APPLICATION OF LÉVY PROCESSES TO UNITISED WITH-PROFITS POLICIES

By

Chenming Bao

Submitted for the Degree of

Doctor of Philosophy

at Heriot-Watt University

on Completion of Research in the

School of Mathematical and Computer Sciences

May 2009.

I hereby declare that the work presented in this thesis was carried out by myself at Heriot-Watt University, Edinburgh, except where due acknowledgement is made, and has not been submitted for any other degree.

Chenming Bao (Candidate)

Prof. Andrew Cairns (Supervisor)

Dr. Terence Chan (Supervisor)

Dr. Tom Fischer (Supervisor)

Dr. Mark Willder (Supervisor)

Date

Contents

List of Tables

List of Figures

Abstract

The objective of this thesis is to develop more realistic long term asset models based on Lévy processes and discuss their applications to risk management of unitised with-profits policies.

We investigate the behaviour of long-term returns of the UK total share return index by testing the common statistical properties for financial data, so-called "stylised facts". We show that for the monthly U.K. share total return indices, the Gaussian return hypothesis is rejected in series of tests. The local distribution of the returns has higher kurtosis and heavier tails than the Gaussian. In addition, the returns series show significant nonlinear autocorrelation, extreme returns appear in clusters.

The first long term asset model purposed in this thesis is the exponential Lévy model with non-Gaussian increment. We describe the Generalised Hyperbolic distributions with their subclasses. They are considered as candidate distributions for the increments of the driving Lévy processes.

We estimate model parameters to the UK share gross total return index using two approaches, maximum likelihood (MLE) and Markov Chain Monte Carlo (MCMC) algorithm. Statistical and graphical goodness-of-fit tests demonstrate that these Lévy driven models give more accurate fits to the historical equity index returns data.

For the liability model we consider long term participating life insurance products specifically unitised with-profits contracts. The payouts of unitised with-profits policies are simulated under a variety of asset models driven by Lévy processes. At first a basic model policy is considered with limited insurer operations and no risk controls. We look into various risk measures of the maturity loss for the insurer

and compare the statistical properties for different non-Gaussian increment Lévy models. It is found that the classical Gaussian model substantially underestimates the risk measures in unitised with-profits policies. The Lévy driven models which have semi-heavy tailed increments are aggregate to normal distributions in the long run.

Then we consider different retrospective bonus mechanisms by varying the participating rate and the smoothing period. As a comparison we use a bonus earning power method with deterministic projected maturity asset share and 25 percent terminal bonus cushion. We study the joint distributions of the maturity asset shares and guarantees under these two bonus mechanism. With similar risk measures, there are larger expected maturity guarantees under bonus earning power method than retrospective bonus. Declaring bonuses on a more frequent basis is then tested, which has the desired effect of reducing the risk measures when declaring monthly bonuses using bonus earning power mechanism.

We make observations on two different investment strategies, a diversified investment strategy and a hedging based investment strategy. The former method tries to reduce the variance of the invest return distribution while the hedging investment strategy, on the other hand, narrows the left tail of the maturity loss distribution by paying an extra amount of expenses.

Finally, the Lévy models are extended by using GARCH(1,1)-m type volatility. Both maximum likelihood estimators and Bayesian estimators using Markov Chain Monte Carlo are presented. The statistical tests on the devolatilised data show that the GARCH model reduces the non-linear autocorrelation in the conditional return processes and furthermore improve the fitting of the asset models. Also, multi-variable models are considered. Stochastic bridges driven by Lévy processes are constructed while the yearly returns follow the Wilkie model.

Acknowledgements

I would like to thank my supervisors: Prof Andrew Cains, Dr Terence Chan, Dr Tom Fischer and Dr Mark Willder, for their invaluable advice and guidance throughout the years. This thesis uses data supplied by Prof David Wilkie. I would like to thank Prof Wilkie for sharing his thorough understanding and experience of financial modelling.

I would like to thank all the staffs and PhD students of the Department of Actuarial Mathematics and Statistics for their support and friendship. It was a great pleasure to work with my colleagues: Jiangchun Bii, Baopeng Lu, Jozef Kollar, Dylan Bowden, Sule Sahine, Sing-Yee Ling and Xingran Fu. Thanks to Dr Ken-Tak Siu, for the discussion on research and application of mathematics, the topics cover much more than only technical questions of mathematics, from which I have learnt much.

I dedicate this thesis to my family. This work could not be possible without the support from my parents. It is my indebtedness to them for their understanding, love and confidence in me. Needless to say, I would like to thank Shuo, it was her love that helped me in the difficult times.

This research has been supported and funded by Heriot-Watt University through the MACS scholarship, the financial and technical support is gratefully acknowledged.

Introduction

The classical Black-Scholes framework in Black and Scholes (1973) and Merton (1973) uses a Brownian motion with Gaussian distributed increments to model the risky asset returns. It has been considered to be a poor fit to the real world data. The empirical distribution of such returns is more peaked and has fatter tails than the Gaussian distribution, which implies that large changes in returns occur with a higher frequency than under normality. In addition it is often skewed with a heavier left tail. The volatility appears time-varying and clustered, returns are serially uncorrelated, but squared returns are serially correlated.

During the last ten years, there has been a growing interest in the use of Lévy processes to model financial market behaviours. However, most of the asset models driven by Lévy processes in the literature are mainly considering short term periods from intra-day to weeks. For actuarial using purposes, we aim to build long term asset models on a monthly basis. They should be able to capture the key statistical features of the marginal distributions of the asset returns. Furthermore, the asset models should adequately explain the long term dependence structure in financial time series, both autocorrelation in asset returns and cross-corrrelations between different classes of assets.

Actuaries are required to calculate the market consistent value of the with-profits liability on the realistic balance sheet. The liability of the guarantee embedded in a with-profits policy can be valued and hedged as a path-dependent financial option. From this point of view, a continuous long term investment model may provide both a good fit to empirical data and mathematical tractability. On the other hand, for risk management, the insurers can adjust their investment and bonus strategies taking into consideration the possibility of future jumps in asset returns.

We construct the driving Lévy processes by describing the law of the increments as family of generalised hyperbolic distributions (GH). The GH distributions were first introduced in Barndorff-Nielsen (1978), they provide empirical good fit to share returns. The long-term time series properties of the investment returns are modeled by using stochastic volatility models and bridge technique based on the Wilkie model. We provide reliable and relatively fast algorithms for calibrating Lévy driven asset models.

The Unitised With-Profits (UWP) policies as long term investment and saving products make up a significant part of the UK's life insurance business. It allows the policyholders to share the benefits of the life office fund while a guaranteed minimum return is provided by the insurer to protect the investors (policyholders) against the financial investment risk. The management of investment guarantees risk embedded in unitised with-profits and other participating life insurance products is now a challenge for both, actuarial profession and researchers.

In order to understand the potential risks of the unitised with-profits liability, we simulate the payouts of unitised with-profits policies under asset models driven by Lévy processes. Briefly, insurers have two strategies to control the risks of UWP policies. They can either control the guarantee liability by adjusting bonus mechanism or manage the asset of the UWP fund by improving the investment strategies.

The rest of the thesis is organized as follows.

In Chapter 1 we investigate the behaviour of long-term returns of the UK share return index by studying some of its statistical properties. We focus on the monthly total share return rate which is essential for the with-profits investment performance.

Chapter 2 provides an extensive review of the general structure, path properties and the decompositions of Lévy processes and the applications in financial modelling. We introduce the driving processes of our model which have generalised hyperbolic-distributed unit increments. Specifically, the exponential Lévy models are proposed as a long term continuous time series model for monthly UK Share gross total return index.

In Chapter 3 we discuss the topic of parameters estimation for the Lévy driven asset models. A maximum likelihood estimation algorithm is constructed for the

generalized hyperbolic distributions. An alternative approach based on the Bayesian estimation called Markov Chain Monte Carlo (MCMC), specifically, the Metropolis-Hesting Algorithm (MHA) is introduced to expose the Bayesian estimators with their joint distributions. We examine the goodness of fit and study the statistical properties of fitted models by simulation.

Chapter 4 considers the key factors in modelling, reserving and valuation of unitised with-profits policies. Special attention will be paid to measuring the financial risk in the UWP policies. UWP policies are modelled by providing the operation rules which include asset allocation, charging, reserving and bonus declaration. The financial risks will be measured by random variables based on the loss distribution of the insurer's maturity payouts. A ten year single premium UWP policy is studied by simulation experiments. We focus on a comparison of the risk measures under Lévy driven models as more realistic models and the Gaussian driven models. The effect of the estimation risk is discussed by using "hyper-models" with parameters estimated by MCMC.

In Chapter 5 several dynamic bonus mechanisms are introduced as tools to control the maturity guarantee liabilities of the UWP policies. We first investigate the risk measures under the retrospective bonus mechanisms by varying the participating rate and the smoothing period. Then a bonus earning power method with deterministic projected maturity asset share and 25% terminal bonus cushion is used. The joint distributions of the maturity asset shares and guarantees are simulated using Lévy driven asset models. Finally we consider declaring bonuses on a more frequent basis.

In Chapter 6 we consider the insurer's investment strategies for UWP policies. First, there are two investment strategies based on "diversification" of the with-profits fund, an invest-and-forget strategy and a re-balanced equity backing ratio strategy are tested. Then some issues of dynamic Delta hedging of the contingent claim will be discussed. Finally, we construct a long term with-profits fund investment strategy based on the internal dynamic hedging.

In Chapter 7 we propose more realistic asset models based on Lévy processes. We extend the exp-Lévy models by replacing the stochastic time by a GARCH(1,1)-m

volatility process to capture the volatility clustering properties in the SGTRI data. Also, the multi-variable models are considered, we demonstrate continuous time models using stochastic bridges driven by Lévy processes in the Wilkie model. Both maximum likelihood estimators and Bayesian estimators using Markov Chain Monte Carlo are presented.

Chapter 1

Empirical properties of asset returns

1.1 Introduction and objectives

In this chapter we investigate the behaviour of long-term UK total share return index by studying some of its statistical properties.

Our data set are based on the same one used in Wilkie (1995). The sample observations of UK financial data includes series of U.K. retail price indices, wage indices, U.K. share price and yield indices, long-term and short-term interest rates, property indices and exchange rates.

Although there exists a considerable volume of literature related to the empirical properties of asset prices and market indices under so-called "fine scales" (Δt less than one week), for example see Cont and Tankov (2004) and Pagan (1996), less attention has been paid to the studies of the monthly and yearly data.

We discuss the key statistical features for the share total return rates. The approach to do this is by testing some of the most common statistical properties that have been proved in a lot of articles published in last ten years and applied on a wide range of financial data so-called "stylised facts". We focus on the monthly return rate series which is essential for measuring the with-profits investment performance.

The chapter is organised as follows. In next section we introduce the data set called SGTRI. In section 1.3, we look at the distributional properties of the monthly

data set, specifically, the classical Gaussian return hypothesis is tested. The sample linear autocorrelation properties is studied in 1.4. And the nonlinear autocorrelation properties of the sample series will be discussed in Chapter 7.

1.2 The Data

We are able to access the same data set used to fit the Wilkie model (Wilkie (1986) and Wilkie (1995)) which has been carefully weighted to be consistent over different time intervals.

The main data set which we consider in this thesis contains a monthly total return (rolled-up) index on U.K. shares calculated gross of tax and assumes the re-investment of dividends. The monthly share price and yield indices from 1929 to 1962 are based on the Actuaries Investment Indices, published by the Institute of Actuaries and the Faculty of Actuaries and privately circulated. After 1962, the indices are based on the FTA All-Share Index. In recent years this has been the Total Return Index for the All-Share Index in the FTSE-Actuaries U.K. Share Indices. See details of the data sources in Wilkie (1995) Appendix F.

The full range of data runs from December 1923 to April 2005, giving 977 end-of-month observations. The starting year of the sample observation is chosen same as in Wilkie (1995). We will also observe the the model by looking at subset of this sample spanning from January 1950 to April 2005.

Generally, given a time scale Δt, the log return of the asset or indices S_t at time scale Δt is defined as:

$$Y(t, \Delta t) = \log S_{t+\Delta t} - \log S_t, \tag{1.1}$$

where Δt can be any time from seconds to decades. In this thesis, time scale Δt is one month. We refer to these share log-returns as the SGTRI (share gross total return index) series and the compounded log return of the monthly share gross total return indices S_t is given by:

$$Y_i = \log S_{i+1} - \log S_i, \quad i = 1, ..., 976.$$

The long term monthly returns have not been analysed often because lack of

consistent data availability. Our contribution starts from next section by observing the key features from the statistical evidence of the UK monthly SGTRI data.

1.3 The distribution of SGTRI

1.3.1 Unconditional distribution & non-Gaussian

Over the short term the unconditional distribution of financial market data such as share return rates, indices, exchange rates and interest rates, is characterised by some common properties called stylised empirical facts. We refer to Cont and Tankov (2004) as a comprehensive review of stylised facts. Here one says "short term" means the time lags of the return processes are less than one week, some authors also call this "fine scales".

Under fine scales, non-Gaussian characterisation has been observed for various financial data, for example see Eberlein and Keller (1995). It is true even for the conditional distribution, such as distribution of returns which is devolatilised (see Chapter 7), has been found to be non-Gaussian, see for example Cont (2001). This rules out the models whose short term marginal distributions are normal (Gaussian) such as Geometric Brownian motion (GBM), although as a special case of a Lévy process, Brownian motion is still widely used as a driving process because of its attractive analytical features.

On the other hand, for the yearly returns, unconditional normality hypothesis can not be rejected, see Cont and Tankov (2004). The Wilkie model (Wilkie (1986) and Wilkie (1995)) gives the annual share price as ratio of the dividends D_t and share yield Y_t where both of their log rate are modelled by time series driven by Brownian motion. Hence the logarithm of the share total return rate is

$$R(t, \Delta t) = \log \left[\frac{P_{t+\Delta t} + \sum_{s \in (t, t+\Delta t]} D(s)}{P_t} \right],$$

where $P_t = D_t / Y_t$, is normally distributed.

We know the shape of distributions for return processes may vary with time scales. To this end, we test the Gaussian distribution hypothesis on the monthly share return process.

A graphical way to show goodness of fit are quantile-quantile (QQ) plots. We show the QQ plot in Figure 1.1, where each cross represents the percentage quantile of the empirical process against the standard normal quantile.

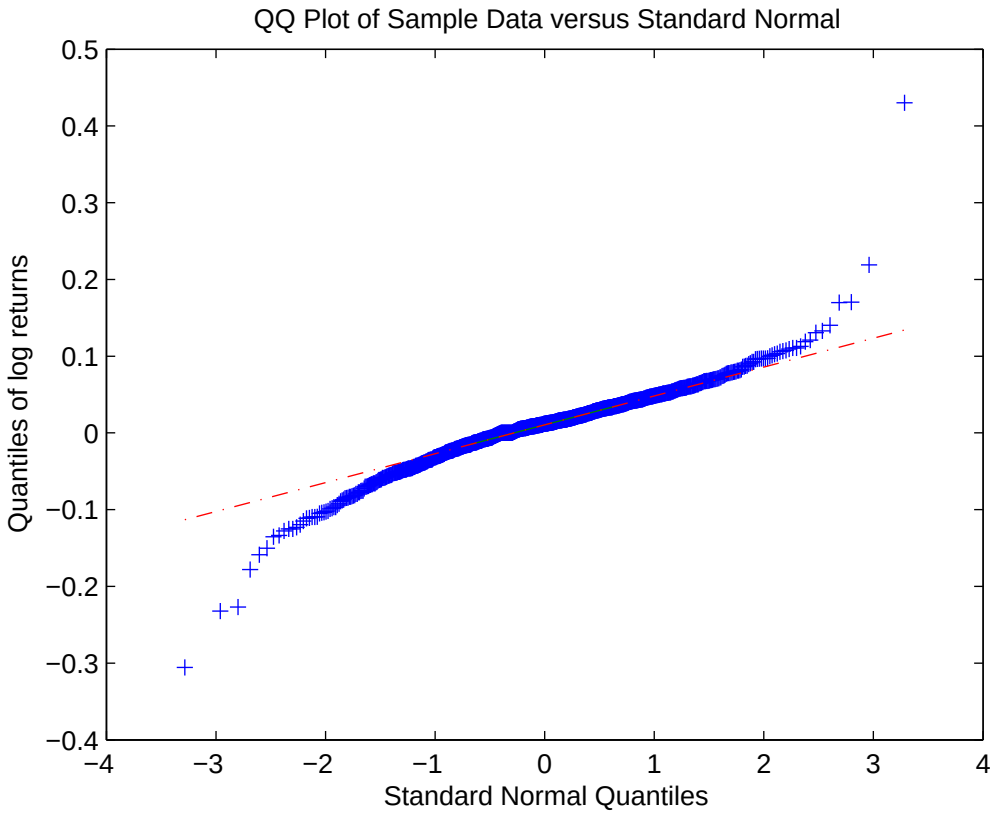

Figure 1.1: QQ plot (Normal)

If the sample observations were unconditional normally distributed, the sample quantile points would be reasonably linear. Compare to the straight line (broken line with dots) in Figure 1.1 which is the theoretical normal-normal quantiles curve, the sample quantile points display a shape that have lower left tail quantiles and higher right quantiles than the "normal-normal" line. It shows the fatter tails on both sides of the unconditional distribution of the sample observations. This fact suggests that the non-Gaussian characterisation such as fatter tail property is still significant for the monthly UK SGTRI log return distribution. To support this and to see other statistical features of the data, some tests are presented below.

Table 1.1 shows the statistics of the monthly returns on the U.K. SGTRI. The full data set from 1923 to 2005 is called "SGTRI23" while the sub-sample from 1950

to 2005 is called "SGTRI50" in the table and the latter test results. The skewness for SGTRI23 is -0.0821 and kurtosis is 11.6782. For a Gaussian distribution, these two statistics should be zero and three respectively. A random variable is said to be "leptokurtic" if its kurtosis is greater than 3. In this test, "leptokurtic" has been found in both sample sets. For both sample sets skewness is close to 0 and the full data set SGTRI23 shows the return process is negative-skewed.

Table 1.1: Statistics for log-return of the SGRTI

Statistics	Mathematical Form	SGTRI 12/23-04/05	SGTRI 01/50-04/05
Mean	$\mathbb{E}[Y]$	0.0086	0.0103
Variance	m_2	0.0024	0.0028
Skewness	$m_3/m_2^{3/2}$	-0.0821	0.0678
Kurtosis	m_4/m_2^2	11.6782	11.7106
Maximum	Y_{max}	0.4300	0.4300
Minimum	Y_{min}	-0.3056	-0.3056

We can use the χ^2 test to check whether the data is normally distributed. Let χ_1^2 denote the test statistic computed with cells of equal probability $1/k$, while χ_2^2 is used for cells of equal width. $\chi^2_{k-1;0.99}$ denotes the 0.99-quantile of the χ^2-distribution with $k-1$ degrees of freedom.

Table 1.2: Statistics for χ^2 tests, k=20

Statistics	$\chi^2_{k-3;0.99}$	χ_1^2	χ_2^2
SGTRI 50	33.41	61.80	69.65
SGTRI 23	33.41	94.33	112.40

Table 1.2 shows that both SGTRI50 and SGTRI23 have higher χ^2 statistics than the theoretical percentage points under the same degrees of freedom. So we can safely reject the null hypothesis that the sample observations are from normally distributed data at significance level 99%.

Eberlein and Keller (1995) studied the daily return process of BASF and Deutsche Bank using χ^2-tests and rejected the Gaussian assumption. In this section we have

the same conclusion by testing the monthly data for UK share returns. However the non-Gaussian characterisation is not so clear for the yearly returns. We test the yearly end-of-June returns whose χ_1^2 statistic with $k = 10$ is 11.5 and p-value 0.8818, therefore cannot find strong evidence to reject the normal distribution hypothesis.

Next we consider the Jarque-Bera test, the statistic with n sample observations is given by

$$JB_n = \frac{n}{6}\left(s^2 + \frac{(k-3)^2}{4}\right) \tag{1.2}$$

where s is the sample skewness, k is the sample kurtosis and n is the sample size. For large n, the statistic JB_n has a χ^2 distribution with two degrees of freedom.

For the UK SGTRI23, the Jarque-Bera statistic is equal to 3063.8 and for the UK SGTRI50 the Jarque-Bera statistic is equal to 2096.6. For both of the sample sets, the p-values are close to zero, hence it rejects the normality hypothesis.

The Lilliefors test is a 2-sided goodness-of-fit test when a fully-specified null distribution is unknown and its parameters must be estimated. The default null hypothesis is that the sample of SGTRI comes from a distribution in the normal family, against the alternative that it does not come from a normal distribution. The Lilliefors test statistic is given by:

$$LS = \max_{x\in\mathbb{R}} |F_{emp}(x) - F_{est}(x)|, \tag{1.3}$$

where $F_{emp}(x)$ is the empirical cdf estimated from the sample and $F_{est}(x)$ is the normal cdf with mean and standard deviation equal to the mean and standard deviation of the sample. For the UK SGTRI23, the Lilliefors statistic is equal to 0.0806 and for the UK SGTRI50 the Lilliefors statistic is equal to 0.0742. Again for both SGTRI23 and SGTRI50, the p-values are close to zero, hence we reject null hypothesis that the data comes from normal distribution family for both sub-samples.

1.3.2 Tails and extreme value

Given the sample observations of a series of returns $x_i,\ i = 1, 2, ..., n$, the minimal and maximal returns are defined as:

$$m_n = \min\{x_i,\ i = 1, 2, ..., n\},$$
$$M_n = \max\{x_i,\ i = 1, 2, ..., n\}.$$

The Fisher-Tippett theorem of extreme value for i.i.d. sequence (cf. McNeil *et al.* (2005)) says if there exist normalizing constants (λ_n, σ_n) and a non-degenerate limit distribution H for the normalised maximum return:

$$\mathbb{P}\left(\frac{M_n - \lambda_n}{\sigma_n} \leq x\right) \overset{n\to\infty}{\longrightarrow} H(x) \tag{1.4}$$

then the limit distribution H has the form:

$$H_a(x) = \begin{cases} \exp\{-(1+ax)^{-1/a}\}, & a \neq 0, \\ \exp\{-e^{-x}\}, & a = 0, \end{cases} \tag{1.5}$$

where the shape parameter a determines three distributions: $a = 0$ for Gumbel, $a < 0$ for Weibull and $a > 0$ for Frechet distributions.

For guarantee liability of the with-profits contracts, either an extreme large negative return or a large positive return increases the embedded risk of the policy. The former scenario leads to a lower asset share at the maturity, makes the with-profits policies be "in-the-money". The later scenario increases the so-called policyholder's reasonable expectations (PRE) , force the insurer increases the guarantee by declaring higher reversionary bonuses. Thus the loss distribution of the with-profits policies is caused by two-sided investment returns rather than only poor investment returns.

We say a distribution or its density function $f(x)$ has a semi-heavy tail if the density function behaves as

$$f(x) \sim \begin{cases} C_-|x|^{\rho_-}\exp\{-a|x|\}, & \text{as } x \to \pm -\infty, \\ C_+|x|^{\rho_+}\exp\{-b|x|\}, & \text{as } x \to \pm +\infty, \end{cases} \tag{1.6}$$

for some $\rho_-,\ \rho_+ \in \mathbb{R}$ and $C_-,\ C_+,\ a,\ b \geq 0$. And we say $f(x)$ has a heavy tail if

$$C|x|^{-\rho}, \text{as } x \to \pm\infty, \tag{1.7}$$

for some $C > 0$ and $0 < \rho < 2$.

The figure 1.2 shows the large movements of the return process over four thresholds. One can see that large movements of the SGTRI data are clustered and for

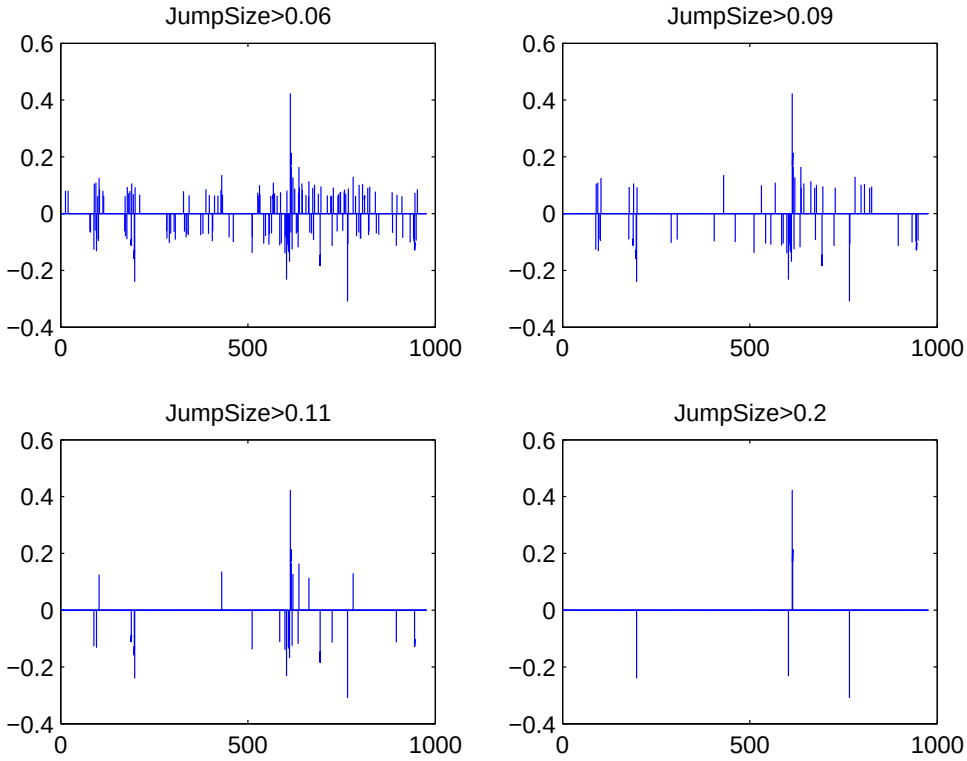

Figure 1.2: Jumps in log-return process of SGTRI23-05

each threshold, there are more negative large movements than positive, e.g., 16 negative returns and 7 positive returns larger than 0.11 while there are 3 negative returns and 1 positive returns larger than 0.20.

1.4 Autocorrelation

Cont and Tankov (2004) point out that for the daily and intra-day return process, except for very small time scales ($\simeq$ 20 minutes), autocorrelations are often insignificant. Here we look at monthly returns, Figure 1.3 shows the linear autocorrelation function (ACF) of the monthly return process with the 95% confidence bounds that the autocorrelation function is zero. Base on the investigation, there is no significant

evidence for functional shape of the SGTRI autocorrelation function. It shows that the SGTRI sample have small positive correlation for lag of 1 which exceed the 95% confidence bound. However, the sample ACF for lag of 2 is negative which should be positive if the autocorrelation for lag of 1 were positive. Cont (2001) points out that

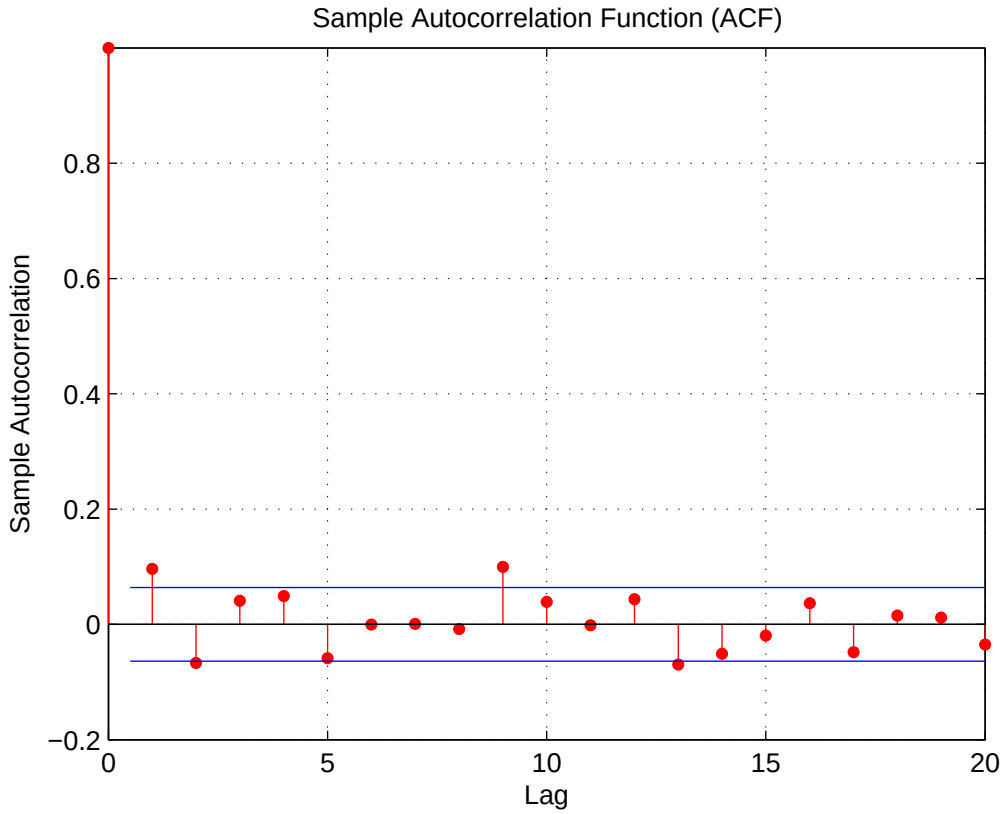

Figure 1.3: Autocorrelation of SGTRI23-05

the autocorrelation of the return processes decay rapidly to zero when time scales increase from 0 to 20 minutes. For time scales $\Delta t \geq 15$ minutes it can be safely assumed to be zero. When time scale Δt of return process is increased, there is linear autocorrelation for weekly and monthly returns. One of the reasons for this is for short term, there are arbitragers who looks for any statistical autocorrelation of the returns, then their investment strategies push the market to be absent of autocorrelation, except for the first 15 minutes due to the "reaction time for new information". We suggest that when the time scales is increased, the economic cycle, government policies and long term investors such as insurance companies' activities dominate the market equilibrium prices. Our statistical conclusion is only based on tests of

monthly UK SGTRI, the long term statistical features may be very dependent on the different markets.

1.5 Concluding Remarks

To sum up, we have studied U.K. share total return indices in this Chapter. The unconditional distribution of the monthly UK SGTRI display non-Gaussian with positive excess kurtosis. The statistical tests and the graphs show the empirical distribution belongs to distribution classes with heavy or semi-heavy tails. The Gaussian marginal distributions are not able to model the extreme value in the SGTRI return series. The linear autocorrelation of the return processes sample is insignificant.

Chapter 2

Lévy Processes and Asset Models

2.1 Introduction and Objectives

In this chapter, we discuss some topics of modelling the financial time series data by using stochastic processes. Special attention will be paid on the models driven by Lévy processes with non-Gaussian distributed increments.

First, we provide an extensive review of general structure, path properties and the decompositions of Lévy processes.

We introduce the driving processes of our model which have generalised hyperbolic-distributed unit increments. As the increment law for monthly total returns, the generalised hyperbolic (GH) distributions were first introduced in Barndorff-Nielsen (1978). These distributions belong to the infinitely divisible distribution family that makes them natural candidates to be used to construct Lévy processes. The subclasses and limiting cases of the generalised hyperbolic distribution have been studied both empirically and theoretically, for instance, normal inverse Gaussian (Barndorff-Nielsen (1998)), hyperbolic (Eberlein and Keller (1995)), variance gamma (Madan and Seneta (1990)) and student-t (Platen and Rendek (2007)). We will see that using these distributions can significantly improve the fit of equity returns in the parameter estimating and goodness of fit tests in Chapter 3.

Then asset models analogous to geometric Brownian motion but driven by Lévy processes are built of which can capture the large sudden movements in the invest-

ment returns. Specifically, the exponential Lévy models are proposed as a long term continuous time series model for monthly UK Share gross total return index. The advantage of this model is its flexible statistical features such as heavier tail, high kurtosis than Gaussian, distributions of the increments can be skewed, while its simple mathematical form offers good analytical tractability. The new Lévy models are continuous on the real line and the return process is locally non-Gaussian distributed. Actuaries might be interested in it since, on the one hand, the international accounting standards have been moving to a fair value approach. Actuaries are now required to calculate the market consistent value of the with-profits liability on the realistic balance sheet. The liability of the guarantee embedded in a with-profits policy can be valued and hedged as a path-dependent option. From this point of view, a continuous long term investment model may provide both a good fit to empirical data and mathematical tractability. On the other hand, for risk management, the insurers can adjust their investment and bonus strategies taking into consideration of the possibility of future "shocks" in share values. The potential extension of the exponential Lévy model will be discussed in Chapter 7, includes the forms of multi-variable and GARCH volatility. The chapter is organised as follows.

In next section, we review the basic definition of the Lévy processes with the pivotal result called Lévy-Khintchine formula. For the observant reader who is interested in Lévy processes, references which cover the detail concepts of Lévy processes are given.

In section 2.3, Lévy processes are constructed by two methods. The first one is to define a infinitely divisible distribution as the increment law. Either by its distribution function or by the Lévy triple in its characteristic function [1]. The second method we use is by applying the stochastic time changing technique, based on the so-called Brownian subordination. These methods are consistent and moreover the time changing skill is the essential of some more complex extension models driven by Lévy processes such as the stochastic volatility (SV).

The GH law with its subclasses and limiting cases are introduced in section 2.4. In section 2.5, exponential Lévy models are constructed. We end this chapter by

[1]See section 2.2 for the definition of Lévy triple.

briefly demonstrating some common ways of extend the independent increment Lévy models in order to capture the volatility clustering in the data.

2.2 Lévy Processes

The theoretical facts we present in this section are heavily rely on Protter (2004) and Applebaum (2004), other monographs on Lévy processes include Bertoin (1996) and Sato (1999). Schoutens (2003) and Cont and Tankov (2004) contain overview of some aspects of research result for applying the Lévy processes in mathematical finance, both for risk management and pricing purposes. The way we show these results is in a "review style" instead of introduction of the Lévy processes. For readers who is not familiar with theory of Lévy processes, we suggest Applebaum (2004) for an introduction with mathematical rigour and Cont and Tankov (2004) for a simplified but more intuitive version of introduction.

We start with the definition of infinite divisibility of the random variable. Let X be a real valued random variable with probability law μ_X. We say that X is infinitely divisible if, for all $n \in \mathbb{N}$, there exist i.i.d. random variables $Y_1^{(n)}, ..., Y_n^{(n)}$ such that

$$X \stackrel{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}, \tag{2.1}$$

where $\stackrel{d}{=}$ denotes equality in distribution. In other word, for each $n \in \mathbb{N}$, the law μ_X has a convolution n-th root that is itself the law of a random variable, i.e.

$$\mu_X = \mu_{Y_1}^{*n}. \tag{2.2}$$

Let $\phi_X(u) = \mathbb{E}(e^{iuX})$ denote the characteristic function of X, where $u \in \mathbb{R}$. From the definition, if X is infinitely divisible, for each $n \in \mathbb{N}$, ϕ_X has an nth root that is itself the characteristic function of a random variable, i.e.

$$\phi_X(u) = [\phi_{Y_1}(u)]^n. \tag{2.3}$$

The general characterisation of infinitely divisible random variables is given by the **Lévy-Khintchine formula**. Denote the characteristic exponent $\Psi(u) = -\log \mathbb{E}(e^{iuX})$.

Follow the Lévy-Khintchine formula, a probability law μ of a real valued random variable is infinitely divisible if and only if there exists a triple (γ, σ, ν), where $\gamma \in \mathbb{R}, \sigma \geq 0$ and ν is a σ-finite measure supported on $\mathbb{R}\backslash\{0\}$ satisfying $\int_{\mathbb{R}} \min(1, x^2)\nu(dx) < \infty$, such that

$$\Psi(u) = iau + \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (1 - e^{-iux} + iux\mathbf{1}_{(|x|\leq 1)})\nu(dx) \tag{2.4}$$

for every $u \in \mathbb{R}$.

The measure ν is called the Lévy measure and the triple (γ, σ, ν) is called the **Lévy triple**.

Next we define the $\mathbb{R}$ -valued Lévy process with *càdlàg* paths. [2]

Let $X = (X(t), t \geq 0)$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that X is a *Lévy process* if:

L1 $X(0) = 0$ *a.s.*;

L2 X has *independent increments*: for every increasing sequence of times $t_0 ... t_n$, the random variables $(X(t_{j+1}) - X(t_j), 1 \leq j \leq n)$ are independent;

L3 X has *stationary increments*: each $X(t_{j+1}) - X(t_j) \stackrel{d}{=} X(t_{j+1} - t_j) - X(0)$

L4 X is *continuous in probability*, i.e. for all $a > 0$ and for all $s \geq 0$

$$\lim_{t \to s} \mathbb{P}(|X(t) - X(s)| > a) = 0$$

From the Lévy-Khintchine formula we can see that for any $t > 0$, X_t is a Lévy process whose law μ belongs to the class of infinitely divisible distributions. Inversely it can be shown that for each infinitely divisible distribution μ there exists a Lévy process X_t such that μ is the distribution of X_t. Thus, for every $t > 0$, we have the characteristic exponent $\Psi_t(u)$ following the Lévy-Khintchine formula. Hence for any $t > 0$,

$$\Psi_t(u) = t\Psi_1(u).$$

Let $Q(dt, dx)$ be a Poisson measure on $(0, \infty) \times \mathbb{R}\backslash\{0\}$ with intensity measure $dt \times \nu(dx)$, we have the so-called **Lévy-Itô decomposition**:

$$X_t = at + B_t + J_t^b + \lim_{\varepsilon \downarrow 0} J_t^{s,\varepsilon} \tag{2.5}$$

[2] *càdlàg* paths denotes the paths that are continuous on the right and always have limits on the left.

where J_t^b and $J_t^{s,\varepsilon}$ represent the big jumps and small jumps respectively:

$$J_t^b = \int_{|x|\geq 1, s\in[0,t]} xQ(dt \times dx) \text{ and} \tag{2.6}$$

$$J_t^{s,\varepsilon} = \int_{\varepsilon\leq|x|<1, s\in[0,t]} x\{Q(dt \times dx) - \nu(dx)dt\}. \tag{2.7}$$

This result argues that a Lévy process is no more than a linear combination of a Brownian motion with drift and a possibly infinite sum of independent compound Poisson processes.

2.3 Constructing the background driving Lévy process

The essentials of our models are the background driving Lévy processes (BDLP) . There are two methods used in this thesis to construct them. They are equivalent and all Lévy processes can be expressed in these two forms respectively.

In the previous section we show the relationship between the class of infinitely divisible laws and the stationary independent increments processes. Hence a Lévy process can be uniquely defined by demonstrating the probability law of its infinitely divisible distributed increments or, equivalently, by giving the characteristic exponent function $\Psi_x(u)$ via the Lévy triple.

From the Lévy-Itô decomposition it is clear that a Lévy process can be constructed by defining respectively the Brownian motion part and the pure-jump process which is independent to the Brownian motion,

$$X_t = -\gamma t + \sigma B_t + J_t,\ t \geq 0$$

where B_t is a standard Brownian motion and J_t is a pure jump process. Absence of J_t, the X_t will be a continuous process which is a Brownian motion with drift $-\gamma$ and standard deviation σ. In this case the Lévy triple is $(\gamma, \sigma, 0)$. On the other hand, the triple $(\gamma, 0, \nu)$ leads to a pure jump process.

One of the other commonly used approaches to build Lévy processes or more general semi-martingales, see Protter (2004), is by the stochastic time changing

technique called Brownian subordination. Monroe (1978) points out that any semi-martingale L_t can be realised as a time changed Brownian motion as

$$L_t = \mu + \beta T_t + W(T_t),\ \mu, \beta \in \mathbb{R}\ and\ t \geq 0, \tag{2.8}$$

where $W(\cdot)$ is a Brownian motion and T_t is a non-negative increasing process of stopping times. In this case, the law of any increment $X_t - X_s,\ t > s \geq 0$ can be represents as a normal mean-variance mixture:

$$f_{X_{\Delta t}} = \int_0^\infty N(x; \mu + \beta\omega, \omega) g_{\Delta t}(\omega) d\omega, \tag{2.9}$$

where N is the normal density function with respect to mean and variance. $g_{\Delta t}$ denotes the density function of the increment law of the subordinator T_t.

We will discuss the detail in next section by construct the processes of which the unit increment follow the generalised hyperbolic law.

2.4 The Increment law

The generalised hyperbolic (GH) distribution was first introduced by Barndorff-Nielsen (1977) and have become popular in financial modelling in the last ten years.

An advantage of the GH distribution is that it embraces many special subclasses and limiting distributions, including hyperbolic (HYP), normal inverse Gaussian (NIG), (skewed) Student-t, variance gamma (VG) and Gaussian where some of them are well studied and have been widely used in modelling financial data. The GH law can captures some of the key statistical features of the financial data, for instance, it can be asymmetric, and the tails can be heavier than those of the normal distribution. Some of these features are inherited by its subclasses and limiting distributions. It belongs to the class of infinitely divisible distributions which can be used as the unit increment law of Lévy processes in financial models.

Further more, the GH distribution is a normal mean-variance mixture to the generalised inverse Gaussian distribution (GIG), see Barndorff-Nielsen (1977). Hence the Lévy process X_t which has GH-distributed unit increment is a time-changing Brownian motion $\mu t + \beta T_t + W(T_t)$ with stochastic time T_t (subordinator) where T_1

follow the generalied inverse Gaussian (GIG) law. Hence the marginal distribution of the process X_t is uniquely decided by it's stochastic time T_t.

The univariate GIG distributions can be parametrised in several ways. We follow Prause (1999) and Barndorff-Nielsen and Shephard (2001), define the GIG law, written $GIG(\lambda, \delta, \gamma)$, by the density

$$\text{gig}(x; \lambda, \delta, \gamma) = \frac{(\gamma/\delta)^\lambda}{2\mathbb{K}_\lambda(\delta\gamma)} x^{\lambda-1} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\}, \tag{2.10}$$

where $\mathbb{K}_\lambda$ denotes a modified Bessel function of the third kind,

$$\mathbb{K}_\lambda(z) = \frac{1}{2}\int_0^\infty y^{\lambda-1} \exp\left(-\frac{1}{2}z(y + y^{-1})\right) dy, \tag{2.11}$$

with the domain of variation of the parameters:

$$\begin{aligned} &\delta \geq 0, \quad \gamma > 0\ , \qquad \text{if } \lambda > 0\ ; \\ &\delta > 0, \quad \gamma > 0\ , \qquad \text{if } \lambda = 0\ ; \\ &\delta > 0, \quad \gamma \geq 0\ , \qquad \text{if } \lambda < 0\ . \end{aligned}$$

Denotes $\alpha = \sqrt{(\beta^2 + \gamma^2)}$, $T_1 \sim \text{GIG}(\lambda, \delta, \gamma)$ and is independent of $\epsilon \sim N(0, 1)$. The random variable $X_1 = \mu + \beta T_1 + T_1^{1/2}\epsilon$ has a generalised hyperbolic distribution, written $X_1 \sim \text{GH}(\lambda, \alpha, \beta, \delta, \mu)$. The Lebesgue density is given by:

$$\begin{aligned} \text{gh}(x; \lambda, \alpha, \beta, \delta, \mu) &= \phi(\lambda, \alpha, \beta, \delta)\left(\delta^2 + (x-\mu)^2\right)^{(\lambda-1/2)/2} \\ &\quad \mathbb{K}_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)\exp(\beta(x-\mu)), \qquad (2.12) \\ \phi(\lambda, \alpha, \beta, \delta) &= \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda-1/2}\delta^\lambda \mathbb{K}_\lambda(\delta\sqrt{\alpha^2 - \beta^2})},\ x \in \mathbb{R} \end{aligned}$$

where $\mathbb{K}_\lambda(z)$ is the modified Bessel function of the third kind.

For convenience, in this thesis, the Lévy process will be denoted by the name of its unit increment law, for instance a "GH process" means a Lévy process whose unit increment is GH distributed.

