

https://dspace.org/

Market-consistent valuation and risk management of guaranteed annuity options

Zheng, Chong

2012-02-03
http://hdl.handle.net/10539/11240

Downloaded from a DSpace repository. https://dspace.org/

Market-Consistent Valuation and Risk Management of Guaranteed Annuity Options

Chong Zheng

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 6 October 2010 (Revised 30 August 2011)

Table of Contents

Declaration 4
Abstract 5
List of Figures 6
List of Tables 7
Chapter 1. Introduction 9
1.1 Guaranteed Annuity Options 9
1.2 Historical Valuation Method and Its Shortcomings 9
1.3 Other Risks 9
1.4 Aim of Research 10
1.5 Limitations of Research 10
1.6 Research Methodology 11
1.7 Dissertation Structure 12
Chapter 2. Financial Reporting Standards & Regulatory Solvency Requirements 13
2.1 International Financial Reporting Standards 13
2.2 Solvency II & Individual Capital Assessment 15
2.3 Linking Economic Capital to Risk Management 19
Chapter 3. Literature Review 20
3.1 Guaranteed Annuity Options 20
3.2 Risk Management of GAO 22
3.3 Interest-rate Models 23
3.4 Equity Models 30
3.5 Integrated Models 35
Appendix 3 – Glossary of technical terms 37
Chapter 4. Stochastic Valuation 39
4.1 Introduction 39
4.2 Real World vs. Risk-Neutral 39
4.3 Equilibrium vs. No-arbitrage 43
4.4 Summary of Model Types 45
4.5 Valuation in an Incomplete Market 45
4.6 Stochastic Valuation Methods 48
Chapter 5. Economic Scenario Generator 50
5.1 Introduction 50
5.2 The ESG at a High Level 50
5.3 Term Structure of Interest Rates 51
5.4 Martingale Test 69
5.5 Bond-portfolio Returns 70
5.6 Equity-portfolio Returns 72
Appendix 5.1 – Building an Interest-rate Tree 80
Appendix 5.2 – Model Tests 87
Appendix 5.3 – Calibration Data 94
Appendix 5.4 – Antithetic Variables 95
Appendix 5.5 – ESG Parameters 97
Chapter 6. Valuation of Guaranteed Annuity Options 100
6.1 Introduction 100
6.2 The Hypothetical Portfolio and Modelling Assumptions 100
6.3 The Model 101
6.4 Results and Interpretation 102
Chapter 7. Risk Measures 104
7.1 Introduction 104
7.2 Two Different Methods of Determining Economic Capital 104
7.3 Determining the Shocks to Apply 106

7.4 Summary 115
Chapter 8. Managing Market Risks 116
8.1 Introduction 116
8.2 Sensitivity of Market-consistent Liabilities to Changes in Market Conditions 116
8.3 Interest-rate Duration Hedge Using a Single Zero-coupon Bond 121
8.4 Interest-rate Hedge Using Bond Gearing 124
8.5 Interest-rate Hedge Using a Basket of Vanilla Swaptions 127
8.6 Swaptions and Equity Options 134
8.7 Summary of Findings 140
8.8 Limitations of the Hedges 140
Chapter 9. Longevity Risk 142
9.1 Sensitivity to Longevity 142
9.2 Hedging Longevity Risk 143
Chapter 10. Conclusions and Future Research 144
10.1 Conclusions 144
10.2 Comparison to Prior Research 145
10.3 Gaps for Future Research 145
Acknowledgement 148
References 149

Declaration

I declare that this dissertation is the result of my own work. To the extent usually and reasonably expected, assistance and peer review were received from my supervisor. The dissertation is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University. I was employed by Liberty Life (South Africa) and Sun Life Financial (Canada) during the period during which the dissertation was written.



Chong Zheng

6th day of October 2010 (Revised 30 August 2011)

Abstract

The aim of the research is to develop a method and tools that facilitate the market-consistent valuation of a hypothetical portfolio of guaranteed-annuity options (GAOs). The fund underlying the GAOs carries a guaranteed minimum rate of return.

This is accomplished by the construction of a stochastic economic-scenario generator. As an illustration, this scenario generator is calibrated to the South African market conditions as at the end of December 2007, although the same principles can be applied to all markets. The proposed model uses a one-factor Black–Karasinski interest-rate model and a Black–Scholes equity model.

Further to this, a suitable hedge for this GAO portfolio against the main market risks (interest rates and equity returns) is sought. This research aims to evaluate a series of potential hedges and investigates the advantages and disadvantages of each.

The sensitivity of this GAO portfolio to improvements in future mortality is also investigated.

List of Figures

Figure 5.1 High-levl structure of the economic scenario generator 51
Figure 5.2: Simulating future term-structures by nested simulations 63
Figure 5.3: Simulating future term-structures by nesting lattices within simulations 64
Figure 5.4: Illustration of the testing of future term structures 66
Figure 5.5: Extrapolation of spot curve by assuming constant forward curve 68
Figure 5.6: A comparison of yield curve generated by the interest-rate tree and Monte–Carlo simulations to the input 68
Figure 5.7: A comparison of implied volatilities of receiver swaptions with 20-year tenor: market, simulated and interest-rate lattice 69
Figure 5.8: Martingale test for 15-year ZCBs 71
Figure 5.9: Martingale test for a bond portfolio with six ZCBs in equal weightings 72
Figure 5.10: Market implied-volatility and its comparison against those calculated using the model and simulations 78
Figure 5.11: Martingale test for the equity portfolio 79
Figure 5.12: Stage 1 interest-rate tree 81
Figure 5.13: Trimming of the tree at the (a) uppermost and (b) lowermost nodes respectively 82
Figure 5.14: Stage 1 trimmed interest-rate tree 83
Figure 5.15: Stage 2 trimmed lattice 84
Figure 5.16: Comparison of simulated ZCB prices (t=5) to the calibration targets 91
Figure 5.17: Comparison of simulated spot yields (t=5) to the calibration targets 91
Figure 5.18: Comparison of simulated ZCB prices (t=15) to the calibration targets 92
Figure 5.19: Comparison of simulated spot yields (t=15) to the calibration targets 92
Figure 5.20: Comparison of simulated ZCB prices (t=25) to the calibration targets 93
Figure 5.21: Comparison of simulated spot yields (t=25) to the calibration targets 93
Figure 7.1: Coefficients for mapping observations from the old axes to the new axes 109
Figure 7.2: Historic observed log changes mapped onto the new axes 110
Figure 7.3: Shocked term structures of interest rates when using the first factor 111
Figure 7.4: Shocked term structures of interest rates when using the second factor 111
Figure 7.5: Shocked term structures of interest rates when using the third factor 111
Figure 7.7: Implied volatility on swaptions where an upward shock corresponding to a 99.5% confidence is applied to the first factor 114
Figure 7.8: Implied volatility on swaptions where a downward shock corresponding to a 99.5% confidence is applied to the first factor 114
Figure 8.2: Sensitivity of GAO reserves to the changes in term structure of interest rates 118
Figure 8.3: Sensitivity of total reserves to the changes in term structure of interest rates 118
Figure 8.5: Comparison of the value at risk between (a) no hedge, and (b) duration hedge with a ZCB of 61.7 years 124
Figure 8.8: Notional amounts of 25-year swaptions of various tenors to be held in the hedge portfolio 130
Figure 8.10: Comparison of VaR of all guarantee liabilities when (a) no hedge is held in respect of the risks; and (b) a hedge consisting of receiver swaptions and cash is held 133
Figure 8.11: Comparison of VaR of all guarantee liabilities when (a) no hedge is held in respect of the risks and (b) rho and vega hedge consisting of receiver swaptions, equity puts and cash 139

