mean percentage error (MPE). We always choose the model with minimum values in the above statistics.

The Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) was developed by (Bollerslev, 1986). GARCH is an extension of the ARCH model which is similar in nature to the extension of an AR to ARMA process. The *GARCH* $(p,\ q)$ model has the same equation as an *ARCH* $(1,1)$ for the log returns the only difference is that the equation for the volatility, includes q new terms, that is:

$$r_t = \sigma_t \varepsilon_t,\ \varepsilon_t \sim N(0,1) \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_q r_{t-q}^2 + \beta_1 r_{t-1}^2 + \cdots + \beta_p r_{t-p}^2 \quad (3.5.2.7)$$

Where $t >$ (max (p, q) and the parameters for the model are $\alpha_0,\ \alpha_1,\ \ldots,\ \alpha_q,\ \beta_1,\ \ldots,\ \beta_p$ for some positive integers p, q.

If $p = 0$, the above $GARCH\ (p,\ q)$ model is reduced to the $ARCH\ (q)$. Thus the $GARCH$ model generalizes the $ARCH$ by introducing values of $\sigma_{t-1}^2,\ \sigma_{t-2}^2$ in the equation:

Let $\{\ r_t\}$ be the mean corrected return, ε_t be a Gaussian white noise with mean zero and unit variance and H_t be the information set or history at time t given by $H_t = \{r_1,\ r_2,\ \ldots,\ r_{t-1}\}$ as in the ARCH model. The process $\{r_t\}$ is a $GARCH\ (1,1)$ model if and only if

$$r_t = \sigma_t \varepsilon_t, \varepsilon_t \sim N\left(0, 1\right) \tag{3.5.2.8}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 r_{t-1}^2 \tag{3.5.2.9}$$

Forecasting using the $GARCH$ model is the same as ūsing the $ARMA$ model. This is because conditional variance of $\{r_t\ \}$ is obtained simply by taking the conditional expectation of the squared mean corrected returns. Assuming a forecasting origin of T; then the l-step ahead volatility forecast will be given by:

$$r_t^2\left(l\right) = E\left[r_{t+1}^2 \middle| r_t\right. \tag{3.5.2.10}$$

$$= \alpha_0 + \sum_{i=1}^{m} \left(\alpha_i + \beta_j\right) E\left(r_{t+l-i}^2 \,\middle|\, r_t\right) - \beta_j \sum_{i=1}^{p} E\left(v_{t+l-i} \middle| r_t\right) \tag{3.5.2.11}$$

Where $r_t^2, \ldots, r_{t+l-m}^2$ and $\sigma_t^2,\ \ldots,\ \sigma_{r+l-p}^2$ are assumed known at time t and the true parameter values α_i and β_j for $i = 1 \ldots m$ are replaced by their estimates. Also, the l-step ahead forecast of the conditional variance in a $GARCH\ (p,\ q)$ model is given by:

$$[r_t^2\left(l\right) = E\left[r_{t+l}^2 \middle| r_t\right] \tag{3.5.2.12}$$

$$= \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \beta_j) E\left(r_{t+l-i}^2 \mid r_t\right) - \beta_j \sum_{(i=1)}^{p} E\left(v_{t+l-i} \mid r_t\right) \tag{3.5.2.13}$$

Where $E\left[r_{t+l}^2 \mid r_t\right]$ for $i<l$ can be given recursively as for $l \geq i$, $E\left(v_{t+l-i} \mid r_t = 0\right)$ for $i<l$, $E\left(v_{t+l-i} \mid r_t\right) = v_{t+l-i}$ for $i \geq l$.

We now consider the methods used for selecting the model with the best fit. As previously stated, the accuracy of the models can be compared using some statistic such as the mean error (ME), the mean absolute error (MAE), the root mean square error. (RMSE), the mean absolute percentage error (MAPE), and the mean percentage error (MPE). We always choose the model with minimum values in the above statistics.

The reason why we opt for an $ARIMA - GARCH$ model instead of any other model like $Cox - Ingersoll\ Rox$ (CIR) model or a basic $ARMA$ or $GARCH$ model is because for a standard generalized autoregressive conditional heteroscedasticity $(GARCH)$ model, the normal distribution does not completely capture completely capture the skewness and leptokurtosis of the financial time series and hence the need for an $ARMA - GARCH$ model which adds another assumption on the z- distribution. Also, the $ARIMA$ model when used alone only forecasts the condition mean (returns) while the $GARCH$ model when used alone only forecasts the conditional variance (volatility) hence when $ARIMA$ and $GARCH$ are combined, the forecast the returns that are adjusted to volatility.
Consider y_t that has unconditional mean zero and follows an $ARMA(p,q) - GARCH(s,r)$ process. An $ARMA - GARCH$ model is defined by

$$[y_t \sim D\left(\mu_t,\ \sigma_t^2\right)] \tag{3.5.2.14}$$

Where

$$\mu_t = \varphi_1 \mu_{t-1} + \cdots + {}_p\mu_{t-p} + (\varphi_1 + \theta_1)\,\mu_{t-1} + \cdots + (\varphi_m + \theta_m)\,\mu_{t-m} \tag{3.5.2.15}$$

and

$$[\sigma_t^2 = \omega + \alpha_1 \mu_{t-1}^2 + \cdots + \alpha_s \mu_{t-s}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_r \sigma_{t-r}^2] \tag{3.5.2.16}$$

$$[\frac{\mu_t}{\sigma_t} \sim i.i.D\,(0,1)] \tag{3.5.2.17}$$

Where $\mu_t = y_t - \mu_t$; D is some density, for instance, normal $\varphi_i = 0$ for $i > p$; and $\theta_j = 0$ for $j > q$.

The **conditional mean** process due to ARMA has essentially the same shape as the **conditional variance** process due to GARCH, just that the lag orders may differ (allowing for a nonzero unconditional mean of y_t should not change this result significantly). Importantly, **neither** has random error terms once conditioned on I_{t-1}, thus **both** are predetermined.

We use this to forecast future interest rates. To estimate the pullback rate α and the long-term average rate β, we run a regression of r_{t+1} against r_t, obtaining the standard error, the gradient and intercept which then are used to solve for σ, α *and* β.

(ii) **Inflation rates**

The inflation rates were as well extracted from the country's Central Bank website from 2010 to 2018. Similarly, they are stochastic nature and portrayed by the mean reverting tendencies and we assume that they are also positive in the future.

Similar to interest rates, we use the ARIMA-GARCH model to forecast future rates as well. According to the (World Data Atlas, 2018), the life expectancy in Kenya was at 67.5 years in 2018 from 66.7 years in 2015 growing at an aggregate rate of 0.41 % per annum. We therefore forecast inflation rates for 50 years, the number of years the individual is expected to live for after joining the scheme.

(iii) **Monthly expenses**

Let Z be the amount needed to sustain similar consumption levels once the individual retires. Z is determined by computing the average monthly consumption expenditure (this was collected from several slums including Mukuru wa Reuben, Mukuru wa Njenga, Viwandani, Kiandutu and Kibera) and then adjusting it for inflation as depicted in the equation below:

$$Z = G\prod_{r=1}^{50}(1+e_r) \tag{3.5.2.18}$$

Where Z is the inflated expenses up to the age of 68, G is the expenditure value at the start of the period and e_r is the forecasted inflation rate for each subsequent period. Assuming that the individual receives their pension income monthly, we determine the total amount required to have been saved by the time the individual retires by discounting the value of the annuity Z from equation 3.5.18 above for 13 years from the age of 68 to 55 such that:

$$Y = Z\alpha_{N=13_{r_t}}^{(12)} \times P_t \tag{3.5.2.19}$$

Where Y is the total discounted amount needed at age 55 to facilitate post retirement expenditure, Z as indicated above, N is the number of years between retirement and death and P_t is the probability of survival. We assume $N = 13$ as the life expectancy of Kenyans remains 68 years; rt are forecasted interest rates.

For each of these years we multiply the accumulated amount by the probability of survival of the individual. Kenya's life tables were sourced from the Global Health Observatory data repository of the World Health Organization (World Health Organisation, 2018)

(iv) **Monthly Savings**

The first step is to calculate the individual's current mobile expenditure. Let X represent the average amount spent per month by the individual on calls and n the total number of respondents. Therefore, the daily mobile expenditure on calls is given by:

$$[Daily\ expenditure\ per\ person = \frac{\sum minutes}{Number\ of\ days} \times 4] \tag{3.5.2.20}$$

The monthly mobile expenditure on calls will be given by:

$$[X = \frac{(\sum Daily\ expenditure)}{n} \times 30] \qquad (3.5.2.21)$$

Where X is the monthly amount spent, 4 is the amount charged currently per minute on airtime and 30 is the number of days in a month.

We then use the forecasted interest rates to accumulate this annuity X monthly for 37 years (this is the period between when the individual makes their first contribution and when they retire). For each of these subsequent years, we multiply the accumulated amount by the probability of survival of the individual at the corresponding age, that is, the accumulated savings amount is given by:

$$[Acc.\ Amt = Ks^{(12)}_{37_{i_t}} \times P_t] \qquad (3.5.2.22)$$

Where K is the amount that needs to be *saved* per month.

Further to this, we modify the pensions model to adjust for entry and exit rates as deduced from the Mbao pension scheme. The modified amount is given by:

$$[Amount = \left(Ks^{(12)}_{37_{i_t}} \times P_t\right) \times \underbrace{exit\ rate \times (1 + entry\ rate)}_{b}] \qquad (3.5.2.23)$$

Where which is retrieved from the Kenyan life table and the rest of the variables are as previously defined.

The entry and exit rates are stochastic in nature and b is used to explain the entry and exit points.

Feasibility testing

Let the daily savings of the individual be denoted by K such that the feasibility equation is given by:

$$Y = Accumulated\ amount$$

$$Za^{(12)}_{N=13_{r_t}} \times P_t = Ks^{(12)}_{37_{i_t}} \times P_t \qquad (3.5.2.24)$$

The value of the accumulated annuity X is initially used as a proxy for K, the savings. Given that there are discrepancies between this figure and the amount Y required, we use solver to determine the amount of X, the monthly expenditure that equates them. This amount is therefore that which requires to be saved monthly i.e. K.

Dividing K by 30 gives us the amount of daily savings required. Dividing this further by the average number of minutes spent on outbound calls daily allows us to determine the extra charge per call which, once deducted, will allow the individual to save sufficiently for retirement. Therefore, the additional charge per minute is given by:

$$[A = \frac{K}{30 \times \frac{minutes}{day}}] \qquad (3.5.2.25)$$

Where A is the additional charge per minute.

(v) Expenditure Patterns

We compute the post-retirement values in order to understand the expenditure patterns for low-using (Bruce et al, 2012). According to (Bruce et al, 2012)'s study, the working-class were asked whether they predicted being able to maintain their quality of life after retirement. Approximately one third of the respondents stated that they did not expect to have enough to maintain their current quality of life. The retirees were asked to compare their living standards before and after retirement. The retirees were also asked whether they expected their current earnings before they retired. The retirees and expected a decrease in the living standards where almost similar with the people who had retired and anticipated an increase in the living standards. Of those who had not retired, 27% reported a decrease in their standard of living while 21% reported an increase in their standard of living. This shows that a decrease in income in associated with retirement. (Bruce et al, 2012) came up with expenditure categories that are directly influenced by retirement (see appendix A). The authors study sought to find out whether the expenditure on the commodity is influenced by any of the first four factors that are specific to retirement, that is, reduction in work-related costs, the increase in health-service consumption, different patterns of time allocation and life-cycle patterns in house financing. We are going to utilize the same criteria to come up with an average probability of reduction.

Furthermore, the expenditure patterns in (Bruce et al, 2012) showed that reduction in work-related costs, increase in health-service consumption, price reductions, home production and increase in leisure expenditure that are directly influenced by retirement where as some of the factors that are indirectly influenced by retirement include home and property insurance, repairs and maintenance such as motor and home-related, credits, corporate payments; most foods except eating out; non-alcoholic drinks; clothing except work-related cost (in the expenditure categories that are directly influenced by retirement); furniture, home appliances, cutlery; electronic equipment; books, gambling, sports equipment (other than golf); non-prescribed pain relievers, sunscreens; air fares, holiday vehicle hire, overseas holidays; pets; personal care (except hair care); personal expenditures; and gifts.We utilized 11% reduction in households and 5% increase in health.