The tails of the GH distribution behaves as

$$\text{gh}(x; \lambda, \alpha, \beta, \delta) \sim \text{const}|x|^{\lambda-1}\exp(-\alpha|x| + \beta x),\ \text{as } x \to \pm\infty. \tag{2.13}$$

The tails of GH and GH subclasses decline as a hyperbolic function as $x \to \pm\infty$. When β is not zero, the two tails of GH law decay at different rate and the distribution is skewed.

In very few cases, the GH process is directly applied in financial modelling. One of the problems is the lack of analytically tractability of the GH process. Note that the GH law is not closed under convolution, i.e, the sum of two GH random variables is no longer GH-distributed. This is a restriction for flexibility of choosing the time scale used in the GH-processes driving asset models. In this case, the return processes are GH distributed only under defined unit time scales. As an example, we make the assumption that the monthly log-return rate is a random variable which has GH distribution. Because of the non-closing under convolution, neither the yearly nor the weekly returns are GH-distributed since. We are not able to approach for instance, the daily return in this particular model and simulation is only available for integer month lags.

Moreover, it is of great difficulty to estimate the parameters of GH process because of the flatness of the GH likelihood function, see for instance Prause (1999).

Some of the subclasses and limiting distributions of the GH law possess same attractive statistical properties such as heavy or semi-heavy tails and asymmetric which GH has, for example HYP, NIG, VG and skew student t distributions. And they have comparatively simpler mathematical form than the GH law. It leads to captivating theoretical properties and analytical tractabilities, for instance, the VG and NIG law are closed under convolution.

In chapter 3 we will estimate the parameters for both the GH law and some of its subclasses and limiting distributions to the UK SGTRI. The likelihood and goodness of fit tests will show that some of the subclasses and limiting distributions performs equally well as the GH distribution.

Some of the limiting distributions and subclasses of GH law are introduced below.

2.4.1 Normal inverse Gaussian

The NIG distribution was introduced by Barndorff-Nielsen (1977). It is able to model asymmetric distributions with possibly semi-heavy tails in both sides.

The density of normal inverse Gaussian law (NIG) is given by

$$\mathrm{nig}(x;\alpha,\beta,\delta_t,\mu_t) = \frac{\alpha\delta}{\pi}\exp\left(\delta(\alpha^2-\beta^2)^{\frac{1}{2}}+\beta(x-\mu)\right)\frac{\mathbb{K}_1(\alpha(\delta^2+(x-\mu)^2)^{\frac{1}{2}})}{(\delta^2+(x-\mu)^2)^{\frac{1}{2}})}, \tag{2.14}$$

where $\mathbb{K}_1(z)$ is the modified Bessel function of the third kind, $x, \mu \in \mathbb{R}$, $0 \leq \delta$ and $0 \leq |\beta| \leq \alpha$.

It is a normal mean-variance mixture with the parameter $\lambda = -1/2$ for the GIG distribution in equation 2.10. Note that for the modified Bessel function of the third kind

$$\mathbb{K}_1(z) = \mathbb{K}_{-1}(z).$$

It is a special case for the GIG distribution which called inverse Gaussian (IG). As the name may suggest, it is the distribution of the first passage time of a Brownian motion start from 0 with drift γ into a finite boundary δ, see Barndorff-Nielsen (1998). The density $\text{ig}(x; \gamma, \delta)$ is

$$\text{ig}(x; \gamma, \delta) = \frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\left(-\frac{1}{2}\frac{\gamma^2}{x}\left(x - \frac{\delta}{\gamma}\right)^2\right). \tag{2.15}$$

The class of inverse Gamma distributions is closed under convolution with same shape parameter γ and δ, see Webber and Ribeiro (2003). Thus a IG process at any time t is distributed according to the IG law. More specifically, if the T_1 has density $\text{ig}(x; \gamma, \delta)$, the inverse Gamma process T_t has the density

$$\text{ig}(x; \gamma, \delta t). \tag{2.16}$$

Consequently, a normal inverse Gaussian process X_t is a time-changed Brownian motion with inverse Gaussian stochastic time $T_t \sim \text{ig}(x; \gamma, \delta t)$. With same shape parameter α and β, the normal inverse Gaussian law is closed under convolution. From 2.16 the normal inverse Gaussian process $X_t \sim \text{nig}(x; \alpha_t, \beta_t, \delta_t, \mu_t)$ with

$$\alpha_t = \alpha; \beta_t = \beta; \delta_t = \delta t; \mu_t = \mu t. \tag{2.17}$$

Table 2.1 lists the mean, variance, skewness and kurtosis of normal inverse Gaussian distributions.

The tails of the NIG distribution behaves as

$$\text{nig}(x; \alpha, \beta, \delta) \sim \text{const}|x|^{-3/2} \exp(-\alpha|x| + \beta x), \text{ as } x \to \pm\infty. \tag{2.18}$$

The tails on both sides are semi-heavy and the NIG law is skewed.

Table 2.1: Distribution moments of the $NIG(\alpha, \beta, \delta, \mu)$.

Statistics	Mathematical Form	$NIG(\alpha, \beta, \delta, \mu)$
Mean	$\mathbb{E}[Y]$	$\mu + \delta\beta(\alpha^2 - \beta^2)^{-\frac{1}{2}}$
Variance	m_2	$\delta\alpha^2(\alpha^2 - \beta^2)^{-\frac{3}{2}}$
Skewness	$m_3/m_2^{3/2}$	$3\beta/\alpha\sqrt{\delta(\alpha^2 - \beta^2)^{\frac{1}{2}}}$
Kurtosis	m_4/m_2^2	$3(\alpha^2 + 4\beta^2)/\delta\alpha^2(\alpha^2 - \beta^2)^{\frac{1}{2}}$

2.4.2 Hyperbolic distribution

For $\lambda = 1$, we have hyperbolic distributions (HYP), whose density is

$$\frac{(\alpha^2 - \beta^2)^{1/2}}{2\delta\alpha\mathbb{K}_1(\delta(\alpha^2 - \beta^2)^{1/2})} \exp\left(-\alpha(\delta^2 + (x - \mu)^2)^{1/2} + \beta(x - \mu)\right), \tag{2.19}$$

where $x, \mu \in \mathbb{R}$, $0 \leq \delta$ and $|\beta| < \alpha$.

The HYP distribution has been used to model the stock returns in Eberlein and Keller (1995). Their log-density is a hyperbola function which offers the semi-heavy tails on both sides. The HYP process is a time-changed Brownian motion with so-called positive hyperbolic (Barndorff-Nielsen (1998)) stochastic time.

The tails of the HYP distribution behaves as

$$\text{hyp}(x; \alpha, \beta, \delta) \sim \text{const} \exp(-\alpha|x| + \beta x), \text{ as } x \to \pm\infty. \tag{2.20}$$

The two tails are both semi-heavy and the HYP law is skewed.

2.4.3 Student-t distribution

Student $t(\nu, \mu)$ is GH limiting distribution with density:

$$f_x(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \tag{2.21}$$

where $\lambda = -\nu/2, \alpha = \beta = 0, \delta = \sqrt{\nu}$ and $\Gamma(x)$ is gamma function.

We include the Student-t here because it is a typical example of so-called α-stable processes (see Mandelbrot (1977)). The distribution with $\alpha \sim (0, 2)$ has infinite second order moments. The empirical study of most stocks and exchange rates (see Cont (2001) or Carr *et al.* (2002)) shows the distribution will tend to be normal

when the time lag increase. However the convolution of the α-stable processes stay in the same class of distribution with same α. The student-t offer heavier tail than the other GH subclasses introduced in this section since the tails decline as a power function, i.e.,

$$f_x(x) \sim const|x|^{-\nu-1}, \quad as\ x \to \infty, \tag{2.22}$$

The two tails are both heavy and the student-t distribution is symmetric.

2.4.4 Variance gamma distribution

The variance gamma law (VG) was introduced by Madan and Seneta (1990) as a model for stock returns.

The variance gamma distribution has three parameters, σ, ν, and θ. The Lebesgue density $vg(x;\sigma,\nu,\theta)$ is given by

$$vg(x;\sigma,\nu,\theta) = \frac{2\exp(\frac{\theta x}{\sigma^2})}{\sigma\sqrt{2\pi}\nu^{\frac{1}{\nu}}\Gamma(\frac{1}{\nu})}\left(\frac{x^2}{\frac{2\sigma^2}{\nu}+\theta^2}\right)^{\frac{1}{2\nu}-\frac{1}{4}} \mathbb{K}_{\frac{1}{\nu}-\frac{1}{2}}\left(\frac{|x|\sqrt{2\sigma^2/\nu+\theta^2}}{\sigma^2}\right) \tag{2.23}$$

where $\mathbb{K}_\nu(z)$ is the modified Bessel function of the third kind (Equation 2.11). The characteristic function of the VG(σ,ν,θ) law is given by

$$\phi_{VG}(\mu;\sigma,\nu,\theta) = (1 - i\mu\theta\nu + \sigma^2\nu\mu^2/2)^{-1/\nu}. \tag{2.24}$$

Variance gamma is a limiting case of generalised hyperbolic distribution with parameters:

$$\lambda = \sigma^2/\nu,\ \alpha = \sqrt{(2/\nu)+(\theta^2/\sigma^4)},\ \beta = \theta/\sigma^2,\ \delta = 0 \text{ and } \mu = 0.$$

Table 2.2 lists the mean, variance, skewness and kurtosis of variance gamma distributions.

A variance gamma process VG_t is pure jump process with characteristic function

$$\phi_{VG_t}(\mu;\sigma,\nu,\theta) = \phi_{VG}(\mu;\sigma,\nu,\theta)^t = (1 - i\mu\theta\nu + \sigma^2\nu\mu^2/2)^{-t/\nu}. \tag{2.25}$$

Let $\sigma_1 = \sqrt{t}\sigma, \nu_1 = \nu/t$ and $\theta_1 = t\theta$, from 2.25 we have

$$\phi_{VG_t}(\mu;\sigma_1,\nu_1,\theta_1) = (1 - i\mu\theta_1\nu_1 + \sigma_1^2\nu_1\mu^2/2)^{-1/\nu_1}. \tag{2.26}$$

Table 2.2: Distribution moments of the $VG(\sigma, \nu, \theta)$.

Statistics	Mathematical Form	$VG(\sigma, \nu, \theta)$
Mean	$\mathbb{E}[Y]$	θ
Variance	m_2	$\sigma^2 + \nu\theta^2$
Skewness	$m_3/m_2^{3/2}$	$\theta\nu(3\sigma^2 + 2\nu\theta^2)/(\sigma^2 + \nu\theta^2)^{3/2}$
Kurtosis	m_4/m_2^2	$3(1 + 2\nu + \nu\sigma^4(\sigma^2 + \nu\theta^2)^{-2}$

Then the VG_t has variance gamma distribution with parameters $\sigma_1, \nu_1, \theta_1$. The VG process is closed under convolution.

One of the distinct features of variance gamma process from other GH subclasses and limiting cases is called finite variation (FV), see Carr *et al.* (2002). According to FV property, VG process can be decomposed to two Gamma process with opposite directions,

$$\mathrm{VG}(t) = G_u(t) + G_d(t) \tag{2.27}$$

where $G_u(t)$ denotes a positive Gamma process and $G_d(t)$ is a negative Gamma process which is independent to $G_u(t)$.

VG_t can be also represented as a time-changed Brownian motion, $W(T_t)$, the stochastic time T_t is a Gamma process $T_t \sim \Gamma(\frac{t}{\nu}) \sim \nu\Gamma(\frac{t}{\nu})$. The density of T_t is given by

$$f_\Gamma(x) = \frac{x^{\frac{t}{\nu}-1}\exp\left(-\frac{x}{\nu}\right)}{\nu^{\frac{t}{\nu}}\Gamma(\frac{t}{\nu}}. \tag{2.28}$$

The tails of the variance gamma distribution behaves as

$$\mathrm{vg}(x;\sigma,\nu,\theta) \sim \mathrm{const}|x|^{\sigma^2/\nu-1}\exp(-\sqrt{(2/\nu)+(\theta^2/\sigma^4)}|x| + \theta/\sigma^2 x), \text{ as } x \to \pm\infty. \tag{2.29}$$

The tails in both sides are semi-heavy and the variance gamma law is skewed.

2.5 Exponential Lévy Models

As asset models analogous to the geometric Brownian motion, we consider here two models called geometric Lévy and exponential Lévy (exp-Lévy), both driven by non-Gaussian increment Lévy processes.

Let $(\mathbb{L})_{t\geq 0}$ be a Lévy process with Lévy triplet (c^2, ν, γ). The geometric Lévy process is given by the stochastic differential equation:

$$dS_t = \mu_t S_{t-} dt + \sigma_t S_{t-} d\mathbb{L}_t$$

where $\mathbb{L}_t$ is a one dimensional Lévy process. This equation has explicit solution given by the Doleáns-Dade exponential:

$$\begin{aligned} S_t &= S_0 \exp\left\{ \int_0^{t\wedge T} \sigma_s d\mathbb{L}_s + \int_0^{t\wedge T} \left(b_s - \frac{c^2\sigma_s^2}{2} \right) ds \right\} \\ &\quad \times \prod_{0<s\leq t\wedge T} (1 + \sigma_s \Delta\mathbb{L}_s) \exp(-\sigma_s \Delta\mathbb{L}_s) \end{aligned} \tag{2.30}$$

where $\Delta\mathbb{L}_t = \mathbb{L}_t - \mathbb{L}_{t-}$ is the jump size at time t, T is the stopping time given by $T = \inf\{t > 0 | \sigma_t \Delta\mathbb{L}_t < -1\}$. This condition ensures that the asset value S_t is always a positive value and it will stop when the default happens, i.e. $S_t < 0$.

In practice one can only have the asset value data in discrete time but not the information of the whole continuous process sample path. The problem arising here is how can we define the $\Delta\mathbb{L}_t$ in time interval (t_i, t_{i+1}) while we only know the value of $\mathbb{L}_{t_i}$ and $\mathbb{L}_{t_{i+1}}$. In this case we can define a semimartingale X_t:

$$X_t = \int_0^{t\wedge T} \sigma_s d\mathbb{L}_s + \sum_{0<s\leq t\wedge T} \{\log(1 + \sigma_s \Delta\mathbb{L}_s) - \sigma_s \Delta\mathbb{L}_s\}.$$

We make the assumption that σ_t is constant, and assume that S_t is always positive. This is true for the market indices as in practice limited liability excludes negative share values. Mathematically it is equivalent to letting the Lévy measure $\nu((-\infty, -1/\sigma_t]) = 0$. Under these assumptions, X_t is a Lévy process and we have the so called exponential Lévy model:

$$S_t = S_0 \exp\{\sigma X_t + \mu_t\}. \tag{2.31}$$

Compare to the geometric Lévy in Equation 2.30, the exponential Lévy in Equation 2.31 is easier to descretise and calibrate. Among the many choices of BDLP we propose the GH process specifically whose distribution of unit increments belongs to the subclasses and limiting cases of GH law.

The GH process is a pure jump process since the lack of the diffusion part in decomposition, hence the exp-GH models has discrete sample path. Geman *et al.*

(1998) argues that the daily share return should be modelled by a pure jump process by introduce the so-called business time scale in the Brownian motion. The business time scale embedded is itself time-changed Brownian motion where a Gamma stochastic time leads to VG process is used in Geman *et al.* (1998).

The exp-GH models have discontinuous sample paths, where the models driven by Brownian motions such as GBM, RSLN (Hardy (2003)) and GARCH-GBM have continuous sample paths. In fact, all these models can be seen as time-changing Brownian motions as in Equation 2.8. The BDLP, which have discontinuous path stochastic time T_t, leads to jump in calender time. The stochastic times for Brownian motion driven models for example, RSLN and GARCH-GBM model, are continuous, these models have continuous sample path.

We will discuss this in Chapter 7 when we try to extend the model by using more complex time series structures. The non-Gaussian increment Lévy processes offer nice statistical features for the local marginal distributions of returns.

The empirical study of the SGTRI data in Chapter 1 shows the ACF of the monthly log-return process has an irregular shape around value of zero. However it is not the case for the volatility of the return process. The square return rate of SGTRI from 1923 to 2005 are positive correlated. It suggests there exist some periods of time in which the fluctuation of investment returns are more active than the others. The large movement (upward or downward) are more likely to be followed by another large movement when during the "quiet" time period, returns are stable and change relatively slow. This gives the evidence of the volatility clustering property for the return processes which can not be captured by the independent increments type Lévy processes. That is, if we consider monthly returns as i.i.d. sample observations, the GH law will give the better fitting performance than the Gaussian law. But if the time effect is taken into account, the sample set is thus a single observation involve a sequence of data.

2.6 Concluding Remarks

In this chapter, we review the relationship between the infinitely divisibility and the Lévy process. Brief outline of the theoretical facts are given and, specifically, the Lévy-Khintchine formula shows the general form of the characteristic exponent of the infinitely divisible laws and the decomposition of the Lévy processes. For the comprehensive introduction includes all the proofs of these theorems we refer to Applebaum (2004).

Lévy processes are constructed either by defining an infinite divisible distribution as its increment law or by the stochastic time changing technique.

The distributions belongs to GH class offer features include leptokurtic, semi-heavy or heavy tails for financial data modelling. The general GH case is not closed under convolution. Thus the asset models driven by GH processes are not analytically tractable under changing of time scales. Normally, explicit form of the density functions only exist at unit time for GH processes. In this regard, we introduced two of the GH subclasses, VG and NIG which are closed under convolution.

Finally, we make some comments on our approach of long-term stochastic investment modelling. For building stochastic investment models, we tried to find a balance between being accurately fitted to the empirical data and being mathematically tractable. Also the choice of asset models is based on the possible practical purposes. The application of the models in this thesis is for long-term with-profits modelling, involving valuation, reserving, investment and risk management. Hence we argue that monthly time lag is an appropriate time scale for most of the with-profits applications.

Given a proper probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, we construct an asset model as a stochastic process $X(\omega, t)$ adapted to the filtration $\mathcal{F}_t$ by considering two aspect of properties. The first one is the capability of capturing distributional features. That is, given the value t, the local dynamics of the asset should be captured by $X_t(\omega)$. In the multi-variate model, this involves the joint and marginal distribution of all the random factors. The second fact is the model's time series structure. It is the functional characteristics of the $X_\omega(t)$ given ω, for example, the dependence features of the return increments over time t.

In Chapter 1 we discuss the stylised facts of returns of the UK share indices. Most of them lie in the two catalogues of the target model properties. We try to capture them, on the one hand, by using more realistic time series structure. There are many long-term stochastic investment models have been suggested by various authors. For instance, RSLN of Hardy (2003), Wilkie model of Wilkie (1995), GARCH volatility models, e.g.McNeil *et al.* (2005) and SV models, e.g. Barndorff-Nielsen and Shephard (2001). Or, on the other hand, models can be improved by choosing the proper driving processes. Lévy processes offer more flexible local distributions than the widely used Brownian motion, both statistically and risk neutrally. It captures the sudden big movements in asset values. Other examples of BDLP involve the α-stable processes of Mandelbrot (1977) such as fractional Brownian motions and more broadly, the semi-martingale models.

Chapter 3

Estimation

3.1 Introduction and objectives

In the last chapter we set up parametric models and made assumptions regarding the types of driving processes. In this chapter we discuss some context of the parameter estimation for the Lévy driven asset models.

The moment calibration method requires the weakest assumptions about the data, i.e. that the observations are identically distributed. Since calibrating the moments to the data is equivalent to fitting the characteristic function at point zero, so it may be used as the (numerical) starting point in some more complex estimation procedures.

The other two criteria we use are likelihood function value and the distance function value. The likelihood function is essential for the estimation procedures. We derive the likelihood function for the log-return increment processes in section 3.2.

The maximum likelihood estimation (MLE) is based on the assumption that the estimated parameters give the highest joint probability of the sample observations set. In section 3.3 we apply MLE to find appropriate parameters for the Lévy asset models. Also we test the goodness of fit and discuss the technical difficulties in MLE.

An alternative approach based on the Bayesian estimation called Markov Chain Monte Carlo (MCMC) is applied in section 3.4. Since it is less commonly applied in

actuarial models, we give a short introduction to the MCMC with further references. The Metropolis-Hesting Algorithm (MHA) is introduced to expose the Bayesian estimators with their joint distributions.

3.2 Likelihood

Under the GH assumption, the log return process has independent increments. The log-likelihood function for the GH distribution is

$$\begin{aligned} L(\lambda,\alpha,\beta,\delta,\mu \mid \underline{x}) &= n\log\phi(\lambda,\alpha,\beta,\delta) + \left(\frac{\lambda}{2}-\frac{1}{4}\right)\sum_{i=1}^{n}\log\left(\delta^2+(x_i-\mu)^2\right) \\ &+ \sum_{i=1}^{n}\left[\log\mathbb{K}_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x_i-\mu)^2}\right)+\beta(x_i-\mu)\right], \quad (3.1) \end{aligned}$$

where n is the sample size and $\underline{x} = x_1, x_2, \cdots, x_n$ are the sample observations of the monthly log-return rates.

The exponential Lévy model with GH increments has 5 parameters $(\lambda, \alpha, \beta, \delta, \mu)$. By fixing λ in log-likelihood Equation 3.1, the log-likelihood of some subclasses of GH law can be calculated. For example the log-likelihood function for NIG increment case can be derived from Equation 3.1 by fixing $\lambda = -0.5$, or, for the HYP case, the $\lambda = 1$.

The variance gamma and the Student t cases will be fitted using the different parametrisation given in Equation 2.23 and Equation 2.21. The log-likelihood function for the exp-Lévy model with VG increments is given by

$$\begin{aligned} L(\mu,\theta,\sigma,\nu \mid \underline{x}) &= 2n(\theta-\sigma^2) + \sum_{i=1}^{n}(x_i-\mu) - n\log\left(\sigma\sqrt{2\pi}\nu^{\frac{1}{\nu}}\Gamma\left(\frac{1}{\nu}\right)\right) \\ &+ \left(\frac{1}{\nu}-\frac{1}{2}\right)\left[\sum_{i=1}^{n}\log(|x_i|) - \frac{1}{2}\log\left(\frac{2\sigma^2}{\nu}+\theta^2\right)\right] \\ &+ \sum_{i=1}^{n}\log\left(\mathbb{K}_{1/\nu-1/2}\left(\frac{|x_i|\sqrt{2\sigma^2/\nu+\theta^2}}{\sigma^2}\right)\right), \quad (3.2) \end{aligned}$$

where n is the sample size and $\underline{x} = x_1, x_2, \cdots, x_n$ are the sample observations of the monthly log-return rates.

And the log-likelihood function for the exp-Lévy model with student-t increments

is given by

$$L(\mu, \sigma, \nu \mid \underline{x}) = n \log \left(\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \right) - \left(\frac{\nu}{2} + \frac{1}{2} \right) \sum_{i=1}^{n} \log \left(1 + \frac{(x_i - \mu)^2}{\nu} \right), \quad (3.3)$$

where n is the sample size and $\underline{x} = x_1, x_2, \cdots, x_n$ are the sample observations of the monthly log-return rates.

According to the mathematical result of the Lévy-Khintchine formula, the Fourier transform of the infinitely divisible distributions (characteristic function), in some cases, is much easier to get than its probability density function. Some infinitely divisible distributions have simple characteristic functions but rather complicated form for density, for example, CGMY, see Carr *et al.* (2002). When it is impossible or computationally expensive to derive the explicit likelihood functions, but given the characteristic function is known, Carr *et al.* (2002) show a Fast Fourier transform can be employed to compute the density function value approximately from a single set of parameters.

In addition to the likelihood functions, the moments of the increment laws can be calculated by differentiating of the characteristic function. Cont and Tankov (2004) discuss the advantages of using so-called generalised method of moments. For the reader who is interested in the generalised method of moments and the statistical properties of the methodology, we refer to Hansen (1982).

3.3 MLE Estimations

The MLE procedure actually is an optimisation problem of the nonlinear log-likelihood equations we derived in section 3.2. The partial derivatives of the log-likelihood function with respect to each parameter are given in Prause (1999). To solve these equations numerically, we can hire the non-linear solver such as Newton-Raphson method or the Steffensen interation, see Press *et al.* (2002). However, the differentiated log-likelihood function contains the differentiated values of the modified Bessel function $\mathbb{K}$ whose value itself is calculated by numerical algorithm. When using the Newton-Raphson solver, second order derivative of the modified Bessel function is required. In order to prevent generating extra error from the numerical approximation algorithms, a simplex searching method is used.

More precisely, the Nelder-Mead simplex method is used to find the maximum value of the log-likelihood funcit on. We find that the log-likelihood function of the GH model is flat, the accuracy of the searching depends on the starting values of the estimators. This fact is also observed for daily data in Eberlein and Keller (1995). The parameters easily converge to a local maximum point when applying the numerical optimisation algorithm. This requires some preliminary knowledge of the possible intervals where parameters lie in. From the analogy of the exponential Lévy model and the GBM model, we know approximately the size of drift and the volatility of stochastic components. In Table (3.1) we give the results of the maximum likelihood estimation of the GH and its subclasses by fixing the parameter λ. An alternative algorithm to get the MLE for GH is the EM method, see McNeil *et al.* (2005).

Table 3.1: Maximum likelihood estimates of generalised hyperbolic distributions and the following subclasses: $\lambda = -3/2$, NIG ($\lambda = -1/2$), hyperboloid($\lambda = 0$), hyperbola ($\lambda = 1/2$), and hyperbolic ($\lambda = 1$). The first line gives the maximum likelihood estimates of all 5 parameters where λ has not been fixed.

λ	α	β	δ	μ	LogLH
Share Gross Total return Jan 1950 - May 2005					
-1.6274	10.2791	-4.9408	0.0668	0.0226	1075.18
-1.5	11.5786	-4.9832	0.0645	0.0227	1075.16
NIG	19.8063	-5.3474	0.0472	0.0235	1074.74
0	23.7088	-5.4744	0.0384	0.0237	1074.31
0.5	27.7326	-5.7687	0.0293	0.0240	1073.71
HYP	31.8806	-5.9859	0.0196	0.0247	1072.91
Share Gross Total return Dec 1923 - May 2005					
-1.1881	12.8222	-4.2157	0.0534	0.0184	1650.90
-1.5	10.1304	-4.1590	0.0588	0.0184	1650.84
NIG	19.2131	-4.3965	0.0416	0.0183	1650.66
0	23.4465	-4.3766	0.0328	0.0183	1650.21
0.5	27.6807	-4.3687	0.0235	0.0182	1649.49
HYP	32.1176	-4.2614	0.0126	0.0179	1648.43

The maximum likelihood parameters for Gaussian, Student t and variance gamma law of the UK SGTRI are given in Table 3.3.

Table 3.2: Maximum likelihood estimates of variance gamma, student t and Gaussian.

	μ	σ	ν	θ	LogLH
Share Gross Total return Jan 1950 - May 2005					
VG	0.02410	0.04927	0.69678	-0.01380	1071.51
Student t	0.01335	0.03726	4.10582		1070.80
Gaussian	0.01030	0.05253			1013.10
Share Gross Total return Dec 1923 - May 2005					
VG	0.01670	0.04691	0.83342	-0.00815	1647.52
Student t	0.01138	0.03380	3.73435		1645.54
Gaussian	0.00865	0.04908			1557.55

Not surprisingly the 5 parameter GH law has the highest log-likelihood. Compared with the Gaussian law, which is 1013.10 and 1557.55 for SGTRI50 and SGTRI23, all the non-Gaussian distributions we have considered so far offer a significant improvement in likelihood. The difference between the subclasses by fixing λ and the general case is small. Student t has lowest log likelihood in these distributions, this is partly because student t is the limiting case of GH law where α and β are zero,i.e., it is symmetric, and cannot capture the different tail decay rate in SGTRI.

Hardy (2003) uses an alternative approach to estimate the parameters for the RSLN models called Markov Chain Monte Carlo. This method has some attractive advantages in application. In the next section, we introduce this estimating method and apply in the model fitting of the UK SGTRI.

3.4 Bayesian Estimation: A Markov Chain Monte Carlo approach

The Markov Chain Monte Carlo methodology provides a framework within which parameters can be estimated using a computer simulation approach. In this section, we describe the Metropolis-Hasting algorithm which is applied to find out the Bayesian estimators of the parameters.

Recall the Bayesian inference methodology, let $D = \{X_1, X_2, ..., X_n\}$ denote the observed data which is the log-returns of the SGTRI in our estimation procedure. θ is the model parameter vector. Contrast to maximum likelihood estimation where the parameters are fixed numbers and the estimators are random variables of the sample observations, whereas the Bayesian method assumes the parameter θ is a vector of random variables.

Before the experiments, we have some information about θ, this gives a prior distribution whose density function is given by $\varphi(\theta)$.

Based on the data D, the conditional likelihood function is calculated by:

$$f(D|\theta) = \prod_{i=1}^{n} f(X_i|\theta), \tag{3.4}$$

where $f(\cdot|\cdot)$ is the conditional probability density function of the data. The joint distribution is given by the prior $\varphi(\theta)$ and the conditional likelihood $f(D|\theta)$:

$$f(D, \theta) = f(D|\theta) \cdot \varphi(\theta). \tag{3.5}$$

In the next step we apply Bayes theorem to determine the distribution of θ conditional on D:

$$f(\theta|D) \quad = \quad \frac{\varphi(\theta) f(D|\theta)}{\int \varphi(\theta) f(D|\theta) d\theta}. \tag{3.6}$$

This is called the posterior distribution of θ. The denominator $\int \varphi(\theta) f(D|\theta) d\theta$ is constant since it is the likelihood of the data D. Thus we have $f(\theta|D) \propto \varphi(\theta) f(D|\theta)$.

The Bayesian estimators are some statistics of the posterior distribution. For example, the mean, percentage quantiles, etc. These can be expressed in terms of the expected value of some function of θ:

$$\mathbb{E}[u(\theta)|D] = \frac{\int u(\theta) \varphi(\theta) f(D|\theta) d\theta}{\int \varphi(\theta) f(D|\theta) d\theta}. \tag{3.7}$$

Analytically calculating $\int u(\theta)\varphi(\theta)f(D|\theta)d\theta$ is impossible in most cases. Alternative approach needs to be hired to find out the value of $\mathbb{E}[u(\theta)|D]$.

The key idea of Markov Chain Monte Carlo is to compute the expectation, $\mathbb{E}[u(\theta)|D]$, by simulating a sequence of random variable $\{X_0, X_1, X_2, ...\}$ which has similar statistical properties as the samples from the posterior distribution of the parameter vector function $u(\theta)$. We note that in Equation 3.6 the numerator can be calculated for certain values of θ, and the denominator is constant and stays unchanged. Thus we can avoid integrating the joint distribution function in the denominator by simply eliminating it, for example, dividing the $f(\theta_1|D)$ by another posterior $f(\theta_2|D)$. To take this advantage, we can build a Markov chain $\{X_0, X_1, X_2, ...\}$ in which the transition kernel $\mathbb{P}(\cdot|\cdot)$ is given by the form of dividing posterior. Properly using the simulation methodology, this Markov chain will eventually reach the stationary distribution. The first m values of simulation before the chain reaches its stationary distribution will be ignored as a burn-in. The X_t for $t \geq m$, as t increases, will look increasingly like dependent samples from the posterior of θ. By Birkhoff's Ergodic Theorem, the expectation $\mathbb{E}[u(\theta)|D]$ can be estimated as the sample mean of $u(\theta)$.

Here we introduce the Metropolis-Hasting Algorithm(MHA) used to build the Markov Chain. At each time t, in order to generate the next state X_{t+1}, we first generate a candidate value Y from a proposal distribution $q(\cdot|X_t)$. It may depend on current X_t. The next step is to decide to accept or reject the candidate Y. If the candidate is accepted, the Y will be the next state X_{t+1}. In contrary, if is rejected, the X_{t+1} stay in the current state X_t. The acceptance probability is given by:

$$\alpha(X_t, Y) = min\left(1, \frac{\varphi(Y)f(D|Y)q(X|Y)}{\varphi(X)f(D|X)q(Y|X)}\right). \tag{3.8}$$

The feature of Equation 3.8 is that we have the ratio between two posterior distributions, the unknown constant in denominator in Equation 3.6 is no longer involved in calculation. For a comprehensive introduce to MCMC and more details of applications, we refer to Gilks *et al.* (1996) and Hardy (2003).

3.4.1 MCMC for Lévy models

Based on the previous work, the maximum likelihood estimators along with the estimated information matrix give the description of the joint distribution of the parameters. We can use this information as the prior distribution of the estimators. On the other hand, we can assume there is no prior information of the parameter. Since in high dimensional cases, the MCMC has the advantage of less prior knowledge dependency than the MLE. To show this, we choose high variance/domain rate priors such as uniform or normal distributions whose possible random outcomes cover the theoretical domain of the parameters. The first state of the Markov Chain is randomly generated from the prior distribution of each parameter. We check the convergence of the Markov Chain by running parallelled ten MCMC experiments each with different random starting values. In all cases, after the burn-in period, the Markov Chain will reach its stable distribution successfully (we check this by observing the first 4 central moments of the simulated Markov Chain samples).

We try to use Normal, Gamma and Beta as the prior or candidate distributions. The reason is that simulation progress is restricted by the quality and speed of random number generators. We give the details of the computer algorithms for Normal, Gamma and Beta random number generators in Section 4.4.1. Specially the using of normal candidate density cases simplify the acceptance probability, see Hardy (2003), the ratio:

$$\frac{q(X|Y)}{q(Y|X)} = 1 \tag{3.9}$$

in Equation 3.8.

3.4.2 Metropolis-Hasting Algorithm for VG

In this section, details of Metropolis-Hasting algorithm(MHA) for the VG parameter estimation for SGTRI50 data will be discussed. This includes the choice of the prior distributions and proposal distributions for each parameter and some issues of the approximation methods we used. The MHA for SGTRI23 is very similar to the one for SGTRI50, the proposal distributions are chosen with slightly different parameters from those for SGTRI50 in order to obtain the proper acceptance probability. The

Gilks *et al.* (1996) points out the efficient interval of the acceptance probability is $(25\%, 40\%)$.

Denote $\theta^{(r)}$ as the current value of parameter θ in estimation. For fitting the VG model to SGTRI data set, recall the density function of VG in Equation 2.23, we use normal prior distributions for the linear drift μ and the stochastic drift θ:

$$\mu \in N(0.05, 0.02^2); \quad \theta \in N(0.0, 0.02^2).$$

The proposal distribution for μ is $N(\mu^{(r)}, 0.06^2)$. The standard deviation of the normal distribution is carefully chosen in order to give an appropriate probability of acceptance. Also the probability of acceptance depends on the distributions of other parameters. As in estimation progress we have acceptance probabilities of around 34 percent.

The proposal distribution of θ is $N(\theta^{(r)}, 0.007^2)$. Again the standard deviation of the proposal distribution is appropriate which give acceptance probabilities of around 35 percent.

The prior distribution for ν is the gamma distribution with prior mean 1.0 and standard deviation 0.5. The prior distribution is dispersed so that the prior density term will have very little effect on the acceptance probability. The proposal distribution is gamma with mean $\nu^{(r)}$ and standard deviation $\nu^{(r)}/2.5$. The acceptance probabilities for ν candidate is approximately 38 percent.

The prior distribution for σ is the gamma distribution with prior mean 0.05 and standard deviation 0.015. The proposal distribution is gamma with mean $\sigma^{(r)}$ and standard deviation $\sigma^{(r)}/8.5$. The acceptance probabilities for σ candidate is approximately 36 percent.

Alternatively, a modified normal distribution can be used as the prior and proposal distributions for σ. First, we have an initial normal distribution with mean 0.05 and variance 0.015^2. Since the support of the σ is $(0, \infty)$, we cut the left tail of the distribution at point c and add those probability density less than cutting point to it's positive symmetric value of 0.0001, in general, with normal mean μ and standard deviation ξ, the probability density of x is given by:

$$f(x, \mu, \xi, c) = \phi\left(\frac{x-\mu}{\xi}\right) + \phi\left(\frac{2c-x-\mu}{\xi}\right), \quad x \geq c, \tag{3.10}$$

where $\phi(\cdot)$ is the density of standard normal distribution. In estimation, we can approximately ignore the second term in the density function 3.10 if it given very low density. For example, when the cutting point c is lay out of the three times of standard deviation interval of the normal distribution, the integration of the second term in equation 3.10 on real line is less than 0.002 which can be ignored safely in calculation.

The proposal distribution of σ is modified normal with initial normal mean $\sigma^{(r)}$ and standard deviation 0.006. The cutting point c of left tail is 0.0001. In simulation, we can just skip those very few generated candidate values less than 0.0001. The acceptance probabilities is approximately 33 percent. We denote the modified normal distribution as $\tilde{N}(\mu, \xi^2, c)$ where μ and ξ is the mean and standard deviation parameter of the initial normal distribution and c is the cutting point.

We tried both gamma prior/proposal and the normal approximation method, the results are very close when the times of simulation increase. Actually if a random variable X is Gamma distributed with shape parameter α and the scale parameter λ, when 2α is an integer, it has the relationship:

$$2\lambda X \sim \chi^2_{2\alpha}.$$

The χ^2_ν is approximately normality when the ν is big enough. The skewness $\sqrt{8/\nu}$ and the excess kurtosis $12/\nu$ of the χ^2_ν distribution can be used as measures of the convergence to Gaussian. Thus if the α is big enough, for example $\alpha \geq 16$, the Gamma prior or proposal distribution is approximately a normal or modified normal distribution.

We use normal prior distribution $N(0.05, 0.02^2)$ for μ and $N(0.0, 0.02^2)$ for θ to fit the exp-VG model to the SGTRI23 data.

The standard deviation of the normal proposal distribution for μ is 0.054 and with mean $\mu^{(r)}$. The probability of acceptance is around 30 percent.

The proposal distribution of θ is $N(\theta^{(r)}, 0.064^2)$. Again the standard deviation of the proposal distribution is appropriate which give acceptance probabilities of around 32.

The prior distribution for ν is the gamma distribution with prior mean 1.0 and standard deviation 0.5. The proposal distribution is gamma with mean $\nu^{(r)}$ and

standard deviation $\nu^{(r)}/2.5$. The acceptance probabilities for ν candidate is approximately 33 percent.

The prior distribution for σ is the gamma distribution with prior mean 0.05 and standard deviation 0.015. The proposal distribution is gamma with mean $\sigma^{(r)}$ and standard deviation $\sigma^{(r)}/8.5$. The acceptance probabilities for σ candidate is approximately 32 percent.

The result

The MCMC estimators for VG model and their correlation is shown in Table 3.3. The first column in the table gives the mean of the simulated parameter samples. The second column shows the standard deviation and the rest four column gives the parameters correlations. In fact the Metropolis-Hasting algorithm we used for the VG model fitting is similar to a Monte Carlo random searching method to find out the maximum likelihood estimators. Thus, if the likelihood function is smooth near its maximum point, the maximum likelihood estimators should be lay in the area near the medium value of Bayesian posterior distribution for each parameters. Using the mediums of the MCMC samples as the parameter, the log-likelihood values are 1647.49 for SGTRI23 data set and 1071.46 for SGTRI50. These are close to the maximum likelihood which is 1647.52 for SGTRI23 and 1071.51 for SGTRI50 in table 3.3.

Figure 3.1 shows marginal density for the MCMC samples of the estimated parameters. The solid lines show the SGTRI50 results and the dotted lines represents the SGTRI23 density of estimators. The figure 3.2 demonstrate the two way joint simulated sample observations for the parameters estimated for SGTRI50 data. The joint distributions for SGTRI23 parameters are given in the figure 3.3.

3.4.3 MHA for GH

The Metropolis-Hasting algorithm we used to estimate the GH parameters is similar to the one we applied on VG estimation. The HYP and NIG distributions are subclasses of the GH. They share the same parametrisation but with fixed λ, i.e., $\lambda = -0.5$ for the NIG and $\lambda = 1$ for the HYP. Normal prior distribution will be

Table 3.3: Markov Chain Monte Carlo of VG.

	Mean	SD	Parameters correlations with: μ	θ	σ	ν
Share Gross Total Return Jan 1950 - May 2005						
μ	0.0248	0.0038	1.0000	-0.8841	-0.3094	-0.3191
θ	-0.0144	0.0042		1.0000	0.2413	0.2796
σ	0.0494	0.0020			1.0000	0.3532
ν	0.7009	0.1109				1.0000
Share Gross Total Return Dec 1923 - May 2005						
μ	0.0172	0.0023	1.0000	-0.8325	-0.2340	-0.3605
θ	-0.0086	0.0028		1.0000	0.1592	0.3048
σ	0.0471	0.0016			1.0000	0.3906
ν	0.8372	0.1082				1.0000

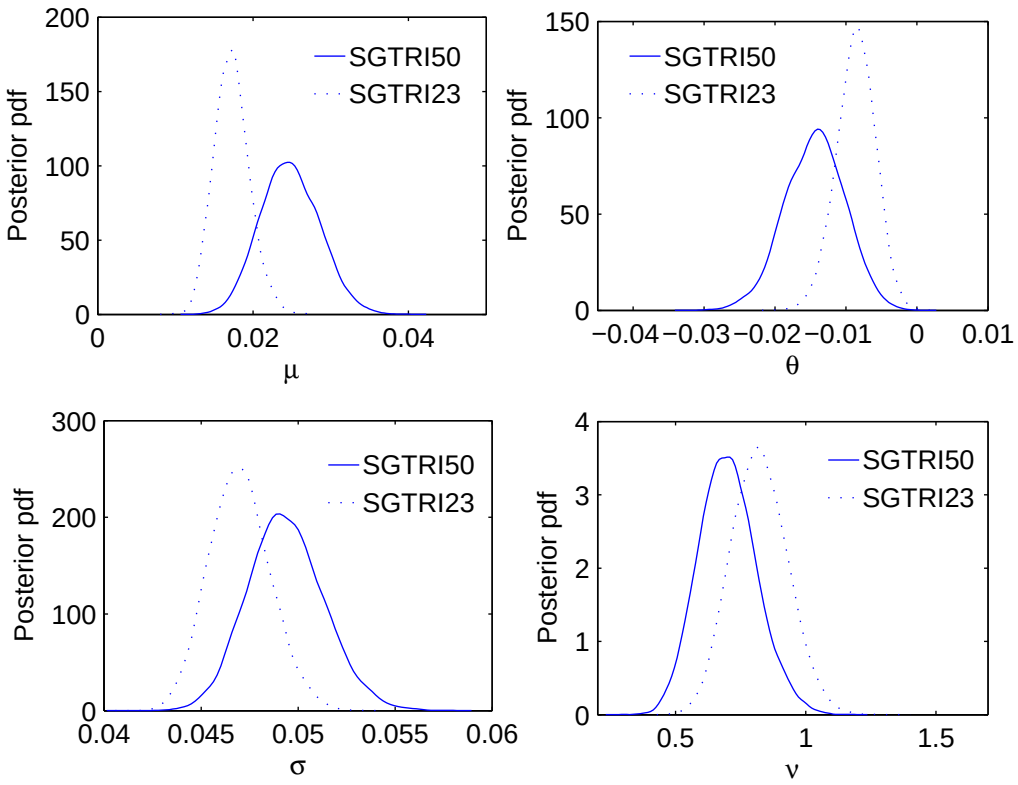

Figure 3.1: Simulated marginal posterior parameter distributions.