List of Tables

Table 5.1 Comparison of market and simulated bond prices and yield curves 87
Table 5.2: Comparison of simulated and actual prices of ZCBs 89
Table 5.3: Comparison of simulated and actual spot rates 90
Table 5.4: Time dependent parameters $a(t)$ and $\theta(t)$ used in the modelling of the term structure of interest rate 97
Table 5.5: Values of x^* for each j and their associated up, middle and down probabilities 98
Table 5.6: The term structure of implied volatilities suggested by the model 99
Table 6.1: Unit-fund asset allocation 101
Table 6.2: Market-consistent value of the policy liabilities 103
Table 7.1: Cumulative variations explained by principle components, expressed as a percentage of total variations 108
Table 7.2: Parameters values and shocks to equity assets 112
Table 7.3: Cumulative volatility of the first five factors in the principle component analysis of changes in the implied volatility surface of swaptions 113
Table 7.4: Parameters values and shocks to the equity volatility 115
Table 7.5: Parameters values and adjusted shocks to the equity volatility 115
Table 8.1: Sensitivity of market-consistent liabilities to changes in the term structure of interest rates when allowing only for movement in the first factor 116
Table 8.2: Sensitivity of market-consistent liabilities to changes in the term structure of interest rates when allowing only for movement in the second factor 116
Table 8.3: Sensitivity of market-consistent liabilities to changes in the term structure of interest rates when allowing only for movement in the third factor 117
Table 8.4: Sensitivity of market-consistent liabilities of the hypothetical liability portfolio to changes in the starting fund values 119
Table 8.5: Sensitivity of market-consistent liabilities of the hypothetical liability portfolio to changes in implied volatility on swaptions 119
Table 8.6: Sensitivity of market-consistent liabilities to changes in equity-option-implied volatilities 120
Table 8.7: The value of guarantees liabilities after a reduction and increase of 50 bps to the term structure of interest rates 121
Table 8.8: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor 122
Table 8.9: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor 122
Table 8.10: Value at risk: changes in the term structure of interest rates when allowing only for the changes in of the third factor 122
Table 8.11: Value at risk: changes in the starting fund values 123
Table 8.12: Value at risk: changes in implied volatility on swaptions 123
Table 8.13: Value at risk: changes in implied volatilities of equity options 123
Table 8.14: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor. 125

Table 8.15: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor. 125
Table 8.16: Value at risk: changes in the term structure of interest rates when allowing only for the changes in of the third factor 126
Table 8.17: Decrease in the expected guaranteed minimum payments. 128
Table 8.18: Notional amounts of 25-year swaptions of various tenors to be held in the hedge portfolio 129
Table 8.20: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor 131
Table 8.21: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor 131
Table 8.22: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the third factor 131
Table 8.23: Value at risk: changes in the starting fund values 132
Table 8.24: Value at risk: changes in implied volatilities of swaptions 132
Table 8.25: Comparison of VaR of all guarantee liabilities when (a) no hedge is held in respect of the risks; and (b) a portfolio of swaptions and cash 134
Table 8.26: The value of guarantees liabilities for a reduction and increase of 200 bps to the term structure of equity implied volatilities 135
Table 8.27: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor 137
Table 8.28: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor 137
Table 8.29: Value at risk: changes in the term structure of interest rates when allowing only for the changes in of the third factor 137
Table 8.30: Value at risk: changes in the starting equity values 138
Table 8.31: Value at risk: changes in implied volatilities on swaptions 138
Table 8.32: Value at risk: changes in implied volatilities on equity options. 138
Table 9.1: Guarantees reserves with and without a limit of 30 years on annuities 142
Table 9.2: The effects of mortality improvements on the value of guaranteed liabilities 143

Chapter 1. Introduction

1.1 Guaranteed Annuity Options

1.1.1 Guaranteed annuity conversion options, or guaranteed annuity options (GAOs), are contracts sold by life-insurance companies. They were particularly popular in the United Kingdom in the 1970s and 1980s, and were still being sold in the 1990s. GAOs were also sold by life insurers in other parts of the world.

1.1.2 A typical GAO contract is an ancillary benefit that has been added to a core retirement-savings policy. The fund account underlying the retirement-savings policy may either be unit linked or with profits. A typical GAO benefit offers the holder of the contract the right to choose, at retirement, one of the following two benefits:

- a cash payment equivalent to the accumulated investment value ('fund account') of the fund underlying the retirement-savings policy, which depends on the investment performance of the fund until the retirement date; or
- an annuity, which is bought using the fund account of the policy, at a guaranteed rate that is specified at the inception of this policy.

1.1.3 The annuity is usually payable throughout the policyholder's remaining lifetime, but other variations exist.

1.2 Historical Valuation Method and Its Shortcomings

1.2.1 In the UK, the cash option has, until very recently, been the option that offered the policyholder the greatest value at maturity date. This is because the guaranteed conversion rate has been lower than those available in the market at the time of maturity and is therefore more expensive to the policyholder. The drop in UK interest rates in the late 1990s caused market annuity rates to fall, and the guaranteed annuity has therefore become more valuable to the policyholder and onerous to the life insurer.

1.2.2 Traditionally, life insurers have reserved for the GAO liabilities using a deterministic approach on a best-estimate basis with a small margin to protect themselves against the risk of inadequate provisions. Because these guarantees were issued far out of the money in the 1970s and 1980s, when inflation and interest rates were high, the guarantee did not come into effect. No guarantee reserve was allowed for. Additional shareholder capital was held to provide another layer of protection against adverse experience. This capital should have reflected the amount of risk that insurers were exposed to by writing the business – the higher the risk, the more capital should be held. Because insurers did not appreciate the riskiness of this product, no action (such as hedging) was taken to reduce the risk, whilst minimal capital was held to cushion against any adverse conditions.

1.2.3 After the 1990s, interest and inflation rates fell significantly and the onerousness of the guarantee forced the world's oldest life insurer, Equitable Life, to close to new business. The reserving margin and capital together were not sufficient to provide adequate protection against the risks of product. This brought about awareness that these reserving and capital management methodologies were no longer appropriate.

1.3 Other Risks

1.3.1 Paragraph 1.2.2 illustrates interest-rate as one of the risks of GAOs. Boyle & Hardy (2003) note that two additional factors influenced the cost and risk of GAOs. Firstly, strong equity-market performance caused the accumulated investment value of the fund to become higher than expected, thereby increasing the value of the unit fund at retirement and the value to which the guarantee applies, increasing the guarantee cost. Secondly, mortality improvements increased the cost of the life annuities. These improvements were not considered when the guaranteed annuity conversion rates of the GAOs were determined at product design. Both these factors contributed to the increased cost of the guarantee.

1.3.2 Hence, the structure of the GAO benefit exposes the life insurer to the following main risks:

- lower than expected interest rate at maturity;
- higher than expected investment growth before maturity; and
- higher than expected longevity.

1.3.3 Boyle & Hardy (*op. cit.*) show that the falling interest rate is by far the largest risk to life insurers. Other business-related risks, such as lapse and expense risks, also exist. These risks are not specific to GAOs and should be handled in a similar way to those of other products.

1.3.4 Other variations of this product may further increase the onerousness and riskiness of the benefits. One such variation is the inclusion of a minimum maturity guarantee on the underlying fund account. This guarantee is usually expressed as a guaranteed minimum rate of return on the allocated premiums. This guarantee can be a nominal or a real rate of return.

1.4 Aim of Research

1.4.1 The aim of the research is to develop the methodology and tools that facilitate the market-consistent valuations of a hypothetical portfolio of GAOs. The fund account underlying the GAO carries a guaranteed minimum rate of return. An example is constructed using a model calibrated to the market conditions in South Africa as at 31 December 2007, but the principles may be used for other economies at other times.

1.4.2 The research also aims to investigate the sensitivity of this market-consistent value to some risk factors, so that the biggest risk factor can be identified and managed appropriately through hedge instruments, leaving the smaller or unhedgeable risks to be absorbed through a risk capital.

1.4.3 A suitable hedge for this hybrid GAO portfolio against the main market risks (interest rates and equity returns) is sought. This research aims to evaluate a series of potential hedge portfolios that consist of simple instruments and investigates the advantages and disadvantages of each.

1.5 Limitations of Research

1.5.1 The focus is on the quantification and the management of market risk. Management of demographic risks such as mortality, longevity, lapse and option take up risks are not

investigated. This is because a lack of a liquid market for trading in these risks limits the ability to effectively manage these risks. While mortality risks (associated with random fluctuations of mortality rates around a best estimate) are diversifiable given a large enough pool of lives, longevity is more systematic in nature and is therefore undiversifiable. For this reason, the sensitivity of the GAO liabilities to longevity is investigated, despite the limited scope for reducing this risk. The author recognises that there may also be secondary effects between market risk and longevity risk – in the event of a fall in interest rates and increase in longevity, the liabilities may be significantly higher than the sum of the two components. This secondary effect is not investigated.

1.5.2 Lapse risks may be related to market performance. Rational and knowledgeable policyholders should not lapse their policies if the guaranteed benefits are in the money. However, some policyholders are not rational and other financial needs may also influence the decisions of the rational policyholders to lapse their policies. The rate of option take up may also be affected by tax treatment of the lump-sum payment. To model these risks fully requires the modelling of behavioural finance with very limited and unreliable data. Like mortality risk, there is no liquid market for the hedging of these risks and so these, too, are not investigated.