Chapter 4

Data Analysis and Findings

4.1 Inflation Rates

For the inflation rates retrieved from the country's Central Bank website. The figure 4.1 is a time series plot of the original inflation rates:

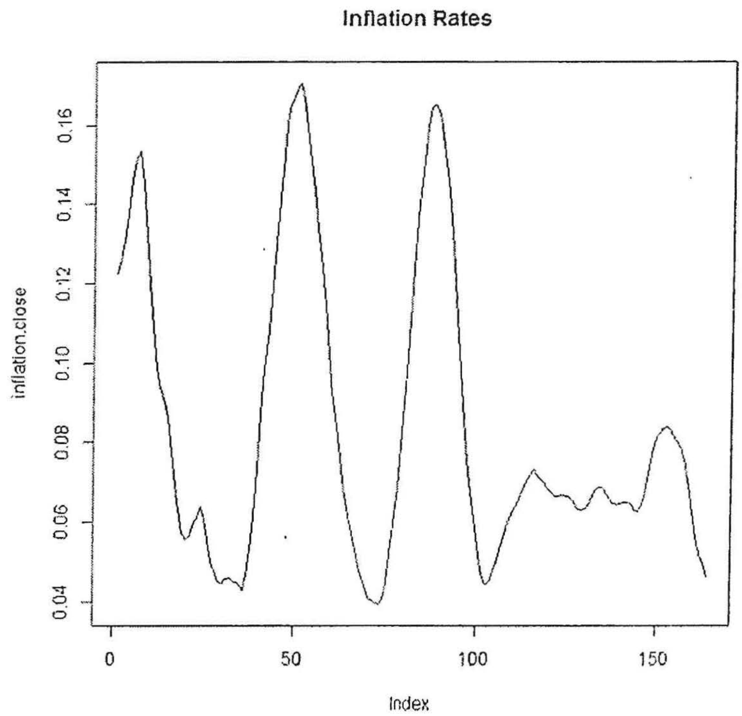


Figure 4.1: Original inflation rates

The figure shows an additive model. The first step is usually to check for the seasonality of the components.

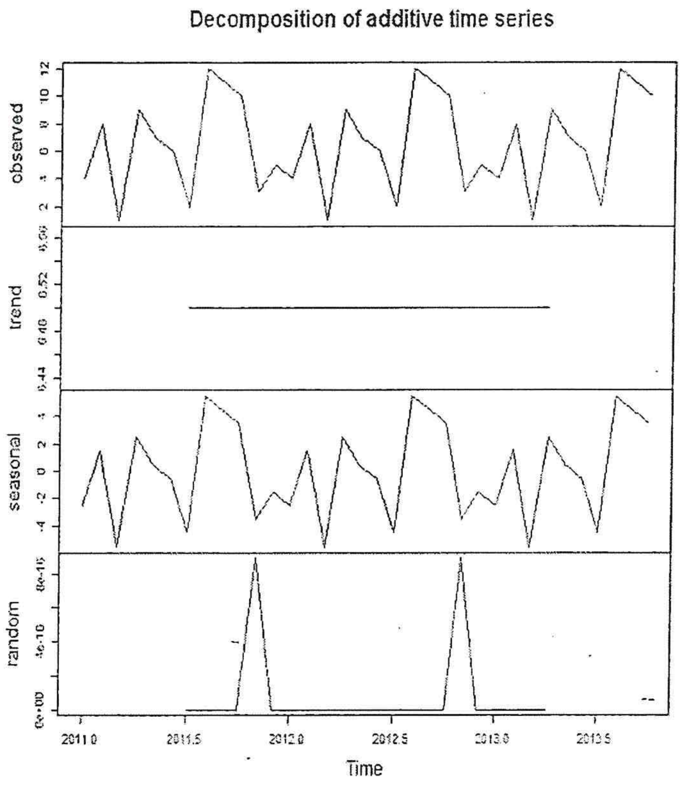
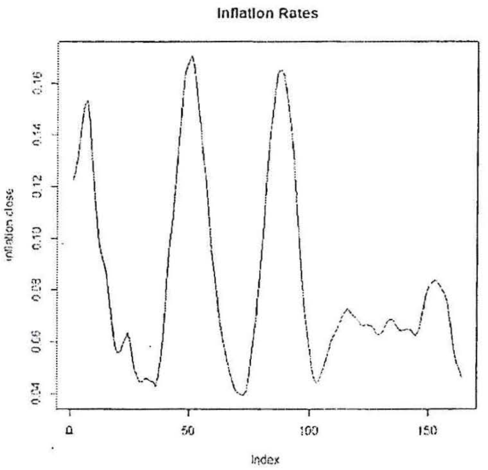


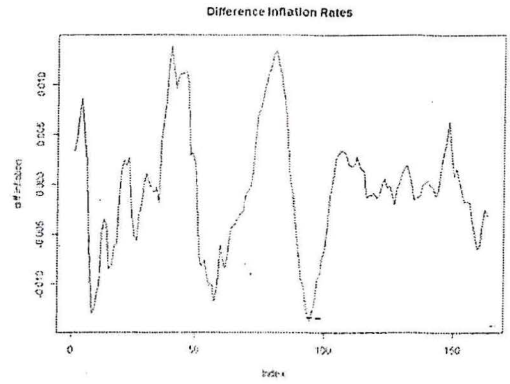
Figure 4.2: Decomposition of the inflation rates

There are four major components of a time series: random, seasonal, cyclical and trend variations. Seasonal variations seek to explain the seasonality of the time series data such that variation happen only on specific periods such as a day, month etc. trend variations seek to explain the predictability pattern of a time series data. Cyclic variations seek to explain the business and economic cycles such as bull and bear that cause major shocks. Random variations are just the white noise disturbances. From the figure 4.2 above the time series analysis of the inflation rates shows that there is a steady trend. The observed and the seasonal look somewhat similar implying that there were barely any seasonal variations between those periods. The random component explains unusual disturbance of the data which might be as a result of the campaign and election period.

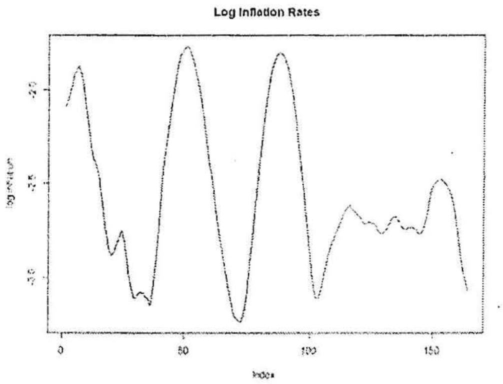
We computed the log and then differenced the values. The figure 4 represents the original inflation rates, the difference inflation rates, the log inflation rates and the difference log inflation rates. The inflation rates figure exhibits exponential growth. The difference inflation rates figure shows that the variance of the series increases as the level of original series increases, and therefore, it is not stationary, the log of the inflation figure exhibits no notable difference as compared to the original inflation rates implying that the original series was already linear. The Difference log inflation figure is more mean-reverting, and variance is constant and does not significantly change as level of original series changes.



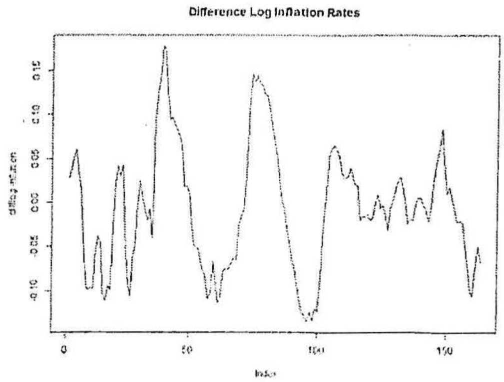
(a) Inflation Rates



(b) Difference Inflation Rates



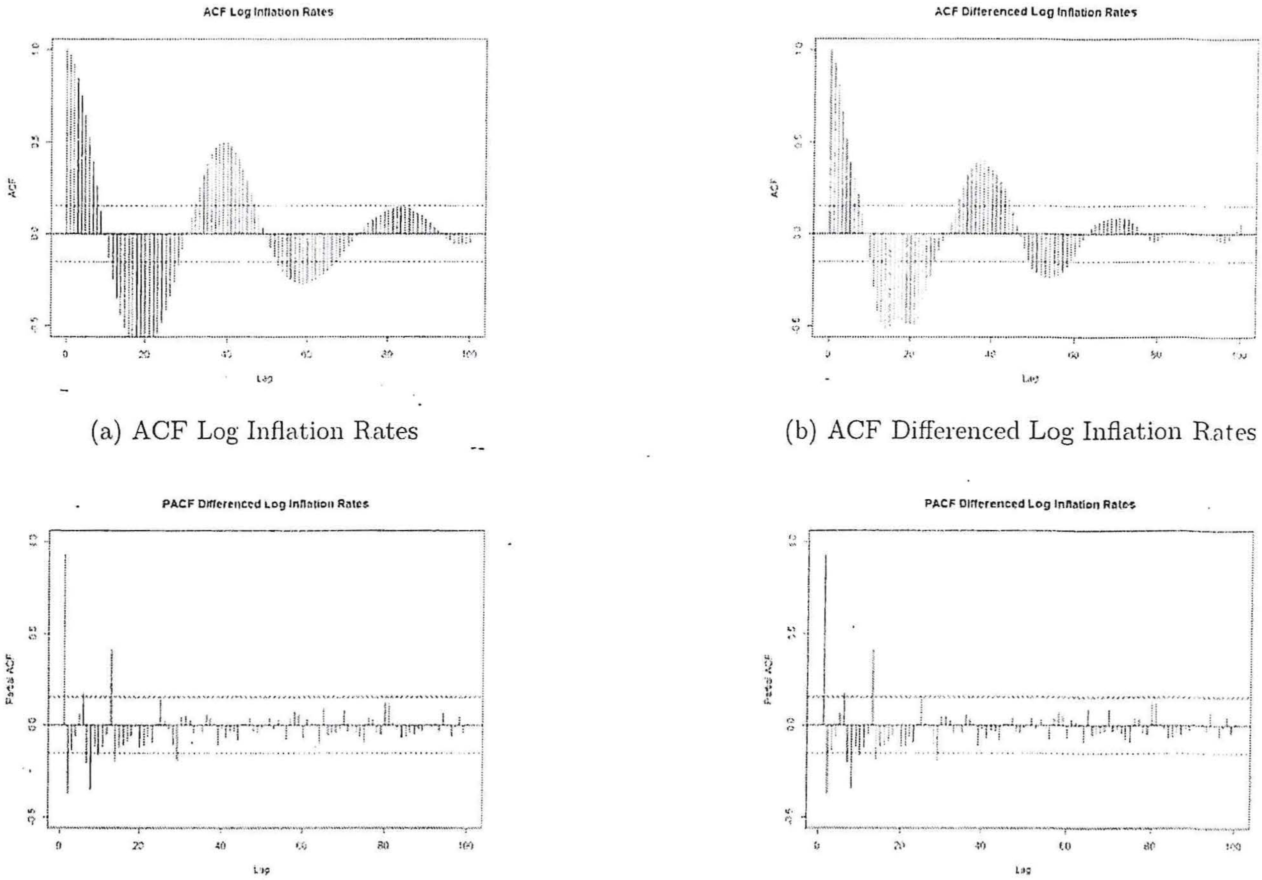
(c) Log Inflation Rates



(d) Difference Log Inflation Rates

Figure 4.3: Comparing the plot of the Original, Logged, Differenced and Diff.log of the Inflation Rate Series

After stationarizing the inflation rates series by differencing, the next step is to determine whether autoregressive or moving average terms are needed to correct any autocorrelation that remains in the differenced series. The best way to do this is by checking the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) plots for any differenced series.



(a) ACF Log Inflation Rates

(b) ACF Differenced Log Inflation Rates

(c) PACF Log Inflation Rates

(d) PACF Differenced Log Inflation Rates

Figure 4.4: Inflation Rates Series-ACF and PACF Plots

	Estimate	Std. Error	z value	Pr(>|z|)
$ar1$	-0.4212	0.0896	-4.7	$2.6e-06$
$ma1$	0.9193	0.0771	11.92	$< 2e-16$
$ma2$	0.6122	0.1174	5.22	$1.8e-07$
$ma3$	0.2622	0.1238	2.12	0.034
$ma4$	-0.2356	0.1107	-2.13	0.033
$ma5$	-0.7351	0.0863	-8.52	$< 2e-16$

Table 4.2: Inflation Rates - ARIMA Model Parameter Estimation

In fitting ARIMA model, p and q should be 2 or less, or the total number of parameters should be less than 3 in view of Box-Jenkins method. The more parameters the greater noise that can be introduced into the model and hence standard deviation. We can compute corrected Akaike Information Criterion (AICc) to find a more defined p and q performed in R programming software where we select the model with the lowest AICc.