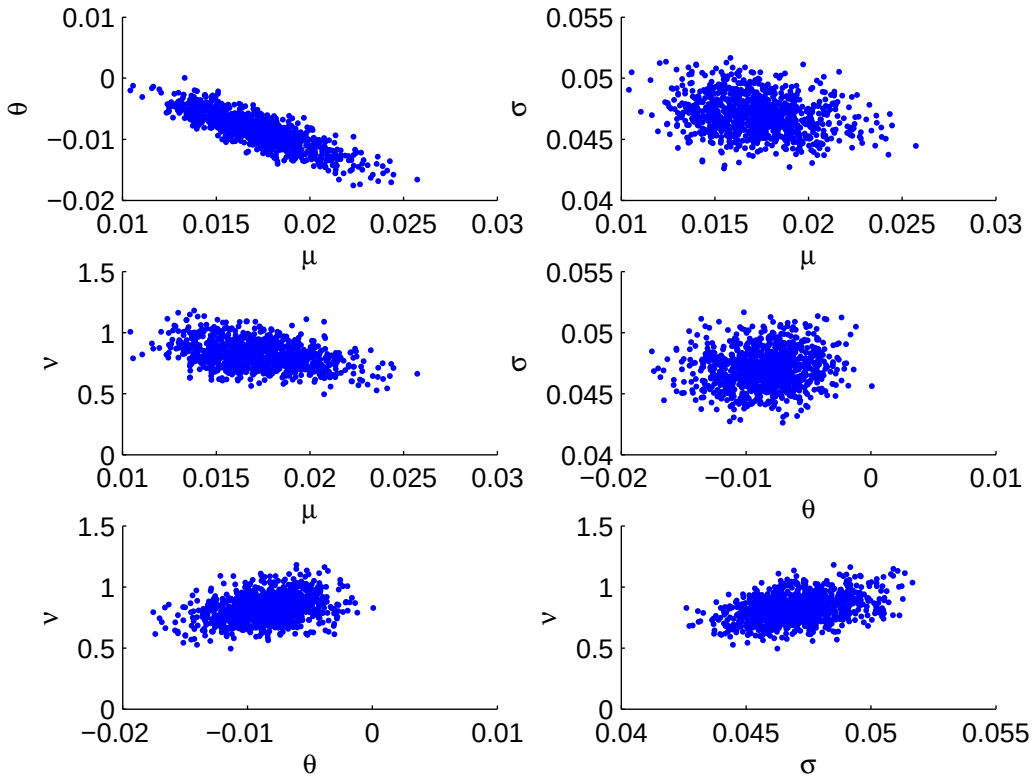

Figure 3.2: Two-way joint distributions for SGTRI50 data.

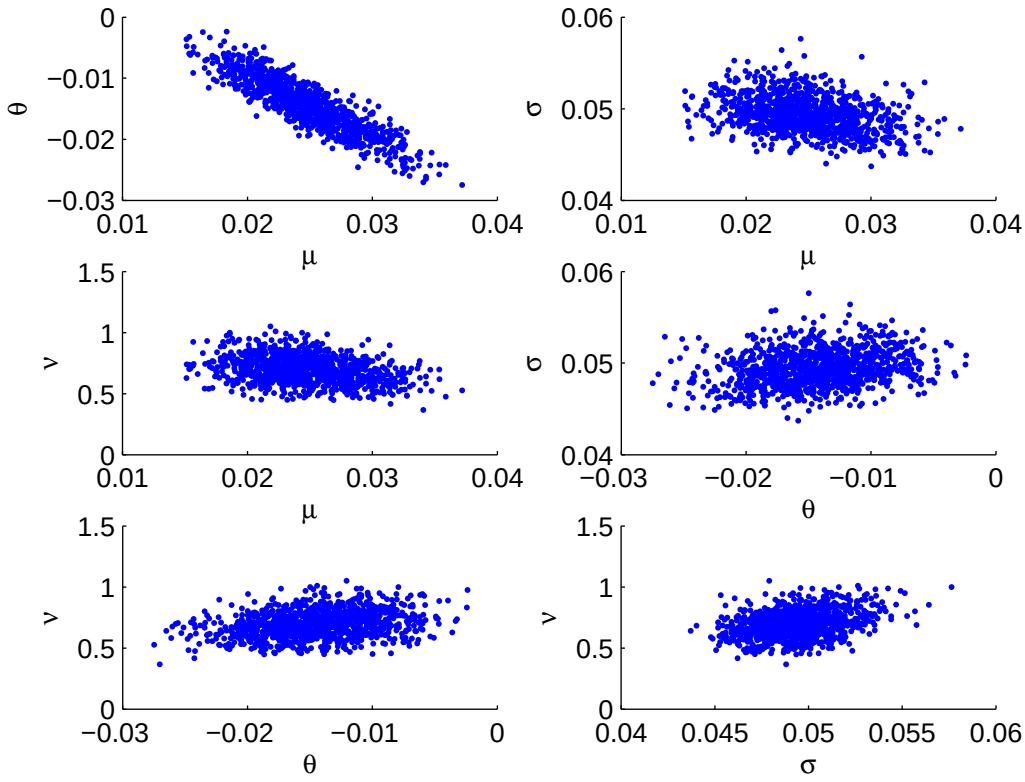

Figure 3.3: Two-way joint distributions for SGTRI23 data.

used for λ with large variance for GH. It is difficult to choose a distribution for α directly because it's domain depends on the value of β, i.e, $|\alpha| > |\beta|$, instead we try to look at the γ which is given by:

$$\gamma = \sqrt{\alpha^2 - \beta^2}, \tag{3.11}$$

and it is on the positive real line. In convenience, only the prior and proposal for γ will be Gamma distributions. For the rest of the parameters, normal and modified normal prior and proposal distributions works well in this example. Table 3.4 gives the prior and proposal distributions and the acceptance probabilities for each model fitted to respectively two data sets, SGTRI50 and SGTRI23.

Table 3.4: Prior&Proposal distributions for GH MCMC

	Prior	Proposal	AP(SGTRI50)	AP(SGTRI23)
GH				
λ	$N(0, 0.7^2)$	$N(0.1, 0.3^2)$	40	38
α	$\Gamma(0.45, 0.05)$	$\Gamma(0.25, 0.25/\alpha^{(r)})$	33	31
β	$N(0.0, 5.0^2)$	$N(\beta^{(r)}, 2.4^2)$	37	35
δ	$MN(0.05, 0.1^2)$	$N(\delta^{(r)}, 0.015^2)$	29	25
μ	$N(0.1, 0.3^2)$	$N(\mu^{(r)}, 0.0075^2)$	27	22
NIG				
α	$\Gamma(0.45, 0.05)$	$\Gamma(0.25, 0.25/\alpha^{(r)})$	41	37
β	$N(0.0, 5.0^2)$	$N(\beta^{(r)}, 2.4^2)$	37	36
δ	$MN(0.05, 0.1^2)$	$N(\delta^{(r)}, 0.015^2)$	29	27
μ	$N(0.1, 0.3^2)$	$N(\mu^{(r)}, 0.0075^2)$	27	22
HYP				
α	$N(20.0, 10.0^2)$	$\Gamma(81, 81/\alpha^{(r)})$	37	36
β	$N(0.0, 5.0^2)$	$N(\beta^{(r)}, 2.200^2)$	37	36
δ	$MN(0.05, 0.1^2)$	$MN(\delta^{(r)}, 0.018^2)$	32	30
μ	$N(0.1, 0.3^2)$	$N(\mu^{(r)}, 0.007^2)$	24	22

The results

We use same prior and proposal for SGTRI50 and SGTRI23. The acceptance probabilities (AP) of the candidate parameters are not sensitive to these two sample sets. The AP(SGTRI50) is often higher than AP(SGTRI23) under same prior and proposal distributions for each parameter. This is simply because SGTRI50 is subsample of SGTRI23 and the latter gives more information.

Table 3.5 shows the results for the GH distribution. We note that the λ estimator of the GH Lévy model has a relatively high standard deviation. λ and δ are extremely strong negatively correlated , i.e., around -0.96. And α is positively correlated to λ, i.e., around 0.89. It suggests that the model can equally perform well in a large range of λ as long as choosing proper α and δ. Similar to VG estimation, the pair, stochastic drift β and linear drift μ are negatively correlated. The joint distributions in the Figure 3.4 gives a graphical view of the joint distribution of parameter vector.

Table 3.5: Markov Chain Monte Carlo of GH

	Mean	SD	Parameters correlations with:				
			λ	α	β	δ	μ
Share Gross Total Return Jan 1950 - May 2005							
λ	-1.2992	0.9486	1.0000	0.8701	-0.0782	-0.9600	0.1096
α	15.2368	7.7061		1.0000	-0.3501	-0.7455	0.3348
β	-5.5449	1.9300			1.0000	0.0385	-0.8850
δ	0.0634	0.0177				1.0000	-0.0633
μ	0.0237	0.0038					1.0000
Share Gross Total Return Dec 1923 - May 2005							
λ	-1.1919	0.7270	1.0000	0.8829	-0.0187	-0.9645	-0.0182
α	14.1982	6.7134		1.0000	-0.2396	-0.7668	0.2000
β	-4.5034	1.4909			1.0000	-0.0674	-0.8655
δ	0.0548	0.0133				1.0000	0.0703
μ	0.0185	0.0026					1.0000

Table 3.6 shows the estimation results for the NIG distribution. Same patterns are found for two sample set. The SGTRI23 estimators have less standard devia-

tions than the SGTRI50 ones. Compare the results for NIG with the general five parameter GH estimation, the SD for α is reduced because of the fixed λ. We use the sample mean of each posterior for the parameters as the "best estimator", the log-likelihood is 1649.43, it's close to the log-likelihood of MLE results.

Table 3.6: Markov Chain Monte Carlo of NIG

	Mean	SD	Parameters correlations with: α	β	δ	μ
Share Gross Total Return Jan 1950 - May 2005						
α	21.2600	3.3122	1.0000	-0.5297	0.7963	0.4235
β	-5.6869	1.8459		1.0000	-0.2018	-0.8733
δ	0.0497	0.0056			1.0000	0.1965
μ	0.0241	0.0035				1.0000
Share Gross Total Return Dec 1923 - May 2005						
α	20.1240	2.5487	1.0000	-0.4873	0.8046	0.3927
β	-4.5457	1.4086		1.0000	-0.2259	-0.8511
δ	0.0430	0.0037			1.0000	0.2215
μ	0.0186	0.0024				1.0000

Table 3.7 shows the estimation results for the HYP distribution. The sample mean of each posterior for the parameters is used as the "best estimator", the log-likelihood is 1647.89.

Table 3.8 shows the MCMC statistics of the Student t distribution. We again use the sample mean of each posterior for the parameters as the "best estimator", the likelihood is 1645.49 for SGTRI23 and 1070.74 for SGTRI50.

Finally, the MCMC statistics of the normal distribution are shown in Table 3.9. The likelihood using sample mean estimators is 1556.93 for SGTRI23 and 1012.87 for SGTRI50.

Table 3.7: Markov Chain Monte Carlo of HYP

	Mean	SD	Parameters correlations with: α	β	δ	μ
Share Gross Total Return Jan 1950 - May 2005						
α	33.0338	2.5774	1.0000	-0.5625	0.7906	0.4497
β	-6.3694	1.9512		1.0000	-0.2703	-0.8832
δ	0.0231	0.0078			1.0000	0.2582
μ	0.0254	0.0037				1.0000
Share Gross Total Return Dec 1923 - May 2005						
α	32.8330	1.9835	1.0000	-0.5462	0.8046	0.4723
β	-4.5458	1.4713		1.0000	-0.3913	-0.8638
δ	0.0146	0.0061			1.0000	0.4048
μ	0.0184	0.0026				1.0000

Table 3.8: Markov Chain Monte Carlo of Student t.

	Mean	SD	Parameters correlations with: μ	σ	ν
Share Gross Total Return Jan 1950 - May 2005					
μ	0.0135	0.0017	1.0000	-0.1726	-0.1900
σ	0.0378	0.0018		1.0000	0.6569
ν	4.4110	0.7728			1.0000
Share Gross Total Return Dec 1923 - May 2005					
μ	0.0114	0.0013	1.0000	-0.1239	-0.1374
σ	0.0343	0.0014		1.0000	0.6768
ν	3.9294	0.5324			1.0000

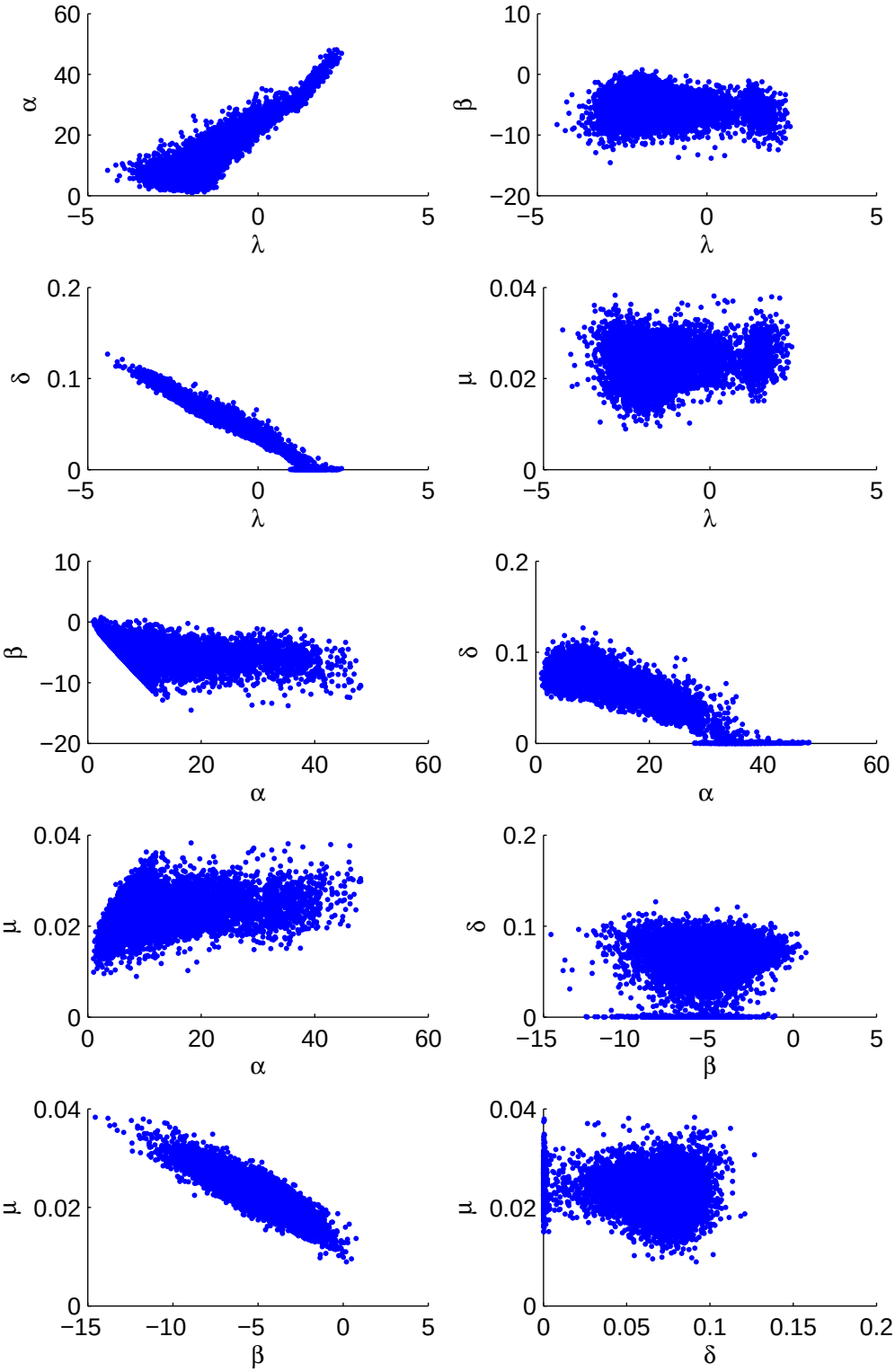

Figure 3.4: Two-way joint distributions for SGTRI23 data.

Table 3.9: Markov Chain Monte Carlo of GBM.

	Mean	SD	Parameters correlations with: μ	σ
Share Gross Total Return Jan 1950 - May 2005				
μ	0.0103	0.0020	1.0000	-0.0104
σ	0.0527	0.0015		1.0000
Share Gross Total Return Dec 1923 - May 2005				
μ	0.0087	0.0016	1.0000	0.0020
σ	0.0491	0.0011		1.0000

3.5 Discussion of Results

We discuss the estimation results from four perspectives. First, we present goodness-of-fit test for the fitted models based on distance of probability function. In order to avoid repeating similar distance tests based on same principle, we present only one of them called Anderson & Darling test. Next, we consider the efficiency of using different return distributions discussed in this chapter by computing so-called information criterion and likelihood ratio tests. We then study some pivotal statistically properties such as skewness and kurtosis embedded in the Lévy models. Finally we discuss some issues of the Bayesian asset models with MCMC estimation called hyper-models.

3.5.1 Goodness of fit test

To test goodness of fit, we use some score functions other than likelihood. Note that for the distance testing procedures such as Kolmogorov and Anderson & Darling test, or the likelihood ratio comparison, all depend on the assumption that observations are i.i.d.

The Anderson & Darling (AD) Statistic is given by

$$AD = max_{x\in\mathbb{R}} \frac{|F_{emp}(x) - F_{est}(x)|}{\sqrt{F_{est}(x)(1 - F_{est}(x))}}, \tag{3.12}$$

where F_{emp} and F_{est} are the empirical and the estimated cumulative density func-

tions (CDFs). Table 3.10 shows the distance test statistic using parameters of the maximum likelihood estimation.

Table 3.10: Distance goodness of fit test for GH, NIG, HYP, VG, Student t and Gaussian law.

	GH	NIG	HYP	VG	t	Gaussian
SGTRI23-05	0.0583	0.0604	0.0622	0.0643	0.09621	36704.1
SGTRI50-05	0.0609	0.0619	0.0622	0.0712	0.10812	6138.4

The poor fit of Gaussian law to the heavy tail in the empirical data leads to extremely large AD distance values. In this example, the Gaussian has a AD distance 36704.1 for SGTRI23-05 and 6138.40 for SGTRI50. For comparison, the student t has an AD distance 0.09621 for SGTRI23-50 and 0.10812 for SGTRI50-05. There is no large difference between the results for GH, NIG, HYP and VG.

3.5.2 Information Criterion & Model Selection

The Akaike information criterion (AIC) uses the entropy analysis to examine the goodness of fit with complexity of the model. The AIC is given by:

$$AIC = 2k - 2\log(L), \tag{3.13}$$

where k is the number of parameters, and L is the likelihood function. The preferred model is that with the lowest AIC value. It means each extra parameter is worthwhile only when the more complex model can improve log-likelihood by more than one. It gives an easy approach to measure goodness of fit and discourages over-fitting.

The Bayesian information criterion (BIC) is similar to AIC. It prefers the model with the lowest value of

$$BIC = k\log(n) - 2\log(L), \tag{3.14}$$

where n is the sample size or the number of observations, k is the number of parameters and L is the likelihood function. The extention of the basic model to a more

complex model with additional parameters is considered to be worthwhile only if the log-likelihood increases given the number of observations. In case that sample size is large, an additional parameter added to the model must lead to a greater improvement of the log-likelihood than in small sample size case.

Table 3.11 shows the value of AIC and BIC of GH and its subclasses and limiting distributions for UK SGTRI from January 1950 to May 2005 and from December 1923 to May 2005. In both cases the exponential Lévy models have lower AIC and BIC than GBM. Among the exp-Lévy models, the NIG has the lowest AIC and BIC for both time intervals. But the difference between exp-Lévy models is small. HYP has less AIC and more BIC than GH, this is caused by BIC's calculation with sample size n. Comparing the four parameter subclasses NIG, HYP and VG with the three parameter Student t, Student t has higher BIC than VG and HYP. It suggests the Student t may be a quite efficient model for UK SGTRI.

The general 5 parameters GH law looks over-fitted to the SGTRI data. In our estimating procedures, fixing parameter λ reduces one dimension in the optimisation problem to find maximum likelihood. The searching procedures for NIG and HYP converge quicker and are more reliable than the general 5 parameters case. The variance gamma model with its own parametrisation (θ, σ, ν) has only 3 parameters to estimate and the estimator of the 4th parameter, drift μ, can be easily calculated by the difference between sample mean and θ.

One of the Bayesian approaches to compare different models is to use Bayes factors. Given two models M_1 and M_2, based on sample data D, the Bayes factor for model M_1 against $M2$ is

$$B(M_1, M_2) = \frac{f(D|M_1)}{f(D|M2)} \tag{3.15}$$

where $f(D|M_i)$ is the conditional likelihood function condition called marginal likelihood for model i. We assume the parameter vector for M_i is θ_i with posterior density $\varphi(\theta)$, the joint likelihood function $f(D, \theta_i|M_i)$ is

$$f(D, \theta_i|M_i) = f(D|\theta_i, M_i)\varphi(\theta). \tag{3.16}$$

The marginal likelihood is $\int f(D, \theta_i|M_i)d\theta$, hence Equation 3.15 can be calculated

Table 3.11: Selection information for GH, NIG, HYP, VG, ST and GBM.

Model	Parameters k	LogLH	AIC	BIC
Share Gross Total return Jan 1950 - May 2005				
GH	5	1075.18	-2140.36	-2117.87
NIG	4	1074.74	-2141.48	-2123.49
HYP	4	1072.91	-2137.82	-2119.83
VG	4	1071.51	-2135.02	-2117.03
ST	3	1070.80	-2135.60	-2122.11
GBM	2	1013.10	-2022.20	-2013.21
Share Gross Total return Dec 1923 - May 2005				
GH	5	1650.90	-3291.80	-3267.38
NIG	4	1650.66	-3293.32	-3273.79
HYP	4	1648.43	-3288.86	-3269.33
VG	4	1647.52	-3287.04	-3267.51
ST	3	1645.54	-3285.08	-3270.43
GBM	2	1557.55	-3111.10	-3101.33

by

$$B(M_1, M_2) = \frac{\int f(D|\theta_1, M_1)\varphi(\theta_1)d\theta_1}{\int f(D|\theta_2, M_2)\varphi(\theta_2)d\theta_2}. \tag{3.17}$$

One can see from Equation 3.17, a likelihood-ratio statistic

$$\lambda(D) = \frac{f(D|\theta_1, M_1)}{f(D|\theta_2, M_2)}$$

is a special case of Bayes factor. Let the parameter vector θ_i to be the maximum likelihood estimator $\hat{\theta}_i$ for M_i, we have the maximum likelihood-ratio statistic.

For the MCMC estimators, the generated Markov Chain $(X_t)_{t\in\mathbb{Z}}$ is sample from the posterior distribution of parameter θ. Thus we use the sample distribution of the Markov Chain $\hat{\theta}(k), k \in \mathbb{Z}$ as the density $\varphi(\theta)$. The margin likelihood for M_i can be estimated using sample mean of the conditional likelihood functions.

$$\begin{aligned} MLH(M_i) &= \int f(D|\theta_i, M_i)\varphi(\theta_i)d\theta_i & (3.18)\\ &= \mathbb{E}_{\theta_i}[f(D|\theta_i, M_i)] & (3.19)\\ &= \frac{\sum_{m<k\leq n}\prod_{X_j\in D} f(X_j|\hat{\theta}_i k, M_i)}{n-m}, & (3.20) \end{aligned}$$

where m is the number of "burn-in"s. Table 3.13 lists the interpretation for Bayes

factors suggested by Jeffreys (1961). The second column gives the corresponding range of logarithm of Bayes factors.

Table 3.12: Jeffrey's' scale of evidence for Bayes factors.

Bayes factors	Logarithm, $LB(M_1, M_2)$	Interpretation
$B(M_1, M_2) \leq 1$	$LB(M_1, M_2) \leq 0$	Negative, support M_2
$1 < B(M_1, M_2) \leq 3$	$0 < LB(M_1, M_2) \leq 1.10$	Weak evidence for M_1
$3 < B(M_1, M_2) \leq 10$	$1.10 < LB(M_1, M_2) \leq 2.30$	Moderate evidence for M_1
$10 < B(M_1, M_2) \leq 30$	$2.30 < LB(M_1, M_2) \leq 3.40$	Strong evidence for M_1
$30 < B(M_1, M_2) \leq 100$	$3.40 < LB(M_1, M_2) \leq 4.61$	Very strong evidence for M_1
$B(M_1, M_2) > 100$	$LB(M_1, M_2) > 4.61$	Decisive evidence for M_1

One must be careful in calculating of the Bayes factors because when the sample size n is large, the likelihood function (not log-likelihood) value may overflow the computer programme used in calculation. We use Matlab in this example to calculate the marginal likelihood functions. Matlab uses the IEEE 64-bit floating-point number system to present all numbers. The maximum number is

$$f_{\max} = (2 - 2^{-52})2^{1023},$$

which is around $\exp\{709.78\}$. Any numbers have absolute value larger than it are treated as infinities. The minimum in the IEEE 64-bit floating-point number system is

$$f_{\min} = 2^{-1074},$$

which is around $\exp\{-744.44\}$. Any numbers have absolute value less than it are treated as zeros. Thus we are not able to calculate marginal likelihood functions directly.

We find the sample medium of the estimated parameters $\hat{\theta}_i^{(m)}$, let constant C be the conditional log-likelihood,

$$C = \log\{f(D|\hat{\theta}_i^{(m)}, M_i)\}.$$

Thus margin likelihood $MLH(M_i)$ can be written as

$$MLH(M_i) = \exp(C)\mathbb{E}_{\theta_i}\left[\frac{f(D|\theta_i, M_i)}{C}\right].$$

Table 3.13 lists the logarithm of margin likelihood of the exp-Lévy models we have discussed. And Bayesian factors for M_1 against M_2 is

$$B(M_1, M_2) = \exp\left(\log(MLH(M_1)) - \log(MLH(M_2))\right).$$

Table 3.13: Logarithm of marginal likelihood for GH, NIG, HYP, VG, Student-t and GBM.

Model	Log(MLH) Jan 1950 - May 2005	Log(MLH) Jan 1923 - May 2005
GBM	1012.47	1555.88
ST	1069.74	1646.01
VG	1070.31	1646.75
HYP	1071.47	1646.94
NIG	1073.27	1648.93
GH	1073.48	1649.25

All non-Gaussian models have decisive support evidence against the GBM model. For both sample sets, GH has higher marginal likelihood than its subclasses NIG, HYP and VG. The largest Bayes factor among the non-Gaussian distributions is $B(GH, VG)$ which is 12.18 for sample from Jan 1923 to May 2005 and 23.81 for sample from Jan 1950 - May 2005. It is a strong support evidence for GH model. Among the subclasses of GH, NIG perform better than HYP and VG. Bayes factor for NIG against HYP is 7.32 for sample from Jan 1923 to May 2005 and 6.55 for sample from Jan 1950 - May 2005, and Bayes factor for NIG against VG is 8.84 for sample from Jan 1923 to May 2005. There is moderate support evidence for NIG. Bayes factor for NIG against VG is 19.29 for sample from Jan 1950 - May 2005 which suggests strong evidence for NIG in this case. GH and NIG give very close marginal likelihood, it suggests there is very weak evidence show one model over the others.

Compare the hyper-models with the models using MLE, there is weak evidence showing that the fix parameter models using MLE estimators perform better than hyper-models using Bayesian estimators with respect to same driving process.

3.5.3 Statistical properties of exp-Lévy models

Skewness and Kurtosis

Recall Table 1.1 lists the skewness and kurtosis of the two sample sets. Although explicit analytical form of the GH moments can be derived by differentiate the moment generating function which is given in Prause (1999), we study the higher moments of the fitted increment distributions include GH, NIG, HYP and VG using Monte Carlo techniques. A sample involve 10000 monthly log-returns is simulated for each of the GH laws and its subclasses described in this chapter. The statistics of the skewness and kurtosis are listed in Table 3.14.

Table 3.14: Statistics for simulated returns

Statistics	Mathematical Form	Gaussian	t	VG	NIG	GH
Mean	$\mathbb{E}[Y]$	0.0086	0.0114	0.0086	0.0087	0.0086
Variance	m_2	0.0024	0.0025	0.0023	0.0023	0.0023
Skewness	$m_3/m_2^{3/2}$	0	0	-0.4037	-0.7058	-0.7126
Kurtosis	m_4/m_2^2	3.0000	N/A	5.5029	7.3674	7.4109

Note that the maximum likelihood estimators of the VG, NIG and GH laws give smaller variance than the sample variance. While on the other hand the student-t distribution over-estimate the variance. This may because that student t has heavy tails on both side as power functions (Equation 2.22) and the VG, NIG and GH have semi-heavy tails which are exponentially decreasing functions (Equation 2.13). And for the SGTRI23 data set which is a relatively small sample for distribution fitting, there are several extreme sample observations. The largest log-return is 0.4300 and the second largest is 0.2190, there is no data point in between, and in this case distributions with power decreasing tails give more weight on likelihood of extreme value than distributions with exponential decreasing tails. The Gaussian distribution gives exactly the sample variance because the maximum likelihood estimators for Gaussian law is same as moment estimators.

Density and Quantile

Figure 3.7 and Figure 3.8 show the density functions of the estimated log return processes with regards to different models. It is obvious that the Lévy driven models fit the data accurately. They all have higher peak in the middle and heavier tails for both sides than Gaussian law. Thus it captures the feature of high kurtosis in empirical data. QQ-plots are shown in Figure 3.5 and Figure 3.6, we find that all these distributions except normal give good fit to left tail. For right tail, VG, HYP, NIG, GH perform better than student t. This is mainly because student t distribution is symmetric, and the sample set of the SGTRI is negatively skewed but exist the positive large movement in January 1975, when the SGTRI went up suddenly in one month by 43%. For those cases when the return data are not strongly asymmetric, student t seems good enough.

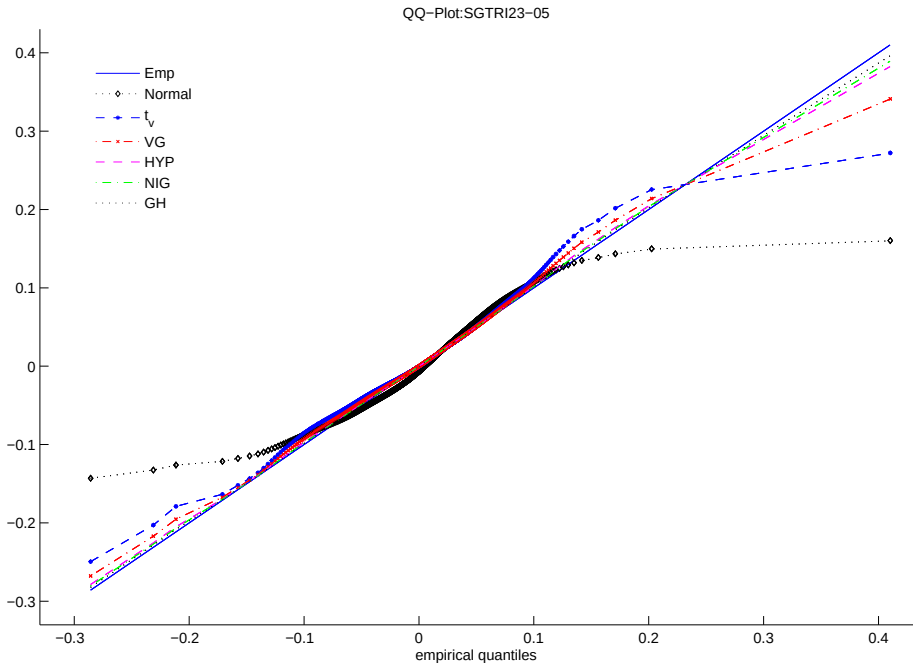

Figure 3.5: QQ plots of SGTRI23-05

3.5.4 MCMC estimator and Hyper-models

One comment about MCMC is that the predictive distribution of the model contains Bayesian estimators which are random variables, given the posterior distribution

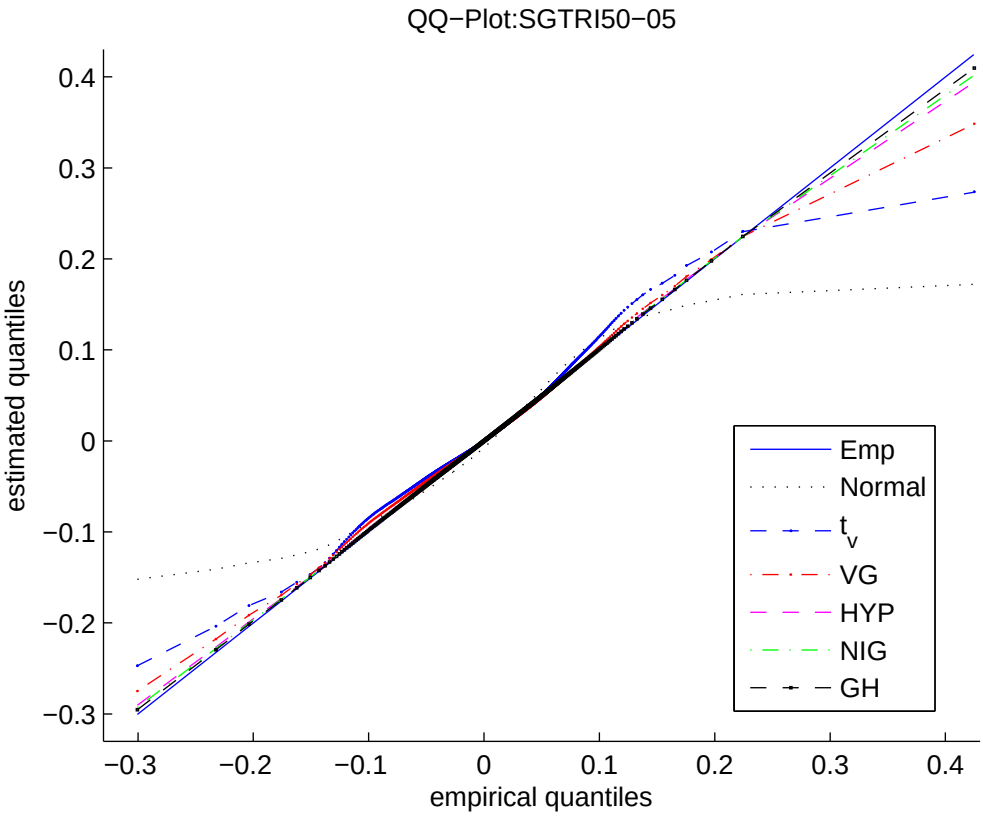

Figure 3.6: QQ plots of SGTRI50-05

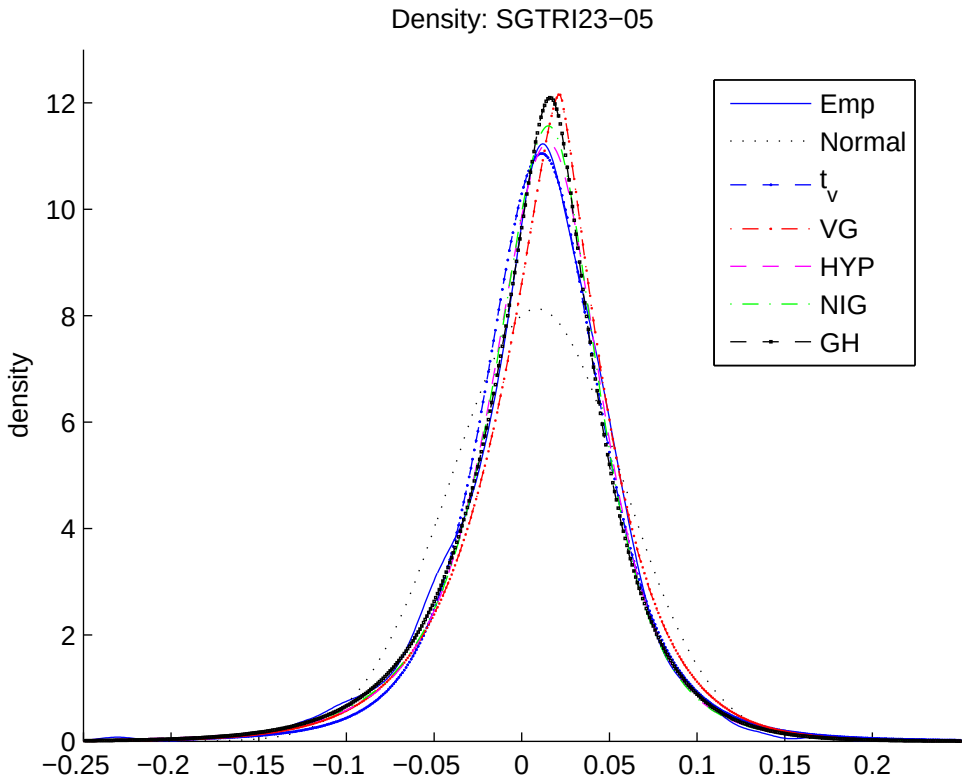

Figure 3.7: Density of SGTRI23-05

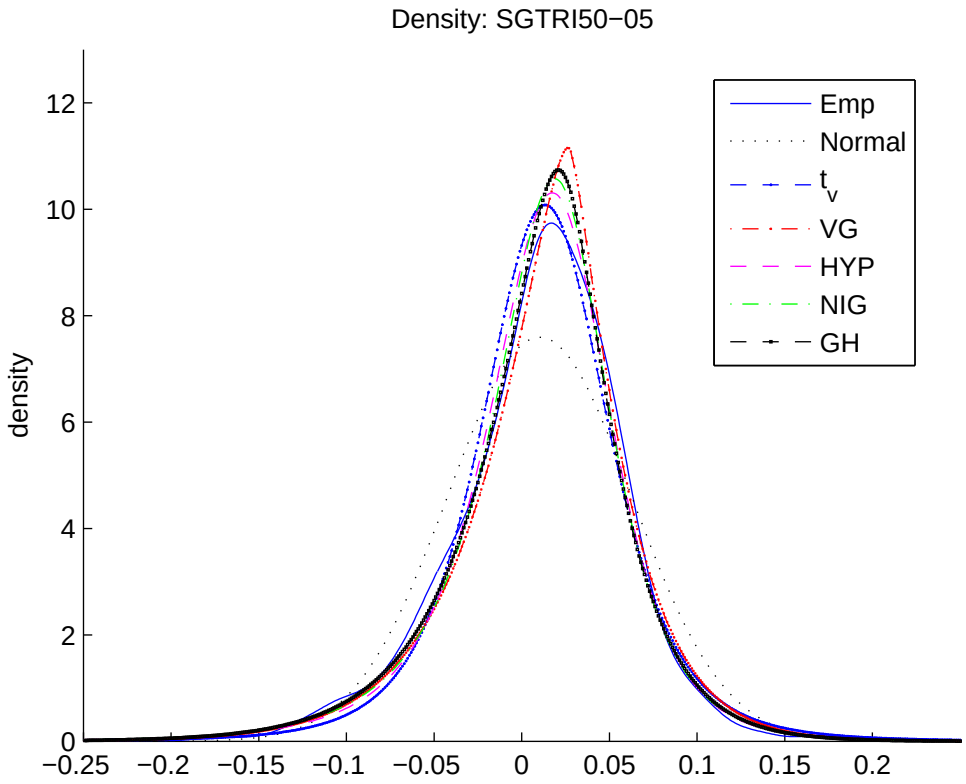

Figure 3.8: Density of SGTRI50-05

$\pi(\theta)$. Denote x as the sample observations used to simulate the posterior sample under the MHA. The marginal density of the possible future outcome z is

$$f(z|x) = \int_{\theta} f(z|\theta, x)\pi(\theta|x)d\theta \tag{3.21}$$

where $f(\cdot|\cdot)$ is the conditional density of z. The predictive distribution can be written as

$$\mathbb{E}_{\theta|x}[f(z|\theta, x)]. \tag{3.22}$$

It is different from using the mean of the posterior distribution as a point estimate for θ:

$$f(z|\mathbb{E}[\theta|x]). \tag{3.23}$$

Using predictive distributions in the simulation leads to the so-called a "hyper-model", see for instance, in Wilkie *et al.* (2005). For each simulation, the values of the parameters to be used are drawn from the posterior distribution. The explicit form of the multivariate distributions of the parameters are not necessary in this case because the sample generated by the MHA can be used as an approximation.

3.6 Concluding Remarks

This chapter estimated the parameter of the exponential Lévy models using likelihood-based estimation approaches. The GH increment Lévy processes are fitted to SGTRI data, both for maximum likelihood estimation and the Bayesian estimation. The results of MLE estimation show a significant improvement in performance of model based on the likelihood function. The five parameter GH case has the highest likelihood but the difference between general case and its subclasses are relatively small.

Markov Chain Monte Carlo has been proposed as an alternative approach to estimate the parameters of Lévy driven models, specifically the Metropolis-Hasting Algorithm. The MCMC estimation results for GH are closed to those estimated by MLE. In addition hyper-models using Bayesian parameters accounts for estimation risk.

The Anderson & Darling statistic is calculated to give a goodness of fit in sense of distance measure. Also we discussed the information criteria in order to see the efficiency of using more complex models.

Chapter 4

Measuring the risks for Unitised With-Profits policies

4.1 Introduction and objectives

As long term investment and saving products, the Unitised With-Profits (UWP) policies make up a significant part of the UK's life insurance market. It allows the policyholders to share the benefits of the life office fund while a guaranteed minimum return is provided by the insurer to protect the investors (policyholders) against the financial investment risk.

We consider the key factors in modelling, reserving and valuation of UWP policies. Special attention will be paid to measuring the financial risk in UWP policies.

In this thesis UWP policies are modelled by providing the operation rules which include asset allocation, charging, reserving and bonus declaration. In practice every insurer in UK issues their own with-profits guide to explain how they operate the with-profits products. A Principals and Practices of Financial Management (PPFM) clarifies, briefly, where and how the fund will be invested. The PPFM is designed to be accessible for all policyholders. In addition, a series of periodic reports review the investment performance and asset allocation strategy analysis of the with-profits fund along with the bonus declaration. Before purchasing with-profits contracts the investors should be informed of the guaranteed minimum return rate and how they will be charged.

The with-profits or other financial investment products contain a variety of guarantees to provide protection against risk. Thus the unitised With-Profits are in the business line of offering policyholder risk management over the long term. Some of the these risks include for example, the mortality risk can be pooled and diversified by selling large amount of policies, whilst some other, such as investment risks, cannot. We consider the major risk embedded in with-profits business is the guarantee risk which caused by the mismatching of the asset and the guarantee.

The financial risks include for example the fluctuation of asset, inflation and interest rates. The performance of fund investment depends on the financial market behaviour. The actuarial risks include unexpected mortality, changes to legislation such as taxation, and financial accounting principals.

The aim of this chapter is to build a mathematical model of the participating life insurance contract, more specifically, the unitised with-profits policy. The financial risks will be measured by some random variables based on the loss distributions of the insurers. A ten year single premium UWP policy will be studied by Monte Carlo simulation experiments. Some of the exp-Lévy models fitted to SGTRI23 will be used as more realistic models to simulate the investment returns in order to observe more accurate risk measures of the maturity loss of the insurer. The effect of the estimation risk of the models is discussed by considering parameter uncertainty in hyper-models.

The chapter is organised as follows. In Section 4.2 we describe liability model which is the unitised with-profits policy. Then we introduce the risk measures we use to quantify the risk embedded in the policy in section 4.3 and present the numerical results of the simulation studies in section 4.4. Finally in Section 4.5 a summary is given.

4.2 Modelling UWP policies

The UWP policies models are characterised by the business activities of the life office by determining investment strategy, bonus rate, charging, guarantee rate and reserves rules of the contracts.

At the outset of the UWP contract, the policyholder pays a premium P to purchase units along with a initial guaranteed value. Charges are normally deducted through bid-offer spread, policy fee, fund management charge or allocation rate. The insurer invests the premium paid in by the policyholder in the financial market. The actual accounting value of the accumulated assets the policyholder owns in the fund is called asset share $AS(t)$.

The payout of the UWP policy, which can be at the end of the policy term or after a death of the policyholder prior to maturity, is the policyholder's asset share subject to the value of the units so-called the guarantee $G(t)$, in other words, the payout of the policy at time t is

$$\max\{G(t), AS(t)\},\ t \leq T. \tag{4.1}$$

In case at maturity the value of the asset share $AS(T)$ is greater than the value of the guarantee $G(T)$, the excess amount will be paid back to the policyholder as a terminal bonus TB which is given by,

$$TB = \begin{cases} AS(T) - G(T) & \text{if } AS(T) > G(T) \\ 0 & \text{otherwise,} \end{cases} \tag{4.2}$$

A practical activity for the surrender payment is the insurer may apply a market adjustment to bring down the payment to the asset share that the surrenders actually own in the fund.