1.5.3 Expenses are also ignored because these tend to be company-specific and are, in the author's experience, small in magnitude relative to the other risks — under normal market conditions, the expenses are only a very small fraction of the annuity payments. Furthermore, the life insurer is exposed to high inflation, which is usually accompanied by high interest rates, which in turn reduces the cost of the annuity. Since any increase in administrative expenses tends to be offset by a reduction in the cost of the annuity, this risk is not investigated.

1.5.4 Other risks may exist. These include the additional risks that companies willingly expose themselves to as a result of taking mismatched positions. An example of this is the investments in credit-risky bonds. These present the company with credit-default risks and credit-spread risks. As these mismatching decisions are usually strategically driven, they are not investigated as a part of this research. As a practical matter, most companies only attempt to manage the largest risks and retain the remainder for potential profitability and also because it is impossible to hedge out all risks without transferring the entire block of business to a third party. The author recognises that these risks should be further researched and addressed through capital. We show in Chapter 2 below that risk management and economic capital are intimately related, but this research deals with risk management and not capital.

1.6 Research Methodology

1.6.1 Some of the concepts outlined below may be unclear but a high-level description and motivation of some of these are given in Chapter 2 to Chapter 4 below.

1.6.2 An economic scenario generator (ESG) is essential to the valuation of GAOs and is constructed in Matlab as a part of this research. This ESG allows the simulation of future interest rates and equity returns in a manner that is arbitrage-free and market-consistent. As an illustration, this ESG is calibrated to the market conditions in South Africa as at 31 December 2007, but the principles defined here may be used for other

economies and at other times as long as the market conditions are within reasonable limits of the model. The 2007 date is used as this was the latest data that was available to the author in the initial phase of the research without incurring significant financial costs. The volatile market conditions between the second half of 2008 and mid 2009 would have also pushed the limitations of the model.

1.6.3 A cashflow projection model is constructed in Microsoft Excel for the projection of the cashflows underlying a GAO contract. The simulations that the ESG produces are input into this model and together they facilitate the market-consistent valuation of the guaranteed liabilities.

1.6.4 Historical market events are analysed to determine a set of market shocks. These represent some events that could reasonably be repeated in future. The market-consistent liability is recalculated under each shock to investigate its sensitivity to the shock.

1.6.5 A set of potential hedges are proposed to reduce the market risks posed by these GAO liabilities. These asset portfolios are also sensitivity-tested to analyse how their market values change relative to the change in the market-consistent liabilities. This shows the effectiveness of the hedge.

1.7 Dissertation Structure

1.7.1 In Chapter 2, the author sets the regulatory backgrounds of the market-consistent valuation of GAOs and their regulatory capital framework in various parts of the world. Because of the rapidly changing regulatory framework, it is necessary to anchor the research to the regulations in effect at the end of 2008. In Chapter 3, the author gives a high-level overview of past research on GAOs and their risk management. A listing of some asset models that are useful for the market-consistent valuation of financial options is also given. Chapter 4 compares and contrasts the different types of stochastic valuations: real-world vs. risk-neutral and equilibrium vs. no-arbitrage. Different stochastic modelling techniques are also discussed. These are closed-form solutions, lattices and Monte-Carlo simulations.

1.7.2 Chapter 5 specifies, in detail, the economic-scenario generator that was constructed for the calculation of the market-consistent value of GAOs, its calibration and validation. Chapter 6 formulates the cashflow-projection model for the valuation of GAOs and shows the results of the valuation. Chapter 7 compares two different measures for the risk management of financial risks. This chapter also describes the methods that are used for determining the shocks to the market conditions that should be applied in order to test out the sensitivity tests of GAOs that would then determine the relative sizes of the various risk factors. In Chapter 8, the author investigates a series of hedge portfolios that utilise only simple assets and investigates their advantages and disadvantages. In Chapter 9, the sensitivity of the GAO to longevity is investigated. The conclusions are given and ideas for future research are identified in Chapter 10.

Chapter 2. Financial Reporting Standards & Regulatory Solvency Requirements

This chapter describes the international financial reporting standards and the regulatory solvency requirements of UK, EU and South Africa. The general directions of these requirements have been based on the principles of market-consistency for the past few years. However, as the development of these standards is a consultative process, their detailed implementation guidelines changed regularly. The requirements discussed below are accurate as at the end of December 2008. Since it is not possible for the author to continuously alter his research to reflect the rapidly changing regulatory framework, the research has been anchored to the regulations in effect at the end of 2008.

2.1 International Financial Reporting Standards

2.1.1 Traditionally, each country or state has set its own generally accepted accounting standards (GAAP) for its companies to follow. GAAP differed between states so that the financial statements of companies from a state are not directly comparable with those from another. For most states, GAAP also allowed subjectivity in the setting of assumptions in the valuation of liabilities and recognition of profits and losses, so that financial statements between companies cannot be easily compared even within a specific state. For most states, the local GAAP requirements also required the use of real-world models for the valuation of liabilities (see Section 4.2 for more details).

2.1.2 For the past few years, the International Accounting Standards Board (IASB) has been preparing what was called 'The Insurance Accounting Project'. The project aims to prepare a new international reporting standard for insurance companies. This is called the International Financial Reporting Standards (IFRS). According to a publication by Ernst and Young[1], the IASB is 'leading a global effort to transform financial reporting that has significant implications for insurance companies worldwide.' The adoption of the new standards is at the discretion of the member states of the IASB. The IASB concluded that the goal of new insurance standards should be to move towards fair-value accounting. This means that insurers should in future, for the purposes of financial reporting, value both assets and liabilities at the 'amount for which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's length transaction.'[2]

2.1.3 However, according to Ernst and Young[1], "the complexity of the issues, opposition and lack of time led to a 'two-phased' approach being introduced for insurance reporting." Phase 1 was completed when IASB issued IFRS 4 Insurance Contracts on 31 March 2004. IFRS 4 requires significantly increased disclosure and some changes to the way insurance contracts are accounted for. In particular, it introduced the requirement to separate contracts into insurance and investment components. Phase 1 also requires that assets be valued at fair value. There is no requirement to value

[1] *IASB IFRS 4 and IAS 39 – Insurance and Investment Contracts. Assessing the Implications of Phase 1 for Insurers*, Ernst and Young, 2004.
[2] *IFRS 4 Insurance Contracts*, International Accounting Standards Board, 2005.

insurance liabilities on the fair-value basis at this stage, although investment-related liabilities need to be valued in accordance with International Accounting Standards 39 (IAS 39).

2.1.4 IAS 39 is the standards set by the International Accounting Standards Committee (IASC), the predecessor of the IASB. IAS 39 has been in existence since 2001, and before that, it existed as IAS 25 since 1987. IAS 39 was amended in 2004, when it was changed to require the valuation of financial instruments at fair value. This is in contrast with the requirements of real-world models under most local GAAP.

2.1.5 The amended IAS 39 and newly issued IFRS 4 were effective from 1 January 2005, for the member states that chose to adopt it. For member states that chose not to adopt these standards, their local GAAP remained applicable. In South Africa, IAS 39 and IFRS 4 were adopted by insurers that are listed on the Johannesburg Stock Exchange as a part of their listing requirements. This, too, took effect on 1 January 2005. Publications by audit firms[1] and rating agencies[3] summarised the requirements under the new reporting standards.

2.1.6 Following the implementation of IFRS 4, financial results of insurers are likely to be more volatile than before. Fitch[3] distinguishes between 'accounting volatility' and 'economic volatility'. Economic volatility stems from economic mismatch inherent in the business and describes the underlying economic reality of the business. Accounting volatility describes asymmetrical accounting treatment, such as the difference in valuation basis used for valuation of assets (fair value) and that used for insurance liabilities (discounted value). To the extent that insurers are not yet required to value their insurance liabilities on a fair-value basis and there is no consensus on the best method for calculating the value of insurance liabilities in the absence of a deep and liquid market, their financial results are likely to be more volatile. Although this accounting volatility is an undesirable feature of IFRS 4, it does have its advantages. Fitch states:

> "Although phase 1 does result in some accounting volatility (e.g. due to assets being largely at fair value whilst liabilities are not) this sub-optimal result is still considered to be vastly preferable to a stability in reported results that is misleading."

2.1.7 Upon implementation of phase 2, in which both assets and liabilities will be expected to be valued on a fair-value basis, the accounting volatility is likely to be decreased. Provided that the assets and liabilities are well matched, economic volatility should also be decreased. It is therefore in the insurer's best interest to minimise its economic volatility by matching its assets and liabilities appropriately.