Model	AIC
0 1 0	-398.77
1 1 0	-722.22
0 1 1	-565.2
1 1 1	-740.5
0 1 2	-655.55
2 1 0	-747.42
2 1 1	-747.49
2 1 2	-740.65

Table 4.1: Inflation Rates Series-AIC values for testing the best ARIMA(p,q) Model

Based on AICc, we select ARIMA (2,1,1). We conduct an *auto.arima*() in order to determine the appropriate parameter estimates. This gives us $ARIMA(1, 0, 5)$ as shown in figure 5.1

Fitting the ARIMA model gives us the parameter coefficients as follows;

We finally conduct a $GARCH(0,1)$ model on the residuals of the $ARIMA(1,0,5)$ and remove the volatility and we obtain the following graph:

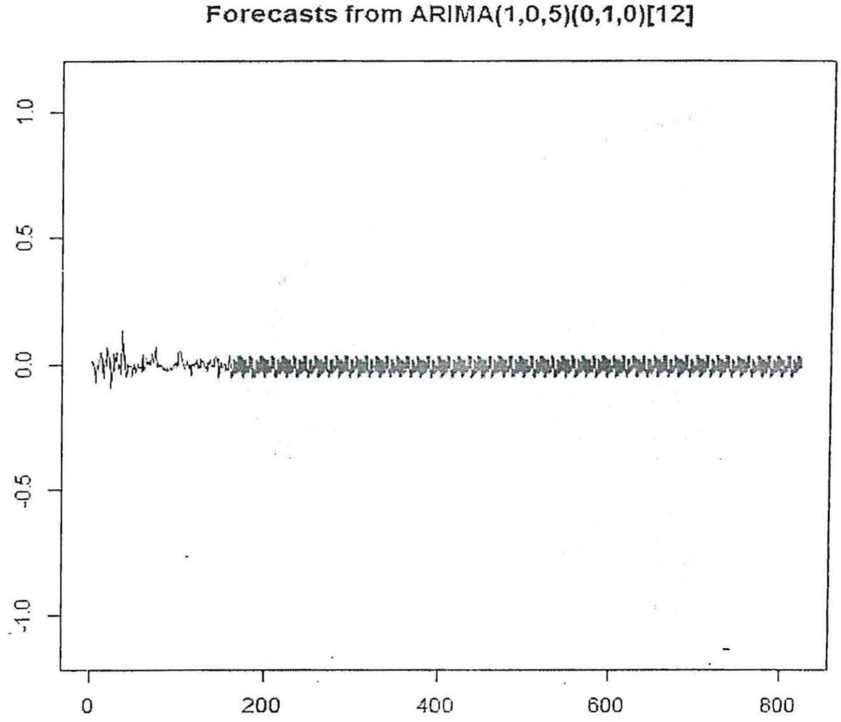


Figure 4.5: (ARIMA(1,0,5)-GARCH(0,1) Model Forecast

The black line shows the actual values where as the blue line shows the forecasted values for period of 162 i.e. monthly data for 13.5 years.

Diagnostic check:

This is done by observing residual plot and its ACF & PACF diagram or by checking the Ljung-Box result. From the Ljung-Box test at lag 100, the p-value is $2.581e^{-08}$ which is less that the significance level of 0.05 we reject the null hypothesis that the selected model is not significant. Since the ACF and PACF of the model residuals show no significant lags, the selected model is appropriate.

Although the ACF and the PACF of residuals and the Ljung-box test have no signifi-

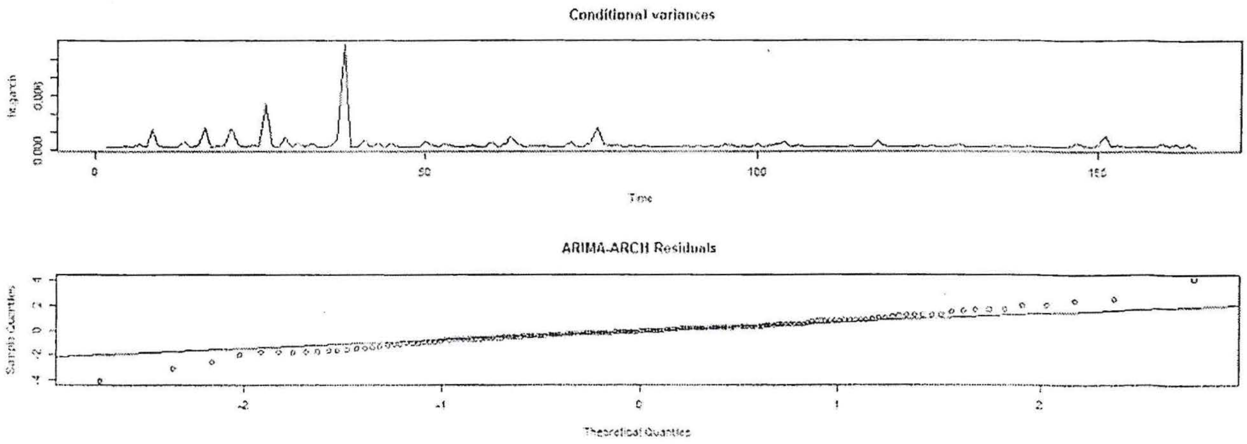


Figure 4.7: Conditional Variance and ARIMA GARCH

4.2 Interest Rates

The methodology here is similar to that of the inflation rates series.The figure 4.8 represents the original plot of the interest rates series:

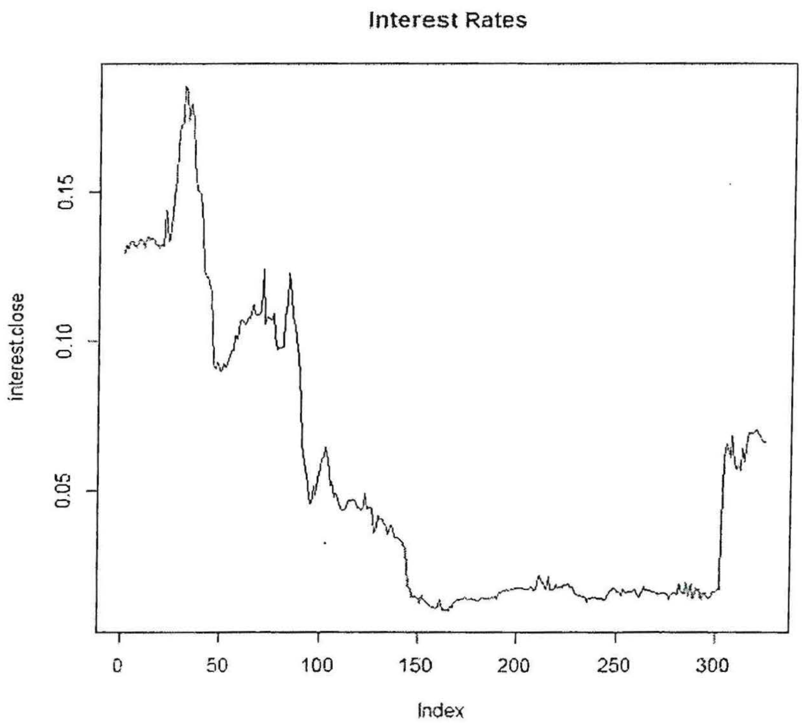


Figure 4.8: Plot of the original interest rates series

The plot shows a multiplicative model. We first analyze the time series components as follows:

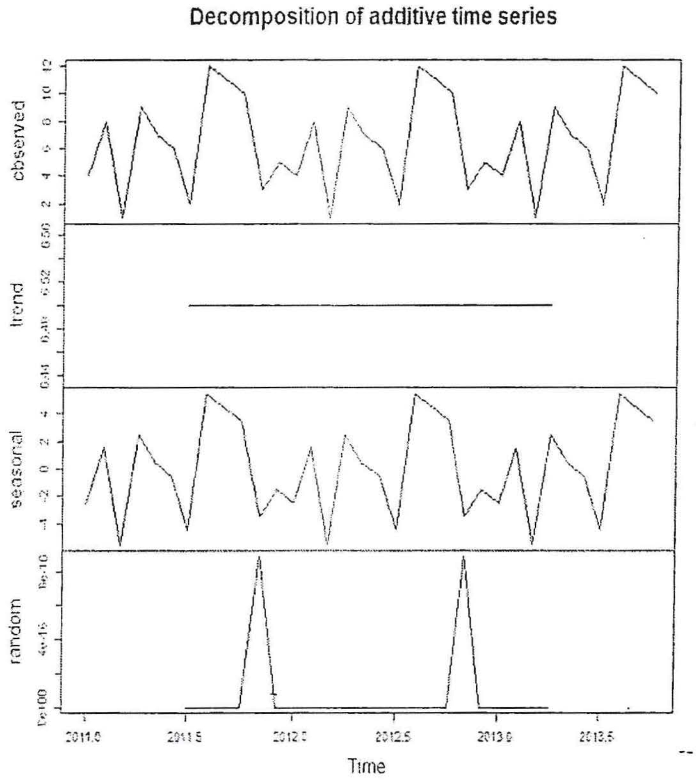


Figure 4.9: Decomposition of the interest rates

Similarly, there is a steady trend in the interest rates series. The observed and the seasonal look somewhat similar implying that there were barely any seasonal variations between those periods.

The random component explains unusual disturbance of the data which might be have been affected by the fluctuation in the inflation rates. We then difference the values as shown by the following graphs:

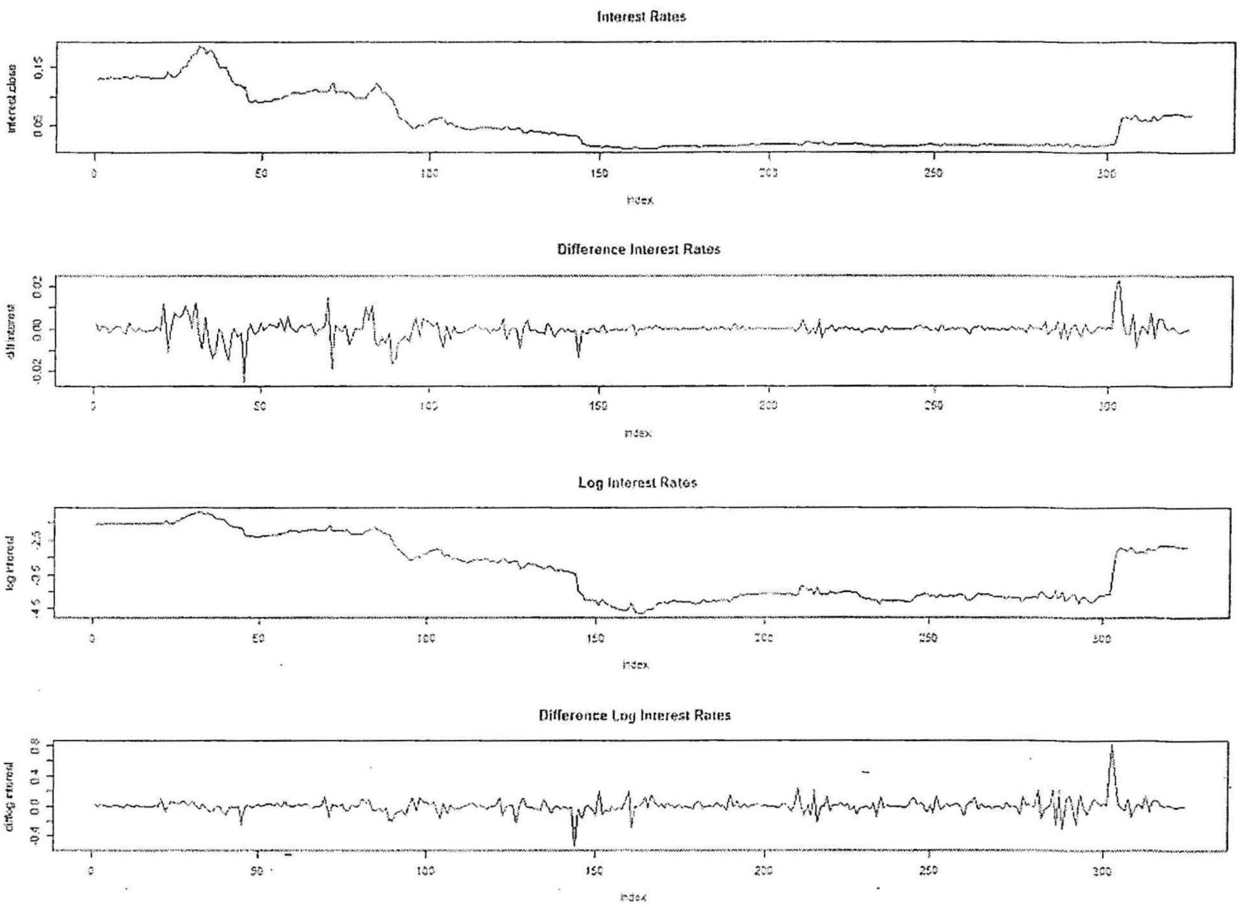


Figure 4.10: Plots for the original,logged,differenced and diff. log of Interest rates

In figure 4.10, the first figure represents the original interest rates series which exhibits exponential growth. The second figure represents the difference interest rates series which shows that the variance of the series increases as the level of original series increases, and therefore, is not stationary, the third figure shows the log of the interest rates series and it exhibits no notable difference as compared to the original graph implying that the original series was already linear. The fourth figure represents the difference log of the interest rates series, it is more mean-reverting, and variance is constant and does not significantly change as level of original series changes.