On death, in addition to the unit value, the UWP contract may contain a sum assured to cover the early death benefits. The policyholder will receive the sum assured if it is greater than both guarantee and asset share. The sum assured is not payable at maturity.

During the term of the policy, the reversionary bonus b_t with a guaranteed minimum rate of bonus g is declared periodically to bring up the guarantee value or unit value $G(t)$. From time to time, the insurer changes the bonus rate depending on the investment performance of the reference portfolios. Normally the insurer declares bonus b_k at start of every policy year k and adjusts the bonus rate yearly. The unit value (guarantee) increases at the bonus rate on a daily basis. The $G(k)$ at k integer years after the issue date is:

$$G(k) = (G(k-1) + CF(k-1))(1 + b_k) \ \text{ with } \ b_k \geq g, \tag{4.3}$$

where CF(k) is cash-flow at time k. This can be the amount of premium paid-in P_k less the deducted charges CH_k at the beginning of policy year k,

$$CF(k) = P_k - CH_k. \tag{4.4}$$

More specifically, $G(t)$ is given by

$$G(t) = G(0) \prod_{i=1}^{\lfloor t \rfloor + 1} (1 + b_i) + \sum_{j=1}^{\lfloor t \rfloor + 1} CF(j) \prod_{k=j}^{\lfloor t \rfloor + 1} (1 + b_k) \quad \text{with} \quad b_k \geq g, \tag{4.5}$$

where b_i is the bonus rate declared at the start of i-th policy year since the policy was set up.

The insurer may declare the bonus more frequently, for example half yearly, quarterly or monthly. In practice, premiums are often paid monthly and the charges would probably be deducted at more frequent intervals. More generally, $G(t)$ can be considered as a function that increases at the rate combining the compound guarantee g and the continuous reversionary bonuses:

$$G(t) = G(0) \exp\left\{\int_0^t b_s ds\right\} + \int_0^t CF(s) \cdot \exp\left\{\int_s^t b_u du\right\} ds \text{ with } \quad b_t \geq g. \tag{4.6}$$

The Equation 4.6 shows that the reversionary bonus is an important factor to control the guarantee. The philosophy of bonus strategies is discussed carefully in Willder (2004) and Chadburn (1997). In summary, bonus should:

- reflect the performance of the investment;
- change in a smooth way from year to year;
- be competitive with other contracts; and
- maintain adequate solvency.

One of the methods to compute bonus is directly linked to investment return with smoothing. It has also been considered in Grosen and Jorgensen (2002), Haberman *et al.* (2003) and Ballotta (2005). The bonus declared at start of the i-th policy year is the greater of the guaranteed rate g and part of the previous year's return on the underlying asset. In other words

$$b_i = \max\left\{g, \beta\left(\frac{AS_i}{AS_{i-n}} - 1\right)\right\}, \ i = 1, 2, ..., 10. \tag{4.7}$$

We shall use this method in the example in this chapter. More advanced bonus mechanisms will be discussed in Chapter 5.

4.2.1 Investment bonus cycle

The liability model of the maturity guarantees in the unitised with-profits policies have some option-like features. Brennan and Schwartz (1976) and Brennan and Schwartz (1979) were the first to apply the modern option pricing theory to the fair value of unit-linked type policies. Wilkie (1987), Willder (2004), Tong (2004), Wilkie *et al.* (2005), Chadburn (1997), Chadburn and Wright (1999) and Kleinow and Willder (2006) considered the with-profits guarantees as a portfolio of shares and put options. The put options are based on the underlying shares and have the same maturity as the with-profits policy. The unitised with-profits contract is thus modeled as a European "maxi option" (see Wilkie *et al.* (2005)). The value of the contract is the cost of guarantees for the maxi options.

Jorgensen (2001) considers participating life insurance policies with guarantees but with default options which is different from the approach in Wilkie (1987) and in Willder (2004). Actually, in Jorgensen (2001), the liability of the insurer (shareholders) selling the participating policy is a long position in a European call type option. This call option costs the shareholders the share value in the fund for the right to share the benefits of the policyholders. When a default (not shortfall) happens, the shareholders are subject to a limited liability and the policyholders bear the loss. The liability model with default options is also considered in Ballotta (2005), Grosen and Jorgensen (2002) and Haberman *et al.* (2003).

In this thesis, we assume the insurer will not default at maturity of the policy. Under shortfall, the office will pay the guarantee amount to policyholders at maturity and bear the loss. The loss of the insurer may be covered by,

- the charges from policyholders;
- the reserves of the insurer;
- a third party, for example, a reinsurer or a financial institute such as the short position in put options that the insurer buy
- and the profits created by other generations in with-profits fund.

Kleinow and Willder (2006) point out that the so-called traditional case has considered where bonuses are set according to the performance of some exogenously

given reference portfolio. The insurer is unable to change the reference portfolio or the way it is invested. However, in the so-called practical case the bonuses are linked to the performance of with-profits fund investments, which is the asset share of the policyholder. The insurer's decisions on investment strategy will directly effect the bonuses and hence the guarantees. On the other hand, the regular bonus declared may effect the investment strategy, see for example, under the "dynamic EBR" and "dynamic hedging" investment strategies in Chapter 6. This is called the investment-bonus cycle in the practical case.

In the practical case a UWP contract can be regarded as a sequence of embedded put options written on the asset share AS_t. As the value of the units (guarantee) increases, these put options become deeper in-the-money, and so are more valuable. Thus with each bonus declaration, the total value or so-called fair value of the UWP contract increases.

4.3 Risk Measurement

The approach to measure risk for insurance products has changed radically over the last twenty years. Risk measures we use in this thesis are statistics based on the (un)conditional loss distribution include Value-at-Risk (VaR), probability of shortfall (POSF)and conditional tail expectation (CTE).

These risk measures are based on the loss distribution, while some of the classical actuarial approaches are "scenario-based" such as the one used in resilience tests which observe the maximum loss of the life office under some possible future financial scenarios.

The probability of shortfall is the probability that the insurer can not declare terminal bonus at maturity. Under the shortfall, the insurer needs to pay the guaranteed units value to the policyholders and bear the loss. The POSF is given by:

$$\text{POSF} = 1 - \mathbb{P}(TB \geq 0). \tag{4.8}$$

The Value-at-Risk at probability level α, VaR_α, is simply the quantile of the maturity loss distribution. It is the smallest value l such that the probability that

the loss $L = \max(G(T) - AS(T), 0)$ exceeds l is no larger than $(1 - \alpha)$:

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R}, \mathbb{P}(L > l) \leq 1 - \alpha\}. \tag{4.9}$$

The major concern of using VaR is that it has poor aggregation properties. Artzner *et al.* (1999) shows that VaR is incoherent. The axiom of coherence is given, for example, in McNeil *et al.* (2005). Briefly, a coherent risk measure should be translation invariant, sub-additive, positive homogeneitic and monotonic.

To this end, we define the conditional tail expectation, CTE, of the maturity loss distribution. For a continuous loss distribution, the CTE is coherent (see Acerbi and Tasche (2002) and Dowd and Blake (2006)) and is given by:

$$\text{CTE}_\alpha = \mathbb{E}[X|X \geq VaR_\alpha], \tag{4.10}$$

where α is the chosen probability. CTE has also been used as risk measures in Hardy (2003), Hardy (2001) and Wilkie *et al.* (2005). The CTE and VaR are known as the quantile measures and are applied widely in actuarial risk management.

4.4 A model office study: Basic case

Consider a 10-year single premium unitised with-profits policy called M0 with zero guaranteed minimum rate and annual reversionary bonus rate b_t. The bonus b_t depends on the performance of the reference portfolio (asset share) and is declared at the start of every policy year with one third of last years investment return in asset share subject to a minimum of zero in case the fund has a negative investment return in the past year. That is, the guarantee will be increased as around a 33.3 percent of participating rate β in equation 4.7,

$$b_t = \max\left\{0.33\left(\log\frac{AS_t}{AS_{t-1}}\right), 0\right\},\ t = 0, 1, \cdots 9. \tag{4.11}$$

We ignore mortality and deduct a one percent guarantee charge on the asset share at start of every year except for the maturity. The whole asset share is invested in shares. Under these assumptions, we use the Monte Carlo simulation to project the maturity payout with three different asset models: GBM, exp-NIG and the exp-VG models. These models are calibrated to the monthly index log-return rate, SGTRI data from 1923 to 2005, SGTRI23.

4.4.1 Algorithms

The simulation study requires that the computer codes generate independent uniform, normal, gamma and inverse Gaussian variates. For the long term, more than 10 years time horizons, the maturity asset share and guarantee are the product of hundreds of stochastic monthly random return rate variates. The results are heavily dependent on the quality of the random number series generated by some algorithm. For example, the results of risk measures which will be observed in the next section are sensitive to the distribution and autocorrelation structure of the random numbers. Failure to take this into account in the algorithm will lead to large systematic errors.

To this end, after careful comparison of various random number generating syntax, we use the generators in our Monte Carlo studies as follows. Uniform variates are generated using ran2 from Press *et al.* (2002), as, thus good performance in random number quality tests. For tests of random numbers we refer to Gentle (2003). The inverse normal cumulative density function is calculated by Hastings' approximation in Glasserman (2004). The incomplete gamma function value is computed using algorithm gcf in Press *et al.* (2002). The advantage of generating the univariate normal and gamma variable using inverse CDF methods is that they are tractable and share the good quality of the ran2 uniform random number, such as quick $\mathcal{L}^p$ convergence with $p > 1$ which is the key factor of some risk measures. The inverse Gaussian variate is generated directly by Michael, Schucany and Haas (MHS) algorithm, see Schoutens (2003), Chapter 8. The VG and NIG processes are generated by Brownian time changing skills.

4.4.2 Using MLE estimators

Table 4.2 shows the statistics and risk measures of the UWP payoff under the GBM, exp-VG and exp-NIG models by $100,000$ Monte Carlo studies. The parameters used in the Monte Carlo simulation of model office M0 is summarized in Table 4.1.

First of all we look at the probability of shortfall, for the GBM case, there is 7.22% probability that the actual asset share will fall below the guaranteed level at maturity, in this case the insurer will pay the guarantee and bear the loss which

Table 4.1: Parameters for simulation

Number of simulations	100,000
Term of the policy	10 years
Single premium at t=0	100
Guarantee rate g	0
Participating rate β	33.3%
Charge	Guarantee charge 1% per annum
Parameters in GBM	
Mean share return, $\mu - 1/2\sigma^2$	0.0087
Standard deviation of share return, σ	0.0491
Parameters in exp-NIG	see table 3.1
Parameters in exp-VG	see table 3.3

is the difference between the actual asset share and the guarantee amount. In this example, the figure seems larger than an insurer would be comfortable with because we suppose a constant 100 percent equity backing ratio (EBR) which is a highly risky investment strategy in practice. The UWP under exponential variance gamma real world model assumption has 1.8 percent higher probability of short fall than the GBM. The POSF under exponential normal inverse Gaussian model is the highest in these models. The same pattern applies for the mean of loss (MOL). The maximum loss together with value at risk at two levels show the tail behaviour of the loss distribution which has the heaviest tail under exp-NIG and has heavier tail under exp-VG than the one under GBM. These results are not surprising and we conclude that more accurate modelling by using Lévy processes leads to higher value of the estimated potential risks of the UWP contract hence higher reserves are required.

Table 4.2: UWP simulation using exponential VG, exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
Loss (GBM)	0.0722	19.1296	9.1592	36.0618	25.6835	45.6409
Loss (exp-VG)	0.0902	21.2661	15.0627	42.0876	32.3346	52.9947
Loss (exp-NIG)	0.0946	21.4323	16.7054	44.1215	33.4324	53.7207

Table 4.3 shows that the asset share at maturity under GBM model, A_T(GBM), exp-VG model, A_T(exp-VG) and exp-NIG model, A_T(exp-NIG), have similar mean and standard deviation (SD). As the VG and the NIG laws have finite second order

Table 4.3: UWP simulation using exponential VG model and GBM model.

	MEAN	SD	L-Kurtosis
A_T(GBM)	291.8689	166.5314	3.0173
A_T(exp-VG)	296.6749	165.3233	3.0226
A_T(exp-NIG)	288.3072	161.2154	3.0181
G_T(GBM)	145.6820	23.9722	3.1022
G_T(exp-VG)	147.0211	23.8659	3.0799
G_T(exp-NIG)	141.1001	22.5542	3.0698

moments, in the long run, the i.i.d. assumption for the log-return process leads the sum of the generalized-hyperbolic-law increments to be normally distributed. It's easy to see this from the L-Kurtosis, which is the kurtosis of the logarithm process, is close to 3. Hence the asset values at maturity under the three models are all, approximately, normally distributed.

In this example, the marginal distribution of the guarantee G_T is observable. The $\log -G_T$ is sum of i.i.d one-side Lévy process with the scaled density same as the asset model's BDLPs truncated at point zero. For example, if the return of the asset share is modelled by a variance gamma process, with density function $\mathrm{vg}(x)$, then the cumulative probability function for the log guarantee process is:

$$g(x) = \frac{1}{p}\int_{-\infty}^{x} \mathrm{vg}(s/p)ds,\ x \geq 0,$$

where p is the percentage bonus calculation weight which is 0.33 in our example. Of course in this particular case $\log -G_T$ is normally distributed. In general, the probability distribution of G_T is dependent on the methodology the insurer used to decide the bonuses.

The advantage of this aggregation normality property of variance gamma process and normal inverse Gaussian process is obvious. It will save the calculation power when projecting the long run future distributions. For example, the insurer who set the reserve or declare bonus as quantile of projected future payoff distribution, even under non-Gaussian return assumption, can use approximately all the well-studied statistical properties of the normal distribution.

For VG and NIG processes, the stochastic bridging skills have been studied and the closed form conditional bridging distributions for intermediate points of the

paths have been derived in, for example, Webber and Ribeiro (2004) and Webber and Ribeiro (2003).

A 2-dimensional copula $C(u_1,\ u_2)$ is a joint cumulative distribution function for random variables $U_1,\ U_2$ in the unit hypercube with uniform marginals,

$$C(u_1, u_2) = \mathbb{P}(U_1 \leq u_1,\ U_2 \leq u_2), \tag{4.12}$$

where $U_1,\ U_2$ have uniform distribution on $[0, 1]$.

The 2-dimensional random vactor (AS_T, G_T) have joint distribution $F(x_1, x_2)$ and marginal distributions $F_1(x)$ and $F_2(x)$. AS_T is the maturity asset share and G_T is the corresponding guarantee amount. Given (AS_T, G_T), (U_1, U_2) is defined as

$$U_1 = F_1(AS_T);\ U_2 = F_2(G_T).$$

Hence, the U_1 and U_2 have uniform distributions on $[0, 1]$.

The paired sample $(AS_T(i), G_T(i))$ is the asset share and guarantee of the i-th outcome scenario generating by Monte Carlo method. The marginal sample probability functions are given by

$$F_A(x) = \frac{\#\{AS_T(i) \leq x,\ i = 1, 2, \ldots, n\}}{n}; \text{and} \tag{4.13}$$

$$F_G(y) = \frac{\#\{G_T(i) \leq x,\ i = 1, 2, \ldots, n\}}{n}, \tag{4.14}$$

where $\#\{\cdot\}$ is the counting measure and n is the sample size.

We present the results in the scatter plot in figure 4.1. The copula $C(F_A(AS_T(i)), F_G(G_T(i)))$ for GBM and exp-NIG both have Gaussian marginals, where F is the sample cumulative density. However from the figure 4.1, we can see that the exp-NIG copula which are the right two graphs, have points spread wider than the left two which are the GBM results. The top left corner of the two graphs in the first row is of interest to us because it is the extreme situation when the asset share is in it's lowest level whilst the corresponding guarantee is in it's highest level. Hence, this scenario represents the policies on which the insurer makes the biggest loss. Zooming in the top left area of the first row's two graphs in the second row, there are no points which appear in this "ruin corner" for GBM outcomes when all of them lie in the bottom right square. More appear in NIG outcomes and all

of them have either higher rank of guarantee or lower rank of asset share than the GBM points.

We argue that although the marginal distributions of the maturity asset share AS_T and guarantee G_T are both approximately log-normal with close statistics under the GBM model and one of the Lévy driven asset model, exp-NIG, the 2-dimensional copulas with transformed margins show the very different dependence structure between the simulation outcomes of these two asset models. In this case GBM understates the size of these extreme losses.

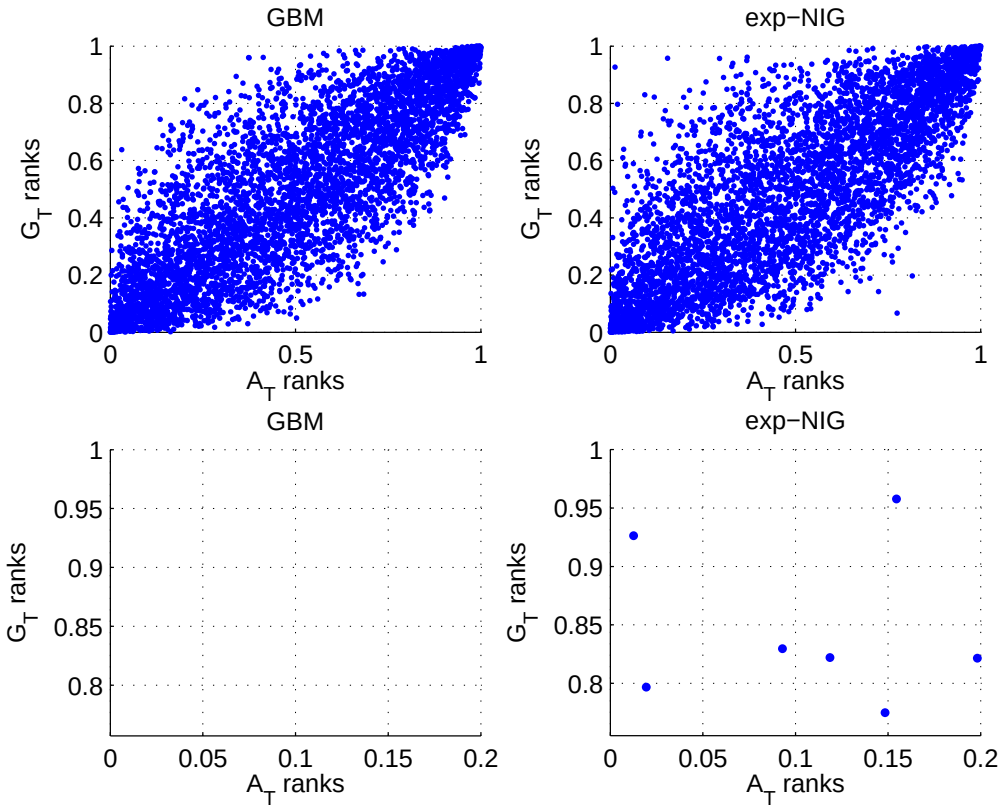

Figure 4.1: Scatter-plot of maturity asset share versus guarantee under GBM and exp-NIG

4.4.3 Using MCMC estimators

So far the parameters of the real world models are assumed to be the MLE estimators and fixed. Under Bayesian estimation, parameters are correlated random variables. Metroplis-Hasting Algorithm is used to generate the quasi-sample observations of

the parameter vector from its multivariate joint distribution. We therefore upgrade parameters of the asset models in every simulation scenario using the sample Markov chain of parameters generated by the MCMC procedure. In every single simulated sample path of UWP asset, the parameters are fixed. For the exp-Lévy models this retains the i.i.d. property of the log return process.

Table 4.4: UWP simulation using exponential VG, exponential NIG and GBM models with MCMC estimators.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
Loss (GBM)	0.0901	21.6498	16.1833	42.8258	32.8897	52.8821
Loss (exp-VG)	0.0966	22.4610	18.4376	45.6333	35.6074	56.0810
Loss (exp-NIG)	0.1175	24.5999	25.2323	53.2655	42.5320	64.1950

We re-consider the model office M0, but with "hyper-" real world models for the asset returns. The Table 4.4 shows the risk measures of the UWP payoff under the hyper GBM, hyper exp-VG and hyper exp-NIG models using the MCMC estimators. All risk measures increase compared with the corresponding ones in Table 4.2 for which the asset model parameters are fixed. The reason for this is obvious. Because the hyper-models include the parameter uncertainty which increase the total variance of the random samples.

As we expected, using hypermodels increases the risk measures in the UWP compare to their corresponding basic version of models with fixed parameters. The use of hyper-GBM model adds an extra 1.79% probability to the figures that simulated using MLE parameters that the terminal bonus is unaffordable. It also substantially increase the mean of loss and other quantile risk measures. The difference of corresponding diagonals in Table 4.4 and in Table 4.2 becomes wider when applying the hyper exp-VG and the hyper exp-NIG. The use of hyper-exp-NIG model adds the largest margins to the corresponding risk measures listed in Table 4.2.

The life office using exp-NIG as real world asset model in measuring the risks for M0 contract will find an additional three percent POSF using MCMC parameters than using MLE parameters. What is more, the quantile measures, both VaRs and CTEs considered in this example, are highly sensitive to the parameter uncertainty assumption.

It can be seen that the uncertainty of the model parameters brings in extra risks. Table 4.5 shows the statistics of the maturity asset shares and guarantees under hypermodels. The maturity asset share under the GBM model, in both MLE parameter approach and the MCMC approach, has highest standard deviation (SD).

Table 4.5: UWP simulation using exponential VG model and GBM model.

	MEAN	SD	L-Kurtosis
A_T(GBM)	300.8993	186.2175	3.0173
A_T(exp-VG)	299.2465	179.1106	3.0226
A_T(exp-NIG)	287.6310	176.9025	3.0089
G_T(GBM)	143.6861	24.1630	3.1022
G_T(exp-VG)	143.6121	23.1654	3.0799
G_T(exp-NIG)	141.9561	22.8366	3.0698

4.5 Concluding Remarks

In this chapter, stochastic asset models have been constructed to study the financial risks of unitised with-profits policies. A single premium model office contract with a 100% EBR and bonuses which are linked to the rate of return has been studied by simulation. Risk measures include probability of shortfall (POSF), mean of loss (MOL), value at risk (VaR) and conditional tail expectations (CTE) are calculated using Monte Carlo simulation under three different asset models. The exp-Lévy models driven by VG and NIG processes which have been proved to be more realistic in last Chapter give higher risk measures.

The hyper models which use the generated sample of Metroplis-Hasting Algorithm show an extra margin of risk measures for all three models while taking the parameter uncertainty into consideration. And the difference between the results of models using Bayesian estimators and MLEs quantifies the estimation risk embedded in using the stochastic investment models. For example, if the stochastic model we are using is highly sensitive to the parameters, it tend to generate large margins of risk measures when using Bayesian estimators instead of MLE.

Chapter 5

Dynamic Bonuses

5.1 Introduction and objectives

In this chapter some more advanced and more complex dynamic bonus mechanisms will be discussed, rather than the simple bonus strategy considered in M0 which is directly linked to the reference portfolio investment returns.

In reality, there is the conflict of interest between policyholders and shareholders. The former group expect high guarantees with their with-profits policies while the later group hope to control the guarantees at a relatively low level.

Briefly, we consider two types of dynamic bonus mechanisms here. One of them is called retrospective bonus strategy. These bonus mechanisms link the reversionary bonus rates to the performance of past investment returns such as the one we used in case M0 in the last chapter. Haberman *et al.* (2003) gives three example bonus schemes to calculate retrospective bonus rates. The first one is calculated as a fixed percentage of the average return rate during the smoothing period, normally the last n years, subjected to a annual guaranteed rate. The second one is based on the geometric average of the last n years returns on the reference portfolio. The third scheme apply a smoothing factor on the bonus which is linked to the reference portfolio returns. Chadburn (1997) and Chadburn and Wright (1999) consider the retrospective bonus strategies similar to the first scheme in Haberman *et al.* (2003) but assume the bonus is linked to the return on consols.

The other type of dynamic bonus mechanisms is called prospective bonus strategy.

The idea of these bonus mechanisms is that we project the future investment return of the asset share or in some cases the reference portfolio if they are not the same, use the projected value as a reference of the target maturity guarantee and the future bonus to build the guarantee.

The bonus earning power (BEP) is the bonus rate that can be declared in the remaining term of the policy in order to lead the guarantee to achieve the objective value. The target guarantee could be any form related to the projected asset share, for instance, it can be risk measures such as VaRs and CTEs or the guarantee will allow a probability of at least a terminal bonus cushion.

In practice, bonuses are declared annually. With the development of modern computer power, it is possible to re-valuate the policies on more frequent bases and re-adjust the bonus rate or at least the suggested bonus rate at any time during a fiscal year. The advantage of this is obvious, the insurer can cut, for example, the monthly bonus earlier than an annual bonus when the investment returns are poor. The re-adjusted bonus rates in a year are probably difficult to be applied on the guarantee based on the practical problem of the with-profits business, i.e., it's too expensive for informing, accounting and reporting of the new bonus changing. However, the suggested bonus rate can give risk managers/insurers a warning of potential probability of insolvency. In this case a possibly more prudent strategies can be taken on the investment, charging and the later bonus declarations.

This chapter is organised as following: In Section 5.2 we consider different retrospective bonus strategies based on the M0 model by varying the smoothing period and the participating rate. The prospective bonuses which is bonus earning power with a risk measure target method will be discussed in Section 5.3. Section 5.4 looks at some issues such as the effect of declaring bonus on a more frequent bases and smoothing the bonus rate. Finally, some summary remarks are given in Section 5.5.

5.2 Retrospective bonus mechanism (Case M1)

Consider a unitised with-profits policy called M1. We use the same assumptions as in Case M0 (see Section 4.4) for the policy and market parameters. The bonus will

be calculated as the weighted arithmetic average of the smoothing period, which is the last n years returns on the asset share, so that

$$b(t) = \max\left\{g, \frac{p}{n}\log\left(\frac{AS(t)}{AS(t-n)}\right)\right\}, \tag{5.1}$$

where $p \in (0,1)$ denotes the participating rate. Thus, the M0 is a special case of M1 when $p = 0.33$ and $n = 1$.

At this stage, the exp-NIG model is chosen to represent the non-Gaussian Lévy driven models as the real world model to compare with the geometric Brownian motion model. It can be seen from Chapter 3, the NIG distribution offers the second highest likelihood among all increment laws considered and is only slightly less than the general 5-parameter GH distribution. It fits accurately to both SGTRI23 and SGTRI50 data sets while it has some attractive mathematical properties. These properties include the closed-under-convolution feature which allow us to extend the model to any time scales and keep the increment distribution in the NIG class. Also the inverse Gaussian random variates can be generated directly by using the MHS algorithm. It reduces the systematic errors caused by the generator algorithm and the numerical approximation in long term model study.

Here we focus on the comparison between the non-Gaussian Lévy processes and the Brownian motion under same stucture of asset model. We also considered the Lévy driven processes other than NIG such as VG, GH and HYP. The results suggest the difference between these models within GH family is insignificant and thus it is not necessary to show all the combinations. Comments on results of other Lévy driven models are given at the end of this chapter.

First we look at risk measures of the loss for M1 for different participating rates p.

Figure 5.1 shows the probability of shortfall (POSF) plotted against the participating rate p, using the range from 0% to 50% with three smoothing period values which are one year, two year and three year. The solid lines are the POSF measure refer to GBM models and the broken lines are the POSF under the exp-NIG models. As we expected, the POSF under exp-NIG case is greater than the one under GBM case with same participating rate p and smoothing period n. Under each level of n and the asset models, the POSF is an increasing function of the participating rate,

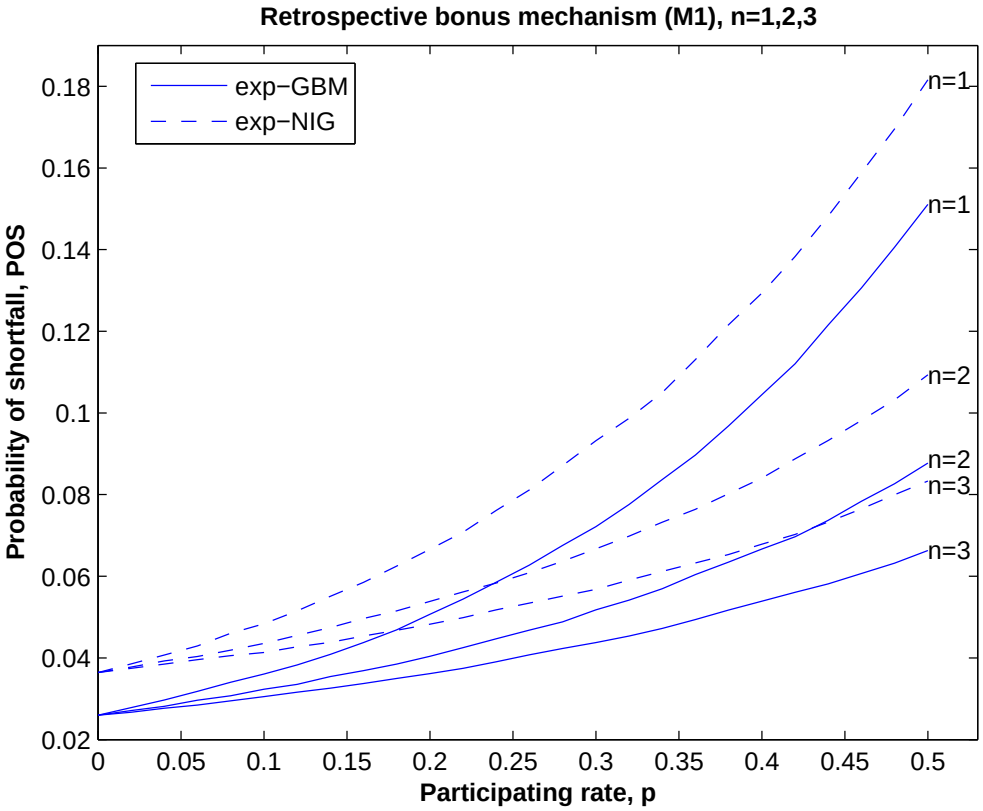

Figure 5.1: Probability of shortfall for office M1

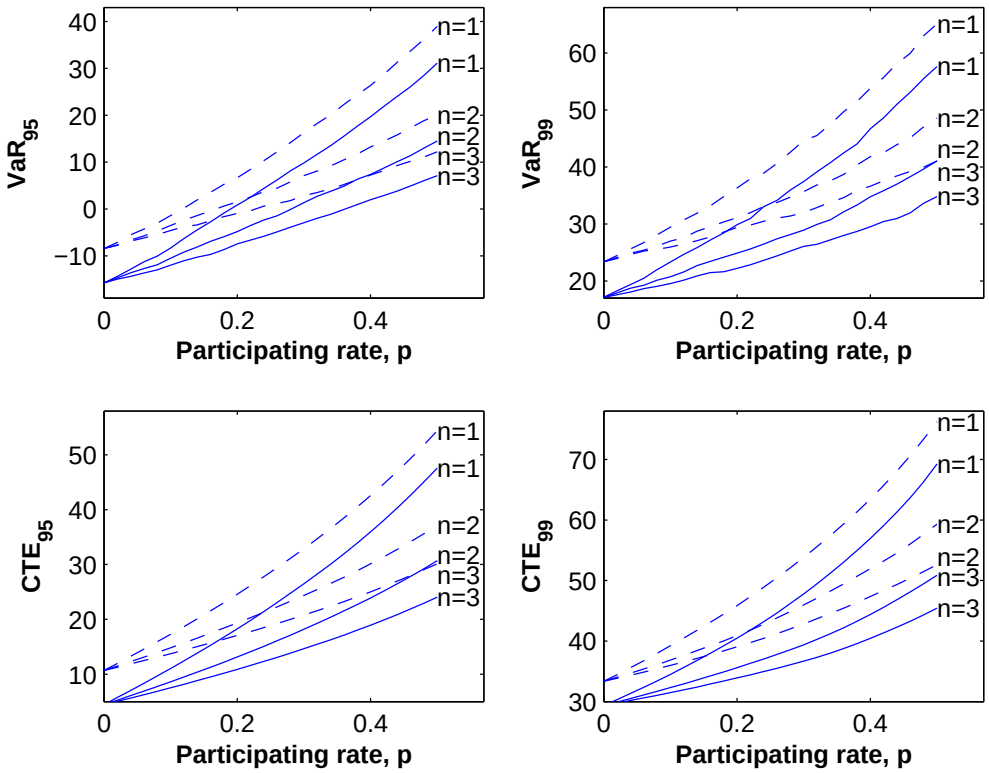

Figure 5.2: Quantile risk measures for office M1

p. This is simply because the higher participating rate will cause a bigger bonus rate in Equation 5.1. It eventually leads to higher maturity guarantee under any investment return scenario.

Then we consider the effect on the probability of shortfall of the UWP policies with different lengths of the smoothing period n. Under the same asset models, POSF is a decreasing function of n. In fact, extending the averaging period n reduces the volatility of the bonus rate in Equation 5.1 , thus the volatility of the maturity guarantee.

Figure 5.2 shows the sensitivity of the quantile measures to the participating rate with different smoothing period n under the two asset models. We observe that the quantile risk measures, which include VAR95, VAR99, CTE95 and CTE99, have the same pattern as the POSF. It is clear that the usual risk measure style which has been found in previous examples between the two asset models apply for the quantile measures. Similar to what is observed in POSF, the risk measures can be reduced by either a rise of smoothing period n or a drop of the participating rate p.

It is worth noting that the VaR lines in the upper two graphs in Figure 5.2 are not as smooth as the ones of CTE shown in lower two figures. This is because in Monte Carlo simulation, the risk measures are estimated by statistics of the sample observations generated from the loss distribution. Since the law of large number, the variance of CTE_α is smaller than the variance of VaR_α.

5.3 Prospective bonus mechanism

In this section, we consider the prospective bonus mechanism specifically the bonus earning power (BEP).

It has also been considered in several different forms in Tong (2004), Willder (2004) and Forfar *et al.* (1989).

The bonus earning power used in this thesis is defined as the constant bonus rate that can be declared in each future bonus declaration date in order to make the current guarantee achieving the projected value of guarantee at the maturity time. This projected maturity guarantee can be any form and linked to the asset share.

In this thesis we will consider three of them:

- The projected maturity guarantee is a percentage quantity of the estimated accumulated asset share AS_T. For instance, Hare *et al.* (2003) project the maturity guarantee as 75% of the estimated asset share value.

- The projected maturity guarantee gives the quantile of the projected asset share distribution at time T. For instance, the projected maturity guarantee can be the 75% quantile of the maturity asset share.

- The projected guarantee value gives certain quantile risk measure for the policies at the maturity. For instance, the BEP can be calculated by making the CTE at 99% level equals to the support capital of the life office.

5.3.1 Bonus earning power (Case M2)

We now give the numerical example of the first bonus earning power method listed above in Case M2. The other two will be discussed later in this chapter. The whole policyholder's fund is assumed to be invested in equities. In order to estimate the asset share at time T, the asset models are required to give the future investment returns. As in model office Case M1, the total return of the equity investment is modelled as a GBM model and one of the Lévy increment exponential model which is exp-NIG. The later model gives a more accurate fitting to the real world data. We adopt the notation used in previous chapters and include some new notation as following,

- $A'_T(t)$: the projected value of the maturity asset share $AS(T)$ at time t, $0 \leq t < T$.

- $G'_T(t)$: the projected value of the maturity guarantee $G(T)$ at time t, $0 \leq t < T$.

- $b'(t)$: the bonus earning power at time t.

- CF_t: the cash-flow at time t.

- TB: the target terminal bonus rate.

First of all the bonus earning power mechanism we considered is using 75% of the estimated asset share as the target guarantee. A regular bonus is declared at the start of each policy year except at maturity. The projected maturity asset share is calculated as the current asset share value $AS(t)$ and any future cash flows increase at investment return rate r, thus the $A'_T(t)$ is given by

$$A'_T(t) = AS(t)\exp\{r(T-t)\} + \sum_{s\in(t,T]} CF_s \exp\{r(T-s)\}. \tag{5.2}$$

The projected guarantee $G'_T(t)$ is given by

$$G'_T(t) = G(t)\exp\{b'(t)(T-t)\}. \tag{5.3}$$

Thus the bonus earning power can be calculated at any time $t,\ 0 \le t < T$, by solving the equation:

$$(1+TB)G'_T(t) = A'_T(t). \tag{5.4}$$

We assume a 33.3% terminal bonus cushion which equivalently makes the $G'_T(t)$ equal to 75% of the projected asset share $A'_T(t)$.

Investment returns were generated by the GBM and exp-NIG model using SGTRI23 parameters. We ignore mortality and surrenders. Thus in this example, the cash-flow CF_S includes only the assumed 1% guarantee charge at start of every year over the policy term.

Table 5.1 shows the risk measures using $100,000$ Monte Carlo studies of a 10-year term policy. We assume the average future investment return r to be 6 percent. The figures in this table seem to be similar to those for Case M0 in Table 4.2 except that we only use the exp-NIG model to be the Lévy driven asset model. There is 8.38% POSF that a deficit will happen under the GBM world. The POSF under exp-NIG model is 2.33% percent higher than the POSF under GBM. The same pattern applies for the mean of loss (MOL). The value at risk at two levels shows the tail behaviour of the loss distribution which has the heavier tail under exp-NIG than under GBM. Hence we have achieved the same conclusion as in M1 that more accurate modelling by using Lévy processes will lead to higher potential risks of the UWP contract hence more reserves are needed.

Table 5.1: UWP simulation using exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
Loss (GBM)	0.0838	16.4430	10.1956	33.3298	24.3496	43.7447
Loss (exp-NIG)	0.1071	19.2516	16.7622	41.4130	32.3705	54.1435

In this example, we find the risk measures are close to the retrospective bonus mechanism case M1 with $p = 0.33$ and $n = 1$. It is now of interest to see the distribution of the guarantee at maturity G_T under these two bonus mechanisms.

Table 5.2: UWP simulation using exponential NIG model and GBM model.

	MEAN	SD
A_T(GBM)	292.8253	167.5538
A_T(exp-NIG)	288.6896	160.7953
G_T(Retro-GBM)	145.6820	23.9722
G_T(Retro-NIG)	141.1001	22.5542
G_T(Pro-GBM)	212.6310	107.2892
G_T(Pro-NIG)	205.2620	101.7103

Table 5.2 and Figure 5.3 give the statistical and graphical description of the the simulated maturity guarantee distributions. The statistics of the maturity guarantee value which accumulated by declaring the retrospective bonus as in Equation 5.1 is given in column "G_T(Retro-GBM)" and "G_T(Retro-NIG)" in Monte Carlo using respectively GBM model and exp-NIG model. The columns "G_T(Pro-GBM)" and "G_T(Pro-NIG)" contain the statistics under these two models for the maturity guarantee value which accumulated by declaring the bonus earning power we described in equation 5.2-5.4.

In this stage, we assume 100% constant equity backing ratio (EBR), no investment decisions based on the investment return and bonus are made at intermediate time over the policy term. Thus, in each simulation scenario, the guarantee calculated by different bonus methods have no effect on the asset share sample path. The maturity asset share depends only on the investment returns and there is no "investment bonus cycle" in this example.

The mean and standard deviation of the maturity guarantees of UWP contracts where the life office declares retrospective bonuses are 145.6820, 23.9722 for the GBM world and 141.1001, 22.5542 for the exp-NIG world. The figures seem close between the two asset models because of the so-called "aggregation normality property" of the exp-Lévy models. Recall, we have studied in Chapter 4 the joint distribution with asset share AS_T using the coupled samples. Here we concentrate on the different bonus mechanisms. We can see the prospective bonus mechanism leads to a higher mean and standard deviation maturity distribution of the Monte Carlo sample under both the two asset models we used. It's more clear to see from Figure 5.3. The upper left subplot gives probability density of the G_T under two bonus mechanism when asset return is modelled by SGTRI23 GBM model. The bottom left subplot shows the same distribution when the asset return is modelled by SGTRI23 exp-NIG model. We can see that the distributions of G_T under bonus earning power mechanism spread wider than the the distribution of G_T under retrospective bonus mechanism. Recall we gave the scatter plots of the AS_T ranks versus ranks of the retrospective bonus accumulated G_T in Figure 4.1. The right two sub-plots in Figure 5.3 show the same rank scatter (We used first $5,000$ scenarios of the simulation). The only difference is the G_T has been accumulated under the bonus earning power mechanism in equation 5.2-5.4.

Comparing to the upper two subplots in Figure 4.1, the ranks points in these cases are more likely to lie in the narrow areas near the cross corner line of the unit square space. There are no points in the first $5,000$ simulation results appears in the upper left 20% region of the graph which gives the lowest 20% asset share and the highest 20% guarantee. Thus we argue that the BEP guarantees are more closely correlated to the investment returns. The reason for this is because the BEP is calculated by the current value of the G_t and AS_t which gives all the information of the past investment returns and bonuses declared. And the future investment return is considered by giving a predicted distribution of the maturity asset share. The retrospective bonuses only take into account the investment returns in the past n years time interval. This can be expressed by a simple example, we consider the situation when investment returns are poor in the early years of the policy term.

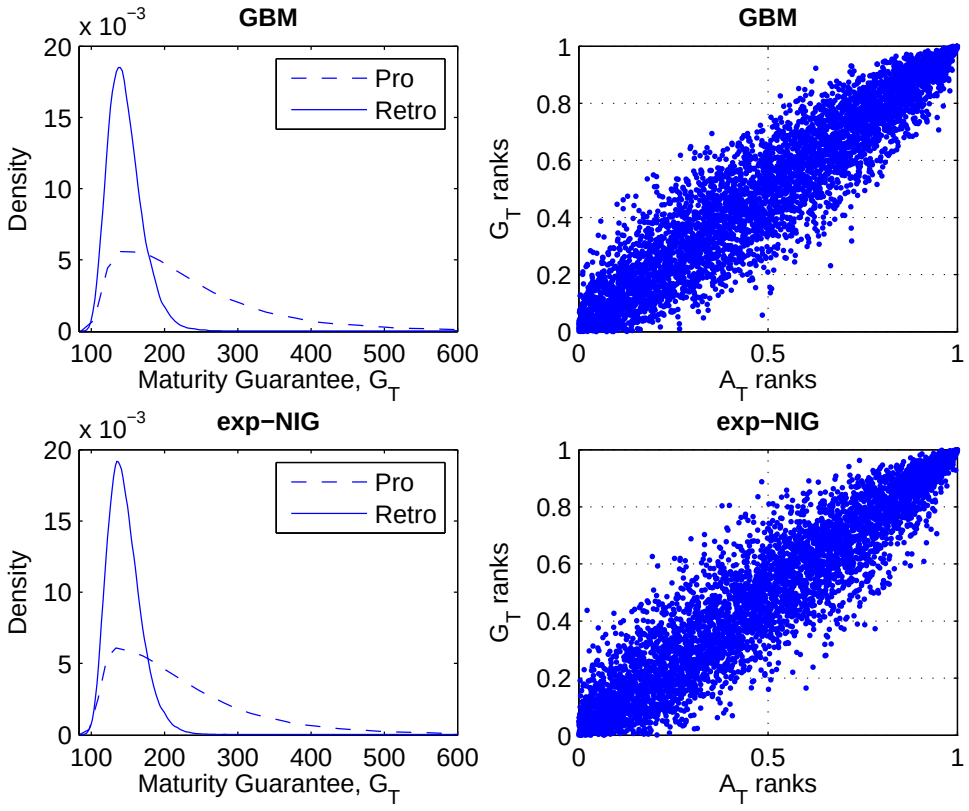

Figure 5.3: Left: density of the maturity guarantee. Right: scatter-plot of maturity asset share versus guarantee under GBM and exp-NIG

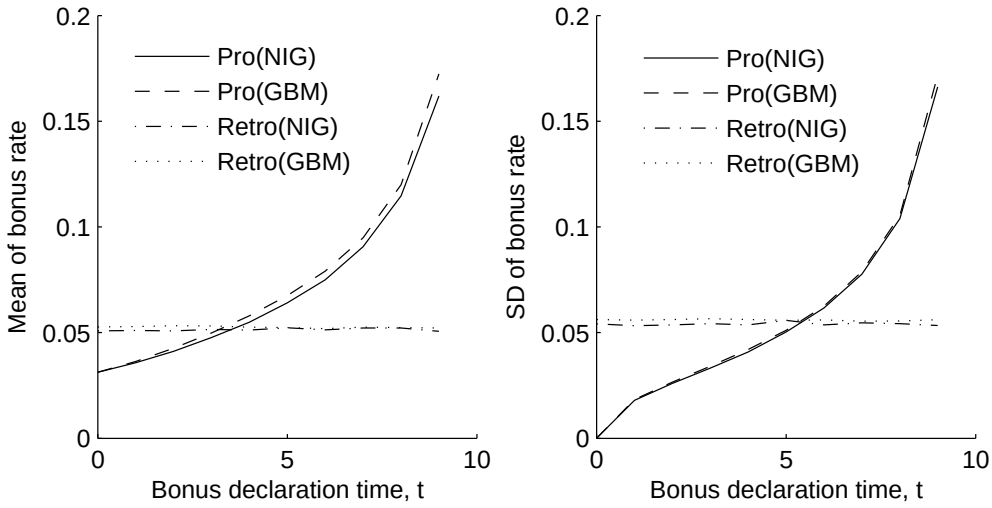

Figure 5.4: Mean and standard deviation of bonus rate

In this case both the retrospective method and the prospective method will try to control the bonus to a relatively low level. Then assets are assumed to have good investment performance in the later years of the term. Under the retrospective bonus rate method, it will start to declare higher bonus than the early years and when the bad returns are past the smoothing period, the method seems to be "memoryless" of the fall of asset share in early years. In contrast, when the investment return is good but the current guarantee is still higher than the current asset share, the projected asset share will be low and this leads to a low bonus earning power.