2.1.8 IFRS 4 specifies that the embedded derivatives of the investment component of liabilities must be valued on a fair-value basis and the movements in this value should be recorded in the income statement. Embedded derivatives that need to be recorded at fair value during phase 1 include life products offering a minimum return on surrender or maturity. Contracts whose insurance components have embedded derivatives are exempt from phase 1. Examples of the contracts exempt include GAOs and guaranteed minimum death benefits – these contracts contain significant elements of mortality risks and are thus considered to be insurance contracts. Regardless of this

[3] *Mind the GAAP. Fitch's View on Insurance IFRS*, Fitch Ratings, May 2004. *http://www.iasplus.com/resource/fitchinsurance.pdf*

exemption, some states have chosen to reserve for these on a market-consistent basis. In South Africa, for example, PGN 110.2[4] required companies to value these on a market-consistent basis. It is expected that phase 2 will require all of these benefits to be fair-valued.

2.1.9 Phase 2 is currently in consultation. The lack of a liquid market for insurance contracts has made the implementation of fair values for insurance liabilities a complex task. The IASB included a sunset clause in Exposure Draft 5[5], indicating that the target date for implementation of phase 2 is 2007, but this was deleted from IFRS 4 phase 1. This deletion implies that there is no implementation date for phase 2. A discussion paper[6] was issued in May 2007 for comments.

2.2 Solvency II & Individual Capital Assessment

2.2.1 Prudential Regulatory Solvency Requirements In the United Kingdom

2.2.1.1 The regulatory solvency requirement in the UK is governed under the Financial Services and Marketing Act 2000 (FSMA 2000). The rules are consolidated within a single Financial Services Act (FSA) Handbook. These are divided into a number of manuals or sourcebooks. The Integrated Prudential Sourcebook (PRU)[7] is one such manual and contains the prudential and notification requirements for insurers. PRU took effect from 31 December 2004. The main thrust of PRU is to set standards for capital management. It distinguishes between two kinds of insurers: regulatory-basis-only and realistic-basis insurers. The former includes insurers that have with-profits liabilities less than £500 million, if they choose to be so classified. The latter includes all insurers that have with-profits liabilities greater than £500 million. An insurer that has with-profits liabilities less than £500 million may choose to be classified in either category.

2.2.1.2 PRU introduces a two-pillar approach. The first pillar, which covers the public solvency information and appears in the FSA returns on the basis of a set of prescriptive rules, is further divided into two 'peaks', namely the 'Regulatory Peak' (Peak 1) and 'Realistic Peak' (Peak 2).

2.2.1.3 Peak 1 is a statutory valuation that satisfies the minimum European Union requirements. It applies to all insurers (both regulatory-basis-only and realistic-basis, although the prescribed methodologies may differ between them), and requires all realistic-basis insurers to value their embedded-derivatives liabilities stochastically. Regulatory-basis-only insurers need not value their embedded derivatives liabilities stochastically.

2.2.1.4 Peak 2, which only applies to realistic-basis insurers, is a realistic valuation that requires the market-consistent assessment of embedded derivatives. This peak is often referred to as the realistic balance sheet.

---

[4] *PGN 110.2: Allowance for Embedded Derivatives*, Actuarial Society of South Africa, Version 2, December 2006.
[5] *ED5 Insurance Contracts* (ISBN 1-904230-27-X), International Accounting Standards Board, 31 July 2003.
[6] *Preliminary Views on Insurance Contracts*, International Accounting Standards Board, May 2007.
[7] *Integrated Prudential Sourcebook*, Financial Services Authority, 31 December 2004.

2.2.1.5 Pillar 2, also known as the 'Individual Capital Assessment' pillar, or ICA, covers confidential assessment of solvency for the FSA, recognising all the risks that an insurer faces, not just the market risks covered by Pillar 1. This is required from all life insurers irrespective of their classification as regulatory-basis-only or realistic-basis. Pillar 2 of the regulation requires that insurers assess all the risks to which they are exposed and hold sufficient capital to cover these risks so that there is 99,5% certainty that each one will remain solvent at the end of one year.

2.2.1.6 As with the realistic assessment of business in Pillar 1 Peak 2, firms should base their ICA calculations on the use of market-consistent techniques. Unlike Pillar 1, where the methodology of valuation is prescribed, Pillar 2 is entirely principles-based. This principles-based approach recognises that different insurers are exposed to different kinds of risks. Insurers are also expected to align their day-to-day running of the business with ICA.

2.2.1.7 Market-consistent valuation methodologies are needed under Pillar 1 Peak 2 and Pillar 2 for realistic-basis insurers, and Pillar 2 for regulatory-basis-only insurers.

2.2.1.8 The actuarial profession in the United Kingdom must also obey guidance notes, which were until December 2006 issued by the Faculty of Actuaries and the Institute of Actuaries. Since then these guidance notes have been adopted by the Board of Actuarial Standards (BAS). GN 45[8] states:

> "Stochastic methods (which include both Monte-Carlo simulation models and closed-form solutions) should normally be used to calculate the market-consistent value of financial options."

2.2.1.9 It also states:

> "A portfolio of GAR [guaranteed annuity rates] has some similarities in form from an economic perspective with a portfolio of swaptions with a range of exercise dates, tenors and strike rates and with quantum equal to the value of the cash fund of the underlying policy on vesting. In most circumstances, the quantum depends upon persistency, take-up rate of pension at vesting, and then market values of the asset shares of the policies and the expected future progress of mortality rates.
>
> "It is therefore appropriate to calibrate stochastic models to interest rate swaptions. Account should be taken of the different profiles of the cash flows from a portfolio of swaps... It is normally necessary to model both the maturity benefit [i.e. the fund account] and the GAR simultaneously using appropriate correlations, although it may be possible to model each separately and combine the results using appropriate analytic techniques."

2.2.1.10 GN 47[9] provides recommended practice for the use of stochastic models for two purposes: market-consistent valuation of the liability and capital assessment. The latter is to ensure that the life insurer meets its liabilities in respect of embedded derivatives to a desired probability level. In respect of the former, it states that

> "Any stochastic approach used for valuing guarantees, options and smoothing when calculating WPICCs [with-profit insurance capital component – part of Pillar 1 Peak 2] should

[8] *GN 45: Determining With-Profits Insurance Capital Component,* The Board for Actuarial Standards (a part of Financial Reporting Council), Version 2.0 (BAS amendment 1), December 2006

[9] *GN 47: Stochastic Modelling of Economic Risks in Life Assurance,* The Board for Actuarial Standards (a part of Financial Reporting Council), Version 2.0 (BAS amendment 1), December 2006

be 'market consistent' and delivers [*sic*] prices of assets and liabilities that can be verified from the market."

2.2.1.11 It also states

"underlying model structures should be arbitrage free. However to the extent that there are small arbitrage opportunities in the simulation produced, this is acceptable provided they are not exploited for the benefit of the results."

2.2.1.12 PRU 7.4, like GN 47, requires

"...any stochastic approach used for valuing guarantees, options and smoothing when calculating WPICCs to be 'market-consistent'".

2.2.1.13 GN 46[10] sets out the recommended practice for the determination of ICA.

2.2.2 Regulatory Solvency Requirements In South Africa

2.2.2.1 South Africa has simpler prudential regulatory requirements than the UK. The Financial Services Board (FSB) is the main regulatory body in South Africa. The professional guidance notes (PGN) are the equivalent of the UK guidance notes. These are issued by the Actuarial Society of South Africa.

2.2.2.2 In respect of the valuation of embedded derivative liabilities, PGN 110[11] requires that maturity guarantees be valued using a real world model. Prior to this, there was no guidance on how to reserve for embedded derivatives.

2.2.2.3 Version 2 of this guidance requires liabilities related to embedded derivatives to be fair-valued. PGN 110.2[12] states:

> *"This guidance note recommends the use of market-consistent stochastic models to quantify reserves required to finance possible shortfalls in respect of embedded investment derivatives. Where there are no traded market instruments from which to calibrate the market-consistent model, the actuary may apply alternative methods and judgement provided that he/she can argue that such derived values used to calibrate the model are probable in the market."*

2.2.2.4 In addition to the reserve, insurers are expected to hold a capital adequacy requirement (CAR) to demonstrate solvency and to reduce the risks of adverse market movements leading to the insolvency of the insurer. The aim of this capital requirement is to ensure that each insurer has a 95% probability of remaining solvent on an ongoing basis. PGN 110.2 also prescribes the rules-based methodology for the calculation of the resilience component of CAR for embedded derivatives.

2.2.2.5 This prescription requires the projection of cashflows and the discounting of shortfalls using economic scenarios generated with a real-world ESG. The discount rate used is determined by the investment strategy of the assets underlying the guarantees reserve and the rate of returns that these assets earn under each scenario. The capital is

[10] *GN 46: Individual Capital Assessment,* The Board for Actuarial Standards (a part of Financial Reporting Council), Version 2.0 (BAS amendment 1), December 2006

[11] *PGN 110: Reserving for Minimum Investment Returns Guarantees*, Actuarial Society of South Africa, Version 1, December 2003.