The ACF of the log interest rates shows a decrease in lags onwards which eventually die down as shown in figure 4.11. The PACF of log interest rates series shows at lag at 4 which is somewhat similar to the ACF of difference log interest rates series. There are no notable lags in the PACF of difference log of interest rates series.

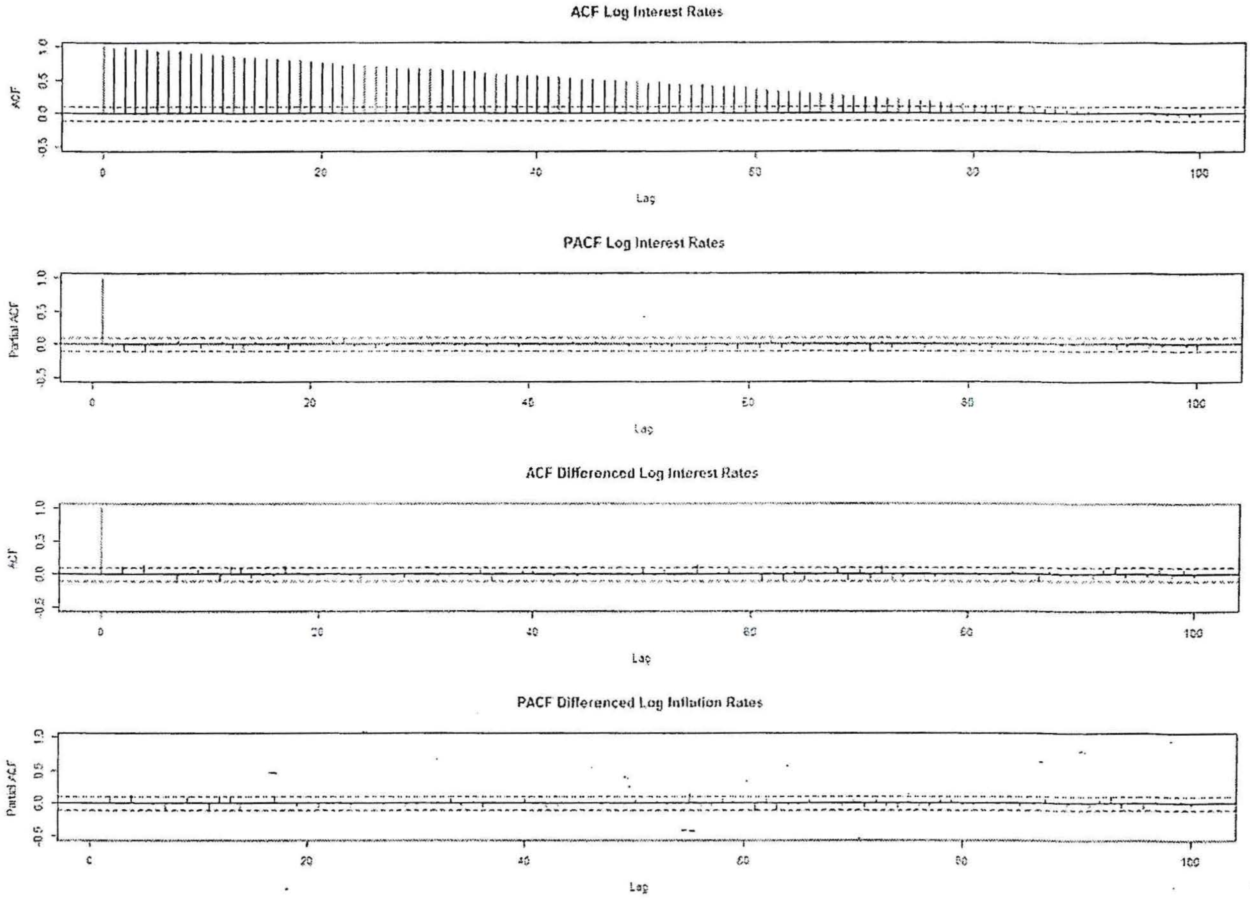


Figure 4.11: Interest Rates Series-ACF and PACF Plots

We can compute corrected Akaike Information Criterion (AICc) to find a more defined p and q performed in R programming software where we select the model with the lowest AICc.

Model	AIC
0 1 0	-398.77
1 1 0	-722.22
0 1 1	-565.2
1 1 1	-740.5
0 1 2	-655.55
2 1 0	-747.42
2 1 1	-747.49
2 1 2	-740.65

Table 4.3: Interest Rates Series-AIC values for testing the best ARIMA(p,q) Model

Based on AICc, we select ARIMA (2 1 1). The simplest method would be to compute the function auto.arima() which gives an accurate estimation of the parameters.

Fitting the ARIMA model gives us the parameter coefficients as follows;

	Estimate	Std. Error	z value	Pr(>|z|)
$ar1$	−0.4212	0.0896	−4.7	$2.6e-06$
$ma1$	0.9193	0.0771	11.92	$< 2e-16$
$ma2$	0.6122	0.1174	5.22	$1.8e-07$
$ma3$	0.2622	0.1238	2.12	0.034
$ma4$	−0.2356	0.1107	−2.13	0.033
$ma5$	−0.7351	0.0863	−8.52	$< 2e-16$

Table 4.4: Interest Rates-ARIMA Model Parameter Estimation

We finally conduct a GARCH(0,1) model on the residuals of the ARIMA(1,0,5) and remove the volatility and we obtain the following graph:

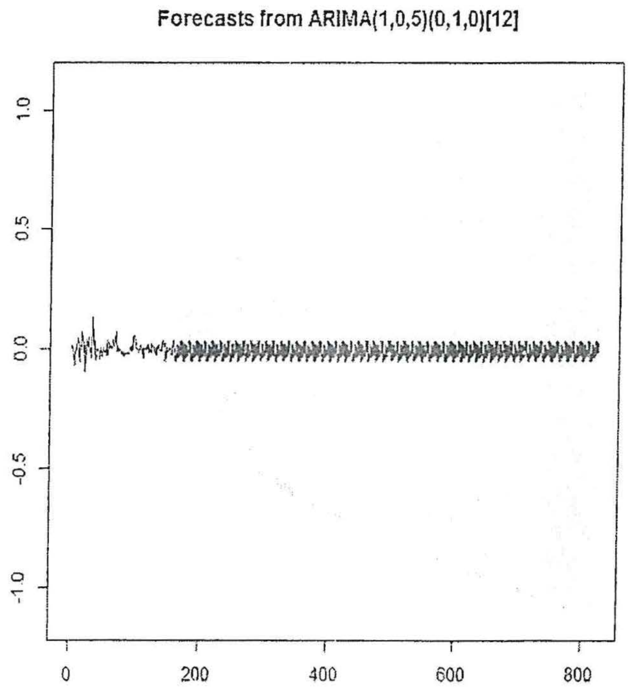


Figure 4.12: ARIMA(1,0,5)-GARCH(0,1) Model Forecast

Diagnostic check

Computing the Ljung-box test, we get:

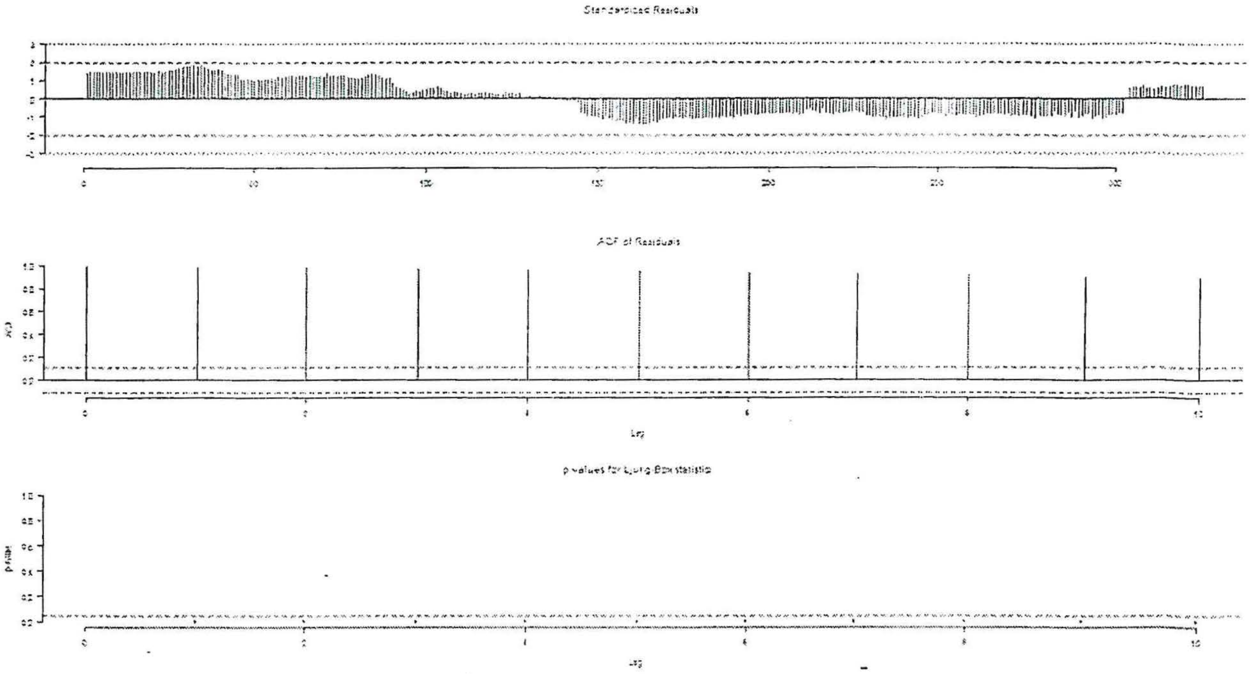


Figure 4.13: Residuals

In order to model volatility, we introduce the ARCH/GARCH model. The plots of the squared residuals plot show cluster of volatility at some points in time. The ACF of the squared residuals seems to die down while the PACF of the squared residuals cuts off after lag 15 even though some remaining lags are significant.

After forecasting ARIMA (1,0,5)-GARCH (0,1), the model converges relatively. We then use the forecasted inflation rates to generate the final amount.

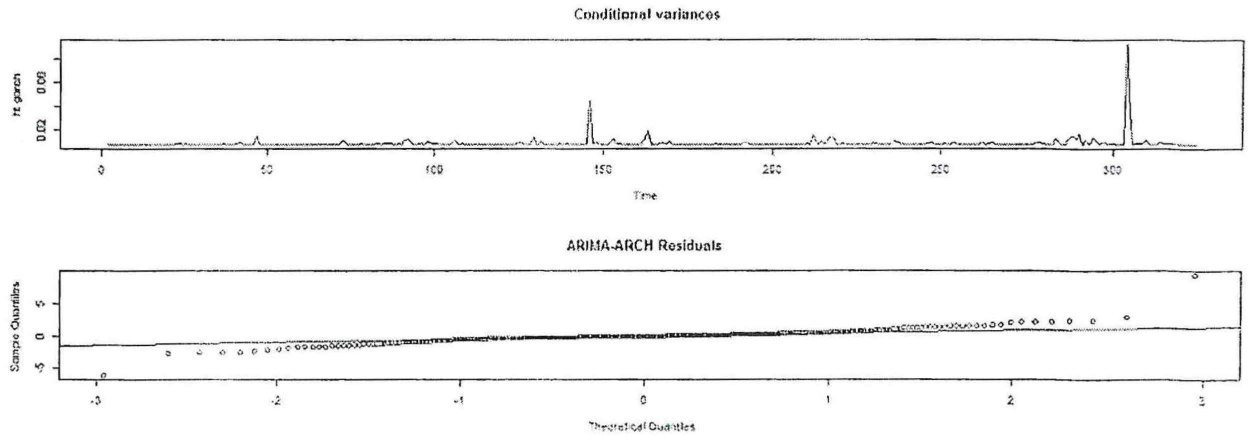


Figure 4.14: Conditional Variance and ARIMA-GARCH Residuals

4.3 Monthly Expenses

Assuming that the individual receives their pension income monthly, we determine the total amount required to have been saved by the time the individual retires by discounting the value of the annuity and multiplying it with the probability of survival. The following graph shows the monthly consumption against time. As time increases, monthly consumption increases steadily up to a certain point majorly because the expenses are subjected to inflation.