The left subplot of Figure 5.4 shows the mean of the bonus over the policy term in Monte Carlo simulation. It can be seen that the mean of bonus calculated using retrospective bonus mechanism maintained the same level at every declaration date while the BEP rises sharply from the policy year 1 to policy year 10. The closer to the maturity, the greater bonuses that the BEP method tend to declare. The right subplot of Figure 5.4 shows the corresponding standard deviation of the bonus referred to the mean of bonus. The SD of the "Retro" cases are stable over time, while the SD of the BEP climbed from 0 at start of the policy to about 0.17 for the final bonus at start of year ten. It suggests that the BEP method we used in this example is highly volatile when the policy is close to maturity. A smooth graduation may be desirable applied to decrease the fluctuation of the bonus.

5.3.2 Alternative prospective bonus mechanisms (Case M3)

The difference between the bonus earning power method used in M2 and the other two of the prospective bonus mechanisms we introduced in this thesis is the later ones rely heavily on the internal asset models used by the life office. In the model office M2, the projected asset share at time t within the policy term, $A'_T(t)$, is an estimation of the maturity asset share value A_T.

Tong (2004) introduced a bonus earning power method where the projected asset share $A'_T(t)$ is a random variable and its possible predicted value is generated by the Wilkie model. The projected guarantee is calculated by the 25% lower quantile of $A'_T(t)$ with a 30% terminal bonus target TB. In other words, the bonus earning power is the constant rate that can be declared as each future bonus that gives a minimum 30% terminal bonus with probability 75%. The actual bonus b_t is the greater value between the guarantee rate g and the bonus earning power at time t, b'_t. In the mathematical form

$$b'_t = \log\left(\frac{Q_{0.25}(A'_T(t))}{G_{(t_-)} \cdot (1+TB)}\right) / (T-t)\,, \tag{5.5}$$

where $Q_{0.25}(x)$ is the 25% lower quantile of random variable x and $G_{(t_-)}$ is the guarantee value before the bonus declaration at time t.

We note that under some asset models when the predicted distribution of the investment return is not available in the close form, such as the Wilkie model used in Chadburn and Wright (1999) and Tong (2004), a Monte Carlo study of the maturity asset share is required at each bonus declaration time for every simulation. It makes this method computational expensive.

Alternatively, the simulation time can be reduced by using the exp-Lévy models. The reason is under these asset models, the future return is i.i.d. and for those whose increment law is closed under convolution, such as Normal; variance Gamma and normal inverse Gaussian, the probability functions of the return process r_{T-t} is known and it gives the distribution of the maturity asset shares $A'_T(t)$. Even for the generalized hyperbolic law whose close form of density under convolution is not analytical, according to the "aggregation normality property", a normal distribution calibrated to long term GH law with same mean and variance can be used as an approximation of the return process with longer than one year term.

We look at the same 10-year single premium UWP contract in M2. The bonus earning power is calculated by targeting the 25% lower quantile of $A'_T(t)$ with a 30% terminal bonus target. The 25% lower quantile of a standard normal distribution is -0.6745. Thus, if the yearly return is normally distributed with mean μ and standard deviation σ, the 25% lower quantile for n-year return process is $n\mu - 0.6745\sqrt{n}\sigma$. The $Q_{0.25}(A'_T(t))$ under the GBM asset model is given by

$$Q_{0.25}(A'_T(t)) = A(t)\exp\{\mu(T-t) - 0.6745\sigma\sqrt{T-t}\}. \tag{5.6}$$

For the exp-NIG case, there are different ways to get the quantiles. One of them is using the normal approximation or using Student-t distribution with the estimated parameters for the same data set as a more accurate approximation. These values can be calculated by re-scaling the standard distribution quantiles on statistical tables. However in this example, we first generate $10,000$ return processes under the exp-NIG for the term from one year to ten years. Then calculate the 25% lower quantile of the return processes and use these values in the bonus earning power calculation.

Table 5.3: 25% lower quantile of the return processes using exponential NIG and GBM models.

t	0	1	2	3	4
M3(GBM)	1.7000	1.6022	1.4888	1.3853	1.2858
M3(NIG)	1.6686	1.5609	1.4575	1.3687	1.2803
M2	1.6478	1.5606	1.4912	1.4185	1.3494
	5	6	7	8	9
M3(GBM)	1.2217	1.1303	1.0725	1.0187	0.9786
M3(NIG)	1.2042	1.1286	1.0682	1.0193	0.9793
M2	1.2837	1.2211	1.1616	1.1050	1.0512

Table 5.3 shows ratio $Q_{0.25}(A'_T(t))/A(t)$ at each time t from 0 to 9. As the comparison, we calculate these ratios for M2 bonus mechanism. It is clear that $Q_{0.25}(A'_T(t))/A(t)$ ratios of M3 and M2 are both decreasing functions against time t. In fact the figures show these two methods are similar. In this particular case, the M3 gives the higher ratio in early years of the policy term than M2 ratios and lower from year 2 to year 9. We can expect a more flat mean bonus rate curve in Figure 5.4 than the M2 case. We will not give the numerical example at this

stage because it is close to the M2 case under exp-Lévy asset models. However this method is asset model dependent which will be different from M2 bonus mechanism under more complicated asset models.

In general, the life office can decide what is the projected guarantee given the predicted distribution of maturity asset share. For example, the projected guarantee with a terminal bonus cushion, $G'_T(t)(1+TB)$, discussed in the bonus earning power method of model M3 is actually the value which achieves the fixed probability of shortfall at maturity time given the distribution of the asset share.

The third bonus earning power method we mentioned earlier in this section is constructed in the same way but the projected guarantee is decided by the life office reserve strategy. If a life office set their reserve by using quantile reserve, such as the value at risk (VaR) or conditional tail expectation (CTE). We explain this by giving the example below. If the life office have support capital FA_t (reserve) at time t. Within the policy term, the actuaries simulate the future return of the investment by using some asset models such as exp-Lévy . Thus they have the distribution of the asset share $A'_T(t)$. The loss of the life office is given by the random variable $\max\{0, G'_T(t) - A'_T(t)\}$. The amount of CTE of the maturity loss at confident level 95% is assumed to be backed as the reserve and it is equals to the support capital. Thus the bonus earning power at time t can be derived by solving:

$$CTE_{95}(\text{Projected maturity loss}) = FA_t.$$

Again, when the asset model used by the life office to predict the maturity asset share distribution is complicated, this method leads to heavy computational expense. It provides more flexible forms of the guarantee accumulation. This is above the scope of this thesis, we point out this as an interesting future research topic.

5.4 Frequency of Bonuses

There are many bonus methods proposed in the previous sections. All of them are assumed to be declared annually. In practice, bonuses are often declared annually or half yearly. In this thesis, continuous asset models and the bridge skills (we will discuss the stochastic bridging skills in Chapter 7) applying on the annual asset

model such as the Wilkie model allow us to investigate the effect of more frequent bonus declaration.

5.4.1 Monthly bonus (Case M4 & M5)

In this section we calculate the risk measures for the same UWP policy M5 as in M3, but with the bonuses declared monthly according to the two bonus strategies described in M2 and M3.

Similar to the yearly retrospective bonus mechanism M0, monthly retrospective bonus rates are calculated as p of the average investment returns within the previous m month. It is given by

$$b_m(t) = \max\left\{g/12, \frac{p}{m}\log\left(\frac{AS(t)}{AS(t-m)}\right)\right\},$$

where the unit of time t is re-scaled to one month. To make it consistent to the M2 model, p is 33% is this example and the m is 12 which is one year. So in other words, the retrospective monthly bonus is 33% of the average return of the past 12 month. Denote x_i as the return of the i-th month, the i-th monthly bonus b_i is given by

$$b_i = 1_{\{\sum_{k=i-12}^{i-1} x_k > g\}} \frac{p}{12}\left(\sum_{k=i-12}^{i-1} x_k\right) + 1_{\{\sum_{k=i-12}^{i-1} x_k \leq g\}} \frac{g}{12}. \tag{5.7}$$

The monthly prospective bonus mechanism will be exactly the same as the bonus earning power we described in model M3, except for re-adjusting the bonus on a more frequent basis. The bonus earning power will be calculated at start of every month by making the projected guarantee with a 25% terminal bonus cushion equal to the projected asset share at maturity.

Table 5.4 shows the risk measures for the UWP contracts with monthly bonus by $100,000$ times of Monte Carlo studies.

First we look at the Retrospective cases, which are in the first (Loss (R-GBM)) and second (Loss (R-exp-NIG)) rows of Table 5.4. It is not surprising that the risk measures are higher in the case when the real world model is an exp-NIG than in the GBM case. What is of interest to see here is the question whether it will be a good strategy for the insurer to declare monthly bonus. Recall the model setting is same as in Case M0 except frequency of bonus declaration. The risk measures for M0 were

Table 5.4: UWP simulation using exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
Loss (R-GBM)	0.0762	19.4613	10.6272	37.0227	26.9169	46.5239
Loss (R-exp-NIG)	0.0992	21.2479	17.9615	44.3168	33.8942	53.7920
Loss (P-GBM)	0.0631	16.0590	4.8509	29.4308	19.7153	39.3614
Loss (P-exp-NIG)	0.0862	17.9741	11.7680	36.5725	26.7724	46.5946

given in Table 4.2. Compare the results in first row in Table 5.4 which gives the risk measures for loss under GBM models with those in Table 4.2, the probability of shortfall increases by 0.4% which is not a significant difference. The mean of loss (MOL) which was 19.1296 for yearly case and here it increases to 19.4613. The tail behaviour of the loss function is observed by the two quantiles VaR_{95} and VaR_{99}. The VaR_{95} is around 1.6 more than the yearly case and VaR_{99} is around 1 more than the yearly case. The CTE95 and CTE99 are larger than yearly case M0 but with very small difference which is around 1 in value. The similar pattern for POSF has been found for the exp-NIG model case whose risk measures are given in second row in Table 5.4. But for the quantile risk measures, VaR and CTE, the yearly and monthly bonus strategies seem to be indifferent under the exp-NIG model.

Table 5.5 gives the mean and standard deviation of the maturity guarantee value under the monthly bonus case. Compare the SDs to the yearly case which are listed in Table 5.2, the guarantee with the monthly retrospective bonus methods under both two asset models are less than in the yearly cases with close mean values. This is obvious because during one calender year the monthly bonus random variables $b_i,\ i = 1, 2, ..., 12$ are correlated stationary time series where the yearly bonus rate is sum of 12 copied of the first monthly bonus rate $12 \cdot b_1$. The variance of the later sum is greater than the former sum.

Here we have achieved the conclusion that the monthly bonus declaration using retrospective bonus mechanism under the exp-NIG and GBM models will not improve the risk management of the UWP maturity guarantee.

The distribution of the rate of bonus earning power method is more complicated and we only study it by looking at the numerical results. We can see from Table 5.4 the probability of short fall is 0.0631 compare to 0.0838 for GBM model. It is

Table 5.5: UWP simulation using exponential NIG model and GBM model.

	MEAN	SD
G_T(Retro-GBM)	140.8732	18.9031
G_T(Retro-NIG)	140.4077	18.0746
G_T(Pro-GBM)	226.6042	119.8085
G_T(Pro-NIG)	217.2320	112.5521

reduced by more than two percent by declaring bonus monthly. For exp-NIG the POSF again reduce to 0.0862 from 0.1071. The VaR measures and CTE measures are all less than they were in yearly case. The quantile reserves required for monthly bonus UWP policy will be less than those with yearly bonuses. Thus, we argue that a life office does improve their risk management of the UWP guarantee by declaring BEP bonus on a monthly time lag. Also, we expect a further improvement when life office declare bonus more often than monthly.

We can see from Table 5.5 and Table 5.2, the mean of maturity guarantee is greater in monthly bonus case and with larger standard deviation than in yearly case.

5.4.2 Any frequency (Case M6)

In this section we look at the bonus strategies which can be declared at any time. The bonus earning power method is the same as described in model M5, except the bonus rate is adjusted at different frequency which is k times per year. The result of risk measures is shown below, probability of shortfall, mean of loss, value at risk VAR95, VAR99 and conditional tail expectations CTE95, CTE99:

These five frequencies are chosen to represent the annually (k=1), half yearly (k=2), monthly (k=12), weekly (k=50) and daily (k=250). Under both GBM and exp-NIG models, the more frequent bonus than annually strategies always provide smaller risk measures than annual cases. Amongst the five cases considered with different bonus declaration frequencies, the monthly bonus case has the lowest POSF under both GBM and exp-NIG real world models. The weekly and daily cases have very close risk measures, for the POSF, MOL, VaRs and CTEs. For GBM model,

Table 5.6: Bonus Frequency

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
GBM						
k=1	0.0838	16.4430	10.1956	33.3298	24.3496	43.7447
k=2	0.0731	16.1782	6.9103	31.1236	21.7760	40.8724
k=12	0.0631	16.0590	4.8509	29.4308	19.7153	39.3614
k=50	0.0656	16.0147	5.3534	30.4274	20.3618	39.6910
k=250	0.0656	15.9772	5.3124	30.3851	20.3243	39.6537
exp-NIG						
k=1	0.1071	19.2516	16.7622	41.4130	32.3705	54.1435
k=2	0.0976	18.7025	14.9218	41.1504	29.9964	50.9629
k=12	0.0862	17.9741	11.7680	36.5725	26.7724	46.5946
k=50	0.0874	18.2157	12.8348	38.0720	27.3720	47.3176
k=250	0.0872	18.2045	12.7645	38.0293	27.3192	47.2775

to the four decimal places shown, there is almost no difference in the risk measures between the weekly and daily cases. It shows that the risk measures first reduce from annually to monthly then figures rise in the weekly case and remain stable at the daily case.

Figure 5.5 shows the POSF and CTE95 and CTE99 under the GBM and exp-NIG models against any frequencies from yearly ($k = 1$) to daily ($k = 250$).

We can see from Figure 5.5, that the POSF at first fell rapidly with k increase in the range from 1 to 14. Then the POSF goes up steadily with k increase in the range from 15 to 50. Varying k from 51 to 250 does not lead to any significant difference in POSF of the contract. In Monte Carlo simulation, only risk measures for integer ks are observed, the POSF reaches a minimum at $k = 14$ for both GBM and exp-NIG model.

The CTE95 and CTE99 in the right subplot in Figure 5.5 show similar pattern against k as in POSF. The CTE95 and CTE99 under both asset models reaches the minimum value $k = 14$. It is clear that by declaring more frequent bonus than yearly the insurer get the advantages of all the risk measures we considered in this example. There is a relatively large improvement when k increases from 1 to 14. The $k = 14$ represents a re-adjustment of bonus rate in every 18 trading days.

Nevertheless, from $k = 15$ to $k = 50$, the risk measures increases and level off

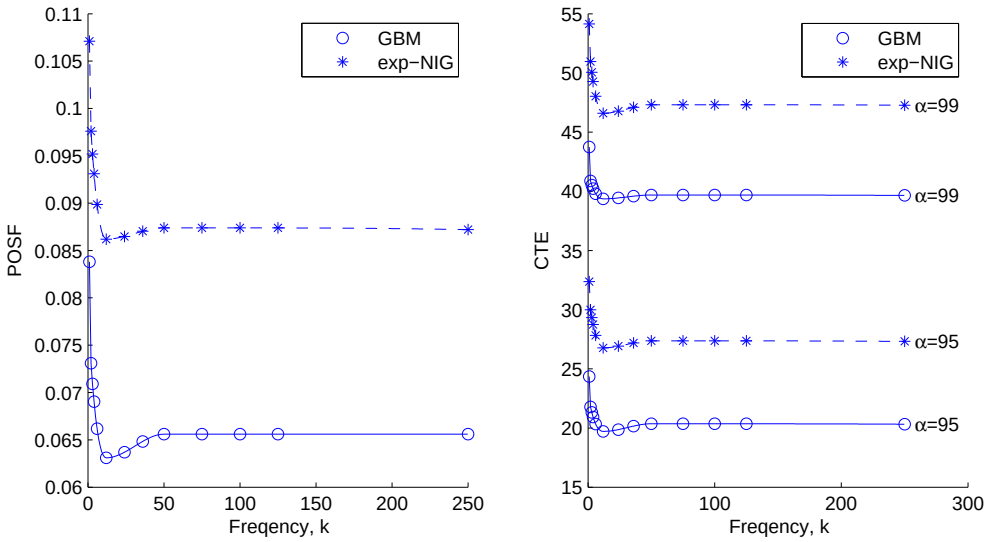

Figure 5.5: Risk measures

onwards. This suggests that, from the risk management point of view, declaring bonus more frequently than monthly may not be worth the expense.

Two Effects

Since in Case M6, a 100 percent EBR investment strategy and a bonus earning power strategy are chosen, two effects arise simultaneously. On the one hand, when a drop appears in value of the reference portfolio, bonus strategy will cut the bonus rate earlier with more frequent bonus than one with yearly bonuses. That reduces the risk of shortfall at maturity. On the other hand, in a particularly good return period, the earlier bonus causes higher guarantees. Moreover, once the bonus had been added to the guarantee it can not be removed, so the two effects are opposite but not symmetric. We are interested in those situations that the asset share is below the guarantee. So the regular bonus mechanism that cut bonus quickly will increase the probability of paying a terminal bonus, the former effect dominates.

An improved approach may be to control the "reflecting" period of the extreme events. When a large fall happens, cut bonus quickly and for a good return period, leave a "waiting" period of time then declare a bonus.

Valuation of Contract

Figure 5.6 shows the behaviour of the mean value of maturity guarantee G_T against the frequency of declaring bonus k. The mean guarantee value at maturity increases in most cases when k increases.

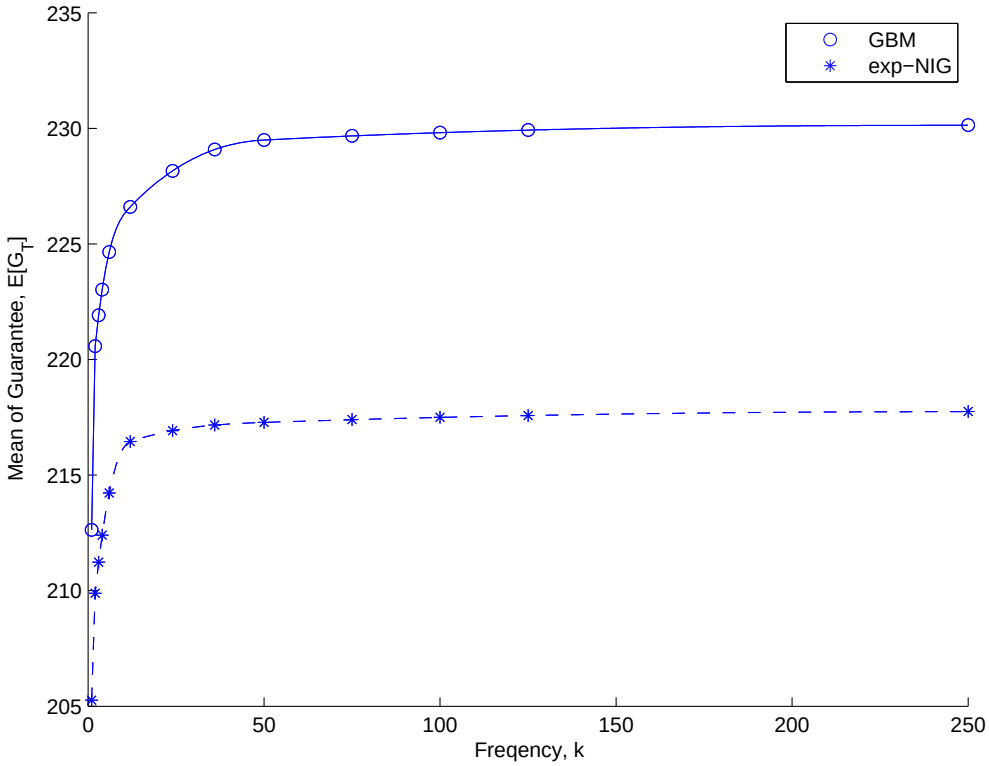

Figure 5.6: Mean of the guarantee

Note that the asset share distribution in this 100% EBR case only depend on the real world underlying asset returns. So the higher the guarantee is, the more valuable is the with-profits policy.

Here we consider the conditional expectation of the guarantee.

$$\mathbb{E}[G_T|G_T \leq AS_T] \text{ and } \mathbb{E}[G_T|G_T > AS_T]$$

Figure 5.7 tells us that the yearly bonus have lowest guarantee value when the policy is affordable. Also it has the highest guarantee when a shortfall happens.

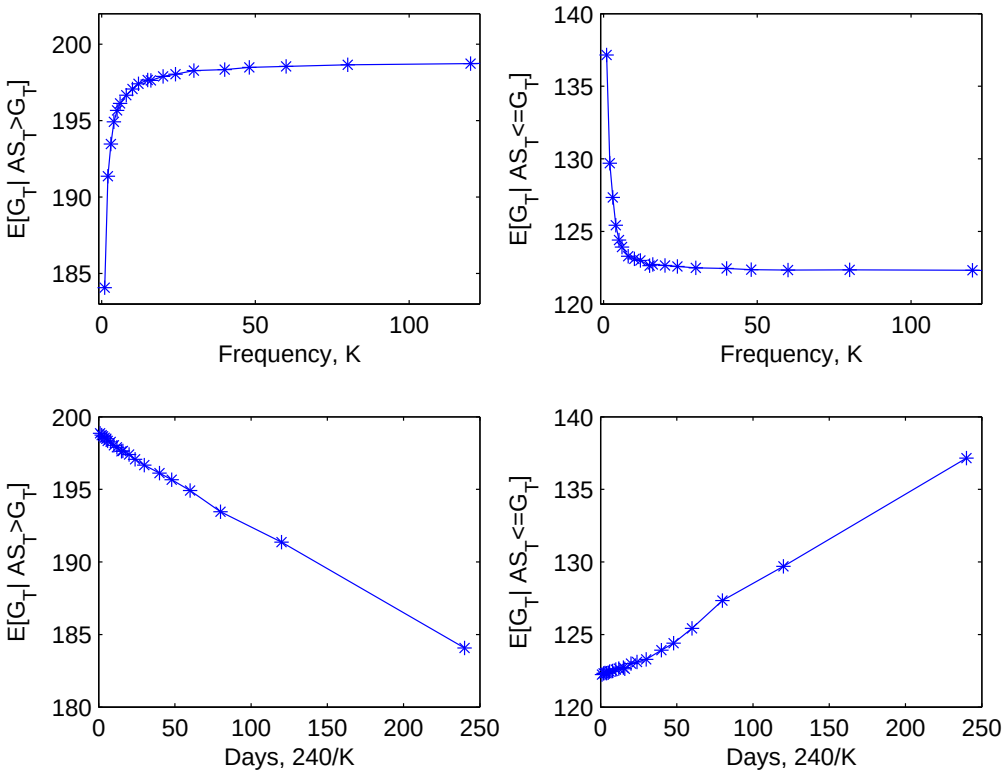

Figure 5.7: Mean of the guarantee

5.5 Concluding Remarks

As a reminder, we list the Case M1 to Case M6 considered in this Chapter:

M1 retrospective bonus calculated as the weighted arithmetic average of the last n years returns on the asset share;

M2 bonus earning power method using deterministic projected maturity asset share with a 25% terminal bonus cushion;

M3 alternative bonus earning power methods using stochastic projected maturity asset share generated by asset models;

M4 retrospective bonus declared on a monthly basis;

M5 bonus earning power declared on a monthly basis;

M6 bonus earning power declared on any frequency.

We can model the policyholder's PRE as the mean of exponential utility functions of the units values (guarantees). Normally the policyholder will expect a higher guarantee at the maturity.

Case M1 considered more general retrospective bonus mechanisms than in Case M0 by varying the participating rate and the smoothing period. The risk measures can be reduced by either increase the smoothing period n or by cutting the participating rate p.

The results in Case M2 using a bonus earning power method with deterministic projected maturity asset share and 25 percent terminal bonus cushion are similar to the risk measures in M0. Next we looked into the distribution of the maturity guarantee and asset share. We find a higher mean guarantee with higher standard deviation. The coupled Monte Carlo sample (AS_T, G_T) are more close related. There seems less possibility that a extreme loss happens in M2.

The BEP methods can be very flexible by making the targeted maturity guarantee as any forms depends on the office internal models. The projected maturity asset share and unit values can be stochastic and path dependent.

There seems to be little advantages in risk controlling by trying to declare retrospective bonus on a monthly basis in Case M4. Nonetheless, the risk measures drop rapidly with the monthly reversionary bonus calculated by BEP.

Chapter 6

Investment Strategy

6.1 Introduction and objectives

In chapter 5 we have discussed the liabilities of the UWP contracts and the bonus strategies. This chapter considers the asset side of the UWP contracts and the investment strategies.

One may ask what kind of services a life office can offer to their UWP policyholders. One of the answers may be the management of capital and the protection of investment risks in the long term. We say, if the performance of a UWP fund has a high correlation with some market assets and ends up with a low guarantee and large terminal bonus, as a policyholder, why not directly invest in the underlying assets instead of paying the charges for investing but without risk management.

One of the investment strategies that can reduce the risks of investment is diversification by managing the EBR of the fund. Modern Portfolio Theory suggests the total return index can be seen as the market portfolio. Thus any diversified portfolios involving total share return and risk free assets are "efficient" and are lying on the so-called "capital market line".

Modern financial theory brings the methodology of hedging for long term capital. By exploiting correlation of various risky investments, using dynamic hedging strategies can lead to improvement of portfolio risk management. Wilkie (1987), Hibbert and Turnbull (2003) and Willder (2004) suggest an option backup method to meet the guarantee. In this case, the hedging portfolio contains European put

options with same duration as the UWP contract. This method is independent of the dynamics of underlying asset, thus we call it "model-independent hedging". An essential requirement of model-independent hedging for the UWP liabilities is the existence of the corresponding tradeable derivatives, in this case, long term put options in the market. By model-independent hedging of the liability, the insurer transfers the risks of the guarantee to a third party who writes these particular put options. By applying the European put-call parity, a "call plus bond" model-independent hedging investment strategy can be constructed.

In practice, there are few long term tradeable put options in the market. Instead, internal dynamic hedging strategy can be used for reducing the guarantee mismatching risk. Normally, a hedging strategy is based on a set of assumptions. For example, the Black-Scholes hedging is based on the assumptions of the log-normal underlying asset, continuous re-balancing the hedging portfolio and no transaction costs. In reality, there are trading restrictions applied, for example, dynamic hedging portfolio is only within discrete time intervals and the insurer is always subject to transaction costs. The other source of error in dynamic hedging includes the difference between real world asset behaviour and the asset model used to calculate the hedging strategy.

The transaction costs are often charged according to the size of deal. Thus the amount of transaction costs is decided by the Gamma (the changing rate of the Delta with respect to the underlying asset) of the UWP contract. Risks are not entirely hedgeable only using the underlying asset and bond in internal hedging. Of course, the matching of the pay-out can be improved by increasing the hedging frequency and control the transaction costs with respect to more realistic option models. We will discuss hedging errors using a numerical example in section 6.3. For more detail see Wilkie *et al.* (2003) and Wilkie *et al.* (2005).

In fact, the insurers do not hedge the long term UWP contracts as vanilla options but they change the EBR from time to time to adjust the possible guarantee risk.

This chapter is organised as follows: In Section 6.2 we consider two investment management strategies based on the diversification of the with-profits fund. An invest-and-forget strategy and a re-balanced EBR strategy are tested. Some issues

of dynamic Delta hedging of the contingent claim will be discussed in Section 6.3. A long term with-profits fund investment strategy based on the hedging strategy is studied in Section 6.4. Finally, some concluding remarks are given in Section 6.5

6.2 Diversification

In this section, we consider two investment strategies based on diversification. We denote them as the model office M7. The UWP fund is assumed to be invested in both equities and zero coupon bonds. The proportion of total fund invested in equities, EBR (Equity Backing Ratio), can be managed by the insurer. And the rest 1−EBR of the fund is invested in bonds. In this stage we use a bond model with constant rate of compound return. If the price of the bond is B_0 at time 0, the bond price at time time t, B_t, is given by

$$B(t) = B(0)\exp\{rt\}, \tag{6.8}$$

where r is the constant compound risk free rate. The investment return from equity is given by the GBM and exp-NIG models fitted to SGTRI23 data.

We assume yearly risk free rate $r = 0.06$ in this thesis. Assume each cash-flow is discounted to the start of its corresponding policy year. This may include premiums, early death claims and surrenders. The recursive expression of the asset share at time $t+1$ is given by

$$AS(t+1) = (AS(t) + CF_t)\left(\alpha_t \frac{S(t+1)}{S(t)} + (1-\alpha_t)\frac{B(t+1)}{B(t)}\right), \tag{6.9}$$

where α_t is the equity backing ratio (EBR) at time t.

At the maturity time T, the asset share value $AS(T)$ is

$$AS(T) = \sum_{t=0}^{9} CF_t \prod_{i=t+1}^{10}\left(\alpha_i \frac{S(i+1)}{S(i)} + (1-\alpha_i)\frac{B(i+1)}{B(i)}\right).$$

We adopt all the assumptions of UWP model as in M0, except for re-adjusting the EBRs of the fund at 70% at start of every policy year.

Comparison to the "rebalanced EBR" case described above, we consider a static EBR case. All cash-flows of the UWP fund are diversified in equity and bond as the

same as the fund EBR at the time it paid in. This means for each cash-flow at time t, the fixed proportion α_t will be invested in/withdrawn from the fund invested in equities and the rest part is invested in/withdrawn from bonds. During the term of the policy no management actions will be taken, the EBR only changes because of the investment performance of the assets. It is an "invest-and-forget" strategy. At maturity, the asset share value $AS(T)$ is

$$AS(T) = \sum_{t=0}^{9} \alpha_t CF_t \prod_{i=t+1}^{10} \left(\frac{S(i+1)}{S(i)} \right) + (1-\alpha_t) CF_t \prod_{i=t+1}^{10} \left(\frac{B(i+1)}{B(i)} \right). \quad (6.10)$$

One would expect a higher mean EBR at the maturity for the investment portfolio in the static EBR case than the rebalanced EBR case because the mean return of the share is higher than the bond.

Table 6.7: 70% Rebalanced EBR UWP simulation using exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
GBM						
Loss(PRO,Y)	0.0237	10.3729	0	9.5757	4.9168	18.7705
Loss(PRO,M)	0.0189	10.2569	0	7.2900	3.8771	16.5415
Loss(RETRO,Y)	0.0252	9.7310	0	12.1876	4.9044	21.7738
Loss(RETRO,M)	0.0231	11.3861	0	10.5893	5.2604	20.0133
exp-NIG						
Loss(PRO,Y)	0.0351	11.5880	0	15.4223	8.1348	25.0782
Loss(PRO,M)	0.0277	11.3513	0	12.5311	6.2886	22.1136
Loss(RETRO,Y)	0.0353	11.5419	0	18.0675	8.1486	27.8343
Loss(RETRO,M)	0.0325	12.6167	0	16.4290	8.2009	25.8845

Table 6.7 shows the risk measures of insurer's maturity loss for the dynamic EBR case and the results of the static EBR case are shown in Table 6.8. The loss distribution of four strategies are observed, which are:

Loss(Pro,Y) yearly bonus earning power method using deterministic projected maturity asset share with a 25 percent terminal bonus cushion (see M2);

Loss(Pro,M) monthly bonus earning power method using deterministic projected maturity asset share with a 25 percent terminal bonus cushion (see M2);

Table 6.8: 70% Static EBR UWP simulation using exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
GBM						
Loss(PRO,Y)	0.0231	8.3494	0	7.0516	3.8574	15.0244
Loss(PRO,M)	0.0156	7.0999	0	3.6289	2.2152	10.1698
Loss(RETRO,Y)	0.0239	8.0512	0	8.8770	3.8485	15.7663
Loss(RETRO,M)	0.0200	7.8572	0	6.3694	3.1429	12.8381
exp-NIG						
Loss(PRO,Y)	0.0371	10.0293	0	13.2862	7.4417	22.5208
Loss(PRO,M)	0.0253	8.6781	0	9.0504	4.3911	16.0437
Loss(RETRO,Y)	0.0362	10.2650	0	14.1294	7.4319	20.9646
Loss(RETRO,M)	0.0304	9.0751	0	11.3806	5.5177	17.8252

Loss(Retro,Y) annual reversionary bonus rates are calculated by retrospective method with participating rate $p = 0.33$ and one year smoothing period; and

Loss(Retro,M) monthly reversionary bonus rates are calculated by retrospective method with participating rate $p = 0.33$ and one year smoothing period.

Same comments of the risk measures between different asset models, bonus mechanisms and bonus frequencies from previous model office observations apply:

- the risk measures are larger under exp-NIG model than under GBM model. It suggests GBM model underestimate most of the risk measures because exp-NIG is proved to be a more realistic model;
- with the same investment strategy and bonus mechanism, the risk measures are reduced when declaring monthly bonus instead of yearly; and
- the improvement of declaring monthly bonuses is more significant with BEP method than the Retrospective method.

The static EBR case is a linear combination of a portfolio including a UWP policy worth 70% of the original one with all funds invested in shares and a cash account. Thus the guarantee under both the retrospective and the bonus earning power methods can be split into two blocks. Since the bonus only add a percentage

of return and bonus earning power has a terminal bonus cushion less than one, the guarantee which is built by bond account investment return will definitely be achieved at the maturity and the excess amount in cash account will reduce the risk in the share investment.

For the re-balancing case, the guarantee liability cannot be split, but a 30% fixed investment in bond still effectively reduce the volatility of the portfolio. The risk measures are low in this example because we assume a constant monthly cash account return rate 0.005 equivalent to 6 percent per year, which seems to be slightly higher than usual assumption and the constant interest rate model is somehow unrealistic for long term bond return model.

However comparing these two strategies, the rebalanced EBR case which the insurer try to keep the fund at a fixed EBR rate each month has higher MOL and quantile risk measures in all the relative sub-cases. And for the two yearly bonus under exp-NIG cases, the static EBR case have higher POSF. The mean of maturity EBRs for the asset share in static EBR case are observed respectively as 76.98 and 76.78 percent for GBM and exp-NIG models. So the static case has a higher return with almost lower risk measures. This is because under the parametric models fitted to SGTRI23 data, the ergodicity properties of the asset model have mean return rate around 0.0086 which is higher than the bond return. In most cases, what the re-balancing EBR strategy do is switching the high return risky asset to lower return bond in order to reduce the total volatility of the investment returns. It suggests the re-balancing strategy may change the EBRs in the wrong direction. If we look at the investment returns for the portfolio, under the re-balancing strategy, the total return is always locally comes from, in our example, 70 percent equity return and 30 percent bond return. With the static EBR, if the investment returns are poor, the portfolio will have more proportion invested in bond. This reduce the total volatility of the return rate of the portfolio, hence the volatility of the bonus. And in the case when returns are good, the total return will have more weight on the equity returns. To this end the static EBR strategy seems to be a more advanced investment strategy than dynamic EBR case. To see this we test the re-balancing strategy at 76 percent of EBR which is same as the maturity EBR for the static

case, results are presented in Table 6.9, higher risks are observed compared to static EBR case. In the next section, an investment strategy based on dynamic hedging will be discussed.

Table 6.9: 76% Rebalanced EBR UWP simulation using exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
GBM						
Loss(PRO,Y)	0.0327	11.5644	0	14.4163	7.5631	23.9434
Loss(PRO,M)	0.0257	11.5308	0	12.0366	5.9268	21.4295
Loss(RETRO,Y)	0.0350	10.7614	0	17.8926	7.5330	27.6435
Loss(RETRO,M)	0.0325	12.7329	0	16.1228	8.2764	25.7091
exp-NIG						
Loss(PRO,Y)	0.0473	12.9612	0	20.5286	12.2613	30.7931
Loss(PRO,M)	0.0369	12.7391	0	17.6160	9.4015	27.3897
Loss(RETRO,Y)	0.0472	12.9941	0	24.2530	12.2664	34.0650
Loss(RETRO,M)	0.0441	14.2310	0	22.3442	12.5517	31.9012

6.3 Hedging for vanilla options

A dynamic hedging strategy is previsible stochastic process that describe the amount invested respectively in share and bond. To be more specific, define the admissible strategies as some previsible processes $\psi(t)$ and $\zeta(t)$. Denote $\Phi(t)$ as the cumulative hedging portfolio, in mathematical form

$$\Phi(t) = \Phi(0) + \int_0^t \psi(u)dS_u + \int_0^t \zeta(u)dB_u, \tag{6.11}$$

where S_t is the share price and B_t is the bond value. At any t, the self-financing condition gives:

$$\Phi(t) = \psi(t)S_t + \zeta(t)B_t. \tag{6.12}$$

The share and bond models used in calculation of the hedging quantities are called the option models and we use some parametric models to simulate the investment performance of the portfolio called the real-world models. The option models are not necessary to be the same as the real-world models. In the model office M8 and

M9, we adopt the Black-Scholes hedging strategy. According to the Black-Scholes framework, GBM models are used to calculate the hedging quantities in this chapter.

The delta, Δ, of a portfolio is defined as the sensitivity of the portfolio to the underlying. Delta-hedging reduces or eliminates the risks by taking the opposite positions on correlated (delta) underlying assets and the contingent claims. One example of delta hedging is the variance-optimal or mean-variance hedging which minimises the squared ℓ^2-distance of the maturity mismatching between the asset and liability,

$$\mathbb{E}((\Phi(T) - G_T)^2),$$

see Schweizer (2001) and Schweizer (1996). Gamma hedging can reduce the costs of hedging and increase the time between the re-balancing time of the hedging. Other common model-dependent hedging strategies include: superhedging, crash-hedging etc. The Black-Scholes methodology which is a delta hedge based on GBM share model and bond model with constant risk-free rate, offers an intuitive mathematical framework and a good approximation.

We tend to use the Black-Scholes option model in calculation. This does not mean the market is Black-Scholes, but use the formulae as an operator which gives some approximated value of delta and hedging quantities for the diffusion-like models.

The option model can be a Lévy model. Under the Lévy driven models, the market is incomplete, see Chan (1999). The risk-neutral martingale measure is not unique and the vanilla options are not perfectly hedgeable. Hubalek and Sgarra (2005) discuss the variance-optimal hedging under the exponential GH models and gives the analytical hedging quantities for the variance-optimal hedging strategy. We do not use these results in this thesis because, first, the complicated mathematical form of the variance-optimal hedging quantities dramatically increase the calculation expenses. And second, it is derived by assuming the exponential i.i.d. GH option models. As we will see in the next chapter, the structures of the real-world models can be more complex to capture some important features such as volatility clustering in the asset returns. It is not possible to derive the analytical forms of hedging quantities for all these Lévy driven asset models. The numerical results for hedging error of the European option in Hubalek and Sgarra (2005) show that the Black-

Scholes approach produces a reasonable hedge for the European call even if real data follows the jump-type exp-GH models.

The underlying asset referred as "share price" in this chapter is the SGTRI indices. The bond value follows the equation 6.8. Wilkie *et al.* (2003) and Wilkie *et al.* (2005) introduce the"vanilla" type option so-called maxi that model the embedded options used in modeling the participating life insurance guarantee. The European type option expires at time T with strike price K and underlying reference portfolio S_t at time t. The payoffs of a maxi option is $\max(S_T, K)$. This is equivalent to a long position in reference portfolio S_t and a European type put option with same underlying asset S_t, strike price K and expires at time T, by put-call parity,

$$\begin{aligned} \text{maxi}(S_t, K, T, t) &= S_t + p(S_t, K, T, t) && (6.13) \\ &= B_t + c(S_t, K, T, t). && (6.14) \end{aligned}$$

The theoretical maxi option hedging portfolio $H(S_t, K, T, t)$ is the sum of the hedging quantities:

$$H(S_t, K, T, t) = H_S(S_t, K, T, t) + H_B(S_t, K, T, t),$$

where $H_S(S_t, K, T, t)$ is the hedging quantities invested in share and $H_B(S_t, K, T, t)$ is the hedging quantities invested in bond.

According to the Black-Scholes formulae, the hedging quantities for the maxi option are:

$$\begin{aligned} H_S(S_t, K, T, t) &= S_t N(d_1) && (6.15) \\ H_B(S_t, K, T, t) &= e^{-r(T-t)} K N(d_2) && (6.16) \end{aligned}$$

where:

$$\begin{aligned} d_1 &= \frac{\ln\left(\frac{S_t}{K}\right) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}}; && (6.17) \\ d_2 &= \sigma\sqrt{(T-t)} - d_1. && (6.18) \end{aligned}$$

The Black-Scholes hedging framework is based on an ideal market which assumes a continuous adjusting of the $H(S_t, K, T, t)$ without transaction costs. If the real world share model is a GBM model with same volatility as the one used in hedging

portfolio calculation, then the continuous hedging would exactly provides the payoff for the maxi option. Thus, if we make

$$\psi(t) = H_S(S_t, K, T, t)/S_t; \text{ and} \tag{6.19}$$

$$\zeta(t) = H_B(S_t, K, T, t)/B_t, \tag{6.20}$$

at any $0 \leq t < T$, the cumulative portfolio at time T would be exactly $\max(S_T, K)$.

In general, the discrete hedging strategy (ψ, ζ) does not give the desired payoffs of the option value at maturity due to the model bias and the trading restrictions. The mismatch part leads to the so-called the hedging errors. The trading restrictions include, for example, the discrete management of the portfolio and expenses of the management (transaction costs). And the hedging quantities are calculated under the assumption that the market returns follow the asset models that used in the equivalent martingale measure calculation (option models). The difference between the option asset model and the "true" investment returns of the underlying is called the model bias (Vega risk). The Monte Carlo method will be used to study the discrete hedging errors in M8 in order to observe the maturity loss of the UWP under dynamic hedging based investment strategy.

The implied volatility is not observable because there is no market price data available for such long term European options. Instead the volatility parameter in the option model 6.15-6.18 is not restricted to be a constant and can be chosen by the insurer, although the Girsanov theorem suggests it should be the same as the statistical volatility calibrated to the return of the SGTRI from 1923 to 2005.