[12] *PGN 110.2: Allowance for Embedded Investment Derivatives*, Actuarial Society of South Africa, Version 2, January 2007.

calculated as the average of the worst 15% of these present values. This is also known as the conditional-tails expectation (CTE). This is approximately equal to the 95th percentile of the distribution of present values. This requirement calls for the use of real-world models for capital calculations.

2.2.2.6 Practically, this method is difficult to implement for insurers who hedged their GAO exposures through the use of options and swaptions. This is because the discounting is done at the rate of return of the underlying asset, thereby necessitating a revaluation of the derivative hedge, at each projection time for each real-world scenario. This requires significant run times that are usually unavailable to insurance companies. See Chapter 7 for more information on runtimes.

2.2.2.7 In April 2008, PGN 110.2 was replaced by PGN 110.3[13]. The new version does not change the methodology related to the calculation of liabilities, but replaced the CTE calculation for capital with a set of sensitivity tests that are applied to the market-consistent value of liabilities. The value of the assets and liabilities are recalculated under each of the test scenarios. These scenarios include instantaneous reduction in equity and real-estate values, a reduction and an increase in the level of future yield-to-maturity.

2.2.3 Solvency II

2.2.3.1 Solvency II is a prudential regulation that is proposed for the European market for better risk management of insurance companies. Like Phase 2 of The Insurance Project, Solvency II is in consultation phase.

2.2.3.2 The aim of Solvency II is to set standards for the measurement and management of capital adequacy of insurance companies operating in the European Union. Like the ICA approach used in the UK, Solvency II aims to use a principles-based capital calculation. This deviates from the rules-based calculation that is currently employed.

2.2.3.3 Solvency II is not only about capital assessment, it is also about the capacity to understand, measure and manage the risks that insurers face. Solvency II promotes better risk management and insurers are expected to align their day-to-day running of their businesses with Solvency II. In other words, regulatory capital and economic capital need to be aligned.

2.2.3.4 Solvency II will also aim to take account of the International Accounting Standards to reduce the burden of reporting requirements, although regulatory and accounting standards may legitimately differ. It is therefore based on a market-consistent approach. The aim is to provide a 99.5% confidence that the company will be solvent after one year.

2.2.3.5 ICA requirements are expected to be replaced by those of Solvency II when the latter is implemented in 2012.

2.2.4 Relationship Between IFRS 4, ICA And Solvency II

[13] *PGN 110.3: Allowance for Embedded Derivatives*, Actuarial Society of South Africa, Version 3, April 2008.

2.2.4.1 IFRS 4, ICA and Solvency II are all based on market-consistent principles. IFRS 4 prescribes the financial reporting standards for many member states or countries, including South Africa, United Kingdom and all continental European countries. ICA and Solvency II prescribe principles-based methods for calculating capital of UK and European Union respectively. ICA is expected to be replaced by Solvency II when the latter is implemented.

2.2.4.2 As mentioned above in paragraph 2.2.3.4, significant synergies can be leveraged by basing Solvency II on IFRS 4.

2.3 Linking Economic Capital to Risk Management

2.3.1 The change of regulatory capital requirements from a prescriptive, rules-based calculation to a principles-based calculation represents an alignment of risk and capital. Where previously all companies used the same calculations to calculate capital to protect themselves against adverse conditions, regardless of their risk management strategies, the principles based calculations reward companies that manage their risks effectively. Companies are required to hold more capital for higher levels of risk. The holding of additional capital has a cost. If appropriately hedged, the amount of capital needed in respect of these risks will be reduced, reducing the cost of capital. This reduction of capital incentivises insurers to better manage their risks.

2.3.2 While hedging reduces the cost of capital, the implementation of the hedge does incur cost. This cost comes in the form of giving up potential upside-risks, as well as the cost of the hedge instruments (either contract fee or active trading costs). Companies must find the hedges that strike a balance between effectiveness and costs, where the latter incorporates the cost of capital, the cost of hedge instruments and the cost of upside gains. In practice, this balance is influenced by the companies' risk appetite.

Chapter 3. Literature Review

3.1 Guaranteed Annuity Options

3.1.1 Work on the reserving and pricing of insurance contracts with financial guarantees has been published as early as Black & Scholes (1973), Brennan & Schwartz (1977) and Ford et al. (1980). Bolton et al. (1997) surveyed 85 companies in the UK. Of the 66 companies that responded, 44 had GAOs. They estimate that the responses represented at least 90% of the total market liabilities. The methods of valuing GAOs at that time were wide-ranging. These included:

- take no account of the guarantee, and
- calculate liability as the greater of the value of the cash option and the value of the guarantee on the deterministic valuation basis.

3.1.2 Some companies made allowance for the effect of future premiums to which the guarantee apply, while others ignored future premiums.

3.1.3 Bolton et al. (*op. cit.*) note that the majority of companies made no allowance for the guarantee in setting investment guidelines. They also explored the inter-relationship of the elements that affected the value of GAOs, such as mortality, the effective rate of interest, type of annuity (such as joint annuities, increasing annuities or annuities with guaranteed periods), annuity frequency, annuities in advance or in arrears, and allowances for expenses. This investigation was done using scenario testing on a deterministic basis. They also note that the majority of the GAO business in the UK was written under with-profits terms. An important but controversial question regarding the management of GAOs was raised: Should the shareholders or the with-profits policyholders be paying for the guarantees?

3.1.4 They concluded that there is no clear answer to this question and that companies should develop their own practice. Equitable Life decided that the policyholders should pay for the guarantee. This was implemented by having different scales of terminal bonus on the with-profits fund: the policyholders who exercise the annuity option received a lower terminal bonus. This practice made a mockery of the GAO and prompted an inquiry by Lord Penrose on request from the House of Commons. Penrose[14] found the practice to be improper.

3.1.5 A stochastic approach using the real-world Wilkie (1995) model was also tested by Bolton et al. (*op. cit.*).

3.1.6 As mentioned above in Section 1.3.1, various authors (Boyle & Hardy, *op. cit.*; Ballotta & Haberman, 2003) identified three main factors which caused the embedded guarantee of the GAO sold by life insurers to become onerous, viz. falling interest rates, longevity and strong equity performance. Boyle and Hardy (*op. cit.*) also show that interest-rate volatility had minimal effect on the value of the guarantee except where the guarantee is very close to at the money.

[14] Right Honourable Lord Penrose (2004). *The Equitable Life Inquiry*. HM Treasury. Available online at http://www.hm-treasury.gov.uk/d/penrose_prelimhs.pdf

3.1.7 Boyle & Hardy (*op. cit.*) then proceeded to quantify each of the three effects and concluded that falling interest is by far the most risky of the three. Three methods of dealing with risks associated with the writing of contracts with financial guarantees were also identified. These are:

- The company sets aside enough capital to ensure that the liabilities under the guarantee will be covered with a high probability. The liabilities are calculated using a stochastic simulation approach.
- The company reinsures the liability with another financial institution such as a reinsurer or investment bank.
- The company sets up a replicating portfolio of traded securities and adjusts the matching position dynamically over time so that the market value of the portfolio corresponds to the liability under the GAO at maturity.

3.1.8 Boyle & Hardy (*op. cit.*) also derives a formula to value the GAO, assuming risk-neutrality and without assuming a specific model for interest rates. As an illustration, they then apply a Hull–White (1990) interest-rate model to calculate prices for the option. This is done under the assumption that mortality risk is independent of the financial risk.

3.1.9 Ballotta & Haberman (2003, unpublished) propose a method to value GAOs by comparing their payoff structure to that of a call option written on a coupon-bearing bond. The model uses a one-factor Heath, Jarrow & Morton (1992) (HJM) framework for the term structure of interest rates. Two alternative formulations of the HJM framework based on different specifications of forward-rate volatilities were used – the first relies on the assumption of constant volatility while the second uses an exponentially decaying volatility structure. Furthermore, under the assumptions of deterministic equity returns and independence between mortality and financial risks, a general pricing framework is proposed and closed-form formulae for the value of simple GAO contracts are obtained. The sensitivity of the price of the option to changes in the key parameters is also analysed.