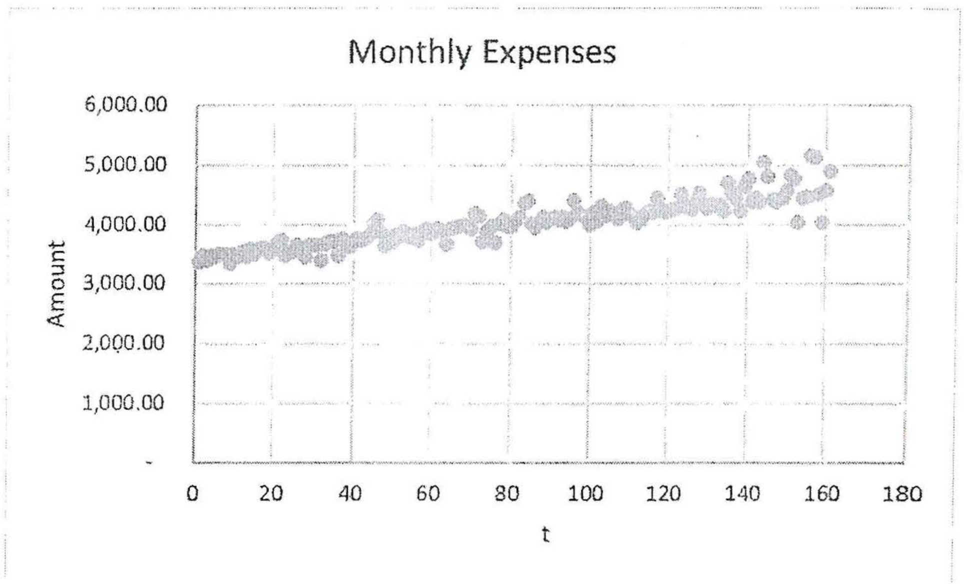


Figure 4.15: Monthly Expenses Scatterplot

According to the expenditure data, the average individual living in Mukuru wa Reuben, Mukuru wa Njenga, Viwandani, Kiandutu and Kibera spends approximately KShs.2009.37 per month. From the inflation forecast model results, we obtain a multiplicative factor of 1.003. This indicates that current expenditure will grow by 100% by the year 2068 such that the individual spending KShs.2009.37 in 2018 will require KShs.4025.45 to maintain the same level of lifestyle then. We estimate that by this year, the individual, assumed to have joined the scheme at the age of 18, will be 55.

4.4 Exit and Entry Rates

To simulate entry and exit rates. the first step is usually to establish the underlying distribution. To test whether it is_id from a normal distribution, plot a histogram of its empirical density and its cumulative distribution as shown below:

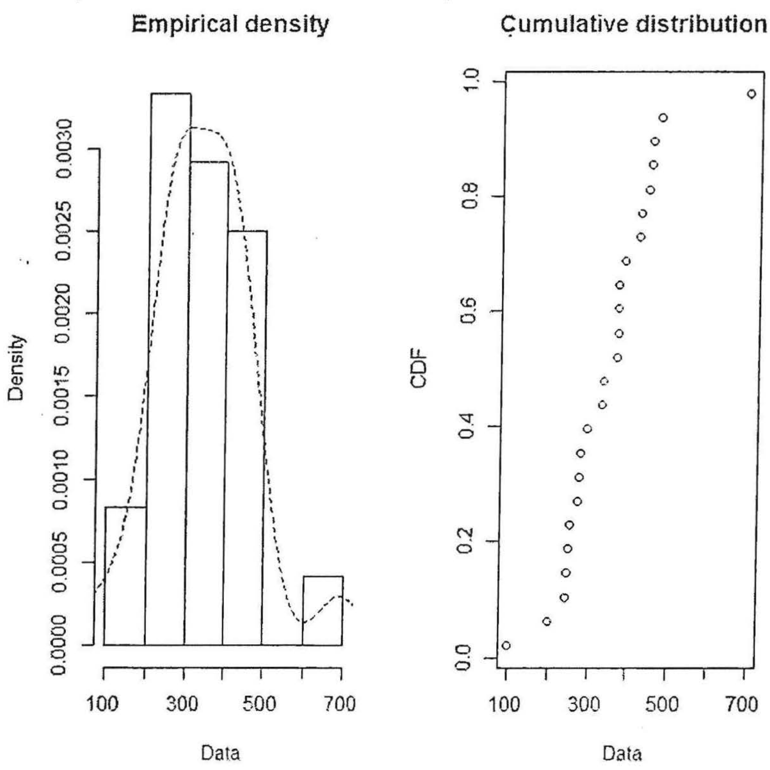


Figure 4.16: Histogram plot of entry processes

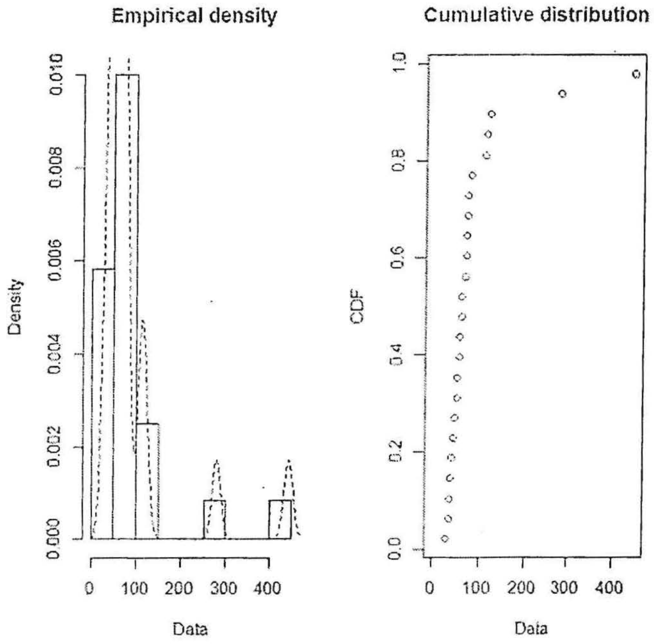


Figure 4.17: Histogram plot of exit processes

The empirical density distribution of the entry processes is more symmetrical as compared to the exit processes. However, this is not proving enough that the entry processes are from the normal distribution, therefore, we text for other statistical distribution using Cullen and Frey.

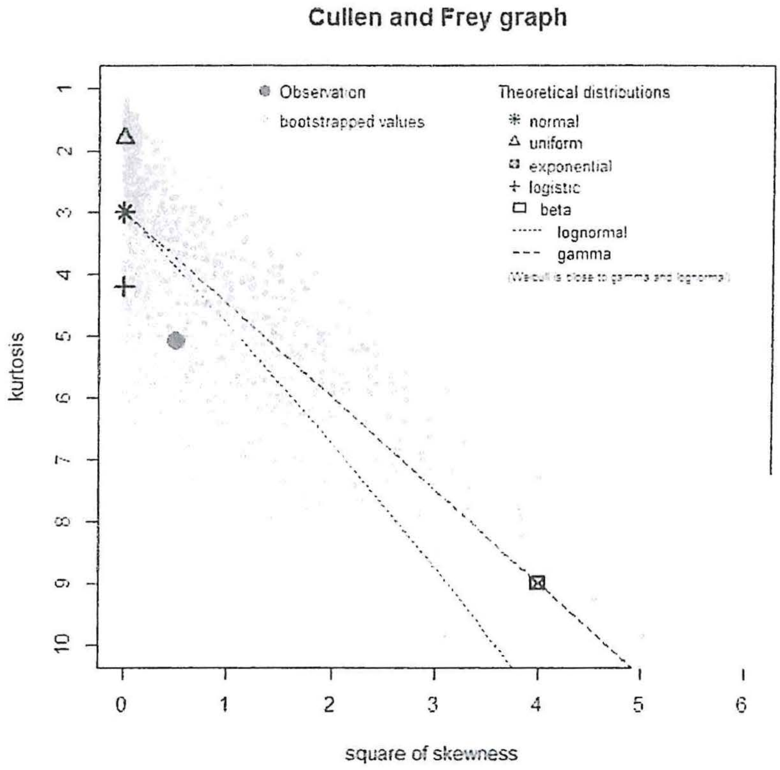


Figure 4.18: Fitting statistical distribution on entry processes

The figure 4.18 shows that the distribution of the entry points is more inclined to the logistic distribution. To be certain that is the distribution, we conduct a goodness-of-fit.

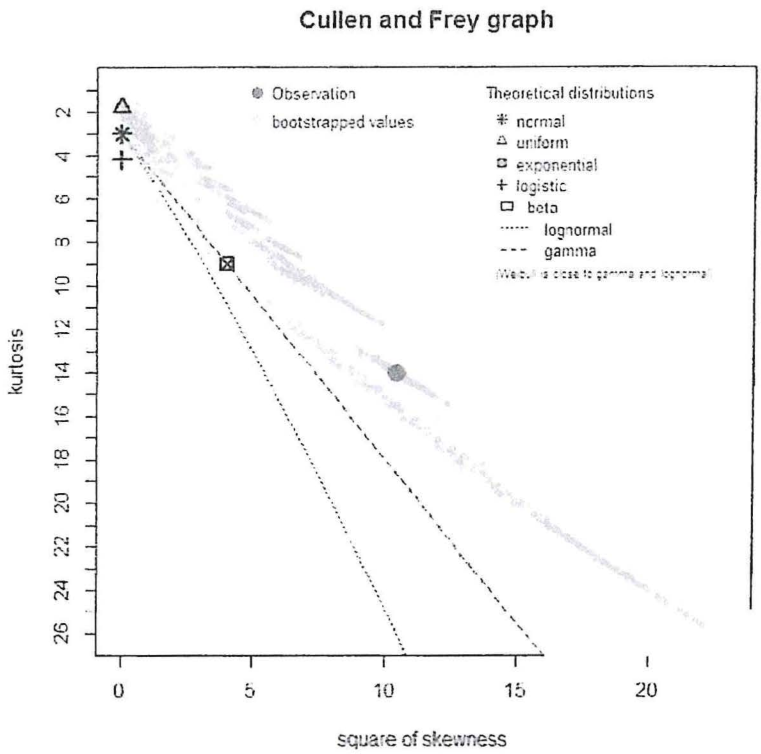


Figure 4.19: Fitting statistical distributions on exit processes

The figure 4.19 shows that the distribution of the exit points is not inclined to any of the distributions in the graph. We therefore conduct a goodness-of-fit to determine which is most suitable.

The following figures show the cumulative density distributions of Weibull, Normal, Lognormal, Gamma, Log-logistic, Pareto and Burr distributions fitted for the entry and exit points.