Let $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$, $t_i, i = 0, 1, 2, \cdots, n$ be the time points that the hedging quantities can be changed within the policy term. The initial hedging quantities $\{\phi(t_0), \zeta(t_0)\}$ are calculated by solving equation 6.19 and 6.20 at time 0. At each successive intermediate time point $t_i = i, 0 < i \leq T$, the value of the hedging portfolio set up at time t_{i-1} is,

$$\Phi(t_i) = \psi(t_{i-1})S_t + \zeta(t_{i-1})B_t. \tag{6.21}$$

The theoretical hedging quantities calculated by Black-Scholes famework at time t is,

$$H(AS_{t_i}, G_T, T, t_i) = H_S(AS_{t_i}, G_T, T, t_i) + H_B(AS_{t_i}, G_T, T, t_i). \tag{6.22}$$

In general, for $i \geq 1$ the real value of the hedging portfolio $\Phi(AS_{t_i}, G_T, T, t_i)$ would not equal the risk neutral value calculated by Black-Scholes formula, $H(AS_{t_i}, G_T, T, t_i)$. Wilkie *et al.* (2005) gives three strategies for setting the hedging quantities at intermediate re-balancing points, which are, investing the correct amount in the share and the balance in the bond; investing the correct amount in the bond; and investing the correct proportions in the share and the bond. We use the third strategy and set the hedging quantity in share as

$$\psi(t_i) = \Phi(t_i) \left(\frac{H_S(AS_{t_i}, G_T, T, t_i)}{H(AS_{t_i}, G_T, T, t_i) S_{t_i}} \right) \tag{6.23}$$

at $t_i,\ 0 < i < n$ and the balance in bond.

Consider a 10 year UWP policy with initial investment asset share of 100. The reversionary bonus rate is constant of 4% p.a. which corresponds to the deterministic maturity guarantee amount of 148. The UWP fund is invested in market portfolio of shares which is SGTRI i.e., the equity backing ratio is 100%. Thus the maturity liability for the insurer is identical to the 10 year vanilla maxi option, which is given by:

$$\max(AS_T, G_T) = \text{maxi} \left(\frac{100}{S_0} S_T, G_T, T, T \right). \tag{6.24}$$

The market price of the UWP contract is the same as the maxi option value at time t_0 which is 119.92 calculated by Black-Scholes formula in Equation 6.15-6.18. The insurer follows the Black-Scholes price and deducts an extra 19.92 charge from the policyholder to construct the initial hedging portfolio by investing the amount of hedging quantities $\phi(t_0)$ in shares and $\zeta(t_0)$ in bond. The insurer re-adjusts the portfolio at start of each hedging steps, Δt. We consider the hedging strategies in three different frequencies which are yearly ($\Delta t = 1$), half yearly ($\Delta t = 1/2$) and monthly ($\Delta t = 1/12$).

The statistics for the reference portfolio (SGTRI) are shown in Table 6.10, followed by statistics for the hedging portfolio and the hedging errors deficit, for three hedging frequencies, yearly, half yearly and monthly under the GBM model and exp-NIG model. As we discussed before reference portfolio at the maturity is approximately log-normally distributed under both GBM and exp-NIG models. In this example, we assume there is no charge, thus, the asset share is the reference portfolio.

Table 6.10: Statistics for reference portfolio, hedging portfolio and hedging error at maturity for model office M8, simulated with GBM and exp-NIG models.

	MEAN	SD	Lowest	Highest
GBM				
Reference	323.7578	185.2086	30.5131	3035.9
$\Delta t = 1$				
Hedging	325.1809	177.0459	108.8807	2945.0
Error	-2.0726	10.2409	-137.2670	27.8666
$\Delta t = 1/2$				
Hedging	326.2054	179.2313	115.7024	2990.8
Error	-1.0481	6.9827	-78.7617	25.2715
$\Delta t = 1/12$				
Hedging	327.0772	181.0154	133.9343	3013.9
Error	-0.1763	2.7955	-28.0483	14.7810
exp-NIG				
Reference	323.0572	179.3749	28.0413	2608.0
$\Delta t = 1$				
Hedging	324.9683	171.9588	91.5550	2589.2
Error	-1.5991	9.8927	-147.1829	29.2088
$\Delta t = 1/2$				
Hedging	325.9977	174.0033	94.6295	2576.2
Error	-0.5697	7.1842	-105.6199	27.0605
$\Delta t = 1/12$				
Hedging	326.8472	175.7475	96.4943	2617.8
Error	0.2798	4.2657	-51.5301	25.8465

First we consider the case when the real world model is GBM with same volatility as the option model which is the statistical volatility of SGTRI23 data, $\sigma = 0.17$. The maturity mean value of the hedging portfolio is slightly higher but close to the mean of reference portfolio. Within the $100,000$ simulated values of the reference portfolio, there are $11,348$ of them less than the guarantee 148.02, that is, a probability around 11 percent. The distribution of the maturity reference portfolio spread widely with a minimum value 30.5131 and the deficit (loss) is 117.51. The minimum values of the yearly hedging fund is 108.88. It reduces the size of deficit, the hedging error is 39.14. With a more frequent hedging, it is 115.70 for the half yearly strategy and 133.93 for monthly case. Also the more frequent hedging has closer maximum value because they rearrange the initial hedging quantities earlier. In this case the Δ or the ratio of hedging quantities H_s/H is close to one in most intermediate times.

The hedging error is the difference between the value of maturity hedging portfolio and the desired payout in equation 6.24. The mean values of the error are negative under the GBM real world model. The standard deviation and the size of extreme values reduced significantly by hedging more frequently. It can be seen the minimum error, that is the largest deficit, is relatively high for yearly hedging. By studying those simulation scenarios which generate large negative hedging errors, we found most of these large errors appear in the case when the maturity maxi options are far out-of-money. The largest possible loss that the insurer would have to pay in this example is the minimum value of hedging portfolio minus the guarantee G_T which is 39.14 for yearly, 32.32 for half yearly and 14.09 for monthly strategies

Under the exp-NIG real world model, the underlying share price has the mean close to the GBM case. The standard deviation is smaller, the distribution of the log-share price are negatively skewed. There are $11,149$ out of $100,000$ simulation outcomes that the reference portfolio is less than the guarantee at maturity, which is round 11 percent in probability. Compared to the GBM case, the minimum values of the yearly hedging fund under exp-NIG real world is 91.55, the size of extreme loss increases to 56.47. The minimum value of hedging portfolio is 94.63 for the half yearly strategy and 96.49 for monthly case. The improvement of extreme loss

control with a more frequent hedging is not significant when using the GBM option model in a exp-NIG real world.

Table 6.11 shows the statistics of the hedging errors and the insurer's losses at maturity. For the yearly rearranged strategy, the statistics are close under the two different real world models.

The loss in the exp-NIG model real world case has fatter left tail than the GBM case, although the GBM have on average larger hedging errors. The insurer will underestimate the quantile based risk measures such as VaRs and CTEs by using the GBM internal model.

The POSF of loss increases when Δ decreases, i.e., the fund is re-balanced over smaller time steps. Meanwhile, MOL, VaR and CTEs under GBM case are reasonably close, the CTE99 and CTE99.5 are respectively 4.81 and 5.88 with monthly, they are acceptable compared with the previous examples. The required quantile reserves for the UWP contract M8 under a monthly rebalanced investment are far lower than those in model office M0-M7.

On the other hand, under exp-NIG model, the decreasing of risk measures are slow. When the total times of re-adjustment of the fund increase from 9 (yearly) to 119 (monthly), the MOL, VaR and CTEs for loss are still more than half of the original value.

The POSF of error is an increasing function of Δ. One may expect a higher probability of "super-hedge" under a more actively managed hedging portfolio.

Note that the more frequent hedging leads to path-dependent but larger, on average, transaction costs. It can be estimated by applying the so-called integral-by-path study (Mallivian calculus). The risk of the cost uncertainty is normally quantified by the gamma, however, the gamma risk is unhedgeable without infinitely many derivatives.

6.4 Hedging based investment strategy, M9

So far we have considered the dynamic hedging strategies for the "vanilla" typed UWP contracts, i.e., the guarantee is deterministic and the reference portfolio is

Table 6.11: Statistics for insurer's loss and hedging error at maturity for model office M8, simulated with GBM and exp-NIG models.

	POSF	MOL	VaR_{99}	CTE95	CTE99	CTE99.5
GBM						
$\Delta t = 1$						
Err	0.5350	8.8325	35.7130	30.0362	47.8583	56.3970
Loss	0.0517	6.0043	9.8206	6.2128	14.8364	18.2400
$\Delta t = 1/2$						
Err	0.5251	5.8975	22.5066	19.2230	29.3022	34.2693
Loss	0.0538	4.4876	7.3645	4.8180	11.2579	13.9190
$\Delta t = 1/12$						
Err	0.5112	2.2214	7.9823	6.8492	10.0683	11.5570
Loss	0.0559	1.9239	3.2829	2.1386	4.8091	5.8844
exp-NIG						
$\Delta t = 1$						
Err	0.5163	8.4629	33.7031	28.4401	44.6779	52.7290
Loss	0.0598	7.3742	12.7729	8.6721	19.0148	23.2901
$\Delta t = 1/2$						
Err	0.4850	5.9737	23.1181	19.6025	31.0867	36.8486
Loss	0.0640	5.9479	10.5019	7.3807	15.7135	19.3329
$\Delta t = 1/12$						
Err	0.4138	3.4870	13.1038	10.9506	17.3734	20.3215
Loss	0.0745	4.1442	7.6804	5.6899	11.5638	14.3216

"external" asset as share return indices. In this section, we consider a dynamic hedging based investment strategy for a UWP policy called "model office M9". The reference portfolio is the asset share and the guarantees are increased by the dynamic reversionary bonuses. As we discussed in Chapter 4, in this particular contract, the guarantee is path-dependant and there exists the so-called "investment bonus cycle". The investment strategy is a function of the guarantees so as the bonuses, meanwhile, the size of bonuses is decided by the performance of investment.

Recall that, in M8, the UWP contract requires an initial charge worth 19.92 which is close to 20% of the asset share. This guarantee seems to be too expensive for the policyholders, although the 4 percent annual bonus rate seems to be reasonable. However, the question here is that when the investment has been rather bad, the policyholder might not expect bonuses to be as high as 4 percent *per annum*. And if the investment has been good in early years of the policy, the policyholder would like to have a high guarantee to protect the profits they have already achieved.

Given a proper probability space $(\Omega, \mathcal{F}_t, \mathbb{P}, \mathbb{Q})$, the $\mathbb{Q}$ is the market equivalent martingale measure (pricing measure), the value of the maxi option is

$$\begin{aligned} \text{maxi}(S_0, G_T, T, 0) &= S_0 + p(S_0, G_T, T, 0) && (6.25)\\ &= S_0 + e^{-rT}\mathbb{E}_{\mathbb{Q}}[(G_T - S_T)_+] && (6.26)\\ &= S_0 + e^{-rT}\int_{\Omega}(G_T - S_T)\mathbf{1}\{G_T > S_T\}d\mathbb{Q}(\omega) && (6.27) \end{aligned}$$

where $p(S_0, G_T, T, 0)$ is the put option whose value is calculated by the second part in the right hand side of equation 6.27. The cost of hedging is decided by the probability that the contract end up with in the money and the size of the corresponding deficit. By using the dynamic bonus mechanism can reduce the value of $G_T - S_T$ in the cases when $G_T > S_T$, the cost of the guarantee could be covered by a lower extra premium.

The support capital for the M9 can be made up by contribution from both policyholders and shareholders. For the simplicity, we consider a mutual life office, the free asset which support the hedging is coming from the guarantee charge paid in by the policyholders.

At policy inception, a single premium P is paid in by policyholder. We denote the nominal asset share before any charges as AS_0' which equals to P. The fix

proportion of the nominal units which is c percent *per annum* will be cashed to pay a guarantee charge. The net present value of asset share for a policyholder AS_0 is the nominal asset share AS'_0 less all the future guarantee charge, $AS'_0(1-c)^T$. We assume the rest part of the asset share, $AS'_0 - AS_0$, will be invested separately from the actual asset share AS_0 and the insurer will use this amount to construct the hedging portfolio. And the net present value of guarantee, G_0, for policyholders is the nominal guarantee G'_0 less all the future guarantee charge, which is $G'_0(1-c)^T$.

At time t, ignore all the future bonus, the minimum projected maturity guarantee for the asset, $G'_T(t)$ is given by

$$G'_T(t) = G_t(1+g)^{T-t}, \quad t = 0, 1, 2, ..., 9, \tag{6.28}$$

where g is the guaranteed minimum bonus rate and G_t is the current net guarantee level.

Thus, at time t, the UWP guarantee $G'_T(t)$ can be achieved by a maxi option, see Wilkie *et al.* (2005), with the expires at time T, exercise price $G'_T(t)$ and current underlying asset with value of AS_t invested in same portfolio as asset share. The payoff at the maturity is $\max(AS_T, G'_T(t))$. (This is identical to a portfolio which invest the asset share in same asset and hold a put option with same maturity and reference portfolio as the maxi option.)

Note that the underlying assets for these options are the policyholder's asset share and not necessary to be any tradable assets in the market. When the EBR is less than 100 percent, the reference portfolio of the maxi option is a combination of shares and bonds.

We denote the $\Phi(S_t, K, T, t)$ as the value of the hedging portfolio at time t with the underlying asset S_t, strike price K and the maturity at T. At time $t = 0$, the $\Phi(S_t, K, T, t)$ is assumed to be the same value as the maxi option with underlying S_t, strike K and maturity T. The free asset (free estate), in our example as the amount of the guarantee charge will be invested in the hedging portfolio $\Phi(S_t, K, T, t)$ to support the expenses replicating portfolio. The amount of this support capital is limited and sometimes the put option backed up for the 100 percent EBR asset share investment is not affordable. In this case, at each portfolio readjustment time t, we adopt the investment strategy as follow:

At time t, if the free asset FA_t is less than the hedging expenses which is the value of the put option,

$$\begin{aligned} FA_t &\leq p(S_t, G'_T(t), T, t); && (6.29) \\ S_t &= AS(t), && (6.30) \end{aligned}$$

where $G_T(t)$ is the minimum maturity guarantee given in Equation 6.28. In this case, we maximise the EBR of the asset share, e_t, under the condition that the put option of the asset share invested in risky asset, S_t, is affordable by support capital. The e_t can be calculated by solving the follow equations:

$$\begin{aligned} FA_t &= p(S_t, K(t), T, t); && (6.31) \\ S_t &= e_t AS(t), 0 < e_t \leq 1; && (6.32) \\ K(t) &= G'_T(t) - (1 - e_t)AS(t)\frac{B(T)}{B(t)}. && (6.33) \end{aligned}$$

$p(S_t, K(t), T, t)$ in Equation 6.31 is a monotonic function with respect to e_t, the solution is unique. The free asset FA_t is the value of the vanilla put option that expires at time T, exercise price $K(t)$ and current underlying asset value $e_t AS(t)$ invested in shares. The exercise price $K(t)$ is calculated as the minimum guarantee $G'_T(t)$ less the part of the guarantee which can be achieved by the $100(1-e_t)$ percent of the asset share invested in bond. The maturity value of the asset share invested in bond is $(1 - e_t)AS(t)B(T)/B(t)$.

We assume that bonuses are declared after the re-arrangement of the portfolio if they happen at same date. Thus at each rearranging time, we need to recalculate the asset share, minimum guarantee and the amount of free assets. We assume in the hedging portfolio, the two components, free asset and policy holders' asset share, always keep the same ratio. For example, at time t, the hedging portfolio is constructed as,

$$\Phi_H(t) = \Phi_A(t) + \Phi_F(t),$$

where $\Phi_A(t)$ is the asset share component of the hedging portfolio and $\Phi_F(t)$ is the free asset component. At time t they have relation $\Phi_A(t)/\Phi_F(t) = C$, then just before the next re-balancing time, $t + \delta_-$, the value of the hedging portfolio is

$\Phi_H(t+\delta_-)$. The free asset component is,

$$\Phi_F(t+\delta_-) = \Phi_H(t+\delta_-)\left(\frac{1}{1+C}\right). \tag{6.34}$$

The minimum guarantee,

$$G'_T(t+\delta) = G_T(t)' \prod_{s\in(t,t_\delta]}\left(\frac{1+b_s}{1+g}\right). \tag{6.35}$$

Thus, at time $t+\delta$, we have the asset share $AS(t+\delta)$, minimum guarantee $G'_T(t+\delta)$ and the free asset $FA(t+\delta)$. The new investment portfolio can be constructed by solving equation 6.31-6.33.

We ignore the mortality risk and the transaction costs.

Table 6.12: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually re-arrange the portfolio, simulated with GBM and exp-NIG models.

	POSF	MOL	VaR_{99}	CTE95	CTE99	CTE99.5
GBM						
Loss(i)	0.0739	10.6999	18.1626	14.3031	20.9960	22.8757
Loss(ii)	0.0233	4.0667	3.8996	1.8951	6.9821	9.0002
Loss(iii)	0.0694	9.6262	16.1922	12.5736	23.9471	30.0889
Loss(iv)	0.0131	5.3155	0.7523	1.3927	6.9432	11.8985
exp-NIG						
Loss(i)	0.0787	12.0091	20.5324	16.6520	24.6054	27.4537
Loss(ii)	0.0299	5.2468	6.1558	3.1376	10.0699	12.7128
Loss(iii)	0.0769	11.3858	19.6254	15.9265	31.3425	40.2826
Loss(iv)	0.0195	7.4873	3.2641	2.9200	13.5040	21.5960

We first check the model office M9 when the real-world model is same as the option model, that is a GBM model with same volatility. This is necessary condition for the Girssanov theorem to be hold. The hedging portfolio is rearranged annually. The loss distribution of four strategies are observed, which are:

Loss(i) annual reversionary bonus rates are calculated by retrospective method with participating rate $p = 0.33$ and one year smoothing period,

$$Loss = \max\{G_T - AS_T, 0\};$$

Loss(ii) annual reversionary bonus rates are calculated by retrospective method with participating rate $p = 0.33$ and one year smoothing period, the free asset will be paid under deficit,

$$Loss = \max\{G_T - AS_T - FA_T, 0\};$$

Loss(iii) bonus earning power method using deterministic projected maturity asset share with a 25 percent terminal bonus cushion (see M2), without support of the maturity free asset;

Loss(iv) bonus earning power method using deterministic projected maturity asset share with a 25 percent terminal bonus cushion (see M2), with the free asset paid under deficit.

The risk measures are listed in Table 6.12. One can see, with the retrospective bonus mechanism, loss(i), the hedging based investment strategy M9 have POSF which is 7.39%. The MOL is 10.6999 and the 99% VaR is 18.1626, the CTEs at level 95%, 99% and 99.5% are 14.3031, 20.9960 and 22.8757. Loss(ii) requires the support of the maturity free asset FA_T. The POSF in Loss(i) of suggests that it has probability of 7.39% that part or whole of free asset will be paid in order to meet the guarantee. The probability of the free assets being insufficient to meet the deficit is 2.33%. Both VaR and CTEs are reduced, which suggest the Loss(ii) has a relatively thin right tail. The difference between VaR_{99} of Loss(i) and Loss(ii) is 14.0139. These figures seem to be acceptable for an insurer, there is a low probability that the liability guarantee is unaffordable.

With the annually bonus calculated by BEP we described in Chapter 5, the hedging based investment strategy M9 have a relatively low POSF, which is 6.94% for Loss(iii) and 1.31% for Loss(iv). If hedging quantities had been theoretically rearranged continuously, the POSF of Loss(iv) would converge to zero in this example. With the support of maturity free asset, the MOL is 5.3155 which is around 2% of the mean maturity asset share. The 99% VaR is close to 0. For hedging based investment strategy M9, if the real-world model is the same as the option model, the maturity loss distributions have thin tails under annual rearrangement.

We now consider the case that the exp-NIG is used to simulate the real world share returns. The statistics of the insurer's maturity loss are given in lower part of Table 6.12. All risk measures we observed in this example increase. Thus we argue that if the real world model is exp-Lévy which, in this case, is an exp-NIG model, the distribution of the hedging error has a fatter right tail than the GBM real world models.

Table 6.13 shows the mean and standard deviation of the maturity asset share and guarantee for model office M9. The mean of the guarantee is 137.7378 under GBM and 137.7448 under exp-NIG. And the standard deviation is less than 19 under both asset models. Thus there is a large probability to achieve maturity guarantee at a range of low value. This is because under the hedging strategy, the volatility of the asset share is reduced. Tt seems that a UWP contract with bonus which declare 33 percent of last year's investment return is not attractive to the policyholders and it is necessary to increase the participating rate p in this case.

Table 6.13: Statistics for asset share and guarantee at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with GBM and exp-NIG models.

	MEAN	SD
Asset Share		
AS_T(Pro-GBM)	279.3927	155.8326
AS_T(Pro-NIG)	278.7333	150.9961
AS_T(Retro-GBM)	280.0110	157.7917
AS_T(Retro-NIG)	279.7270	152.2732
Free Asset		
FA_T(Pro-GBM)	18.9932	1.3383
FA_T(Pro-NIG)	18.9766	1.3042
FA_T(Retro-GBM)	18.9343	1.4081
FA_T(Retro-NIG)	18.8915	1.3856
Guarantee		
G_T(Pro-GBM)	204.1193	99.9518
G_T(Pro-NIG)	203.9140	96.7131
G_T(Retro-GBM)	137.7378	18.8508
G_T(Retro-NIG)	137.7448	18.2015

Hedging portfolio and EBRs

The left subplot in Figure 6.8 shows the mean proportion of asset share e_t that invested in the asset linked to market portfolio (risky asset). The star line (GBM(P)) shows the e_t simulated by using GBM model, the guarantee is increased by declaring BEP bonus. If we assume the SGTRI as the market portfolio, the UWP fund invested in share indexes and the risk free bond is lying on the "Capital Market Line" which is the Markowitz efficient frontier of the market. And the proportion e_t, of which the asset share invested in the hedging portfolio, is the ratio of the portfolio risk divided by the market risk, σ_e/σ_M. By solving the equation 6.31-6.33, the projected guarantee at policy inception is affordable by the free asset, i.e., $e_t = 1$ at time $t = 0$.

The solid line (GBM(R)) shows the e_t simulated by using GBM asset model, and the broken line (exp-NIG(R)) shows the e_t simulated by using exp-NIG asset model, both of guarantees are increased by retrospective bonus with participating rate $p = 1/3$ and $n = 1$. The mean proportion of the asset share that can be afforded to invest in share is decrease through out the policy term. At start of final policy year at time $t = 10$, under GBM model, there are 97.96 percent of the asset share on average is invested in share indices. The figure decrease to 97.39 percent under exp-NIG model. More-on, there is 98.26 percent of the probability for GBM case and 97.82 percent of the probability for exp-NIG case that the EBRs for the final year is one. That is more than 97 percent of probability that the whole asset share can be hedged using the accumulated free asset at start of the final policy year. We can see that in most cases the asset share can be invested in shares.

The dot line (exp-NIG(P)) shows the e_t simulated by using exp-NIG model,the guarantee is increased by declaring prospective bonus. In this case, the e_0 at the is same as the GBM case because we use the same option model and have same initial asset share, free asset and guarantee. The mean of e_{10} at start of the 10th policy year is 0.97819, which is slightly lower than the GBM case. There is 97.73 of the probability that the EBRs for the final year is one.

The right subplot in Figure 6.8 shows the mean of actual EBR of the asset share at start of each policy year. The star line (GBM(P)) shows the actual EBR simulated

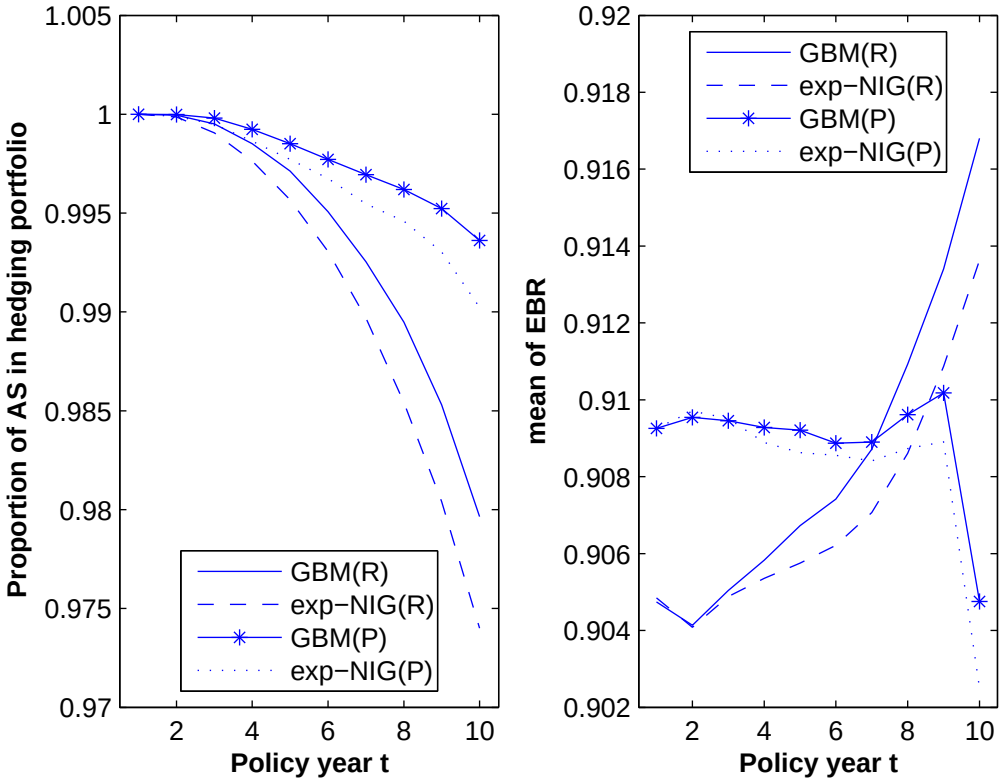

Figure 6.8: Proportion of AS invested in hedging portfolio and the EBR for UWP fund

by using GBM model, the guarantee is increased by declaring prospective bonus.

At the time when policy setup it is 0.9093 for both the GBM and exp-NIG real world models because of the the same option model, initial asset share, free asset and guarantee. With GBM real world model, the mean of the actual EBR at start of the 10th policy year is 0.9048. The dot line (exp-NIG(P)) shows the actual EBR simulated by using exp-NIG model. The mean of actual EBR at start of the 10th policy year is 0.9025, which is almost no difference from the GBM case.

The solid line (GBM(R)) and the broken line(exp-NIG(R)) show the actual EBRs simulated by using GBM and exp-NIG models respectively. The actual EBRs are increasing function of policy year t in these cases and the exp-NIG(R) case has lower EBR throughout the policy term.

The mean of actual EBR is an increasing function of time t from 2 to 10 under the retrospective bonus rate. However the scale of axis in the graph is so small

varying from 0.9041 to 0.9168, actually the mean EBR has very little fluctuations. The mean EBR has a large downward shift at time 10 for the BEP cases. This is because both of the mean and standard deviation of BEP increases rapidly in the last two years, see Figure 5.4. Thus it costs more than the early year in the policy term and less proportion of the asset is invested in shares.

Varying the volatility parameter

It has been discussed in previous sections that the insurer can adjust the parameters used in option models. Keeping the risk free rate r fixed at 0.005 per month, we are now varying the volatility parameter σ in the option model used to calculate the hedging quantities in M9. Values of 0.10, 0.15, 0.20, 0.25, 0.30 are chosen for σ. The statistics of the maturity asset share and guarantee are shown in Table 6.14.

Table 6.14: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with GBM and exp-NIG models.

SD σ	Mean of asset share	SD of asset share	Mean of guarantee	SD of guarantee
GBM				
0.10	305.7682	165.5516	209.3829	105.9056
0.15	300.7732	160.7955	205.8275	102.6305
0.20	294.3428	149.6081	201.8564	95.4082
0.25	288.0676	140.9068	198.5109	89.3703
0.30	283.8702	135.2877	197.3210	85.0436
exp-NIG				
0.10	305.0581	160.7793	209.4266	102.4133
0.15	300.0367	155.4265	205.3594	99.3432
0.20	293.8104	145.7810	201.4774	92.2735
0.25	288.7323	137.2934	199.0686	86.7884
0.30	283.0620	130.1693	196.9989	81.5008

We can see from Table 6.15 that the insurer can achieve smaller risk measures by using a larger volatility parameter in the option models. With the higher volatility, it costs more to hedge the guarantee of the policy. It leads to lower deltas and EBRs of the UWP fund, so as the expected return of the fund and expected guarantees.

Table 6.15: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with exp-NIG models, BEP.

SD(σ)	POSF	MOL	VaR_{99}	CTE95	CTE99	CTE99.5
GBM						
0.10	0.0155	6.4331	1.7765	1.9943	9.7066	16.0232
0.15	0.0145	6.3472	1.0995	1.8407	9.1392	15.8012
0.20	0.0113	4.6333	0.2999	1.0471	5.3771	9.5315
0.25	0.0080	4.1275	0	0.6604	3.3569	6.4885
0.30	0.0051	3.3760	0	0.3444	1.7218	3.5699
exp-NIG						
0.10	0.0236	9.5333	5.5419	4.4997	19.9432	30.6708
0.15	0.0212	7.6127	3.9388	3.2278	14.6562	22.8530
0.20	0.0174	6.6684	2.3541	2.3206	11.2022	18.1186
0.25	0.0130	5.9121	0.8485	1.5371	7.8172	13.3448
0.30	0.0087	5.5464	0	0.9651	4.9967	9.2900

Table 6.14 shows the mean and standard deviation of the maturity asset shares and the guarantees.

For the same volatility parameter, the risk measures simulated using exp-NIG models are always greater than those under GBMs. For example, the statistics of $\sigma = 0.15$ under the GBM model are less than those of $\sigma = 0.20$ under the exp-NIG model and greater than $\sigma = 0.25$ under the exp-NIG model. It shows if the insurer would like to achieve the close results as the 15-percent-volatility hedging simulated using the GBM real-world in the M9, he/she can add a margin between 5 to 10 percent to the volatility in option models if the "true" investment returns follow the exp-NIG.

In M9, varying the volatility in the option model improves the insurer's risk management, meanwhile it reduces the possible high return and guarantee of the policyholder. In this way, varying option model parameters can be seen as a kind of hidden charge or cost.

6.4.1 Terminal bonus cushion and participating rates

So far we have used a 25 percent of the projected maturity asset share as the terminal bonus cushion. We vary the terminal bonus cushion in bonus earning power calculation in equation 5.5 from 12% to 30%.

Table 6.16: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with GBM and exp-NIG models.

TB	Mean of asset share	SD of asset share	Mean of guarantee	SD of guarantee
GBM				
0.3	300.0583	157.3442	191.7824	93.0096
0.28	299.7137	158.7743	197.0484	96.6703
0.26	298.6500	155.8425	201.8319	98.2006
0.24	297.1099	154.9314	206.0451	100.8629
0.22	297.6734	156.9109	211.8519	104.9679
0.20	295.6406	152.7616	215.8301	105.2472
0.18	294.6797	151.8487	220.9944	107.9667
0.16	292.0854	150.1379	224.7394	109.9103
0.14	291.5596	149.1847	230.5945	113.3518
0.12	289.9302	147.4597	235.2463	114.5312
exp-NIG				
0.3	299.4081	152.8185	191.8194	89.9358
0.28	298.9875	153.4560	196.5895	93.5583
0.26	298.1225	151.8737	201.4593	94.9847
0.24	297.8678	151.0099	206.7260	97.9949
0.22	296.7417	150.9905	211.4789	100.6686
0.20	295.3509	150.2151	216.0279	103.2177
0.18	294.4245	146.9580	221.0398	104.3724
0.16	292.2576	146.2080	225.2544	106.8056
0.14	291.0030	144.0876	230.0995	108.1660
0.12	289.9136	143.9428	235.0550	111.3830

Table 6.16 shows the statistics of the maturity asset shares and guarantee. The lower the terminal bonus cushion, the lower the mean of the asset share. This is because the lower TB leads to higher projected guarantee in each intermediate time point throughout the policy. It increases the costs of the hedging portfolio, so less proportion of the asset share will be investment in share. Thus the asset share has

lower mean and less SD with a lower TB. The mean of the maturity guarantee, for both real-world models, is rapidly increasing in the range of investigation TBs. We can expect the lower TB and the lower asset share effect the projected guarantee simultaneously, with the former increase the reversionary bonus rate and the latter bring down the projected maturity asset share. It seems the increasing of the mean of guarantees is more significant than the reducing of the mean of the asset share. One reason for this is because we assume a 6 percent yearly risk free return which seems to be higher than in practice. The difference in expected investment returns between shares and bonds are close.

Table 6.17: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with exp-NIG models, BEP.

TB	POSF	MOL	VaR_{99}	CTE95	CTE99	CTE99.5
GBM						
0.3	0.0074	2.8858	0	0.4271	2.1355	4.2685
0.28	0.0095	4.7522	0	0.9124	4.5621	8.5901
0.26	0.0119	5.0868	0.4562	1.2107	6.1847	10.9235
0.24	0.0142	6.0182	1.1501	1.7092	8.6598	14.9338
0.22	0.0184	6.4462	2.3304	2.3722	11.2317	18.3978
0.20	0.0228	7.1294	3.8361	3.2510	14.6036	23.1836
0.18	0.0292	7.2700	5.4657	4.2457	17.7890	27.4730
0.16	0.0337	8.0694	8.1873	5.4387	21.4447	31.7239
0.14	0.0444	8.2841	10.7968	7.3563	25.9314	36.9208
0.12	0.0556	8.2181	13.8813	9.1751	28.3626	39.0205
exp-NIG						
0.3	0.0120	5.8826	0.4934	1.4118	7.1938	12.5739
0.28	0.0150	5.9653	1.5720	1.7896	8.7944	14.6661
0.26	0.0187	7.0984	2.6071	2.6548	12.6276	20.4702
0.24	0.0220	7.9856	4.3283	3.5137	15.9070	25.1263
0.22	0.0253	8.5926	5.3388	4.3479	19.1420	30.2218
0.20	0.0327	8.4629	7.8606	5.5347	21.7937	32.5967
0.18	0.0374	9.3739	10.6546	7.0117	26.2039	37.5189
0.16	0.0448	9.1009	12.9371	8.1544	27.8033	38.8505
0.14	0.0538	9.6178	15.5765	10.4408	33.0560	46.0445
0.12	0.0630	9.4007	16.9693	11.7726	33.2136	45.4144

The risk measures are listed in Table 6.17. We can see all risk measures under both of the two asset models are increased with lower TB. The quantile risk measures

include VaR_{99} and CTEs are highly sensitive to TB. Under the GBM real world, the 30 percent terminal bonus cushion strategy has the POSF very close to zero and all quantile risk measures are relatively low. With the TB varying from 0.28 to 0.12, the risks measures raise up quickly. With exp-NIG models, the quantile risk measures are much larger than the figures in corresponding GBM case. We can see actually the model parameter TB in BEP bonus control the risks of hedging and the model bias.

Next we vary the participating rate p. In model office M0 to M7, the participating rate is assumed to be 33 percent.

Table 6.18: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with GBM and exp-NIG models.

p	Mean of asset share	SD of asset share	Mean of guarantee	SD of guarantee
GBM				
0.30	299.5046	157.9309	132.5411	16.4657
0.35	298.5785	159.0623	141.3335	20.5428
0.40	296.7958	155.8530	150.4389	24.8874
0.45	294.3746	154.4075	159.8796	29.8276
0.50	293.6105	155.9446	170.0810	35.2615
0.55	290.0748	150.9077	180.4213	40.6139
0.60	287.0213	148.6899	191.2687	46.6226
0.65	281.9405	145.2967	201.7778	52.4120
0.70	277.9931	140.8753	212.9759	58.9053
0.75	272.3033	135.8590	223.6197	64.5505
exp-NIG				
0.30	298.9050	152.8545	132.5904	15.9393
0.35	297.6670	153.0647	141.2917	19.7854
0.40	296.4259	152.7103	150.4342	24.0957
0.45	295.2280	151.0880	160.1420	28.8716
0.50	292.2692	150.5215	169.9562	34.0225
0.55	290.4331	149.4495	180.6862	39.6643
0.60	285.2387	142.9560	190.8931	44.5485
0.65	282.1919	142.2827	201.9961	50.9533
0.70	278.7846	139.2701	213.1016	57.0111
0.75	272.1692	132.5011	223.5919	62.3650

Table 6.19: Statistics for insurer's loss at maturity for model office M9, hedging based investment strategy, annually rearrange the portfolio, simulated with exp-NIG models, BEP.

p	POSF	MOL	VaR_{99}	CTE95	CTE99	CTE99.5
GBM						
0.30	0.0189	3.8348	2.8697	1.4496	5.9988	7.9707
0.35	0.0263	4.2994	4.6998	2.2615	7.6828	9.7343
0.40	0.0363	4.9607	6.8449	3.6015	10.0519	12.2453
0.45	0.0523	5.6303	9.2397	5.8962	12.8305	15.1732
0.50	0.0717	6.3336	11.6596	8.4764	15.6886	18.4349
0.55	0.0990	7.5058	15.6975	12.0119	19.8983	22.7503
0.60	0.1424	8.5643	19.2456	15.8725	23.8356	26.9706
0.65	0.1939	9.9765	24.6164	20.5435	29.6476	33.2027
0.70	0.2614	11.8121	30.6535	26.3647	36.4472	40.5423
0.75	0.3505	13.8949	38.1856	33.4844	46.0236	51.4953
exp-NIG						
0.30	0.0242	4.8592	4.8151	-0.6717	8.6558	11.1899
0.35	0.0332	5.5866	6.9969	2.5727	11.0061	13.7432
0.40	0.0466	6.4553	9.9267	5.9876	14.4601	17.5469
0.45	0.0587	7.2481	12.4756	8.3950	17.1115	20.2936
0.50	0.0826	7.8151	15.2493	11.4271	20.2917	23.8698
0.55	0.1114	9.0316	19.1656	15.2209	25.4646	29.9313
0.60	0.1506	10.1690	23.8420	19.4053	29.8077	33.9553
0.65	0.2024	11.4163	28.2713	23.7858	34.5857	38.9422
0.70	0.2653	13.1582	34.1395	29.3857	41.2294	46.1850
0.75	0.3500	15.0733	41.3944	36.0659	49.3263	54.7728

Compare with diversified EBR and 100 % EBR cases

By studying the model office M9, we find that a investment strategy based on hedging can reduce the maturity risks of the insurer, under both GBM and exp-NIG real-world models. So far we have discussed three types of investment strategies, the 100% EBR investment strategies, the diversified EBR strategies and the hedging based ones.

The hedging based investment strategies need the support capital which, in this example, we assume it is the guarantee charge deducted from the asset share. The charges income is the free capital for the insurer which in our assumption is deducted from the asset share. The insurer can use this extra capital to improve their investment portfolios. Consider the same support capital for the 100% EBR cases and the dynamic EBR cases. There are two simple ways to invest the free asset, in bond or in share, after it is once deducted from the asset share. When a shortfall happen, the maturity value of the free asset portfolio will be paid to the policyholders to cover part of the loss. Table 6.20 shows the risk measures of insurer's loss at maturity when supported by the free asset.

For each investment strategy, investing free asset in bond scenario always gives lower risk measures than investing free asset in share. Since in most cases when shortfall happens, the share returns are poor and it leads to low free asset value at maturity.

6.5 Concluding Remarks

As a reminder, we repeat the description of Case M7, M8 and M9 considered in this Chapter:

M7 investment strategies based on diversification, EBR less than one;

M8 hedging strategies for a 10-year vanilla type UWP contract with fixed annual bonus rate; and

M9 partial hedging based strategies which has limited support free capital, only part of UWP fund can be hedged.

Table 6.20: UWP simulation using exponential NIG and GBM models.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
100% EBR	M0					
Loss (GBM)	0.0722	19.1296	9.1592	36.0618	25.6835	45.6409
Loss (exp-NIG)	0.0946	21.4323	16.7054	44.1215	33.4324	53.7207
100% EBR	free asset in share					
Loss (GBM)	0.0463	17.8760	2.0865	28.2637	16.5532	38.9054
Loss (exp-NIG)	0.0633	20.3493	6.2396	36.9275	24.9427	47.6710
100% EBR	free asset in bond					
Loss (GBM)	0.0340	14.1086	0	18.6391	9.5939	28.2182
Loss (exp-NIG)	0.0487	16.4276	0	26.6988	16.0097	36.2980
Static EBR(76%)	free asset in share					
Loss (GBM)	0.0173	7.6941	0	7.5988	2.6622	18.5990
Loss (exp-NIG)	0.0246	9.7862	0	14.4110	4.8148	25.5879
Static EBR(76%)	free asset in bond					
Loss (GBM)	0.0105	5.9982	0	0.4699	1.2596	10.2208
Loss (exp-NIG)	0.0164	8.0560	0	6.8302	2.6424	16.6423
Dynamic EBR(70%)	free asset in share					
Loss (GBM)	0.0074	3.9833	0	0	0.5895	5.1460
Loss (exp-NIG)	0.0134	6.0802	0	3.2360	1.6295	10.8071
Dynamic EBR(70%)	free asset in bond					
Loss (GBM)	0.0032	6.1302	0	0	0.3923	1.9617
Loss (exp-NIG)	0.0064	9.0275	0	0	1.1555	5.7776

Supporting by the maturity free estate to the UWP fund, the risk measures using 100% EBR investment strategy are reduced. Actually adding the maturity free asset just shifts the location of the loss distribution to the left by a random variable FA(t).

Compare the other two investment strategies, diversified investment strategy (dynamic EBR) and hedging strategy M8. The former method tries to reduce the variance of the invest return distribution while sacrificing the possible high return from risky asset investment. Hedging investment strategy, however, narrows the left tail of the maturity loss distribution by paying an extra amount of expenses, which is called the "market price of risk", see Hull (1999). Both of the loss distributions under these two investment strategies have less volatility than 100% EBR case.

The hedging based investment strategy M9 in our example is actually a combination of these two investment strategies. Sometimes the 100% UWP fund invested in share is too expensive to hedge. Thus we try to reduce the market risk of the equity investment part of asset share by paying the fund cost supplied by free asset and using a dynamic EBR strategy under principle of given the "affordable highest expected return". The EBR of the UWP fund in M9 is more flexible than in M7, in the case when the contract is "out-of-money", there is large EBR in the fund which leads to higher expected future return in this case.

The model offices M8 and M9 hedge the liabilities in a bond-share market and the insurer does not use derivatives because there is few long term options in the market. If we assume that there exist some long-duration tradeable options written on share indices, a direct alternative investment strategy would be to use the free asset to buy put options to achieve the maturity guarantee on the final step instead of rearrange the hedging quantities. For example, assume now there exist one year term put options on same asset as asset share invested in, this would remove the random hedging errors at maturity. In model office M9, if we can buy one year term put $p(S_t, K(t), T, t)$ in equation 6.31 at $t = 9$ (start of 10th policy year), then the maturity guarantee will be achieved with probability one.

Chapter 7

Extension of Lévy models

7.1 Introduction and objectives

In the previous chapters, the risk in the return of asset investment was observed by Lévy driven asset models. The risk management of life offices selling the UWP policies were studied by applying different bonus mechanisms and frequencies, investment strategies include dynamic EBR rearrangement and simple hedging based asset allocations.

However, as mentioned before, there is other big category of stylised facts for financial data have not been tackled so far in the Lévy models described in this thesis, which are the stochastic volatility and clusters of the extreme returns.

Evidence of so-called "long range dependence" in autocorrelation function of the daily returns has been documented in many studies (see Kirman (2006) for the definition of long/short range dependence). Ding *et al.* (1993) show the absolute value of S&P500 returns has the long-memory property in which the sample autocorrelation function of absolute returns decays very slowly and remains significant even at long time lags.

For the monthly SGTRI data, the nonlinear autocorrelation of the return process include the absolute return and squared return series have positive autocorrelations (see section 7.2). Thus i.i.d. increment Lévy models are not able to capture these features, some extension are required.