3.1.10 Ballotta & Haberman (2006) extend on their work (2003, unpublished) by investigating the fair valuation of GAOs in the presence of stochastic mortality. They disapprove of the fair-valuation directives promoted by IASB — these focus only on the financial risk affecting life insurers — and note that UK historical experience shows that very long-term products like GAOs are significantly exposed to unanticipated changes over time in the mortality rates of the reference population. For this reason, fair valuation techniques proposed by IFRS should be integrated with an accurate assessment of future mortality rates. They use the same financial model as in their previous work (*op. cit.*), but they apply the reduction factor model (Renshaw & Haberman, 2003) to the hazard rate, which describes the mortality process:

$$\mu(x+z,t+z) = \mu(x+z,0)e^{(\alpha+\beta(x+t))(t+z)+\sigma_i Y_{t+z}}$$

where:

- $\{Y_t : t \geq 0\}$ is a stochastic process (assumed to be an Ornstein-Uhlenbeck process), which introduces random variations in the forecast trend,
- $\mu(x+z,0) = a_1 + a_2 R + e^{b_1+b_2 R+b_3(2R^2-1)}$

- $R = \frac{x + z - 70}{50}, x \geq 50.$

3.1.11 Ballotta & Haberman (2006) assume that investors are indifferent to mortality risk, so that there is no difference between a real-world and risk-neutral stochastic model for mortality. They make this assumption as there is inherently an incomplete market for mortality with which a risk-neutral model can be calibrated. Since there is no closed-form solution to this model, they calculate the GAO value by Monte–Carlo methods. They conclude that under the base scenario (which includes stochastic financial and mortality models), the GAO value is increase by 26%. They also make a series of sensitivity tests by changing each parameter in the financial and mortality models whilst holding all else constant.

3.1.12 Lee (unpublished) and Cairns (unpublished) discuss the actuarial valuation of GAOs using real-world stochastic interest-rate models. Wilkie, Waters & Yang (2003) and Zhou (unpublished) compare the results of the actuarial valuation to the financial valuation of GAOs. The latter is calculated as the value of a portfolio of simple instruments such as equities and bonds that delta hedges the portfolio of GAOs. This valuation is not entirely appropriate as it ignores the vega risk. Lee (*op. cit.*), Cairns (*op. cit.*) and Zhou (*op. cit.*) also simplified the modelling by assuming no correlation between equities and interest rates – clearly an inappropriate assumption.

3.1.13 Biffis & Millossovich (2006) use an affine jump-diffusion process to represent the short-rate process, which represents the risk-free returns. A similar process is used to describe the returns from a risky asset. They also use an affine process with stopping times that follow a doubly-stochastic Poisson process to model dynamic and stochastic survival probabilities. A doubly-stochastic Poisson process models the survival probability as two nested stochastic processes. The first is a Poisson process to provide a count of the number of deaths up to time t. The second process describes the intensity function of the Poisson process. A combination of these processes provides tractability in the calculation of fair-value GAO liabilities. Luciano & Vigna (2005) provide a high level discussion of using affine process and doubly-stochastic process to model mortality. They note that the theory behind affine process and doubly-stochastic process is enormous and is beyond the scope of their paper. For the same reason, this dissertation will not discuss these processes except to alert the reader to their existence. Interested readers are encouraged to read Luciano & Vigna (*op. cit.*) and Duffie (2002) for more in-depth discussions.

3.2 Risk Management of GAO

3.2.1 Boyle & Hardy (*op. cit.*) examine a number of conceptual and practical issues involved in the dynamic hedging of interest-rate risk. The issues are as follows:

- The asset-price dynamics cannot be completely correctly specified.
- It is not possible to rebalance the hedge portfolio continuously.
- There are transactional costs involved in trading.

3.2.2 Pelsser (2003) notes that dynamic delta hedging using only simple market asset creates what is known as a 'feedback loop', which can be extremely dangerous. This is caused by the pursuing of the same investment strategy by many companies, so that the effect of the dynamic rebalancing of their collective hedge portfolios is no longer

negligible to the market as a whole. To hedge a typical investment guarantee, such as those that guarantee a minimum lump sum at maturity (which is effectively a put option), companies would need to sell the underlying asset of the put option. If the market falls, then more of the same asset will need to be sold in order to maintain a delta-neutral asset position. This forces the asset prices even lower, especially if the insurance companies have large positions that need rebalancing. The cycle of selling is then repeated. The 1987 market crash was blamed on computer systems that followed automated delta-hedging strategies without recognising the feedback loop. The same problem exists for GAOs.

3.2.3 Despite these issues, Boyle & Hardy (*op. cit.*) examine a conceptual dynamic-hedging portfolio for the equity and interest-rate risks underlying GAOs. This portfolio includes a long position in the equities underlying the fund account of the GAOs' base contract, a long position in zero-coupon bonds that mature after the vesting date of the GAO (or other equivalent fixed-income instruments) and a short position in a zero-coupon bond (ZCB) that matures on the vesting date.

3.2.4 A static hedge utilising swaptions is also investigated for hedging the interest-rate risk. Many authors, such as Boyle & Hardy (*op. cit.*), Yang (unpublished), Wilkie et al. (2003) and Pelsser (*op. cit.*), likened a portfolio of GAOs to a portfolio of interest-rate vanilla swaptions with different option dates and tenors. They concluded that this portfolio represents an ideal hedge for such liabilities, but does not account for the changes in the value of the fund account that will ultimately determine the notional to which the GAO will apply. As a result, this static hedge will need to be readjusted regularly and this could be very costly.

3.2.5 Bolton et al. (1997) explored the use of swaps and swaptions to hedge GAO risks. They note that the use of swaptions also presents other problems. These are as follows:

- Policyholders may choose their retirement date, so that the vesting date of the swaption cannot be determined exactly.
- The terms of the annuity and the fixed term of the swaption may not correspond exactly.
- The swaption portfolio is subject to credit risk.
- Some liabilities may have option dates that are longer than those available in the swaptions market.

3.2.6 Milevsky & Promislow (2001) discuss the natural hedging of GAOs by selling life insurance. Boyle & Hardy (*op. cit.*) note that there are limitations to this method because it may not be possible to sell insurance and GAO policies to the same type of policyholders. The insurer also needs to have a good estimate of the distribution of future mortality — if this were true, then GAOs would not have created so many problems in the first place.

3.3 Interest-rate Models

Vital to the market-consistent valuation of complex assets and liabilities is a robust interest-rate model. Many interest-rate models exist in academic literature, such as Vasicek (1977), Brennan–Schwartz (1979, 1980, 1982 and 1983), Cox–Ingersoll–Ross (1985), Ho–Lee (1986), Black–Derman–Toy (1990), Black–Karasinski (1991), Longstaff–Schwartz (1992) and Hull–White (1993a). There are too many models to discuss fully. Cox, Ingersoll & Ross

(*op. cit.*) states "The literature in the area is voluminous, and a comprehensive survey would warrant a paper in itself." The important landmarks in interest-rate modelling are described below.

3.3.1 Term Structure of Interest Rates

3.3.1.1 The theories of the term structure of interest rates refer to those that explain why yields vary with maturity. These are the market-expectations hypothesis, liquidity-preference hypothesis and market-segmentation theory. These theories are regularly quoted in academic literature, such as Brennan & Schwartz (1979, 1980, 1982 and 1983), Cox, Ingersoll & Ross (1985), Audley, Chin & Ramamurthy (2002), Hull (2006) and many others.

3.3.1.2 The market-expectations hypothesis states that bonds are priced so that the implied forward rates are equal to the expected spot rates. Generally, this theory implies that:
- the return on holding a long-term bond to maturity is equal to the expected return on repeated investment in a series of the short-term bonds; and
- the expected return over the next short period is the same for bonds of all maturities.

3.3.1.3 The liquidity-preference hypothesis (Hicks, 1946) agrees with the importance of the expectations hypothesis, but recognises the risk preferences of the market participants. It states that the risk aversion of market participants causes the forward rates to be systematically higher than expected spot rates by an amount that increases with maturity. This term premium is the increment that is required to induce the investors to hold longer-term bonds.

3.3.1.4 The market-segmentation theory (Culbertson, 1957) asserts that investors have strong maturity preferences and bonds of different maturities trade in separate and distinct markets. There is no reason for the term premiums to be positive or increasing functions of maturity.

3.3.2 Vasicek

3.3.2.1 Vasicek (1977) derives a general form of a term structure model for interest rates in an equilibrium framework. He makes three assumptions.
- The instantaneous spot interest rate follows a continuous Markov process (i.e. a diffusion process).
- The price of a bond depends only on the instantaneous spot rate over its term.
- The market is efficient.

3.3.2.2 Under the first assumption, the process can be described by a stochastic differential equation (SDE) of the form:

$$dr(t) = f(r(t),t)dt + \sigma(r(t),t)dZ\,, \tag{3.1}$$

where:
- $r(t)$ is the instantaneous short rate;
- Z is a Wiener process; and
- $f(r(t),t)$ and $\sigma^2(r(t),t)$ are the instantaneous drift and variance of the $r(t)$ process respectively.