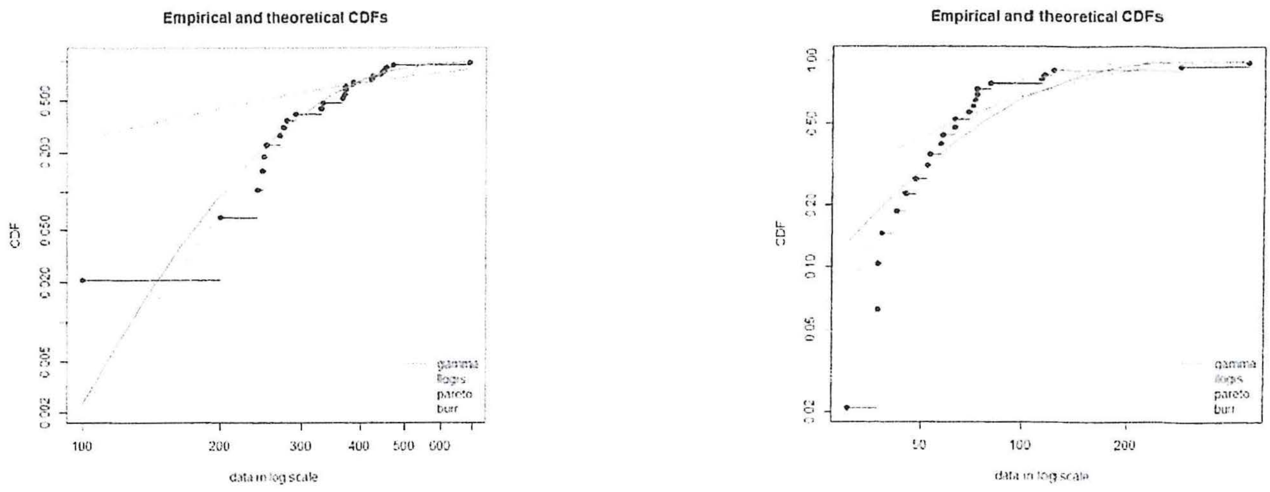


Figure 4.20: Weibull, Normal and Lognormal distrubutions for the entry and exit processes

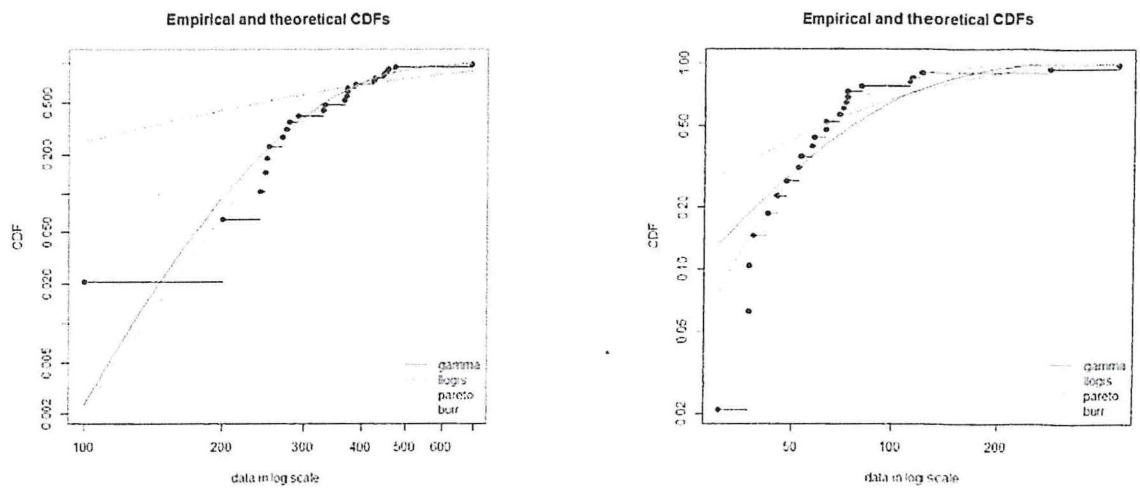


Figure 4.21: Gamma, Log-logistic, Pareto and Burr distributions for entry and exit processes

We conduct a goodness-of-fit test which will be more precise on the distributions of the entry and exit processes.

	Weibull	Norm	Lnorm	Gamma	Llogis	Pareto	Burr
Kolmogorov-Smirnov test	0.1245	0.1007	0.1269	0.1127	0.1195	0.4186	0.1067
Cramer-von Mises	0.0598	0.0507	0.0721	0.0531	0.0526	1.0615	0.4570
Anderson-Darling	0.4806	0.4153	0.5576	0.4193	0.3831	5.1711	0.3338
Akaike's Information Criterion	300.6693	300.4023	302.0945	300.1038	299.8623	332.5828	300.8830
Bayesian Information Criterion	303.0254	302.7584	304.4506	302.4599	302.2184	334.9389	304.4171

Table 4.5: Goodness-of-fit test for entry points

	Weibull	Norm	Lnorm	Gamma	Llogis	Pareto	Burr
Kolmogorov-Smirnov test	0.2509	0.3306	0.2146	0.2673	0.1421	0.2997	0.112
Cramer-von Mises	0.3915	0.7432	0.1621	0.3410	0.0563	0.5099	0.0381
Anderson-Darling	2.2415	3.9412	1.0394	1.9658	0.5760	2.7340	0.2465
Akaike's Information Criterion	264.4411	287.4324	251.5006	260.5718	248.3461	267.2744	245.24
Bayesian Information Criterion	266.7972	289.7885	253.8567	262.9279	250.7022	269.6305	248.77

Table 4.6: Goodness-of-fit test for exit points

From the goodness-of-fit test, we are most interested in the AIC criterion where we choose the least value. For the entry points the least AIC value is at 299.8623 for the log-logistic distribution while for the exit points, the least value is at 245.2410 for the burr distribution.

We then compute the backward probabilities for the Markov chain process since entry and exit processed are stochastic in nature, more so the Markov process. These together with the distributions we have fitted will be used to simulate values for the entry and exit points.

4.5 Monthly Savings

The accumulated amount is computed by accumulating the monthly amount needed to be saved and multiplying it by the probability of survival, adjusted for the exit and the accumulated entry rates. The following graph shows the monthly savings against time.

Figure 4.22: Monthly Savings Scatterplot

As time increases, the monthly savings are somewhat the same except for a few outliers. From the data collected, the study finds that on average, individuals call for 7.6 minutes a day. Assuming that they are Safaricom Network users, their daily expenditure on calls is estimated at an average of KShs 30.55 (916.75 monthly).

4.6 Expenditure Patterns

In order to be able to maintain their level of commodity consumption, people who retire need to increase (or at least maintain) their previous level of income (after deducting housing costs). In fact, however, they decrease their income and expenditure. From the income and expenditure analysis, the fact that households do not in fact maintain their income and expenditure suggests that they should be dissatisfied with their financial situation and more likely to report financial hardship. However, financial satisfaction increases and experience of financial hardship decreases as the people adjust to their current standards of living. The older people might have less income volatility and hence fewer financial stress events, and yet still have a relatively low standard of living. The decrease in expenditure patterns after retirement is usually offset by a few factors such Furthermore, the expenditure patterns in (Bruce et al, 2012) showed that reduction in work-related costs, increase in health-service consumption, price reductions, home production and increase in leisure expenditure that are directly influenced by retirement where as some of the factors that are indirectly influenced by retirement include home and property insurance, repairs and maintenance such as motor and home-related, credits, corporate payments; most foods except eating out; non-alcoholic drinks; clothing except work-related cost (in the expenditure categories that are directly influenced by retirement); furniture, home appliances, cutlery; electronic equipment; books, gambling, sports equipment (other than golf); non-prescribed pain relievers, sunscreens; air fares, holiday vehicle hire, overseas holidays; pets; personal care (except hair care); personal expenditures; and gifts.

Chapter 5

Conclusions and Recommendations

5.1 Conclusion

The main intention of this research was to develop a model that recalculates the amount required to be saved daily considering the discounting assumptions and find out the spending patterns for those in retirement. This modeling task required:

(i) Mobile phone expenditure data which was collected using an app called Bandicoot;

(ii) Entry and exit processed which were derived from the Mbao Pension Scheme;

(iii) Probabilities of survival which were derived from Kenya's life table;

(iv) The pensions data was collected from Eagle Africa;

(v) Life expectancy;

(vi) Inflation and interest rates were extracted from Kenya's Central Bank database.

We first forecasted the inflation and interest rates since the data was inadequate. We did this using ARIMA-GARCH model since ARIMA(p,q) models the mean (returns) while the GARCH (m,s) model models the volatility in the returns. We then backward-simulated the entry and exit processes since the data was also insufficient and we needed to go back in time.We first forecasted the inflation and interest rates since the data was inadequate. We did this using ARIMA-GARCH model since ARIMA(p,q) models the mean (returns) while

the GARCH (m,s) model models the volatility in the returns. We then backward-simulated the entry and exit processes since the data was also insufficient and we needed to go back in time.

The probabilities of survival were computed by finding the average probability of the males and females. With these six parameters, we computed the monthly consumption and the monthly savings. The monthly consumption is a function of the inflated expenditure discounted and multiplied by the probability of survival.

The monthly savings is a function of the average amount that needs to be saved accumulated in the future then multiplied by the probabilities of survival, the exit and entry rates accumulated. The results show that individuals require to be charged at least KShs.3.41 above the current rate per call should they join the scheme in order to raise an amount sufficient to sustain their expenditure post retirement. Of importance is to note that this amount has been calculated from the amount charged on calls only. Should this cost be spread out among other expenditure such as mobile data charges, mobile money charges etc., then it could be less.

Using the model, we also considered the amounts needed using other telecommunication networks, such as Airtel and Telkom. Airtel's current rate per call is KShs.2 per minute across all networks. The additional cost as estimated by the model would bring this charge to KShs.5.65 per minute. Telkom charges KShs.2 for on-net calls and KShs.3 to other networks per minute which would mean that the total charge of the scheme would range between KShs.5.41 and KShs.6.41 per consecutively.

5.2 Limitations of the study

(i) The greatest limitation of the study is the unavailability of the data needed and therefore we had to forecast and simulate our own.

(ii) The other limitation is lack of works by scholars who have really implemented such a model in the sense that we had to do a lot of research in terms of a modified model which had ARIMA-GARCH fitted variables and backward-simulated variables.

(iii) Finally, not every student in the study was responsive because they had to download Bandicoot application. This was a challenge because it required data or Wi-Fi to install and some did not have enough memory space on their phones as well.

5.3 Recommendation

The following is the recommendation we would like to postulate as a result of this study:

(i) Costs should not just be deducted on calls but also on other aspects of the mobile phone such as text message charges, data charges, and mobile money transaction costs.

References

1. Akaike, H. (1974). A new look at the statistical model identification. In Selected Papers of Hirotugu Akaike (pp. 215-222). Springer, New York, NY

2. Ando, A., & Modigliani, F. (1963). The" life cycle" hypothesis of saving: Aggregate implications and tests.*The American economic review*, 53(1), 55-84.

3. Bengtson, V. L., & Oyama, P. S. (2007). Intergenerational solidarity: Strengthening economic and social ties.*New York: United Nations Headquarters.*

4. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity.Journal of econometrics, 31(3), 307-327.

5. Bruce B., & Mendolia, S. (2012).*Living standards after retirement: perceptions and expenditure patterns.* Social Policy Research Centre.Child and Youth Finance International Research Working Group. (2012). Children and youth as economic citizens: review of research on financial capability, financial inclusion, and financial education.*Amsterdam: Research Working Group Report, Child and Youth Financial International.*

6. Burnham, K. P., & Anderson, D. R. (1998). Model selection and inference: a practical information-theoretical approach.New-York: Springel-Verlag.

7. Caroline Van Dullemen (2007). Pensions in Developing Countries: A quest for modern format for intergenerational solidarity.
United Nationshttp://www.un.org/esa/socdev/unyin/documents/egm_unhq_oct07_vandullemen.pdf.

8. Davis, E. P. (2005). The role of pension funds as institutional investors in emerging markets.

9. Demirguc-Kunt, A., & Klapper, L. (2012). *Measuring financial inclusion: The global findex database.* The World Bank.

10. Deshpande, R. (2012). What do youth savers want? Results from market research in four countries. *Save the Children.*

11. Ferreira, F. H., & Walton, M. (Eds.). (2005). *World development report 2006: equity and development* (Vol. 28). World Bank Publications.

12. Friedman, M. (1957). The permanent income hypothesis. In *A theory of the consumption function* (pp. 20-37). Princeton University Press

13. Hu, Y. W., & Stewart, F. (2009). Pension Coverage and informal sector workers.

14. Kabare, K. (2018). The Mbao Pension Plan: Savings for the Informal-Sector.

15. Keizi, L. K. (2006). Barriers to pension scheme participation by workers in the informal economy.

16. Kenya, F. S. D. (2016). FinAccess Household Survey. *Nairobi: FSD Kenya.*

17. Keynes, J. M. (2018). *The general theory of employment, interest, and money.* Springer.

18. Kilara, T., & Latortue, A. (2012). Emerging perspectives on youth savings. *CGAP, Washington DC.*

19. Maloney, W. F., & Saavedra-Chanduvi, J. (2007). The informal sector: What is it, why do we care, and how do we measure it. *Informality: exit and exclusion*, 21-42.

20. Mbuthia, B. N. (2016). *A mobile solution for savings by the youth in Kenya* (Doctoral dissertation, Strathmore University).

21. Modigliani, F., & Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross-section data. *Franco Modigliani, 1*, 388-436.

22. Njuguna, A. G. (2012). Critical success factors for a micro-pension plan: An exploratory study. *International Journal of Financial Research, 3*(4), 82.

23. Olafsson, A., & Pagel, M. (2018). *The retirement-consumption puzzle: New evidence from personal finances* (No. w24405). National Bureau of Economic Research.

24. Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars.Journal of Statistical Software, 27(4), 1-32.

25. Perry, G. E., Arias, O., Fajnzylber, P., Maloney, W. F., Mason, A., & Saavedra-Chanduvi, J. (2007).*Informality: Exit and exclusion.* The World Bank.