In order to model the volatility clustering property in the SGTRI data, the ex-

ponential Lévy model can be extended by adding in time series properties in the long term. Two possible approaches are GARCH/ARCH type models and stochastic volatility type models, driven by the Lévy processes. Eberlein *et al.* (2001) studied volatility models include GARCH, exponential AR, composite and implied volatility in combination with the hyperbolic process and tested their performance on daily German stock index data fitting. Barndorff-Nielsen and Shephard (2001) proposed a stochastic volatility model which is a Ornstein-Uhlenbeck process driven by inverse Gaussian Lévy process. Both of these two models can be represent as the time-homogeneous Lévy processes models.

Also, there are existing asset models that have been proved to be successfully capture key features of the returns of the long-term financial asset. For a summary and comparison of the statistical properties of some of these models we refer to Lee and Wilkie (2000). Other example of long term asset model is the Barrie & Hibbert Economic Scenario Generator (see Hibbert *et al.* (2001)). We consider the multi-variable Wilkie model described in Wilkie (1995). The total return process in Wilkie model are driven by the normal distributed random factors. We will discuss the possibility to use GH processes in the Wilkie model.

The estimation of these models will be discussed with the parameters fitted to SGTRI series. And the risk measures for a 10 years single premium UWP policy loss will be investigated. The chapter is organised as follows.

In section 7.2, the empirical non-linear autocorrelation, specifically, the volatility behaviour of SGTRI is studied. Then we build ARCH/GARCH type stochastic volatility models in combination with the basic exp-Lévy model and describe the estimation procedure in section 7.3. In section 7.4 we apply the stochastic bridges driven by Lévy processes in multi-variable models. We test the statistical properties of the de-volatilized residuals of the fitted models and show backtest of VaR riskmeasures in section 7.5. Finally, in section 7.6, the riskmeasures for insurer are simulated using these more realistic asset models.

7.2 Volatility clustering

A typical empirical behaviour of daily stock returns is that the large movement (upward or downward) are more likely to be followed by another large movement. There are some periods of time in which the fluctuation of asset value or economic indices is more active than the others when during the "quiet" time period, returns are stable and change relatively slow. It gives the evidence of the volatility clustering property for the return processes. The empirical study of the SGTRI data in Chapter 1 shows the autocorrelation of the monthly SGTRI log-return processes is insignificant only with small positive value for lag of one. We plot the autocorrelation function of the squared returns of monthly SGTRI and absolute returns of SGTRI in Figure 7.1. The subplot on the left hand side is the autocorrelation values for squared SGTRI23 monthly log returns. The autocorrelation function stays positive and decays slowly with the first 5 smallest time lags are higher than the horizontal line which is the bound of the 95% confidence interval of 0 autocorrelation hypothesis. Same pattern can be found in the right hand side subplot which is the autocorrelations of absolute SGTRI23 monthly log returns that the autocorrelations for the time lags from 1 month to 16 months are lying outside of the 95% confidence interval.

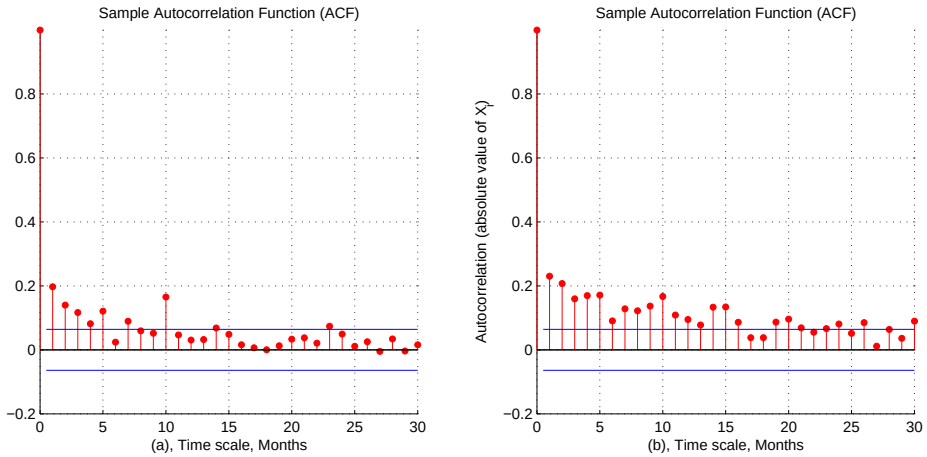

Figure 7.1: Left: Sample autocorrelation function for squared SGTRI23 monthly log returns. Right: Sample autocorrelation function for absolute SGTRI23 monthly log returns. Horizontal solid lines give the 95% confidence interval.

7.3 Stochastic volatility models and Lévy process

We consider the similar univariate asset model as the exp-Lévy models. By $X = (X_t)_{t\geq 0}$ we denote the return process of share indices $S = (S_t)_{t\geq 0}$, i.e.

$$S_t = S_0 \exp(X_t). \tag{7.1}$$

The return process X_t is of the form

$$dX_t = \mu dt + \sigma_{t-} dL_t, \tag{7.2}$$

where L_t is a Lévy process with stationary, independent increment starting at 0 and σ_{t-} denotes a previsible process referred as the instant volatility.

For simplicity, the σ_t can be modelled as discrete-time time series model. One possible discrete time stochastic volatility model is the GARCH(1,1)-m process, see Eberlein *et al.* (2001).

7.3.1 GARCH(1,1)-GH

Given a time scale Δt, we work with the discrete analogue of the Equation 7.2

$$\Delta X_t = \mu \Delta t + \sigma_t \Delta L_t, \tag{7.3}$$

The volatility σ_t is piecewise constant between integer time points and is modelled by GARCH(1,1)-m processes,

$$\sigma_t^2 = c + a\sigma_{t-\Delta t}^2 (\Delta L_t - m)^2 + b\sigma_{t-\Delta t}^2, \tag{7.4}$$

where parameters c is the weighted long term mean of the σ_t,

$$c = (1 - a - b)\sigma_L^2, \tag{7.5}$$

where σ_L^2 is the long term mean of the volatility process and $m = \mathbb{E}(\Delta L_t)$. Also under the monthly time scale we assume that $\mathrm{Var}(\Delta L_t) = 1$

The domain of the parameters is,

$$\sigma_1 \geq 0, c \geq 0, a \geq 0, b \geq 0 \quad and \quad a + b \leq 1. \tag{7.6}$$

The driving process L_t has the GH distributed increment law or it's subclasses and limiting cases include the normal, normal inverse Gaussian, hyperbolic, variance Gamma and student-t distributions.

The GARCH(1,1)-GH model is given by:

$$S_t = S_0 \exp\{\mu t + \sigma_t(\beta T_t + W(T_t))\}, \tag{7.7}$$

where T_1 has GIG$(\lambda, \delta, \gamma)$ law given in 2.10. The driving process

$$L_t = \beta T_t + W(T_t)$$

is GH process. We also consider two subclasses of the GARCH(1,1)-GH model which are fixing the GARCH(1,1)-NIG model when λ is -0.5, and GARCH(1,1)-HYP model when λ is 1.

The GARCH(1,1)-VG model is given by:

$$S_t = S_0 \exp\{\mu t + \sigma_t(\theta T_t + (1 - \theta^2\nu)W(T_t))\}, \tag{7.8}$$

where T_t is a gamma process which T_1 is Gamma(ν,1) distributed and σ_t is given by equation 7.4. We note the VG process has 3 parameter $\{\theta,\ \sigma,\ \mu\}$ and in GARCH(1,1)-VG model, the σ parameter is replaced by the term $1 - \theta^2\nu$. It makes the standard deviation of the de-volatized VG process L_1 to be 1.

The GARCH(1,1)-GBM model is given by:

$$S_t = S_0 \exp\{\mu t + \sigma_t W_t\}, \tag{7.9}$$

where σ_t is the GARCH(1,1)-m type stochastic volatility process.

The linear drift term μt of the GARCH(1,1)-GH model in equation 7.7 is taken outside of the GARCH volatility part. The reason we do not make the drift a GARCH process is that the GARCH typed linear drift leads to positive linear autocorrelation of the return for all lags (because the Lévy processes term have zero autocorrelation and the ACF of $\mu\sigma_t$ will be positive). It is inconsistent with the shape of empirical autocorrelation function in Figure 1.3.

To apply the GARCH volatility on the time changing Brownian motion term, on the other hand, leads to another model setting choice,

$$S_t = S_0 \exp\{\mu t + \beta T_t + \sigma_t W(T_t)\}, \tag{7.10}$$

We prefer the form in Equation 7.7 because one may like to make the process σ_t more intuitive by referring it to the standard deviation of the returns instead of only acting as a scale parameter of the time changing Brownian motion as in Equation 7.10.

7.3.2 Likelihood Function of the GARCH(1,1)-GH

Next, we look at the likelihood function of the GARCH(1,1)-GH models. The observations are given by monthly log-return rate $x_1, x_2, ...x_n$ in order. Here the difference from the estimation of the independent increment models in Chapter 3 is that the monthly log-returns $x_1, x_2, ...x_n$ are not i.i.d. random variables but the whole sample series is a single correlated sample sequence. The general form of the GARCH-Lévy log-return increment $\log(S_t/S_{t-1})$ is:

$$x_t = \mu + \sigma_t(L_t - L_{t-1}). \tag{7.11}$$

Given parameters of the return increment $\Lambda = \{\lambda, \alpha, \beta, \delta, a, b\}$, the likelihood function for observations is:

$$L(\Lambda) = \prod_{i=1}^{n} f(x_i - \mu|\Lambda, \mathcal{F}_{i-1}). \tag{7.12}$$

where $f(y|\Lambda, \mathcal{F})$ is the density of x_i.

The initial condition $(\mathcal{F}_0)$ of the volatility σ_0 is assumed to be the standard deviation of the sample observations and x_0 is assumed to be the mean of sample observations. At time i, the σ_i is upgraded by applying the equation 7.4. Thus, the likelihood function 7.12 can be calculated recursively. The contribution to likelihood of the i-th observation is,

$$\frac{1}{\sigma_i}\text{gh}\left(\frac{x_i - \mu}{\sigma_i}|\Lambda, \mathcal{F}_{i-1}\right), \tag{7.13}$$

where $\text{gh}(y|\Lambda)$ is the density of GH law in equation 2.12.

7.3.3 Maximum Likelihood Estimation

For each model, adding a GARCH(1,1) stochastic volatility model leads to two additional parameter, a and b. The parameter c is calculated by

$$c = (1 - a - b)\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}. \tag{7.14}$$

The maximum likelihood estimation results for the GARCH(1,1)-GH, GARCH(1,1)-NIG and GARCH(1,1)-HYP are shown in Table 7.1. The variance and correlation between the estimators are given approximately by calculating the information matrix numerically.

Table 7.1: Maximum likelihood estimates of GARCH(1,1)-GH and the following subclasses: NIG ($\lambda = -1/2$) and HYP ($\lambda = 1$). The first line gives the maximum likelihood estimates of all 7 parameters where λ has not been fixed.

λ	α	β	δ	μ	a	b	LogLH
Share Gross Total return Jan 1950 - May 2005							
-2.3066	0.5272	-0.3543	1.6999	0.0256	0.1463	0.7757	1100.32
NIG	1.2562	-0.3671	1.1687	0.0261	0.1479	0.7721	1099.90
HYP	1.7628	-0.3805	0.6741	0.0266	0.1468	0.7723	1099.28
Share Gross Total return Dec 1923 - May 2005							
-1.8626	0.6973	-0.2298	1.6504	0.0179	0.1818	0.7768	1706.51
NIG	1.1886	-0.2319	1.2269	0.0180	0.1820	0.7767	1706.31
HYP	1.6528	-0.2303	0.6814	0.0179	0.1818	0.7772	1705.72

Table 7.2: Maximum likelihood estimates of GARCH-VG and GARCH-GBM.

	a	b	μ	ν	θ	LogLH
Share Gross Total Return Jan 1950 - May 2005						
GARCH-VG	0.1384	0.7762	0.0261	0.5015	-0.3346	1098.37
GARCH-GBM	0.1288	0.8040	0.0103			1064.11
Share Gross Total Return Dec 1923 - May 2005						
GARCH-VG	0.1685	0.7860	0.0171	0.5164	-0.2069	1704.86
GARCH-GBM	0.1296	0.8360	0.0086			1644.80

The seven parameter GARCH(1,1)-GH model has the highest log-likelihood, which is 1100.32 and 1706.51 for SGTRI 50-05 and SGTRI 23-05 data. The MLE results for GARCH(1,1)-VG and GARCH(1,1)-GBM models are shown in Table 7.2. Compared with the GARCH(1,1)-GBM, which is 1064.11 and 1644.80 for SGTRI 50-05 and SGTRI 23-05, all the GARCH type stochastic volatility models with

non-Gaussian marginal returns offer a considerable improvement in likelihood. The difference between the subclasses include NIG and HYP by fixing λ and the general case is small. Recall the likelihood of the constant volatility GBM model is 1013.10 and 1557.55 for SGTRI 50-05 and SGTRI 23-05. And the likelihood of the constant volatility GH model is 1075.18 and 1650.90 for SGTRI 50-05 and SGTRI 23-05. It is obvious that using GARCH stochastic volatility models can capture the autocorrelation structure of the SGTRI data.

We can see that the estimation values of b in GARCH-GH, GARCH-HYP, GARCH-NIG and GARCH-VG are all very close to 0.77. Moreon, the a and b estimations are stable for each sample set. Changing the types of driving processes within GH family has little effect on volatility time series.

7.3.4 MCMC for GARCH-GH models

The Metropolis-Hastings algorithms we used to estimate the parameters of the GH-BDLP in GARCH(1,1)-GH are very similar to those we applied in Chapter 3 Section 3.4.

The parameters a and b are constrained to lie in $(0, 1)$. The prior distributions used for the parameter a and b are Beta$(0.6, 2.4)$ and Beta$(3, 1.275)$ giving the prior mean of 0.2 and 0.7 and standard deviation of 0.2 and 0.2 respectively for a and b.

The proposal distributions are also Beta, with $C_a \in$ Beta$(0.2, 0.8)$ for a and $C_b \in$ Beta$(1.167, 0.5)$ for b. The proposal distributions have the same means as the prior distributions with bigger standard deviations. The acceptance rates are approximately 33 percent and 36 percent for SGTRI23 sample and approximately 32 percent and 34 percent for SGTRI50 sample data respectively.

For the prior and proposal distributions of the GH MCMC estimators are similar to those we used in Chapter 3, only with different distribution parameters.

The HYP and NIG distributions are subclasses of the GH. They share the same parametrisation but with fixed λ, i.e., $\lambda = -0.5$ for the NIG and $\lambda = 1$ for the HYP. Normal prior distribution will be used for λ with large variance for GH.

Table 7.3 gives the prior and proposal distributions and the acceptance probabilities for each model fitted to respectively two data sets, SGTRI50 and SGTRI23.

Table 7.3: Prior&Proposal distributions for GH MCMC

	Prior	Proposal	AP(SGTRI50)	AP(SGTRI23)
GH				
λ	$N(0, 0.60^2)$	$N(\lambda^{(r)}, 0.6^2)$	26	22
α	$\Gamma(0.25, 0.25)$	$\Gamma(0.25, 0.25/\alpha^{(r)})$	30	26
β	$N(0.0, 0.5^2)$	$N(\beta^{(r)}, 0.3^2)$	19	19
δ	$MN(0.05, 0.1^2)$	$N(\delta^{(r)}, 0.07^2)$	48	46
μ	$N(0.1, 0.3^2)$	$N(\mu^{(r)}, 0.01^2)$	21	20
a	Beta$(0.6, 2.4)$	Beta$(0.2, 0.8)$	32	33
b	Beta$(3, 1.275)$	Beta$(1.167, 0.5)$	34	36
NIG				
α	$\Gamma(0.25, 0.25)$	$\Gamma(0.25, 0.25/\alpha^{(r)})$	41	37
β	$N(0.0, 0.5^2)$	$N(\beta^{(r)}, 0.3^2)$	38	35
δ	$MN(0.05, 0.1^2)$	$N(\delta^{(r)}, 0.07^2)$	30	27
μ	$N(0.1, 0.3^2)$	$N(\mu^{(r)}, 0.01^2)$	29	24
a	Beta$(0.6, 2.4)$	Beta$(0.2, 0.8)$	33	35
b	Beta$(3, 1.275)$	Beta$(1.167, 0.5)$	34	37
HYP				
α	$\Gamma(0.45, 0.05)$	$\Gamma(0.25, 0.25/\alpha^{(r)})$	37	36
β	$N(0.0, 0.5^2)$	$N(\beta^{(r)}, 0.3^2)$	39	35
δ	$MN(0.05, 0.1^2)$	$N(\delta^{(r)}, 0.07^2)$	31	28
μ	$N(0.1, 0.3^2)$	$N(\mu^{(r)}, 0.01^2)$	30	25
a	Beta$(0.6, 2.4)$	Beta$(0.2, 0.8)$	33	35
b	Beta$(3, 1.275)$	Beta$(1.167, 0.5)$	34	37

We use same prior and proposal for SGTRI50 and SGTRI23. The acceptance probabilities(AP) of the candidate parameter are not sensitive to these two sample sets. The AP(SGTRI50) is often higher than AP(SGTRI23) under same prior and proposal distributions for each parameter.

7.3.5 The results

Table 7.4 shows the results for the GARCH(1,1)-GH distribution. The most of the λ samples lie within a negative interval around -1.1 for SGTRI23. We note that the λ estimator of the GARCH-Lévy model has a relatively high standard deviation. Table 7.5 shows the correlation matrix of the simulated samples. The driven process parameter $\{\lambda, \alpha, \beta, \delta, \mu\}$ is loosely correlated to the volatility structure parameters $\{a, b\}$. It suggests that round the proper MCMC estimator of the stochastic volatility model parameters a and b, the estimators are robust. On the other hand, one can see the dependence structures which are modelled by GARCH time series cannot be captured by the driving process. If the volatility clustering effect can be modelled by using GH processes alone, changing GARCH parameters a, b will lead to a driving process paramization shift. λ and δ are extremely strong negatively correlated , i.e., around -0.96. And α is positively correlated to λ, i.e., over 0.85. It suggests that the model can equally perform well in a large range of λ as long as choosing proper α and δ.

Table 7.6 shows the estimation results for the GARCH(1,1)-NIG. Same patterns applies for two sample set. The SGTRI23 estimators often have less standard deviations than the SGTRI50 ones. Compare with the general seven parameter GH estimation, the SD for α is reduced because the fixed λ. We use the medium of each posterior sample as the "best estimator", for SGTRI23 the likelihood is 1705.98.

Table 7.7 shows the estimation results for the HYP distribution. We use the medium of each posterior sample as the "best estimator", for SGTRI23 the likelihood is 1704.56.

The MCMC estimators for GARCH(1,1)-VG model and their correlation is shown in Table 7.8. Also it shows loose correlation between the volatility model parameters a and b and the driven VG process parameters.

Table 7.4: Markov Chain Monte Carlo of GARCH-GH.

	λ	α	β	δ	μ	G_α	G_β
Share Gross Total return Jan 1950 - May 2005							
Mean	-1.6378	1.0710	-0.4067	1.5598	0.0273	0.1628	0.7173
SD	1.5990	0.5472	0.1275	0.5165	0.0044	0.0482	0.0945
Share Gross Total return Dec 1923 - May 2005							
Mean	-1.1298	1.0541	-0.2470	1.4521	0.0200	0.1934	0.7523
SD	1.8526	0.5616	0.0791	0.5497	0.0027	0.0364	0.0503

Table 7.5: Parameters Correlations for G-GH

	λ	α	β	δ	μ	a	b
Share Gross Total Return Jan 1950 - May 2005							
λ	1.0000	0.8590	0.0186	-0.9560	-0.0060	0.0317	-0.0333
α		1.0000	-0.3031	-0.7309	0.2868	-0.0472	0.0236
β			1.0000	-0.0609	-0.9179	0.1561	-0.1141
δ				1.0000	0.0641	-0.0128	0.0537
μ					1.0000	-0.1236	0.1071
a						1.0000	-0.8115
b							1.0000
Share Gross Total Return Dec 1923 - May 2005							
λ	1.0000	0.8767	0.0380	-0.9632	-0.0373	0.0095	-0.0087
α		1.0000	-0.1197	-0.7798	0.0963	-0.0320	0.0162
β			1.0000	-0.0705	-0.8956	0.0406	0.0300
δ				1.0000	0.0744	0.0082	0.0115
μ					1.0000	0.0047	-0.0714
a						1.0000	-0.9167
b							1.0000

Table 7.6: Markov Chain Monte Carlo of G-NIG

	Mean	SD	Parameters correlations with:					
			α	β	δ	μ	a	b
Share Gross Total Return Jan 1950 - May 2005								
α	1.4391	0.2298	1.0000	-0.5779	0.7860	0.5051	-0.1219	0.0737
β	-0.4054	0.1263		1.0000	-0.2079	-0.9153	0.0717	-0.0378
δ	1.1933	0.1579			1.0000	0.2367	0.0180	0.0964
μ	0.0270	0.0043				1.0000	-0.0318	0.0146
a	0.1413	0.0455					1.0000	-0.8108
b	0.7590	0.0966						1.0000
Share Gross Total Return Dec 1923 - May 2005								
α	1.2368	0.1640	1.0000	-0.3573	0.8117	0.2682	-0.0924	0.0387
β	-0.2409	0.0712		1.0000	-0.1012	-0.8700	-0.0074	0.0688
δ	1.1804	0.1316			1.0000	0.0913	0.0611	0.0320
μ	0.0180	0.0024				1.0000	0.0547	-0.1131
a	0.1866	0.0365					1.0000	-0.8996
b	0.7587	0.0529						1.0000

Table 7.9 lists the logarithm of margin likelihood of the GARCH-GH models we have discussed.

Table 7.7: Markov Chain Monte Carlo of G-HYP

	Mean	SD	Parameters correlations with:					
			α	β	δ	μ	a	b
Share Gross Total Return Jan 1950 - May 2005								
α	1.8810	0.1778	1.0000	-0.5604	0.7904	0.4625	-0.1772	0.0773
β	-0.3984	0.1153		1.0000	-0.2485	-0.9011	0.1070	-0.0748
δ	0.6833	0.1959			1.0000	0.2554	-0.0251	0.0951
μ	0.0269	0.0040				1.0000	-0.0620	0.0513
a	0.1483	0.0465					1.0000	-0.8051
b	0.7475	0.0966						1.0000
Share Gross Total Return Dec 1923 - May 2005								
α	1.7210	0.1394	1.0000	-0.4341	0.8178	0.3582	-0.1659	0.0471
β	-0.2504	0.0783		1.0000	-0.1873	-0.8914	0.0214	0.0539
δ	0.6414	0.1748			1.0000	0.1752	-0.0277	0.0556
μ	0.0182	0.0026				1.0000	0.0212	-0.0991
a	0.1839	0.0346					1.0000	-0.8881
b	0.7621	0.0498						1.0000

Table 7.8: Markov Chain Monte Carlo of GARCH-VG.

	Mean	SD	Parameters correlations with:				
			μ	θ	ν	a	b
Share Gross Total Return Jan 1950 - May 2005							
μ	0.0247	0.0036	1.0000	-0.9014	-0.4013	-0.0023	0.0490
θ	-0.3000	0.0813		1.0000	0.4450	-0.0021	0.0618
ν	0.5183	0.0984			1.0000	-0.0529	-0.0639
a	0.1852	0.0481				1.0000	-0.8845
b	0.6964	0.0921					1.0000
Share Gross Total Return Dec 1923 - May 2005							
μ	0.0190	0.0024	1.0000	-0.8784	-0.3489	0.0398	-0.0299
θ	-0.2088	0.0632		1.0000	0.3692	-0.0239	-0.0012
ν	0.5421	0.0922			1.0000	-0.0552	-0.0323
a	0.1763	0.0336				1.0000	-0.9159
b	0.7544	0.0542					1.0000

Table 7.9: Logarithm of marginal likelihood for GH, NIG, HYP, VG, Student-t and GBM.

Model	Log(MLH) Jan 1950 - May 2005	Log(MLH) Jan 1923 - May 2005
VG	1093.96	1696.67
HYP	1094.46	1696.89
NIG	1095.16	1697.26
GH	1095.10	1697.00

7.4 Multi-Variable model

In this section, we look at the multi-variable models. The Wilkie model is widely used by actuaries in UK. It is designed in order to capture the long term features of the assets returns. The Wilkie model is defined on annual intervals. To the purpose of the risk management for the UWP which could be based on monthly or more frequent time intervals, for example, the bonus declaration and hedging, the "stochastic bridges" are built in order to extend the model to a continuous model between the "gaps" of the each year.

Before constructing the stochastic bridges for the Wilkie model, we need to first calibrate the yearly Wilkie model to the full series of the data which is SGTRI23.

7.4.1 Update the Wilkie Model paramizations

The parameters given in Wilkie (1995) is fitted to the data from 1923 to 1995. Here we adopt the same notation system described in Wilkie (1995). We upgrade the parameters and fit the first three indices which are the inflation rate $I(t)$, the share yield $Y(t)$ and the dividend $D(t)$. This is because the Wilkie model has a cascade structure and in order to get the share total investment return rate, which is

$$r(t, \Delta t) = \frac{S_{t+\Delta t} + \sum_{s \in (t, t+\Delta t]} D(s)}{S_t}, \tag{7.15}$$

requires the models for the inflation rate, the share yield and the dividend. Table 7.10 gives the upgraded maximum likelihood estimators of the Wilkie model up to the 2004 data. We use the same estimation method and boundary conditions

as described in Wilkie (1995). Except for the dividend process $D(t)$, the original Wilkie model used a long series of data before 1922 to run in the $DM(t)$ series. We were not able to get those data and we estimate starting value $DM(0)$ as an extra parameter to maximise the likelihood. Table 7.10 shows the estimation results. First the suggest parameter values given in Wilkie (1995) are listed in column "Wilkie95". Then we re-calibrate the model to 1923-1994 data and the estimators are given in column "1923-1994". The 1923-1994 results are very close to the results in Wilkie (1995). The last column "1923-2004" shows the parameter estimators for 1923-2004 data set.

Table 7.10: Maximum likelihood estimates of Wilkie model.

Parameter	Wilkie95	1923-1994	1923-2004
Retail Price Index			
QA	0.58	0.5773	0.5791
QMU	0.047	0.0473	0.0447
QSD	0.0425	0.0425	0.0403
Share Yield			
YW	1.8	1.7940	1.6782
YA	0.55	0.5492	0.6235
YMU	0.0375	0.0377	0.0367
YSD	0.155	0.1551	0.1542
Dividend			
DW	0.58	0.5821	0.5797
DD	0.13	0.1370	0.1577
DMU	0.016	0.01569	0.0133
DY	-0.175	-0.1765	-0.1516
DB	0.57	0.5721	0.6110
DSD	0.07	0.0671	0.0653

7.4.2 Wilkie Model with Lévy bridge

There is no analogue in continuous time to cascade time series structure of the Wilkie model. The Brownian bridges of the Wilkie model is discussed in Wilkie *et al.* (2003) and Wilkie *et al.* (2005). We construct the stochastic bridge using

Lévy processes for the Wilkie model.

Suppose that we have simulated values of the process X_t at time t_i and t_k. The increment $X_{t_k} - X_{t_i}$ is random variable with distribution F, density f. We are interested in interpolate the value of X_t at any time t between time t_i and t_k with conditional distribution of the increment $X_t - X_{t_i}$, given the values of X_{t_k} and X_{t_i}.

Set time scale to be yearly and starting from July of each year, Y_t is the total share return process simulated by the Wilkie model at each integer period. Assume X_t is the total share log-return process. Continuous stochastic bridges driven by Lévy processes can be built between each integer time step with the condition that

$$X_t = Y_t, \text{ when } t = 1, 2, 3 \cdots ;$$

and X_t is continuous.

One of the method to interpolate the X_t is to combine two stochastic processes by adjusting the linear drift term. For $t_i \leq t \leq t_j$, suppose Z_t is some continuous Lévy process and we have simulated the value of Y_t at time t_i and t_j and Z_t between time period (t_i, t_j). The value of X_t at time t_i is given by

$$X_{t_i} = Y_{t_i}.$$

We then compare the increments of these two processes, calculate the difference, h:

$$h = (Y_{t_j} - Y_{t_i}) - (Z_{t_j} - Z_{t_i}); \tag{7.16}$$

and calculate the drift adjustment, μ_{t_i}:

$$\mu_{t_i} = \frac{h}{t_j - t_i}. \tag{7.17}$$

We then add the drift adjustment μ_{t_i} and increment of Z_t to the X_t,

$$X_t = X_{t_i} + Z_t - Z_{t_i} + \mu_{t_i}(t - t_i), \ t_i \leq t \leq t_j.. \tag{7.18}$$

Then we have X_t is locally a Lévy process between time period (t_i, t_j) with $X_{t_i} = Y_{t_i}$ and $X_{t_j} = Y_{t_j}$.

Re-calibrate the model

The parameters of the bridging-Lévy processes need to be re-calibrated to the same data base that used to estimate the Wilkie model. $\underline{x} = x_0, x_2, \cdots, x_{984}$ are the

sample observations of the monthly share total return index starting in June 1922 and include the data up to June 2004, gives 985 values. x_0 is the total share return index at the end of and x_{984} is the index at the end of June 2004. Note that this sample set is different from the SGTRI23 data which the monthly total share returns lie in the time horizon from December 1923 to April 2005. There are 82 yearly log-return rates

$$y_i = \log\left(\frac{x_{12i}}{x_{12(i-1)}}\right), \quad i = 1, 2, ..., 82. \tag{7.19}$$

The residuals of the linear interpolation in $i - th$ year are

$$\tilde{z}_{i,j} = \log\left(\frac{x_{12(i-1)+j}}{x_{12(i-1)+j-1}}\right) - \frac{y_i}{12}, \ j = 1, 2, ..., 12. \tag{7.20}$$

Let $k = 12(i-1) + j$, we define the series of the residuals

$$z_k = \tilde{z}_{i,j}. \tag{7.21}$$

The statistics of the monthly returns residuals z are shown in Table 7.11. The skewness for SGTRI23 is 0.18443 and kurtosis is 11.65396.

Table 7.11: Statistics for log-return of the SGTRI

Statistics	Mathematical Form	x_k 06/22-06/04
Mean	$\mathbb{E}[Y]$	0.00000
Standard Deviation	$m_2^{(1/2)}$	0.04645
Skewness	$m_3/m_2^{3/2}$	0.18443
Kurtosis	m_4/m_2^2	11.65396
Maximum	Y_{max}	0.40652
Minimum	Y_{min}	-0.29405

Table 7.12 shows the estimation result for the variance gamma, student t and Gaussian distributions. The estimation of μ and θ in VG case is close to zero. Hence we can simplify the model by setting $\mu = 0$ and $\theta = 0$, we show the estimation of the σ and ν parameters in VG(s) case in Table 7.12. The log-likelihood for the VG case is 1707.96 and 1706.92 for the VG(s) case. With simple calculation, the AIC and BIC for VG(s) is higher than the VG case. The 2-parameter VG(s) case

is acceptable in the likelihood sense. The student t performs better than VG in the residual fitting. This may because the conditional residual is approximately symmetric and Student t is more flexible in modelling the tail behaviours. Gaussian estimation gives the lowest log-likelihood.

Table 7.12: Maximum likelihood estimates of conditional monthly residuals using variance gamma (VG), simplified variance gamma (VG(s)), student t and Gaussian.

	μ	σ	ν	θ	LogLH
Share Gross Total return Dec 1923 - May 2005					
VG	0.00281	0.04487	0.80229	-0.00288	1707.96
VG(s)	0	0.04497	0.81061	0	1706.92
Student t	0.00108	0.03202	3.72265		1711.55
Gaussian	0.00000	0.04645			1624.51

The estimation results for generalised hyperbolic, normal inverse Gaussian and distributions are shown in Table 7.13. Compare the result with those in Table 3.1, one can see that parameters μ and β in all three cases are relatively small. The distributions could be simplified by letting $\mu = 0$ and $\beta = 0$ and the estimation of the parameters are given in GH(s), NIG(s) and HYP(s) cases in Table 7.13.

Table 7.13: Maximum likelihood estimates of generalised hyperbolic , normal inverse Gaussian ($\lambda = -1/2$) ,hyperbolic ($\lambda = 1$) distributions and their simplified version by making $\mu = 0$ and $\beta = 0$

λ	α	β	δ	μ	LogLH
basic					
-1.58387	7.31710	-1.64105	0.05709	0.00347	1712.49
NIG	19.13362	-1.75969	0.03958	0.00364	1711.82
HYP	33.34054	-1.80226	0.01287	0.00360	1709.05
$\beta = 0$ and $\mu = 0$					
-1.72630	4.92131	0	0.05960	0	1711.20
NIG	19.07894	0	0.03972	0	1710.36
HYP	33.15751	0	0.01256	0	1707.70

The "linear drift adjustment" method we discussed above is straightforward, easy to simulate and re-calibrate. We can also add more complex time series structures and conditional autocorrelations to the bridge processes. The volatility of the bridge processes can be stochastic such as GARCH(1,1)-m. Time is discretised into 12 time steps in a year, $t = t_0, t_1, ..., t_{12} = t+1$ up to the time $t+1$. The monthly log-return process X_t is

$$X_{t_i} = \mu_t + \sigma_t(L_{t_i} - L_{t_{i-1}}),\ i = 1, 2, ..., 12, \tag{7.22}$$

where L_t can be any suitable driving Lévy processes, and σ_t is the GARCH(1,1)-m volatility time series given in Eqn 7.4. Let μ_t to be the yearly linear drift adjustment simulated by the Wilkie model, we show result of parameter estimation in Table 7.14 and Table 7.15.

Table 7.14: Maximum likelihood estimates of conditional monthly residuals using variance gamma (GVG), simplified variance gamma (GVG(s)), GARCH student t and Gaussian.

	μ	ν	θ	a	b	LogLH
VG	0.0011	0.3739	-0.0495	0.2040	0.7609	1783.20
VG(s)		0.3724	0	0.2046	0.7604	1782.85
Student t	-0.0007	9.8623		0.2514	0.7299	1782.29
Gaussian	-0.0012			0.2087	0.7548	1760.36

Table 7.15: Maximum likelihood estimates of generalised hyperbolic , normal inverse Gaussian ($\lambda = -1/2$) ,hyperbolic ($\lambda = 1$) distributions and their simplified version by making $\mu = 0$ and $\beta = 0$

λ	α	β	δ	μ	a	b	LogLH
-3.47455	0.059314	-0.0593	2.3461	0.0010	0.2147	0.7594	1787.12
NIG	1.3619	-0.0521	1.5175	0.0011	0.2181	0.7551	1785.40
HYP	1.7719	-0.0495	1.0467	0.0011	0.2197	0.7533	1784.91
$\beta = 0$ and $\mu = 0$							
-3.42085	0.0000	0	2.3291	0	0.2154	0.7586	1785.84
NIG	1.3671	0	1.5276	0	0.2192	0.7541	1785.00
HYP	1.7739	0	1.0555	0	0.2206	0.7525	1784.54

Multi-variable model by SDE

Other method to model the multi-variables long term is the continuous investment model driven by Lévy processes due to Chan (1998). We introduce this model here as an alternative of the multi-variable time series model such as the Wilkie model. Calibrating and applying the Lévy SDE models are out of the scope of this thesis, we point it as a further research topic.

Let Z_1, Z_2, Z_3 and Z_4 be four independent Lévy processes. The model has a cascade structure. We start by modelling the consumer price indices:

Retail prices index and inflation Let $Q_t = \exp\{P_t\}$ be retail prices index, where the process P takes form:

$$\begin{aligned} dP_t &= R_t dt \\ dR_t &= -a_1 R_t dt + \phi(t)dt + \sigma_1 dZ_1(t) \end{aligned} \tag{7.23}$$

where R_t is the force of inflation, $a_1 > 0$, $\sigma_1 \in \mathbb{R}$ and ϕ is a deterministic positive periodic function with period $h > 0$. Solving the stochastic differential equation (7.23), we have the explicit formula for R_t:

$$R_t = e^{-a_1 t} R_0 + \int_0^t e^{-a_1(t-s))} \phi(s) ds + \int_0^t \sigma_1 e^{-a_1(t-s)} dZ_1(s).$$

Dividend yield Let Y_t denote the share dividend yield.

$$\begin{aligned} Y_t &= Y_* \exp\{X_t + \zeta R_t\} \\ dX_t &= -a_2 X_t dt + b_1 dt + \sigma_2 dZ_2(t) \end{aligned} \tag{7.24}$$

where $Y_* = Y_0 e^{-(X_0+\zeta R_0)}$, $a_2 > 0$, $\sigma_2 \in \mathbb{R}$ and we have explicit solution for equation(7.24)

$$X_t = X_0 e^{-a_2 t} + b_1 \left(\frac{1 - e^{-a_2 t}}{a_2} \right) + \int_0^t \sigma_2 e^{-a_2(t-s)} dZ_2(s).$$

Share dividend Let D_t be the share dividends.

$$d(\log D_t) = \left(b_2 + \beta\lambda \int_0^t e^{-\lambda s} R_{t-s} ds + \gamma R_t \right) dt + \eta_2 dZ_2(t) + \eta_3 dZ_3(t) \tag{7.25}$$

where $b_2, \beta, \lambda, \eta_2, \eta_3 \in \mathbb{R}$. Interchanging the order of integration one obtains:

$$\lambda \int_0^t \int_0^s e^{-\lambda u} R_{s-u} du ds = \int_0^t (1 - e^{-\lambda(t-u)}) R_u du$$

therefore from 7.25:

$$D_t = D_* \exp\left\{ \eta_2 Z_2(t) + \eta_3 Z_3(t) + \beta \int_0^t (1 - e^{-\lambda(t-u)}) R_u du + \gamma \int_0^t R_u du + b_2 t \right\}$$

where D_* is a constant determined by D_0 and R_0 in a similar manner to Y_*.

Share price The share price S_t in this model is related to the dividends and the yield by $S_t = D_t / Y_t$. Applying Itô's lemma one gets:

$$\frac{dS_t}{dD_t} = \frac{Y_t - D_t \cdot dY_t / dD_t}{Y_t^2}$$

then,

$$\begin{aligned} dS_t &= \frac{dD_t}{Y_t} - \frac{dY_t \cdot D_t}{Y_t^2} \\ &= \frac{d(\log D_t) \cdot D_t}{Y_t} - \frac{dY_t \cdot D_t}{Y_t^2} \\ &= S_t \cdot d(\log D_t) - S_t \cdot d(logY_t) \\ &= S_t \cdot \left\{ \left(b_2 + \beta\lambda \int_0^t e^{-\lambda s} R_{t-s} ds + \gamma R_t \right) dt + \eta_2 dZ_2(t) + \eta_3 dZ_3(t) - dX_t - \zeta dR_t \right\}. \end{aligned}$$

This is consistent with a geometric Lévy model of share prices. We obtain,

$$dS_t = c_t S_t dt + S_t(\delta_1 dZ_1(t) + \delta_2 dZ_2(t) + \delta_3 dZ_3(t))$$

where

$$c_t = b_1 + b_2 + \beta\lambda \int_0^t e^{-\lambda s} R_{t-s} ds + \gamma R_t + a_1 \zeta R_t + a_2 X_t - \zeta\phi(t),$$
$$\delta_1 = \zeta\sigma_1,\ \delta_2 = \eta_2 - \sigma_2,\ \delta_3 = \eta_3$$

Consol yield The yield on consols C_t:

$$C_t = \xi\rho \int_0^t e^{-\rho s} R_{t-s} ds + C_* e^{V_t},$$
$$dV_t = -a_4 V_t dt + \sigma_4 dZ_4(t), V_0 = \nu.$$

The explicit form for V_t can be solved in a similar way to R_t.

7.5 Goodness of fit

7.5.1 Devolatilized residuals

Ideally, one would expect the devolatilized residuals (ΔL_t) defined in Equation 7.3 to be independent, identically distributed random variables. Besides, we would like to know whether these devolatilized residuals are normally distributed. If the volatilized residuals are normally distributed, it is not necessary to use Lévy processes with non-Gaussian margin as driven processes.

Thus, we test the devolatilized residuals in order to observe:

- whether using GARCH volatility models reduce the non-linear autocorrelations substantially for Devolatilized residuals; and
- whether using GARCH volatility models eliminate the non-Gaussian property in residuals.

Given monthly log-return rate $x_1, x_2, ...x_n$ and the parameter estimations of the model, we calculate the devolatilized residuals as

$$\tilde{z}_i = \frac{x_i - \tilde{\mu}}{\tilde{\sigma}_t}, \tag{7.26}$$

where $\tilde{\sigma}_t$ can be calculated recursively by Equation 7.4 corresponding to parameters for each model with different driving processes in this chapter.

In Figure 7.2, we plot the sample autocorrelation functions for squared devolatilized residuals corresponding to:

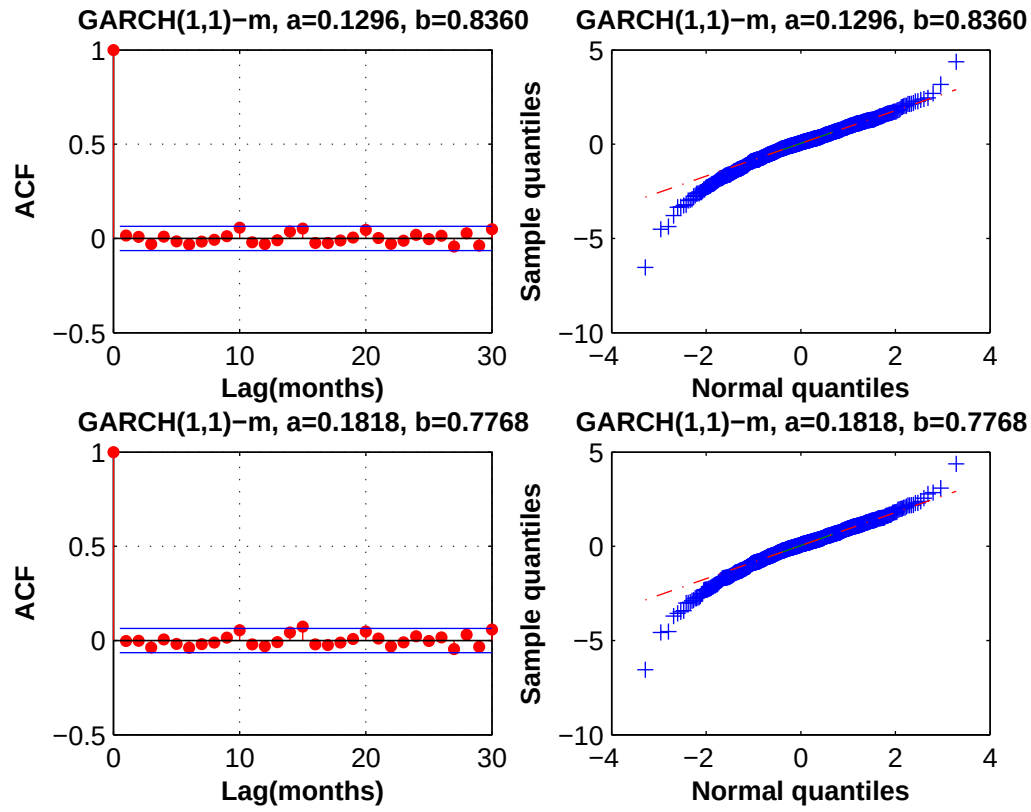

Figure 7.2: Sample autocorrelation function for squared devolatilized residuals of SGTRI23 monthly log returns and QQ plots

GARCH(1,1)(a) GARCH(1,1)-m with $a = 0.1296$ and $b = 0.8360$, this is the case for SGTRI23 data under Gaussian innovations assumption.

GARCH(1,1)(b) GARCH(1,1)-m with $a = 0.1818$ and $b = 0.7768$, these are the estimators for GH driven case, we can see from Table 7.13 that all non-Gaussian GARCH models estimators are similar.

The dotted lines show asymptotic 95% confidence bounds under the null hypothesis that the autocorrelation is zero. One can see from Figure 7.2 that both of the GARCH volatility models reduce the autocorrelation of squared residuals.