3.3.2.3 The last assumption implies that the there are no transaction costs, information is freely available to all investors simultaneously, all investors prefer more wealth to less, all investors have homogeneous expectations and that there is no riskless arbitrage.

3.3.2.4 Under these assumptions, he shows that the expected rate of return on any bond in excess of the spot rate is proportional to its standard deviation. The development of the returns model is based on a no-arbitrage argument. A partial differential equation (PDE) for the process of prices is derived:

$$\frac{\partial P}{\partial t}+(f+\sigma q)\frac{\partial P}{\partial r}+\frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2}-rP=0,$$

where:

- P is the price of the bond; and
- q is the market price of risk.

3.3.2.5 Once the short-rate process has been characterised and the market price of risk q is specified, bond prices can be obtained by solving the PDE subject to the boundary condition that the price of a zero-year ZCB has a unit-value.

3.3.2.6 As a specific case of the model, Vasicek (*op. cit.*) specified the following model for the interest rate process:

$$dr(t)=\alpha(\gamma-r(t))dt+\sigma dZ.$$

3.3.2.7 This model form corresponds to (3.1) with $f(r(t),t)=\alpha(\gamma-r(t))$ and $\sigma(r(t),t)=\sigma$. This form of interest rate model is subsequently referred to as the Vasicek model. The short rate $r(t)$ follows an Ornstein–Uhlenbeck process with $\alpha>0$. This process, results in a stationary process, unlike a random walk (Wiener process), which is an 'unstable process' and after a long time results in infinite values. The instantaneous drift $\alpha(\gamma-r(t))$ keeps pulling the process towards its long-term mean γ with a magnitude proportional to the deviation from the mean. The stochastic element has a constant standard deviation σ. This causes the process to fluctuate around the level γ erratically but continuously.

3.3.2.8 Under this interest-rate process, the bond prices follow the following process:

$$P(t,s,r(t))=\exp\left[\frac{1}{\alpha}\left(1-e^{-\alpha(s-t)}\right)\left(R(\infty)-r(t)\right)-(s-t)R(\infty)-\frac{\sigma^2}{4\alpha^3}\left(1-e^{-\alpha(s-t)}\right)^2\right],t\le s,$$

where

- $P(t,s,r)$ is the price, at time t, of a bond that matures at time s, where the instantaneous rate at time t is $r(t)$; and
- $R(\infty)=\gamma+\frac{\sigma q}{\alpha}-\frac{\sigma^2}{2\alpha^2}$.

3.3.2.9 The term structure of interest rates takes on the form:

$$R(t,T)=R(\infty)+\left(r(t)-R(\infty)\right)\frac{1}{\alpha T}\left(1-e^{-\alpha T}\right)+\frac{\sigma^2}{4\alpha^3 T}\left(1-e^{-\alpha T}\right)^2,T\ge 0,$$

where $R(t,T)$ is the T-year yield at time t.

3.3.2.10 The model form is undesirable for various reasons:

- This standard deviation σ is independent of the level of the short rate at the time. This is not realistic as interest rates are typically more volatile when rates are high and vice versa.
- The model does not preclude interest rates from becoming negative.

3.3.2.11 Vasicek (*op. cit.*) acknowledges the imperfections of the model:

> "It is not claimed that the process given by the equation represents the best description of the spot rate behaviour. In the absence of empirical results on the character of the spot rate process, this specification serves only as an example."

3.3.2.12 While the model is arbitrage-free in the sense that no bond or option prices produced by the model will permit arbitrage, it is, nonetheless not arbitrage-free in the context of actual market prices. This is because the model produces a term structure as an output but does not accept the current term structure as an input. The Vasicek model generates prices that are inconsistent with the current term structure.

3.3.3 Brennan–Schwartz

3.3.3.1 Brennan & Schwartz (1979, 1980, 1982 and 1983) recognise that the traditionally theory of term structure stipulates the relationship between the forward rates in the term structure and the corresponding expected future short rates of interest. In particular, the expectations hypothesis asserts that forward rates are equal to the expected future spot rates. The liquidity premium hypothesis, in contrast, asserts that forward rates always exceed the corresponding expected future short rates by a liquidity premium. This is required to compensate investors for the greater capital risk in longer-term bonds. They also note that the market segmentation hypothesis may be regarded as a modification of the liquidity premium hypothesis to allow for positive or negative liquidity premium on longer-term bonds by recognising that investors with longer-term horizons do not necessarily consider longer-term bonds as being more risky, so that prices of bonds of different maturities are determined by investors' preferences of different horizons.

3.3.3.2 The result is that forward rates may bear no systematic relationship to expected future short rates. They recognise that the limitation of liquidity premium and market segmentation hypothesis is that they lack specificity: the relationship between liquidity premium to maturity is not specified so that there are many undetermined parameters in the model.

3.3.3.3 The authors also recognised that the typical interest-rate models at the time modelled the short rates are Markov models: the current instantaneous rate contains all information about future interest rates, so that the value of a bond of any maturity may be written as a function of this instantaneous rate. They argued that this is at odds with reality — a simple one factor model cannot capture the multidimensionality of the term structure as it only captures the parallel shift of interest rates.

3.3.3.4 They model prices of bonds of different maturities by allowing for change in the instantaneous rate to depend on only on the current short rate, but also the long term rate of interest based on consol yields, so that the long term rate and instantaneous rate follow a joint Gauss–Markov process.

3.3.3.5 The two factor models assume that the two key factor dynamics can be written in terms of two correlated Brownian motions Z_1 and Z_2 with a correlation of ρ. The dynamics of the short rate and long rate processes are given by:

$$dr(t) = (a_1 + b_1(l(t) - r(t)))dt + r(t)\sigma_1 dZ_1 \text{ ; and}$$

$$dl(t) = l(t)(a_2 + b_2 r + c_2 l(t))dt + l(t)\sigma_2 dZ_2 ,$$

where Z_1 and Z_2 represent standard correlated Wiener processes.

3.3.3.6 The volatility underlying each of the two processes is assumed to be proportional to the current value of the respective rate. The short rate has a tendency to drift towards the current value of the long rate. The model calibration is by statistical estimation.

3.3.3.7 Benhamou (unpublished) point out the following as serious flaws of the Brennan–Schwartz model:

- The model does not have an easy solution for pricing of simple instrument.
- The model can permit exceptionally large changes in finite time so that rates can go to infinity.

3.3.4 Cox–Ingersoll–Ross

3.3.4.1 Cox, Ingersoll & Ross (CIR) (*op. cit.*) point out that there are four strands of thoughts with regard to interest rate models. The first three are the three theories of the term structure of interest rates (i.e. market-expectations hypothesis, liquidity-preference hypothesis and market-segmentation theory). CIR note that there are two difficulties associated with the previous hypotheses: they are theories that give reasons to explain why forwards do not necessarily equal to the expected future spot rates; and all of these theories are given in ex-ante terms and that they need to be linked to ex-post realisations to be testable. The attempts to deal with these two difficulties constitute the fourth strand of thought.

3.3.4.2 They consider this problem in general equilibrium theory, and their approach contains elements of all four strands of thoughts. In particular, they consider individuals with homogenous expectations but with different and specific preferences about the timing of their consumptions and so have a preferred habitat. An equilibrium model is a complete intertemporal description of a continuous time competitive economy. The expected utility of individuals is modelled, allowing for deferred consumption in the future. The interest rates and expected rates of return on contingent claims are adjusted until all wealth in invested. A full summary of their derivation is outside of the scope of this research. Readers are encouraged to visit the original literature (Cox, Ingersoll & Ross, *op. cit.*) for a fuller description.

3.3.4.3 In the same paper, CIR also proposed a one factor model for the term structure of interest rates, which they derived through equilibrium assumptions of a production process. They make assumptions about a competitive economy that produces only a single good. The development of the interest rate is given by the SDE:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r^{0.5} dZ .$$

3.3.4.4 If $\kappa > 0$ and $\theta > 0$, this corresponds to a first order autoregressive process where the random moving interest rate mean reverts to a central location θ. The parameter κ

determines the speed of this mean reversion. It can be shown that if $\sigma^2 > 2\kappa\theta$, the process r can reach zero. If $\sigma^2 \leq 2\kappa\theta$, then the process will not reach zero. In any case, an initial nonnegative interest rate can never become negative.

3.3.4.5 As is the case with the Vasicek (*op. cit.*) model, the CIR model is an equilibrium model. It generates yields that are arbitrage free in the sense that no arbitrage opportunities will be projected in future, but the model is not arbitrage-free in the context of matching actual market prices.