26. Raichura, S. K. (2008). Analytical review of the pension system in Kenya.*Mombasa: OECD/IOPS.*

27. Sablik, T. (2016). Life Cycle Hypothesis.*Econ Focus*, (3Q/4Q), 8-8. (Duesenberry, 1949).

28. Schwarz, G. (1978). Estimating the dimension of a model.The annals of statistics,6(2), 461-464.

29. Shankar, S., & Asher, M. G. (2011). Micro-pensions in India: Issues and challenges. *International Social Security Review,64* (2), 1-21.

30. Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control.*Journal of political Economy,89* (2), 392-406.

31. Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving.*Journal of political Economy,112* (S1), S164-S187.

32. Umberson, D., & Karas Montez, J. (2010). Social relationships and health: A flashpoint for health policy.*Journal of health and social behavior,51* (1_suppl), S54-S66.

33. World data atlas. Url: https://knoema.com/atlas/Kenya/topics/Demographics/Population-forecast/Life-expectancyhttps://knoema.com/atlas/Kenya/topics/Demographics/Population-forecast/Life-expectancy retrieved on 31st March, 2019.

34. Zou, L., Tlapek, S. M., Njenga, G., Appiah, E., Opai-Tetteh, D., & Sherraden, M. S. (2015). Facilitators and obstacles in youth saving: Perspectives from Ghana and Kenya.*Global Social Welfare,2* (2), 65-74.

35. Keynes, John Maynard, 1883-1946. (1936). The general theory of employment, interest and money. London :Macmillan.

Appendix

Appendix A:Codes for inflation Rates Forecasting ARIMA(1,0 GARCH(0,1) Model

Load Libraries

```
library(fUnitRoots)
library(lmtest)
library(forecast)
library(FitAR)
library(rugarch)
library(quantmod)
library(lattice)
library(timeSeries)
library(rugarch)
library(tseries)
library(rmgarch)
library(fGarch)
library(xts)
library(zoo)
library(parallel)
```

Reading the data

```
inflation<-read.csv("C://Users//Asunpta.Lagat//Desktop//asumpta//inflation.csv",header=TRUE)
```

Decompose into time series components

```
timeseriescomponents <- decompose(inflation)
timeseriescomponents
plot(timeseriescomponents)
```

Store adj inflation rates in the original file

```
inflation.close<-inflation$rates
inflation.close
```

Plot original inflation rates

```
plot(inflation.close,type='l',main='Inflation Rates')
```

Differencing the original series

```
diff.inflation<-diff(inflation.close)
diff.inflation
```

Plot differences of original series

```
plot(diff.inflation,type='l',main='Difference Inflation Rates')
```

Take log of the original series and plot the log price

```
log.inflation<-log(inflation.close)
log.inflation
plot(log.inflation,type='l',main='Log Inflation Rates')
```

Differencing log inflation rates and plot

```
difflog.inflation<-diff(log.inflation)
difflog.inflation
plot(difflog.inflation,type='l',main='Difference Log Inflation Rates')
```

Obtaining the Autocorrelation Function (ACF) and the Partial Autocorrelation Function(PACF) of Log Inflation Rates

```
acf.loginflation<-acf(log.inflation,main='ACF Log Inflation Rates',lag.max=100,ylim=c(-0.5,1))
acf.loginflation
pacf.loginflation<-pacf(log.inflation,main='PACF Log Inflation Rates',lag.max=100,ylim=c(-
0.5,1))
pacf.loginflation
```

Obtaining the Autocorrelation Function (ACF) and the Partial Autocorrelation Function(PACF) of Differenced Log Inflation Rates

```
acf.diffloginflation<-acf(difflog.inflation,main='ACF Differenced Log Inflation
Rates',lag.max=100,ylim=c(-0.5,1))
acf.diffloginflation
pacf.diffloginflation<-pacf(difflog.inflation,main='PACF Differenced Log Inflation
Rates',lag.max=100,ylim=c(-0.5,1))
pacf.diffloginflation
ts.inflation<-difflog.inflation
ts.inflation
```

Detemine stationarity of data

```
ndiffs(ts.inflation)
tsstationary¡-diff(ts.inflation, differences=0)
tsstationary
library(ggplot2)
plot(tsstationary,type='l')
acf(ts.inflation,lag.max=34)
```

Remove seasonality

```
timeseriesseasonallyadjusted ¡- ts.inflation- timeseriescomponents$seasonal
par(mfrow=c(2,1))
acf(tsstationary, lag.max=34)
pacf(tsstationary, lag.max=34)
```

Fit the model

auto.arima(ts.inflation, trace=TRUE)

fitARIMA¡-arima(tsstationary, order=c(1,0,5),seasonal = list(order = c(0,1,0), period = 12),method="

fitARIMA

This gives us the ARIMA(1,0,5) as shown in below:

```
Now re-fitting the best model(s) without approximations...

 ARIMA(1,0,5) with zero mean      : -754.4

 Best model: ARIMA(1,0,5) with zero mean

Series: ts.inflation
ARIMA(1,0,5) with zero mean

Coefficients:
        ar1    ma1    ma2    ma3    ma4    ma5
      0.526  0.678  0.649  0.713  0.738  0.392
s.e.  0.115  0.125  0.122  0.098  0.094  0.149

sigma^2 estimated as 0.000528:  log likelihood=384.5
AIC=-755.1   AICc=-754.4   BIC=-733.4
>
```

Figure 5.1: Auto.arima

significance of coefficients

coeftest(fitARIMA)

par(mfrow=c(1,1))

acf(fitARIMA$residuals)

Forecast future values

par(mfrow=c(1,1))

f.arima¡-predict(fitARIMA,n.ahead = 660)

f.arima

futurVal <- forecast(fitARIMA,h=660, level=c(99.5))

futurVal

plot(futurVal)

Diagnostic Check using Ljung Box test

```
library(LSTS)
Box.Ljung.Test(tsstationary, lag = NULL, main = NULL)
ts.diag(tsstationary) Box.test(resid(fitARIMA),type="Ljung",lag=0,fitdf=1)
Box.test(resid(fitARIMA),type="Ljung",lag=20,fitdf=1)
Box.test(resid(fitARIMA),type="Ljung",lag=100,fitdf=1)
```

Squared Residuals

```
res.arima211<-fitARIMA$res
res.arima211
squared.res.arima211<-res.arima211^2
squared.res.arima211
par(mfcol=c(3,1))
plot(squared.res.arima211,main='Squared Residuals')
acf.squared211<-acf(squared.res.arima211,main='ACF Squared Residuals',lag.max=100,ylim=c(-0.5,1))
pacf.squared211<-pacf(squared.res.arima211,main='PACF Squared Residuals',lag.max=100,ylim=c(-0.5,1))
```

GARCH model garch01<-garch(res.arima211,order=c(0,1),trace=F)

```
garch01
loglik01<-logLik(garch01)
loglik01
summary(garch01)
```

Generate 1-step forecast, 100-step forecast, and plot of forecast

```
f.garch<-predict(garch01,n.ahead=660)
f.garch
```

Forecast ARIMA(1,0,5)-GARCH(0,1)

```
library(TTR)
```

```
library(AnalyzeTS)
library(TSA)
arima.garch<-forecastGARCH(fitARIMA,garch01,r=6,,trace=TRUE)
arima.garch
library(fGarch)
model = garchFit(formula = garch(1, 1), data = res.arima211, cond.dist = "norm", include.mean = TRUE)
fcst=predict(model,n.ahead=660)
mean.fcst=fcst$meanForecast
mean.fcst
```

Exporting the data to Excel

```
write.csv(fcst, "C://Users//Asumpta.Lagat//Desktop//asumpta//fcst.csv")
write.csv(f.arima, "C://Users//Asumpta.Lagat//Desktop//asumpta//farima.csv")
write.csv(f.garch, "C://Users//Asumpta.Lagat//Desktop//asumpta//fgarch.csv")
```

Appendix B:Codes for interest rates Forecasting ARIMA(1,0,5 GARCH(0,1) Model

Load Libraries

```
library(fUnitRoots)
library(lmtest)
library(forecast)
library(FitAR)
library(rugarch)
library(quantmod)
library(lattice)
library(timeSeries)
library(rugarch)
library(tseries)
library(rmgarch)
```

```
library(fGarch)
library(xts)
library(zoo)
library(parallel)

        Reading the data
interest¡-read.csv(”C://Users//Asumpta.Lagat//Desktop//asumpta//interest.csv”,header=TRUE)

        Decompose into time series components
timeseriescomponents <- decompose(interest)
timeseriescomponents
plot(timeseriescomponents)

    -   Store adj interest rates in the original file
interest.close<- interest$rates
interest.close

        Plot original interest rates
plot(interest.close,type='l',main='Interest Rates')

        Differencing the original series
diff.interest <-diff(interest.close)
diff.interest

        Plot differences of original series
plot(diff.interest,type='l',main='Difference Interest Rates')

        Take log of the original series and plot the log price
log.interest <-log(interest.close)
log.interest
```

plot(log interest,type='l',main='Log Interest Rates')

Differencing log inflation rates and plot

difflog.interest <-diff(log.interest)

difflog.interest

plot(difflog.interest,type='l',main='Difference Log Interest Rates')

Obtaining the Autocorrelation Function (ACF) and the Partial Autocorrelation Function(PACF) of Log Interest Rates

acf.loginterest <-acf(log.interest,main='ACF Log Interest Rates',lag.max=100,ylim=c(-0.5,1))

acf.loginterest

pacf.loginterest <-pacf(log.interest,main='PACF Log Interest Rates',lag.max=100,ylim=c(-0.5,1))

pacf.loginterest

Obtaining the Autocorrelation Function (ACF) and the Partial Autocorrelation Function(PACF) of Differenced Log Interest Rates

acf.diffloginterest <-acf(difflog.interest,main='ACF Differenced Log Interest Rates',lag.max=100,ylim=
0.5,1))

acf.diffloginterest

pacf.diffloginterest <-pacf(difflog.interest,main='PACF Differenced Log Interest Rates',lag.max=100,yl
0.5,1))

pacf.diffloginterest

ts.interest <-difflog.interest

ts.interest

Detemine stationarity of data

ndiffs(ts.interest)

tsstationary<-diff(ts.interest, differences=1)

tsstationary

library(ggplot2)

plot(tsstationary,type='l')

acf(ts.interest,lag.max=34)

Remove seasonality

timeseriesseasonallyadjusted ¡- ts.interest - timeseriescomponents$seasonal

par(mfrow=c(2,1))

acf(tsstationary, lag.max=34)

pacf(tsstationary, lag.max=34)

Fit the model

auto.arima(ts.interest, trace=TRUE)

fitARIMA<-arima(tsstationary, order=c(0,0,0),seasonal = list(order = c(0,1,0), period = 12),method=

fitARIMA

fitARIMA1<-arima(tsstationary, order=c(1,0,5),seasonal = list(order = c(0,1,0), period = 12),method=

fitARIMA1

```
> #fit the model
> auto.arima(ts.interest, trace=TRUE)

 Fitting models using approximations to speed things up...

 ARIMA(2,0,2) with non-zero mean : -592.5
 ARIMA(0,0,0) with non-zero mean : -594.9
 ARIMA(1,0,0) with non-zero mean : -591.5
 ARIMA(0,0,1) with non-zero mean : -592.4
 ARIMA(0,0,0) with zero mean     : -596.3
 ARIMA(1,0,1) with non-zero mean : -591.8

 Now re-fitting the best model(s) without approximations...

 ARIMA(0,0,0) with zero mean     : -596.3

 Best model: ARIMA(0,0,0) with zero mean

Series: ts.interest
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.00918:  log likelihood=299.2
AIC=-596.3   AICc=-596.3   BIC=-592.5
> |
```

Figure 5.2: Auto.arima()

Significance of coefficients

coeftest(fitARIMA)

coeftest(fitARIMA1)

par(mfrow=c(1,1))

```
acf(fitARIMA$residuals)
```

Forecast future values

```
par(mfrow=c(1,1))
f.arima<-predict(fitARIMA,n.ahead = 660)
f.arima
futurVal <- forecast(fitARIMA,h=660, level=c(99.5))
futurVal
futurVal1 <- forecast(fitARIMA1,h=660, level=c(99.5))
futurVal1
plot(futurVal)
```

Diagnostic Check using Ljung Box test

```
library(LSTS)
Box.Ljung.Test(tsstationary, lag = NULL, main = NULL)
ts.diag(tsstationary)
Box.test(resid(fitARIMA),type="Ljung",lag=0,fitdf=1)
Box.test(resid(fitARIMA),type="Ljung",lag=20,fitdf=1)
Box.test(resid(fitARIMA),type="Ljung",lag=100,fitdf=1)
```