The two plots in right column display the QQ plots of the empirical quantiles of devolatilized residuals against quantiles of standard normal distribution. Both of the plots show inverse "S" shape of the quantiles curves which suggest the devolatilized residual distributions have heavier tails than Gaussian distribution.

The results of Jarque-Bera test for GARCH(1,1)(a) and GARCH(1,1)(b) devolatilized residuals, SGTRI and Bridge residuals are shown in Table 7.16. The JB statistics for GARCH(1,1)(a) and GARCH(1,1)(b) are 437.83 and 424.95, the p-values are close to zero and the Gaussian hypothesis are rejected for both cases. The distribution of Bridge residuals has similar JB statistic as the SGTRI23. Non-Gaussianity has been found in all these four samples.

Table 7.16: Jarque-Bera statistics for GARCH(1,1)(a) and GARCH(1,1)(b) devolatilized residuals, SGTRI and Bridge residuals

Sample	GARCH(1,1)(a)	GARCH(1,1)(b)	SGTRI23	Bridge residuals
JB statistic	437.83	424.9	3063.78	3076.11

The BDS test given in Brock *et al.* (1987) is one of the most widely used tests to detect non-linear structure in time series. For parameters $m \in \mathbb{N}$ and $\epsilon > 0$, the BDS statistic is defined as

$$V_{m,T}(\epsilon) = \sqrt{T}\frac{C_{m,T}(\epsilon) - C_{1,T}(\epsilon)^m}{\sigma_{m,T}(\epsilon)}, \tag{7.27}$$

and

$$C_{m,T}(\epsilon) = \sum_{1 \le t < s < T_m} \mathbb{I}_{s,t}(\epsilon)\frac{2}{T_m(T_m - 1)}, \tag{7.28}$$

where $T_m = T - m + 1$ and $\mathbb{I}_{s,t}(\epsilon) = \mathbb{I}_{[-\epsilon,\epsilon]}(\max_{i\in\{0,\ldots,m-1\}}|X_{t+i} - X_{s+i}|)$. $\sigma_{m,T}(\epsilon)$ is a consistent estimator of the asymptotic standard deviation $C_{m,T}(\epsilon) - C_{1,T}(\epsilon)^m$

$$K_T(\epsilon) = \sum_{1\le t<s<r\le T_m} \frac{2(\mathbb{I}_\epsilon(t,s)\mathbb{I}_\epsilon(s,r) + \mathbb{I}_\epsilon(t,r)\mathbb{I}_\epsilon(r,s) + \mathbb{I}_\epsilon(s,t)\mathbb{I}_\epsilon(t,r))}{T_m(T_m-1)(T_m-2)}, \tag{7.29}$$

$$\begin{aligned}\sigma^2_{m,T}(\epsilon) &= 4(K_T(\epsilon)^m + 2\sum_{j=1}^{m-1} K_T(\epsilon)^{m-j} C_{1,T}(\epsilon)^{2j} \\ &\quad +(m-1)^2 C_{1,T}(\epsilon)^{2m} - m^2 K_T(\epsilon) C_{1,T}(\epsilon)^{2m-2}).\end{aligned} \tag{7.30}$$

Under the null hypothesis that $\{X_t\}_{t\in\mathbb{N}}$ is *i.i.d.*, the BDS statistic $V_{m,T}(\epsilon) \to N(0,1)$ for any $\epsilon > 0$ and $m = 2, 3, 4, \ldots$. We use $m = 5$ and $\epsilon = 2\sqrt{X_t}$ and show the results of the test in Table 7.17.

Table 7.17: DBS test on independent returns and devolatilized residuals

Sample	BDS statistic	BDS p-value
SGTRI23	0.7971	42.54%
GARCH(1,1)(a)	0.1563	87.58%
GARCH(1,1)(b)	0.0516	95.88%
Bridge residuals	0.8877	37.47%

Both GARCH(1,1)(a) and GARCH(1,1)(b) models reduce the BDS statistics for the residuals in which the GARCH(1,1)(b) case gives the highest p-value. This is a strong evidence show that the devolatilized residuals under GARCH(1,1)(b) model are *i.i.d.* .

We study the moment statistics of the driving processes. For each process we simulate the monthly returns and calculate numerically their mean, variance, skewness and kurtosis. The results are shown in Table 7.18.

The empirical kurtosis of the devolatilized residuals is 5.9239 for GARCH(1,1)(a) and 5.8874 for GARCH(1,1)(b) which are higher than for the Gaussian distributions. Among the de-volatilized return processes simulated by various Lévy processes, GH, NIG and VG seem to perform better than the Gaussian and t in modelling the kurtosis.

Table 7.18: Statistics for simulated driving processes

Statistics	Mathematical Form	Gaussian	t	VG	NIG	GH
Mean	$\mathbb{E}[Y]$	0.0000	0.0000	-0.2069	-0.2000	-0.2008
Variance	m_2	1.0000	1.0025	1.0003	1.0682	1.0327
Skewness	$m_3/m_2^{3/2}$	0	0	-0.4037	-0.4073	-0.4126
Kurtosis	m_4/m_2^2	3.0000	N/A	4.9053	5.0129	5.2107

7.5.2 Backtesting of Riskmeasures

We have used GARCH volatility models to get approximately *i.i.d.* devolatilized returns. There are 976 observations for SGTRI23 data x_i. We use the most recent 500 observations to backtest VaR. At time t, where t is from 476 to 975, at time $t+1$ the $VaR_\alpha(X_{t+1}|\mathcal{F}_t)$ under different α for log-returns can be estimated by simulating X_{t+1} using the model with parameters calibrated to data in $\mathcal{F}_t$ at time t. Thus initial model parameters are estimated using first 476 observations. The model is upgraded by re-calibrating the parameter vector θ at time $t = 100$, $t = 200$, $t = 300$ and $t = 400$.

The number of observations that $x_{t+1} > VaR_\alpha(X_{t+1}|\mathcal{F}_t)$ during those $N = 500$ tests is denoted by n. The actual frequency is n/N, while the expected probability, p, is $1 - \alpha\%$. We use a likelihood ratio statistic to test the riskmeasures. The null hypothesis is that the expected probability is equal to p. Under the null hypothesis, n is sample from Binomial distribution with size N and probability p, the likelihood ratio statistic is given by

$$LR = 2[(N-n)\log(1-n/N)+n\log(n/N)]-2[(N-n)\log(1-p)+n\log(p)], \quad (7.31)$$

is asymptotically χ_1^2 distributed.

The statistics of VaR_{95} and VaR_{99} backtesting are shown in Table 7.19. For $\alpha = 95$, the p-values are acceptable for all GARCH models. However, there is strong evidence to reject the null hypothesis at for GARCH-GBM and GARCH-t models at level $\alpha = 99$.

Table 7.19: VaR violation backtesting for SGTRI

	$\alpha = 95$			$\alpha = 99$		
	n/N	LR	p-value	n/N	LR	p-value
GBM	0.058	0.6421	42.29%	0.028	10.9940	0.09%
ST	0.054	0.1643	68.52%	0.022	5.4191	1.99%
VG	0.048	0.0426	83.64%	0.018	2.6126	10.60%
HYP	0.044	0.3942	53.00%	0.016	1.5383	21.49%
NIG	0.042	0.7107	39.91%	0.014	0.7187	39.66%

7.6 Risk Measures for With-Profits

With the Lévy driven stochastic volatility models and stochastic bridges fitted to total return data, we have been able to simulate the distributions of the insurer's losses using more realistic asset models incoperate the volatility clustering.

We start by re-investigating risk measures for model office M0, all assumptions are the same as given in Chapter 4, besides seven new asset models are applied:

G-GBM geometric Brownian motion with GARCH(1,1)-m volatility;

G-VG variance Gamma with GARCH(1,1)-m volatility;

G-NIG normal inverse Gaussian with GARCH(1,1)-m volatility;

HG-GBM hyper geometric Brownian motion with GARCH(1,1)-m volatility;

HG-VG hyper variance Gamma with GARCH(1,1)-m volatility;

HG-NIG hyper normal inverse Gaussian with GARCH(1,1)-m volatility; and

Wilkie Wilkie model.

Risk measures for M0 are shown in Table 7.20.

Recall that the riskmeasures for i.i.d. increment models are listed in Table 4.2. All the quantile measures include VaR_{95}, VaR_{99}, CTE95 and CTE99 increase under the GARCH volatility models. Among these models, G-NIG and HG-NIG have some extremely high risk measures. As we pointed out before, the M0 is a very "risky"

Table 7.20: Risk measures for G-GBM; G-VG; G-NIG; HG-GBM; HG-VG; HG-NIG and Wilkie model.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
G-GBM	0.0767	25.4780	12.7016	50.1521	35.9530	68.4691
G-VG	0.0836	41.0957	25.4993	83.5768	61.0030	108.7327
G-NIG	0.1287	42.1246	45.1424	95.9892	76.4579	118.8470
HG-GBM	0.1011	31.6413	32.2367	67.2835	51.6422	89.4564
HG-VG	0.1203	36.8878	36.7702	83.9015	65.5224	104.2108
HG-NIG	0.1647	51.2319	65.4562	112.8109	106.3019	184.2572
Wilkie	0.0763	19.3344	10.0181	38.1612	27.0326	49.1271

Table 7.21: Mean, standard deviation and kurtosis for (log) maturity asset shares and guarantees.

	Asset share			Guarantee		
	MEAN	SD	L-Kurtosis	MEAN	SD	L-Kurtosis
G-GBM	294.9001	182.6894	4.3011	142.0154	25.5415	5.9470
G-VG	278.3153	145.2343	10.9132	139.9257	21.7008	5.2149
G-NIG	273.5725	152.0734	12.1396	139.1113	20.4003	6.5303
HG-GBM	292.1378	201.3235	4.2561	141.9981	27.1684	4.1758
HG-VG	278.0892	155.5319	9.3219	139.1204	20.6450	4.9072
HG-NIG	268.4721	163.3400	15.9209	139.9457	51.9850	9.6911
Wilkie	274.4747	127.6874	3.0321	148.1267	20.5009	3.0379

contract and there is no risk management operations either on the guarantee by adjusting bonus mechanism or on the asset by controlling the EBR of UWP fund.

Table 7.21 shows the mean, standard deviation of the simulated maturity asset shares and guarantees. And we also list the kurtosis of the logarithm of the simulated maturity asset shares and guarantees. The plots of sample kurtosis against the first n simulation results with respect to losses under different asset models are shown in Figure 7.3. If the kurtosis of the maturity asset shares and guarantees does not exist, the sample kurtosis is an increasing function of the sample size n and goes to infinity eventually. The results show the kurtosis for logarithm of maturity asset shares converge and exist under all Lévy driven GARCH models. As one may expected, the kurtosis of log maturity asset shares and log guarantees under Gaussian innovation Wilkie model are close to 3. The L-Kurtosis of the maturity asset shares for G-GBM and HG-GBM are both greater than 4 which shows the distributions of the maturity asset share has heavier tails than the Gaussian distribution. The kurtosis of the log maturity asset shares are relatively high under GARCH-VG, GARCH-NIG, Hyper-GARCH-VG and Hyper-GARCH-NIG models.

We now pack in some risk management operations. First, a bonus earning power mechanism is used as we did in model office M2. We also allow more frequent bonus declaration on a monthly basis (M5). The Wilkie model is used with six sub-classes of the stochastic bridges:

Wilkie-B Wilkie model with Brownian bridge;

Wilkie-V Wilkie model with variance-Gamma bridge;

Wilkie-N Wilkie model with normal inverse Gaussian bridge;

Wilkie-GB Wilkie model with Brownian bridge with GARCH(1,1)-m volatility;

Wilkie-GV Wilkie model with variance-Gamma bridge with GARCH(1,1)-m volatility; and

Wilkie-GN Wilkie model with normal inverse Gaussian bridge with GARCH(1,1)-m volatility.

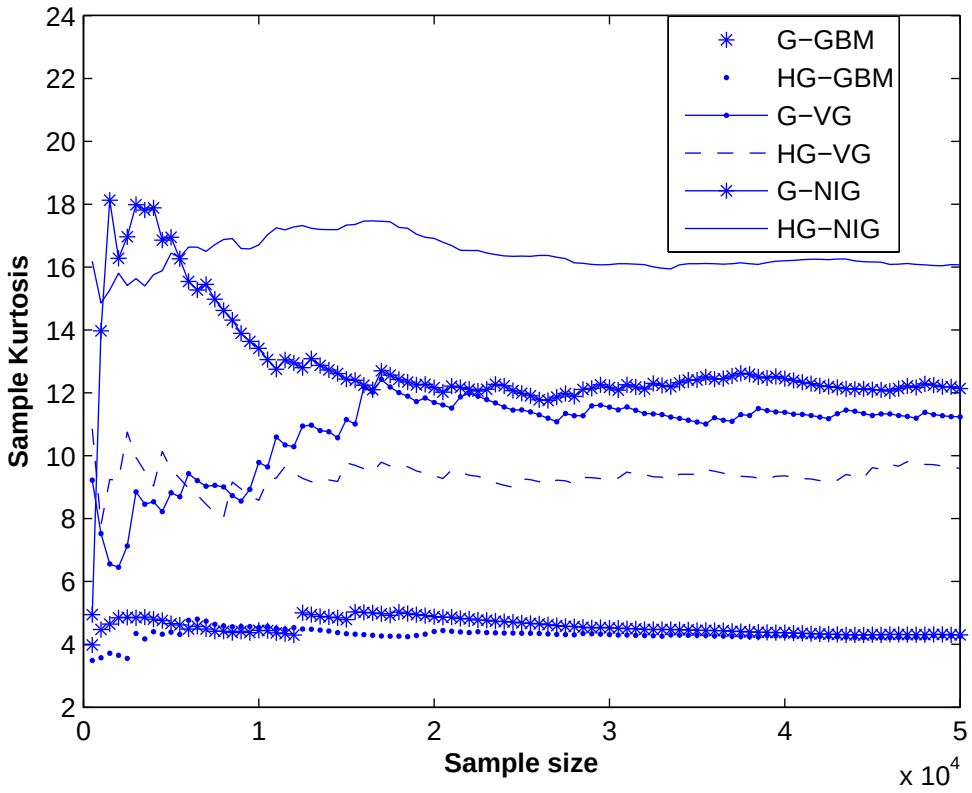

Figure 7.3: Kurtosis for logarithm of maturity asset shares

The simulated risk measures are shown in Table 7.22. Recall that in M2 we assume a 25% terminal bonus cushion and average futrue return rate as 6%.

Table 7.22: Risk measures for G-GBM; G-VG; G-NIG; HG-GBM; HG-VG; HG-NIG and Wilkie model.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99
Bonus earning power, yearly						
G-GBM	0.0787	24.3191	11.8804	47.6860	35.3073	72.2112
G-VG	0.0918	37.7959	25.3485	80.2824	60.0187	109.2451
G-NIG	0.1330	38.1069	40.4472	88.6196	71.8961	116.1389
HG-GBM	0.0800	23.7436	12.4228	46.9096	34.7379	70.8738
HG-VG	0.1227	33.9034	33.2135	79.1361	62.1692	104.1818
HG-NIG	0.1655	45.4986	59.8052	103.2285	92.3338	141.6455
Wilkie	0.1108	19.7567	17.8407	43.9315	34.2220	59.4083
Bonus earning power, monthly						
G-GBM	0.0674	21.4301	7.5805	40.5823	28.2420	59.4348
G-VG	0.0743	35.2697	15.8715	72.6945	49.2361	95.2042
G-NIG	0.1242	35.7580	36.0803	84.8165	65.6626	107.6502
HG-GBM	0.0683	22.7242	7.9405	42.6785	29.9432	62.6459
HG-VG	0.1076	33.6019	29.4309	75.9391	57.1875	94.3519
HG-NIG	0.1539	42.7059	54.4958	97.7897	83.8436	126.2656
Wilkie-B	0.0710	16.8268	7.7323	32.0212	22.6238	42.2788
Wilkie-V	0.0711	16.8656	8.0538	31.6589	22.5982	42.6773
Wilkie-N	0.0711	17.0401	7.7932	31.8656	22.7662	42.9232
Wilkie-GB	0.0737	17.9014	8.9185	33.5571	24.7972	48.5407
Wilkie-GV	0.0750	17.5209	8.6978	32.5993	24.9610	49.9339
Wilkie-GN	0.0749	17.3104	9.3666	32.5247	23.8628	44.0139

One can see the risk measures for model office M2 under GARCH type models follow a similar pattern to the M0 in Table 7.20.

There are some differences in the patterns of risk measures under the Wilkie model in M2 compares to the patterns of risk measures under the Wilkie model in M0. The POSF, VaRs and CTEs for M2 are much higher than for M0. It is because the mismatch of internal asset model used to calculate the bonuses and real world model. In M2, we assume the future investment returns for the asset share are i.i.d. with mean r. However, the yearly log-returns of the shares in the Wilkie model are negatively autocorrelated, so-called "mean-reverting" property. In this case the i.i.d. internal asset model fail to give reasonable projected asset share in

Table 7.23: Risk measures for G-GBM; G-VG; G-NIG; HG-GBM; HG-VG; HG-NIG and Wilkie model.

Bonus earning power, yearly, M2				
G-GBM	294.9001	182.6894	212.4313	124.3549
G-VG	278.5132	145.6741	203.4719	95.0584
G-NIG	273.2639	152.8793	202.4320	92.9849
HG-GBM	292.0432	201.4234	213.2702	138.6234
HG-VG	278.0892	155.5319	204.3588	96.2887
HG-NIG	268.1146	161.4850	202.3264	101.6761
Wilkie	274.2235	127.4662	202.4383	80.7026
Bonus earning power, monthly, M5				
G-GBM	294.9001	182.6894	227.8965	132.3585
G-VG	278.5132	145.6741	220.1097	107.1806
G-NIG	273.2639	146.8793	215.0894	100.5625
HG-GBM	292.0432	201.4234	225.9166	142.5266
HG-VG	278.0892	155.5319	217.7145	108.1980
HG-NIG	268.1146	161.4850	213.7352	109.3955
Wilkie-B	273.9812	126.9117	214.0767	88.9247
Wilkie-V	273.9812	126.9117	213.9256	88.7298
Wilkie-N	273.9812	126.9117	214.2444	89.2226
Wilkie-GB	273.9812	126.9117	214.7112	89.6580
Wilkie-GV	273.9812	126.9117	214.4406	89.7539
Wilkie-GN	273.9812	126.9117	214.6166	89.6432

the bonus earning power mechanism. For example, in simulation under the Wilkie models, the returns which are high in early years in the lifetime of the contract have higher probability to decrease and goes back to long term mean in the remaining of the contract duration than under i.i.d. models. Also the 6% average of future asset share return assumption seems to be too high for the Wilkie model.

The risk measures reduce when increase the frequency of declaring bonuses from annual BEP to monthly. The risk measures under the Wilkie models with different bridge processes are very close, even when a GARCH volatility model is added in the bridge. On the other hand, declaring bonuses on monthly basis in M5 for the Wilkie model with bridges reduce its riskmeasures sharply.

We now consider hedging based investment strategy M9 under more realistic asset models described in this chapter. The simulated risk measures are shown in Table 7.24. The case"(i)" for each asset model, for example "G-GBM(i)", represents the risk measures for insurer's loss without free asset at maturity and the case "(ii)" represents the risk measures for insurer's loss with free asset.

Table 7.24: Risk measures for insurer's payout at maturity, hedging based investment strategy, annually re-arrange the portfolio, simulated with GARCH volatility models and Wilkie model.

	POSF	MOL	VaR_{95}	VaR_{99}	CTE95	CTE99	CTE99.5
G-GBM(i)	0.0503	13.0248	0.0425	17.9418	13.4201	38.9447	55.3223
G-GBM(ii)	0.0089	17.8199	0	0	3.1720	15.8597	31.5897
G-VG(i)	0.0628	16.5050	3.8977	29.7225	20.5188	51.3779	66.0527
G-VG(ii)	0.0190	16.7475	0	10.7885	6.3641	28.3512	41.2843
G-NIG(i)	0.0901	16.0545	9.2262	36.2751	25.8474	55.8789	70.0212
G-NIG(ii)	0.0311	14.6855	0	16.6647	9.1344	34.3354	47.3010
HG-GBM(i)	0.0674	12.4924	3.3586	20.4865	17.1218	47.5174	70.0250
HG-GBM(ii)	0.0127	17.6832	0	1.7028	4.4915	26.0614	46.3907
HG-VG(i)	0.0774	14.1180	6.0317	29.1795	20.4336	47.0667	60.6602
HG-VG(ii)	0.0220	13.3443	0	9.4002	5.8715	25.8197	37.7305
HG-NIG(i)	0.1191	17.0236	13.6224	45.1037	32.5221	67.4909	85.0819
HG-NIG(ii)	0.0434	17.0334	0	25.1909	14.7850	46.5871	61.8081
Wilkie(i)	0.0878	10.6353	6.5211	21.7028	16.4093	33.0361	41.0791
Wilkie(ii)	0.0095	9.3136	0	0	1.7696	8.8479	16.4621

Within each group of models, the pattern of different driving processes are similar to what we have observed in the original M9 case in Chapter 6. The models with

NIG driving processes have highest risk measures in GARCH and hyper-GARCH models. The Gaussian driven models often under estimate the risk measures.

Compare to the case M2, the hedging based investment strategy brings down all the risk measures of the insurer's payout. In G-GBM, G-VG, HG-GBM and HG-NIG, the (ii) case have close MOL or even larger MOL than the (i) case. This is because with the hedging strategies, there is large amount of small negative values of losses, we called hedging errors. The free assets in M9 are assumed to be invested in same EBR as the asset shares. Thus they can cover these small losses but have less effect on reducing the large value of losses.

As we have discussed in Chapter 6, the cutting of risk measures in M9 is caused by extra hidden charges in the policy. Policyholders pay a proportion of their asset shares as the costs of hedging.

7.7 Concluding Remarks

We tried to model the volatility clustering in the SGTRI data by using GARCH(1,1)-m volatility models in this Chapter. Both MLE and MCMC estimation of the models indicate improvement of the likelihood functions. We consider the Lévy bridges in the Wilkie models. The form of bridge is still very simple while we adjust the linear drift within each year for the monthly returns. Some more advanced method could be used to improve the bridges, for example, we can assert the periodic economic cycle within each year by using non-linear drift term.

The GARCH(1,1) model in Equation 7.4 is a discrete time series on monthly basis. It works fine in simulation of the monthly returns and volatilities until we change the time scales, say, for daily returns and daily volatilities. In this case the continuous stochastic volatility model such as the OU processes introduced in Schoutens (2003) may be much easier to apply. Although there are literatures considered the continuous GARCH model, for instance, Drost and Werker (1996), most of the work is based on strong mathematical assumptions and hardly applied in risk management.

Model selection

One can see in case M0, M2, M5 and M9, the risk measures under HG-NIG model are often in the highest levels compare to the others. It is because the hyper-models with the Bayesian estimations add in margin of so-called estimation risk. Recall that the MCMC estimations of the GARCH-GH models have high standard deviations (see Table 7.4). And the GH processes, include it's subclasses have flat log-likelihood functions. This leads to high estimation risks incorporate in the hyper-GARCH-GH models. We suggest that sensitivity tests of the outcomes, in this situation, be constructed to check the varying of risk measures or other statistics by changing of the model parameters. Since the model selection is out of the scope of this thesis, it could be one of further research topics.

Student-t

One may note that we only fit the student-t distribution to the historical data in this thesis but haven't applied it in modelling the asset returns in Monte Carlo simulation. As we have discussed in Chapter 2 that student t distribution have heavy tails (power function) on both sides. The existence of moments which are higher than first order depends on the estimated value of shape parameter ν.

The risk measures are statistics of the random samples generated by Monte Carlo simulation. We should always ensure that the sample statistics do indeed converge to the theoretical quantities we are supposed to estimate. The Chebyshev's Inequality and the Weak Law of Large Numbers require variance of the random variable to be finite. And most statistical estimation procedures and hypothesis tests are based upon the central limit theorem which (in some cases) requires the high-order moments of the random variable to be well defined. The results based on heavy-tailed student-t processes are not as reliable as the semi-heavy tailed distributions especially when combined with non-linear autocorrelation.

Beyond the models

The Lévy models discussed in this thesis are based on historical data and statistical methods. We look at the statistical properties of the historical sample paths of the assets and construct stochastic models which can capture these properties. Then the information in historical data is distracted by the parameter estimation procedure. On the other hand, however, the historical data represent only the information from the past. There are market data which involve information about the forcasts or the believes of the future. These data is based on the historical information and will effect the actions of the market players. One example is the volatility data in the model estimations. The historical volatility is calculated by the past returns and the implied volatility is calculated by the market prices of the assets. Compare with these two estimators, the implied volatility may be more logical to use than a historical volatility in short term cases because the implied volatility is the market-consensus forcast of the future volatility based on historical information. One may expect the time series asset models to be stationary over the time. Thus in a long run, the historical long term volatility is a better estimation.

We use the mathematical models to explain the market behaviours and give suggestions of possible future scenarios. However, one shouldn't only rely on outcomes of models since the financial market is a far more complex system than mathematical models. It is more than necessary to add in personal views and judgements of market rather than only run the simulations on computer based on the models. As Wilkie (1995) pointed out that "An important feature of the way I believe this model should be used is that those using it should form their own opinions about the choice of appropriate mean values." in its inflation model parameter estimation.

Chapter 8

Conclusions and further research

The contribution of this thesis is threefold. First, we investigate the empirical properties of the historical UK share total return indices. Second, we construct long term stochastic investment models driven by Lévy processes and demonstrate how to fit these models to financial data by using either MLE or MCMC methods. Finally, we show the applications of the Lévy asset models in risk management of unitised with-profits policies.

In section 8.1 we briefly summarise the main conclusions of this thesis. Then some possible extensions to this research are discussed in Section 8.2.

8.1 Conclusions

Chapter 1 investigated the statistical properties of long-term returns of the monthly UK share total return index. The main conclusions were:

- The empirical unconditional distribution of the log-return process displays non-Gaussian properties with positive excess kurtosis;
- The statistical tests and the graphs show the empirical distribution belongs to distribution classes with heavy or semi-heavy tails;
- The Gaussian marginal distributions are not able to model extreme values in the SGTRI return series; and

- The data shows a small positive linear autocorrelation of monthly returns for first lag.

Chapter 2 provided an extensive review of the general structure, path properties and the decompositions of Lévy processes and the applications in financial modelling. Brief outline of the theoretical facts were given and, specifically, the Lévy-Khintchine formula shows the general form of the characteristic exponent of the infinitely divisible laws and the decomposition of the Lévy processes. Then Lévy processes were constructed either by defining an infinite divisible distribution as its increment law or by the stochastic time changing skills. We introduced the driving processes of our model which have generalized hyperbolic distributed unit increments.

In Chapter 3, We provided reliable and relatively fast algorithms for calibrating Lévy driven asset models. The main conclusions were:

- The GH increment Lévy processes were fitted to SGTRI data, both for maximum likelihood estimation and the Bayesian estimation using Markov Chain Monte Carlo;
- The results of MLE estimation show a significant improvement in performance of modelling data return distributions based on the likelihood function;
- The five parameter general GH case has the highest likelihood but the difference between the general case and its subclasses are relatively small;
- The MCMC estimation results for GH are close to those estimated by MLE;
- Estimation risk were considered by using Hyper models; and
- The Anderson & Darling statistics were calculated to give a goodness of fit in sense of distance measure.

Chapter 4 considers the key factors in modelling, reserving and valuation of unitised with-profits policies. We investigated a model office called M0 and list the main results as following:

- UWP policies were modelled by providing the operation rules which include asset allocation, charging, reserving and bonus declaration;

- Risk measures included probability of shortfall (POSF), mean of loss (MOL), value at risk (VaR) and conditional tail expectations (CTE) were calculated using Monte Carlo simulation under three different Lévy asset models;

- The exp-Lévy models driven by VG and NIG processes which have been proved to be more realistic show higher risk measures; and

- The hyper-models which use the generated sample of the Metropolis-Hesting Algorithm show an extra margin of risk measures for all three models while taking the parameter uncertainty into consideration.

In Chapter 5 some dynamic bonus mechanisms were introduced as tools to control the maturity guarantee liabilities of the UWP policies. The model office studies and results are summarised as follows:

M1 Case M1 considered retrospective bonus calculated as the weighted arithmetic average of the last n years returns on the asset share. The risk measures can be reduced by either increasing the smoothing period n or by cutting the participating rate p.

M2 The results in Case M2 using a bonus earning power (BEP) method with deterministic projected maturity asset share and 25 percent terminal bonus cushion are similar to the risk measures in M0. Comparing to retrospective bonus method, the BEP method offers some considerable advantages. We looked into the distribution of the maturity guarantee and asset share. We found a higher mean guarantee with higher standard deviation. The coupled Monte Carlo sample (AS_T, G_T) are more close related. There seems less possibility that an extreme loss happens in M2. Normally the policyholder will expect a higher guarantee at the maturity.

M3 We show BEP methods can be highly flexible by making the targeted maturity guarantee as any forms depend on the office internal models. The projected maturity asset share and unit values can be stochastic and path dependent.

M4 There seems to be little advantages in risk measures by trying to declare retrospective bonus on a monthly basis in Case M4.

M5 On the other hand the risk measures drop rapidly with the monthly reversionary bonus calculated by BEP in Case M5.

M6 In M6 we consider bonus earning power declared on any frequency. A significant improvement in risk controls was found when the frequency of declaring bonus increase from yearly to monthly. The risk measures were not sensitive to frequency of bonus declaration which more than once a month.

In Chapter 6 we consider the insurer's investment strategies for UWP policies. The main conclusions for model office Case M7, M8 and M9 are summarised as follows:

M7 investment strategies based on diversification, an invest-and-forget strategy and a re-balanced EBR strategy are tested. Both of the strategies reduce the risk measures of the maturity loss of the insurer ;

M8 hedging strategies for a 10-year vanilla type UWP contract with fixed annual bonus rate. Compare the other two investment strategies, diversified investment strategy M7 and hedging strategy M8. The former method tries to reduce the variance of the invest return distribution while sacrificing the possible high return from risky asset investment. Hedging investment strategy, however, cuts the left tail of the maturity loss distribution by paying an extra amount of expenses. Both of the loss distributions under these two investment strategies have less volatility than 100% EBR case; and

M9 Partial hedging based strategies which has limited support free capital, only part of UWP fund can be hedged. The EBR of the UWP fund in M9 is more flexible than in M7, in the cases when the contract is "out-of-money", there is large EBR in the fund which leads to higher expected future return in this case.

In Chapter 7 we extend the exp-Lévy models by replacing the GIG stochastic time by a GARCH(1,1)-m volatility process to capture the volatility clustering properties in the SGTRI data. Also, multi-variable models are considered, we added continuous time models using stochastic bridges driven by Lévy processes in the annual Wilkie

model. Both maximum likelihood estimators and Bayesian estimators using Markov Chain Monte Carlo are presented. We re-considered model office M0, M2, M5 and M9. The main conclusions are summarised as follows:

M0(r) All GARCH volatility models have higher risk measures than the corresponding constant volatility asset models. The distributions of maturity asset share and guarantee under GARCH volatility models are no longer Gaussian.

M2(r) the risk measures for model office M2 under GARCH type models follow a similar pattern to the M0. The POSF, VaRs and CTEs for M2 under Wilkie model are much higher than for M0. It is because the mismatch of internal asset model used to calculate the bonuses and real world model. The yearly log-returns of the shares in the Wilkie model are negatively autocorrelated, from the "mean-reversion" property. In this case the i.i.d. internal asset model fail to give reasonable projected asset share in the bonus earning power mechanism.

M5(r) The risk measures are reduced when increasing the frequency of declaring bonuses from annual BEP to monthly. The risk measures under the Wilkie models with different bridge processes are very close, even when a GARCH volatility model is added in the bridge. On the other hand, declaring bonuses on monthly basis in M5 for the Wilkie model with bridges reduce its risk measures sharply.

M9(r) The pattern of different driving processes are similar to what we have observed in the original M9 case in Chapter 6. The models with NIG driving processes have highest risk measures in GARCH and hyper-GARCH models. The Gaussian driven models often under estimate the risk measures. Compare to the case M2, the hedging based investment strategy brings down all the risk measures of the insurer's loss. In G-GBM, G-VG, HG-GBM and HG-NIG, the (ii) case have close MOL or even larger MOL than the (i) case. The cutting of risk measures in M9 is caused by extra hidden charges in the policy. Policyholders pay a proportion of their asset shares as the costs of guarantee and fund management.

8.2 Suggestions for further research

Throughout this thesis, we use a constant force of interest rate model. A natural path of further research is to bring in the interest rate dynamic by using stochastic interest rate models. One can either construct/use a separate interest rate model driven by Lévy processes, or use multi-variable asset models which incorporate the interest rate models and give the cross correlations between different asset classes.

Also we would like to build a continuous multi-variable model driven by Lévy processes. We discussed a model proposed in Chan (1998) which is defined by stochastic differential equations. This model has its advantage in many theoretical applications while is a big challenge to fit the model to the data. We would be interested to see further research on this topic.

Some of the possible combinations of asset models and risk control strategies were not investigated in the thesis because of the volume of the work. Some of them are worth being tested, for example:

- A life office both declares bonuses and hedges on monthly basis;
- A hedging strategy for UWP policies in any time intervals and other hedging strategies; and
- Some BEP under risk measures based on the management purpose such as those we described in M3.

We consider the guarantee liability of UWP model by assuming that there is no surrender and early death payment. This is majorly because the mortality risk and surrender risk are different from the financial investment risk we focus on. However, there is no reason why these risks could not be taken into account in more complicated liability models. Based on the existing result in this thesis, a mortality model both as a deterministic mortality table or a stochastic mortality model can be asserted in each simulation study. For the surrenders of the contract, guarantee surrender values can be considered during the term of the UWP policies. This could be interesting because in this case the UWP liability model would be "American option style" under the allowance of surrenders.

In this thesis, we focus on a single generation of UWP policies, within which each "unit" have the same value of guarantee and asset share. A possible extension would be to consider a multi-generation liability model of UWP policies. For a single generation of UWP policies, we assume that the insurer never defaults. However, for multi-generation liability model, we can consider the UWP "fund-wise", insurers face the risk not only from shortfall of each generation of UWP policies but from insolvency of the whole fund if they cannot pay the guarantee at any maturity. We can even include shareholders to share the profits if the insurer is not mutual. Thus the liability model in this case is more complicated and the author would be interested to see further research on whether insurers can pool these risks by selling policies with different maturities.

As we discussed in Chapter 7, more work is required to investigate the robustness and reliability of the asset models. For example, sensitivity tests are required on estimation and simulation of the Lévy models.

Last but not the least, in order to improve the efficiency of the Monte Carlo methods of the asset models, one can apply variance reduction techniques and quasi-Monte Carlo techniques in simulation studies.

Bibliography

Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance*, **26**, 1487–1503.

Applebaum, D. (2004). *Lévy processes and stochastic calculus*, volume 93 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge. ISBN 0-521-83263-2.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228.

Ballotta, L. (2005). A Lévy process-based framework for the fair valuation of participating life insurance contracts. *Insurance Math. Econom.*, **37**(2), 173–196. ISSN 0167-6687.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Royal Society of London Proceedings Series A*, **353**, 401–419.

Barndorff-Nielsen, O. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Statist.*, **5**(3), 151–157. ISSN 0303-6898.

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance Stoch.*, **2**(1), 41–68. ISSN 0949-2984.

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **63**(2), 167–241. ISSN 1369-7412.

Bertoin, J. (1996). *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge. ISBN 0-521-56243-0.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**(3), 637–659.

Brennan, M. J. and Schwartz, E. S. (1976). The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics*, **3**, 195–213.

Brennan, M. J. and Schwartz, E. S. (1979). *Pricing and Investment Strategies for Guaranteed Equity-Linked Life Insurance*. Monograph No. 7. The S. S. Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania.

Brock, W., Dechert, W. and Scheinkman, J. (1987). A test for independence based on the correlation dimension. *Working Paper University of Wisconsin at Madison*.

Carr, P., Geman, H., Madan, D. B. and Yor, M. (2002). The Fine Structure of Asset Returns. *Journal of Busines*, **Volume 75 Number 2**, 305–32.

Chadburn, R. G. (1997). The use of capital, bonus policy and investment policy in the control of solvency for with-profits life insurance companies in the uk. *City University Actuarial Research Paper*, **95**, 1–29.

Chadburn, R. G. and Wright, I. D. (1999). The sensitivity of life office simulation outcomes to differences in asset model structure. *City University Actuarial Research Paper*, **120**, 1–58.

Chan, T. (1998). Some applications of levy processes to stochastic investment models for actuarial use. *Astin Billetin*, **28 No.1**, 77–93.

Chan, T. (1999). Pricing contingent claims on stocks driven by Lévy processes. *The Annals of Applied Probability*, **9**(2), 504–528. ISSN 1050-5164.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, **1**, 223–236.

Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL. ISBN 1-5848-8413-4.

Ding, Z., Granger, C. W. J. and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, **1**(1), 83–106.
URL: *http://www.sciencedirect.com/science/article/B6VFG-45JK6Y4-6/1/5553df87f8bbe916f03eb8c68a03fa75*

Dowd, K. and Blake, D. (2006). After var: The theory, estimation, and insurance applications of quantile-based risk measures. *Journal of Risk & Insurance*, **73**, 193–229.

Drost, F. C. and Werker, B. J. M. (1996). Closing the garch gap: Continuous time garch modeling. *Journal of Econometrics*, **74**(1), 31–57.
URL: *http://ideas.repec.org/a/eee/econom/v74y1996i1p31-57.html*

Eberlein, E., Kallsen, J. and Kristen, J. (2001). Risk management based on stochastic volatility. Technical report, Journal of Risk.

Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli 1*, pages 281–299.

Forfar, D. O., Milne, R. J. H., Muirhead, J. R., Paul, D. R. L., Robertson, A. J., Robertson, C. M., Scott, H. J. A. and Spence, H. G. (1989). Bonus rates, valuation and solvency during the transition between higher and lower investment returns. *Transactions of the Faculty of Actuaries*, **40**, 490–562.

Geman, H., Madan, D. and Yor, M. (1998). Asset prices are brownian motion: only in business time.

Gentle, J. E. (2003). *Random Number Generation and Monte Carlo Methods*. Springer.

Gilks, W., Richardson, S. and Spiegelhalter, D. (eds.) (1996). *Markov Chain Monte Carlo in Practice*. Chapman&Hall/CRC.

Glasserman, P. (2004). *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York. ISBN 0-387-00451-3. Stochastic Modelling and Applied Probability.

Grosen, A. and Jorgensen, P. L. (2002). Life insurance liabilities at market value: An analysis of insolvency risk, bonus policy, and regulatory intervention rules in a barrier option framework. *Journal of Risk and Insurance*, **69**, 63–91.

Haberman, S., Ballotta, L. and Wang, N. (2003). Modelling and valuation of guarantees in with-profit and unitised with-profit life insurance contracts. *Working paper, Faculty of Actuarial Science and Statistics, City University, London*, pages 1–26.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, **50**, 1029–1054.

Hardy, M. (2003). *Investment Guarantees: Modelling and Risk Management for Equity-Linked Life Insurance*. Wiley (New York).

Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, **5**, 41–53.

Hare, D. J., Craske, G., Crispin, J. R., Desai, A. J., Dullaway, D. W., Earnshaw, M. A., Frankland, R., Jordan, G. A., Kerr, M. G., Manley, G. M., McKenzie, J. L., Rae, R. A. and Saker, M. C. (2003). The realistic reporting of with-profits business. *The With-Profits Working Party of the UK Actuarial Profession*, pages 1–63.

Hibbert, J., Mowbray, P. and Turnbull, C. (2001). A stochastic asset model & calibration for long-term financial planning purposes. Technical report, Barrie & Hibbert Limited.

Hibbert, J. and Turnbull, C. (2003). Measuring and managing the economic risks and costs of with-profits business. *British Actuarial Journal*, **9**, 725–777.

Hubalek, F. and Sgarra, C. (2005). Esscher transforms and the minimal entropy martingale measure for exponential Lvy models. Technical report, University of Aarhus.

Hull, J. C. (1999). *Options, Futures, and Other Derivatives*. FT Prentice Hall.

Jeffreys, H. (1961). *The Theory of Probability (3e)*. Oxford.

Jorgensen, P. L. (2001). Life insurance contracts with embedded options. *Working paper, University of Aarhus, Denmark*, pages 1–21.

Kirman, A. (2006). *Long Memory in Economics*. Springer, 1st edition. ISBN 354022694X.
URL: *http://amazon.com/o/ASIN/354022694X/*

Kleinow, T. and Willder, M. (2006). The effect of management discretion on hedging and fair valuation of participating policies with maturity guarantees. *forthcoming in Insurance: Mathematics & Economics*, pages 1–14.

Lee, P. and Wilkie, A. (2000). A comparison of stochastic asset models. *Proceedings of the 10th AFIR Colloquium, Tromsoe*, pages 407–445.

Madan, D. B. and Seneta, E. (1990). The variance gamma model for share market returns. *Journal of Business*, **63**, 511–524.

Mandelbrot (1977). *Fractals: Form,Chance and Dimension*. Freeman,N.Y.

McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative risk management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ. ISBN 0-691-12255-5. Concepts, techniques and tools.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics*, **4**(Spring), 141–183.

Monroe, I. (1978). Processes that can be embedded in brownian motion. *The Annals of Probability*, **6**(1), 42–56.

Pagan, A. (1996). The economics of financial markets. *Journal of Empirical finance*, **3**(1), 15–102.

Platen, E. and Rendek, R. (2007). Empirical evidence on student-t log-returns of diversified world stock indices. (194).
URL: *http://ideas.repec.org/p/uts/rpaper/194.html*

Prause, K. (1999). *The generalized hyperbolic model: Estimation, Financial Derivatives, and Risk Measures.* Ph.D. thesis, University of Freiburg.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2002). *Numerical Recipes in C++.* Cambridge University Press.

Protter, P. (2004). *Stochastic integration and differential equations.* Stochastic modelling and applied probability. Springer. Second Edition.

Sato, K. (1999). *Lévy processes and infinitelly divisible distributions.* Cambridge University Press.

Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives.* John Wiley & Sons, Ltd.

Schweizer, M. (1996). Approximation pricing and the variance optimal martingale measure. *The Annals of Probability*, **24**, 206–236.

Schweizer, M. (2001). A guided tour through quadratic hedging approaches. In *Option pricing, interest rates and risk management*, Handb. Math. Finance, pages 538–574. Cambridge Univ. Press, Cambridge.

Tong, W. (2004). *Reserving for maturity guarantees under unitised with-profits policies.* Ph.D. thesis, Heriot-Watt University.

Webber, N. and Ribeiro, C. (2003). A monte carlo method for the normal inverse gaussian option valuation model using an inverse gaussian bridge. *Computing in Economics and Finance*, **5**(5).

Webber, N. and Ribeiro, C. (2004). Valuing path dependent options in the variance-gamma model by monte carlo with a gamma bridge. *Journal of Computational Finance*, **7**, 81–100.

Wilkie, A. D. (1986). A stochastic asset model for actuarial use. *Transactions of the Faculty of Actuaries*, **39**, 341–373.

Wilkie, A. D. (1987). An option pricing approach to bonus policy. *Journal of the Institute of Actuaries*, **114**, 21–77.

Wilkie, A. D., Owen, M. P. and Waters, H. R. (2005). Notes on options, hedging, prudential reserves and fair values. *British Actuarial Journal*, **11**(2), 199–293. ISSN 1357-3217.

Wilkie, A. D., Waters, H. R. and Yang, S. (2003). Reserving, pricing and hedging for policies with guaranteed annuity options. *British Actuarial Journal*, **9**, 263–391.

Wilkie, D. (1995). More on a stochastic asset model for actuarial use. *British Actuarial Journal*, **1**, 777–945.

Willder, M. (2004). *An option pricing approach to charging for maturity guarantees given under unitised with-profits policies.* Ph.D. thesis, Heriot-Watt University.

Index