Squared Residuals

```
res.arima211<-fitARIMA$res
res.arima211
squared.res.arima211<-res.arima211^2
squared.res.arima211
par(mfcol=c(3,1))
plot(squared.res.arima211,main='Squared Residuals')
acf.squared211<-acf(squared.res.arima211,main='ACF Squared Residuals',lag.max=100,ylim=c(-
0.5,1))
pacf.squared211<-pacf(squared.res.arima211,main='PACF Squared Residuals',lag.max=100,ylim=c(-
0.5,1))
```

GARCH model

```
garch01<-garch(res.arima211,order=c(0,1),trace=F)
garch01
loglik01<-logLik(garch01)
loglik01
summary(garch01)
```

Generate 1-step forecast, 100-step forecast, and plot of forecast

```
f.garch<-predict(garch01,n.ahead=660)
f.garch
```

Forecast ARIMA(1,0,5)-GARCH(0,1)

```
library(TTR)
library(AnalyzeTS)
library(TSA)
arima.garch<-forecastGARCH(fitARIMA,garch01,r=6,,trace=TRUE)
arima.garch
library(fGarch)
model = garchFit(formula = garch(1, 1), data = res.arima211, cond.dist = "norm", include.mean = TRUE)
fcst=predict(model,n.ahead=660)
mean.fcst=fcst$meanForecast
mean.fcst
```

Exporting the data to Excel

```
write.csv(fcst, "C://Users//Asumpta.Lagat//Desktop//asumpta//fcsti.csv") write.csv(f.arima,
"C://Users//Asumpta.Lagat//Desktop//asumpta//farimai.csv") write.csv(f.garch, "C://Users//Asun
```

Appendix C: Codes for Backward Simulation of Entry and Exit Rates using Backward Markov Process and Survival Simulation

```
install.packages("MonteCarlo")
install.packages("data.table")
install.packages("ggplot2")
install.packages("dplyr")
install.packages("foreign")
install.packages("tidyr")
install.packages("triangle")
install.packages("fitdistrplus")
install.packages("logspline")
install.packages("MASS")
install.packages("actuar")
install.packages("survsim")
install.packages("HMM")
install.packages("markovchain")
install.packages("TraMineR")
install.packages("RcppHMM")
library(RcppHMM)
library(fitdistrplus)
library(logspline)
library(MonteCarlo)
library(data.table)
library(ggplot2)
library(dplyr)
library(foreign)
library(tidyr)
library(triangle)
library(MASS)
library(actuar)
```

```
library(survsim)
library(HMM)
library(markovchain)
library(TraMineR)
rm(list=ls())
life<-read.csv("C://Users//Asumpta.Lagat//Desktop//Asumpta//life.csv",header=TRUE)
```

Let e be entrants

```
e<-life$entrants
```

Let ex be exits

```
ex<-life$exits
```

Finding and fitting the distribution

```
plotdist(e, histo = TRUE, demp = TRUE)
descdist(e, boot=1000)
fit.weibull <- fitdist(e, "weibull")
fit.norm <- fitdist(e, "norm")
fit.lnorm <- fitdist(e, "lnorm")
fit.gamma<- fitdist(e, "gamma", start = list(shape = 1, scale = 500))
fit.llogis <- fitdist(e, "llogis", start = list(shape = 1, scale = 500))
fit.pareto <- fitdist(e, "pareto", start = list(shape = 1, scale = 500))
fit.burr<- fitdist(e, "burr", start = list(shape1 = 0.3, shape2 = 1,rate = 1))
par(mfrow = c(2, 2))
plot.legend <- c("weibull", "normal", "lognormal")
plot.legend1 <- c("gamma","llogis","pareto","burr")
denscomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext
= plot.legend)
denscomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext
= plot.legend1)
qqcomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext =
plot.legend)
```

```
qqcomp(list(fit.gamma,fit.llogis,fit.pareto.fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext
= plot.legend1)
cdfcomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext
= plot.legend)
cdfcomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext
= plot.legend1)
ppcomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext =
plot.legend)
ppcomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext
= plot.legend)
summary(fit.norm)
summary(fit.lnorm)
summary(fit.gamma)
summary(fit.llogis)
summary(fit.pareto)
summary(fit.burr)
```

Quantile

```
quantile(fit.llogis, probs = 0.05)
```

Goodness of fit

```
gofstat(list(fit.weibull,fit.norm,fit.lnorm,fit.gamma,fit.llogis,fit.pareto,fit.burr),fitnames = c("weibull","
"gamma","llogis", "pareto", "burr"))
```

fitting the distribution

```
fit_glm <- glm(entrants ., data = life)
```

fit_glm

```
res <- as.data.frame(t(fit_glm$coefficients))
res
summary(fit_glm)
```

Get the backward probabilities for a Markov Process

```
trans_prob<-c(0.04,0.05,0.04,0.04,0.03,0.05,0.03, 0.03,0.03,0.02,0.03,0.01,0.03,0.04,0.03,0.04,0.08,0.04,(
0.05,0.05,0.04,0.06,0.05)
e.seq <- seqdef(life[,1:2], e)
tr <- seqtrate(e.seq)
m.e<-mean(e)
sd.e<-sd(e)
states<-c("entrants","exits")
residues<-paste(1:24)
transition<-trans_prob
dimnames(transition)<-list(from=states,to=states)
emission<-matrix(c(rep(1/24,24),rep(1/10,24),1/2),nrow=2,byrow=TRUE)
x<-structure(list(transition,emission),class="HMM")
backward(x,life)
hmm = initHMM(c("entrants","exits"), c("e","ex"), transProbs=trans_prob),emissionProbs=emission
print(hmm)
```

Sequence of observations

```
observations = c("e","e","ex","ex")
```

Calculate backward probablities

```
logBackwardProbabilities = backward(hmm,observations) logBackwardProbabilities
backwardProb¡-(exp(logBackwardProbabilities))
```

Simulating data

```
fit.unif <- fitdist(e, "unif")
fit.unif
dist.ev <- "llogistic"
anc.ev <- 1
beta0.ev <- 5.268
dist.cens <- "llogistic"
anc.cens <- 1
```

```
beta0.cens <- 5.368
z¡-list(c("unif",100,691))
prob¡-backwardProb
m.prob¡-mean(prob)
m.prob
x <- list(c("e",m.prob))
beta <- list(-1.0, -4.63)
simulated¡-simple.surv.sim(84, 365, dist.ev, anc.ev, beta0.ev,dist.cens, anc.cens, beta0.cens, z ,
beta , x )
```

Exporting the simulated entrance processes

```
write.csv(simulated, "C://Users//Asumpta.Lagat//Desktop//Asumpta//simulated_entrance.csv")
install.packages("MonteCarlo")
install.packages("data.table")
install.packages("ggplot2")
install.packages("dplyr")
install.packages("foreign")
install.packages("tidyr")
install.packages("triangle")
install.packages("fitdistrplus")
install.packages("logspline")
install.packages("MASS")
install.packages("actuar")
install.packages("survsim")
install.packages("HMM")
install.packages("markovchain")
install.packages("TraMineR")
install.packages("RcppHMM")
library(RcppHMM)
library(fitdistrplus)
library(logspline)
library(MonteCarlo)
library(data.table)
```

```
library(ggplot2)
library(dplyr)
library(foreign)
library(tidyr)
library(triangle)
library(MASS)
library(actuar)
library(survsim)
library(HMM)
library(markovchain)
library(TraMineR)
rm(list=ls())
life<-read.csv("C://Users//Asumpta.Lagat//Desktop//Asumpta//life.csv",header=TRUE)
```

Let e be entrants.

```
e<-life$entrants
```

Let ex be exits

```
ex<-life$exits
```

Finding and fitting the distribution

```
plotdist(ex, histo = TRUE, demp = TRUE)
descdist(ex, boot=1000)
fit.weibull <- fitdist(ex, "weibull")
fit.norm <- fitdist(ex, "norm")
fit.lnorm <- fitdist(ex, "lnorm")
fit.gamma<- fitdist(ex, "gamma", start = list(shape = 1, scale = 500))
fit.llogis <- fitdist(ex, "llogis", start = list(shape = 1, scale = 500))
fit.pareto <- fitdist(ex, "pareto", start = list(shape = 1, scale = 500))
fit.burr<- fitdist(ex, "burr", start = list(shape1 = 0.3, shape2 = 1,rate = 1))
par(mfrow = c(2, 2))
plot.legend <- c("weibull", "normal", "lognormal")
```

plot.legend1 <- c("gamma","llogis","pareto","burr")

denscomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend)

denscomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend1)

qqcomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend)

qqcomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend1)

cdfcomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend)

cdfcomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend1)

ppcomp(list(fit.weibull,fit.norm,fit.lnorm), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend)

ppcomp(list(fit.gamma,fit.llogis,fit.pareto,fit.burr), xlogscale = TRUE, ylogscale = TRUE,legendtext = plot.legend)

summary(fit.norm)

summary(fit.lnorm)

summary(fit.gamma)

summary(fit.llogis)

summary(fit.pareto)

summary(fit.burr)

Quantile

quantile(fit.burr, probs = 0.05)

Goodness of fit

gofstat(list(fit.weibull,fit.norm,fit.lnorm,fit.gamma,fit.llogis,fit.pareto,fit.burr),fitnames = c("weibull"," "gamma","llogis", "pareto", "burr"))

Fitting the distribution

```
fitglm <- glm(exits ., data = life)
fitglm
res <- as.data.frame(t(fitglm$coefficients))
res
summary(fitglm)
```

Get the backward probabilities for a Markov Process

```
trans_prob<-matrix(c(0.06,0.13,0.21,0.03,0.02,0.02,0.03,0.01,0.03,0.05,0.05,0.02,0.04,0.03,0.02,0.02,0.0
0.02,0.02,0.03,0.03,0.02,0.03,0.03),nrow=2,byrow=TRUE)
trans_prob
ex.seq <- seqdef(life,2:2,labels=c("entrants"))
tr <- seqtrate(ex.seq)
m.ex<-mean(ex)
sd.ex<-sd(ex)
states<-c("entrants","exits")
residues<-paste(1:24)
transition<-trans_prob
transition
dimnames(transition)<-list(from=states,to=states)
emission<-matrix(c(rep(1/24,24),rep(1/10,-5),1/2),nrow=2,byrow=TRUE)
emission
x<-structure(list(transition,emission),class="HMM")
x
backward(x,life)
initHMM(ex,3,method="llogistic")
hmm = initHMM(c("entrants","exits"), c("e","ex"),transProbs=trans_prob,emissionProbs=emission)
hmm
print(hmm)
Sequence of observations
observations = c("e","e","ex","ex")
Calculate backward probablities
logBackwardProbabilities = backward(hmm,observations)
logBackwardProbabilities
```

```
backwardProb<-(exp(logBackwardProbabilities))
```

Simulating data

```
fit.unif <- fitdist(ex, "unif")
fit.unif
dist.ev <- "llogistic"
anc.ev <- 1
beta0.ev <- 5.268
dist.cens <- "llogistic"
anc.cens <- 1
beta0.cens <- 5.368
z<-list(c("unif",30,444))
prob<-backwardProb
m.prob<-mean(prob)
m.prob
x <- list(c("ex",m.prob))
beta <- list(-1.0, -4.63)
simulated <- simple.surv.sim(84, 365, dist.ev, anc.ev, beta0.ev,dist.cens, anc.cens, beta0.cens,
z , beta , x )
simulated
```

exporting the simulated entrance processes

```
write.csv(simulated, "C://Users//Asumpta.Lagat//Desktop//Asumpta//simulated_exit.csv")
```

Appendix D: Probability of Survival

Prob of dying	Prob of survival			
	male	female		
<1 year	0.051235	0.043353	0.047294	0.952706
1-4 years	0.025941	0.025118	0.025529	0.974471
5-9 years	0.012941	0.009059	0.011	0.989
10-14 years	0.008412	0.006765	0.007588	0.992412
15-19 years	0.015059	0.01	0.012529	0.987471
20-24 years	0.020941	0.015471	0.018206	0.981794
25-29 years	0.029294	0.033353	0.031324	0.968676
30-34 years	0.040353	0.045294	0.042824	0.957176
35-39 years	0.056294	0.058294	0.057294	0.942706
40-44 years	0.061882	0.057235	0.059559	0.940441
45-49 years	0.065235	0.054765	0.06	0.94
50-54 years	0.075588	0.060706	0.068147	0.931853
55-59 years	0.085059	0.067294	0.076176	0.923824
60-64 years	0.114176	0.089235	0.101706	0.898294
65-69 years	0.161529	0.130059	0.145794	0.854206

Table 5.1: Probabilities of survival