3.3.5 Ho–Lee

3.3.5.1 The Ho–Lee model (Ho & Lee, 1986) is the first model that allows for the matching of the initial term structure, so that the theoretical prices of bonds match those in the market. It is based on the Black–Scholes (1973) paradigm. The bond prices are modelled so that their movements are arbitrage free. A risk-neutral implementation was used. The model is presented with a discrete binomial-tree implementation with time spacing of one year. A bond price process, rather than a short-rate process, is specified by Ho & Lee (*op. cit.*). It has the following equivalent SDE for the short-rate process:

$$dr(t) = \theta dt + \sigma dZ .$$

3.3.5.2 The model allows the average change in short rate to drift with parameter θ and random volatility σ. The shortcomings of this model include:

- negative interest rates are not precluded; and
- there is no mean reversion so that rates can increase without bound if $\theta > 0$ and decrease without bound if $\theta < 0$.

3.3.6 Black–Karasinski

3.3.6.1 Black & Karasinski (1991) presents a model that does not allow negative interest rates. The short rates are modelled as a lognormal process:

$$d\log(r(t)) = \phi(t)\left[\log\mu(t) - \log r(t)\right]dt + \sigma(t)dZ ,$$

where:

- $\mu(t)$ is the target rate;
- $\phi(t)$ is the mean reversion speed; and
- $\sigma(t)$ is the local volatility in the expression for the local change in $\log(r)$.

3.3.6.2 The model outputs:

- the yield curve that gives the yield on a ZCB of each maturity;
- the volatility curve that gives the current yield volatility on a ZCB of each maturity; and
- the cap curve that gives the price of an at-the-money differential cap.

3.3.6.3 They recognise that all of the outputs of the model are observable in the market as they all correspond to market prices. They implement the model in two different settings: one where the inputs are known and another where the outputs are known.

3.3.6.4 When the outputs are known, the model is intended to be a risk-neutral one and the inputs ($\mu(t)$, $\phi(t)$ and $\sigma(t)$) are chosen to give known outputs that match market prices. They state:

> "We do not claim that our inputs imply a reasonable process for the short rate, except as a rough estimation. We choose inputs that give reasonable outputs, though they may be somewhat unreasonable themselves...by generating a 'risk-neutral' distribution... This is not a true distribution, but it nonetheless gives correct option prices."

3.3.6.5 When the inputs are known, the model serves as an equilibrium model. In this case, the inputs ($\mu(t)$, $\phi(t)$ and $\sigma(t)$) are chosen to reflect a reasonable progression and the output is equilibrium option prices that may be different to the current market prices.

3.3.6.6 In their paper, the model is implemented as binomial trees. The implementation lacks degrees of freedom so that it is not possible to have equal time spacing. When the mean reversion $\phi(t)$ is positive, the time spacing between nodes decreases over time.

3.3.6.7 The Black–Karasinski model is discussed in more details in Section 5.3.4.

3.3.7 Hull–White

3.3.7.1 Hull & White (1990) extends the Vasicek (1977) and CIR (1985) models so that they are both consistent with the current term structure of interest rates and either the current volatilities of all spot interest rates or the current volatilities of all forward interest rates.

3.3.7.2 Hull & White (*op. cit.*) proposes an extension to the Vasicek (*op.cit.*) model that is very tractable. The model has the form:

$$dr(t) = \big(\theta(t) + a(t)(b - r(t))\big)\,dt + \sigma(t)dZ\,.$$

3.3.7.3 It has the same form as the standard Vasicek (*op.cit.*) model, but the parameters are all deterministic functions of time. This model is therefore also known as the extended Vasicek model. This provides the necessary degrees of freedom to match the market prices.

3.3.7.4 The parameters of the process followed by the short-term interest rate and European bond option prices can be determined analytically. They fitted the extended-Vasicek model to the initial term structure of interest rates and the initial term structure of interest-rate volatilities. Interest-rate option prices are calculated and compared with those generated by a variety of different one- and two-factor models. They conclude that the extended-Vasicek model provides a good analytic approximation to the European option prices given by the other models.

3.3.7.5 The method for fitting the models to market data that is presented in Hull & White (1990) can only be used on the extended-Vasicek and extended-CIR models. Similarly, the methods proposed in Ho & Lee (*op.cit.*) and Black–Karasinski (*op. cit.*) are also model-specific. The trinomial tree method presented Hull & White (1993a) is more general and may be applied to all short-rate models that are one factor and Markov in nature, such as the extended-Vasicek (Hull & White, 1990), extended-CIR

(Hull & White, 1990), Black–Karasinski (*op.cit.*). This procedure is explained fully in Appendix 5.1.

3.3.8 Two Factor Models

Each of the above mentioned models with mean-reversion properties can be extended to include more factors. This allows the projection of more complex term structure and volatility patterns at the cost of pragmatism. The literature in this area is not reviewed but the author acknowledges their existence.

3.3.9 LIBOR Market Model

The interest rate models described above all belong to a more general class of short-rate models. The LIBOR market model, or Brace–Gatarek–Musiela (1997) model, is a different class of model that is capable of being fitted to much more complex volatility structures than short-rate models. Models in this class generally have more degrees of freedom. This type of model is discussed more fully in Section 5.3.3.

3.4 Equity Models

An equity model is also needed for projecting equity returns for the purpose of projecting the fund value at retirement. Again, these are too many to discuss here and a selection of models are described below.

3.4.1 Black–Scholes

3.4.1.1 Black & Scholes (1973) propose a model that is capable of pricing simple European call and put options. Under some assumptions, the model yields a closed-form solution for the valuation of these options. These assumptions are as follows:

- The risk-free rate of return is known and is constant.
- The share price follows a random walk in continuous time with a variance that is proportional to the square of the stock price so that the share price at the end of any finite interval is lognormal.
- The share pays no dividends.
- The option is European.
- There are no transaction costs in buying or selling the share or the option (except for the price of the option).
- It is possible to borrow any fraction of the price of the share to buy it or hold it.
- There are no penalties in short selling.

3.4.1.2 The second assumption implies that the share price follows the following process:

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ .$$

3.4.1.3 Since the model is implemented in a risk-neutral world, the share is expected to earn a risk-free rate of interest. In this case, $\mu = r$ so that the process becomes:

$$dS(t) = rS(t)dt + \sigma S(t)dZ .$$

3.4.1.4 The share price follows a geometric Brownian motion (i.e. it has a lognormal distribution):

$$Log\frac{S(T)}{S(t)} \sim N\left(r(T-t), \sigma^2(T-t)\right).$$

3.4.1.5 Under these assumptions, an European call option on a share that does not pay dividends has the following price:

$$C = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),$$

where:

- K is the strike price of the option;
- S is the current price of the share;
- T is the expiry date of the option;
- t is the date of the valuation of the option;
- Φ is the cumulative standard normal distribution function;
- $d_1 = \frac{log\frac{S}{K} + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma(T-t)^{\frac{1}{2}}}$; and
- $d_2 = d_1 - \sigma(T-t)^{\frac{1}{2}}$.

3.4.1.6 The price of an European put option on a share that does not pay dividends is calculated as:

$$P = Ke^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1).$$

3.4.1.7 More details of the Black–Scholes implementation by Monte–Carlo simulations and its practical relevance are given in Section 5.6.1.

3.4.2 The Regime-switching Lognormal (RSLN) Model

3.4.2.1 Hardy (2004) notes that the model proposed by Black–Scholes (1973), when implemented under real-world probabilities, does not generates a thick enough tail when compared with historical returns. In particular, the model assumes that the returns over non-overlapping periods are independent. This provides a reasonable approximation over shorter time intervals but does not appeal to longer term problems.

3.4.2.2 She proposes a model that incorporates stochastic volatility — the parameters of the lognormal model take on any one of n sets of discrete values and switch between these values randomly. The switch between regimes is Markov – it is determined by the prevailing regime but not on the history of previous regimes. The model maintains some of the tractability of the lognormal normal that underlies the Black–Scholes model, but more accurately captures the extreme observed behaviour.

3.4.2.3 The rationale behind the regime-switching framework is that the market may switch from time to time between different economic environments, such as a stable low-volatility regime with moderately expected returns, and a more unstable high-volatility regime with very low or negative returns.

3.4.2.4 Consider the case where there are n different such economic regimes, then the process followed by the share price conditional on the economy being in regime η_t, is:

$$dS(t)\big|\eta(t) = \mu_{\eta(t)}Sdt + \sigma_{\eta(t)}SdZ,$$

so that

$$Log\frac{S(t+1)}{S(t)}\bigg|\eta(t) \sim N\left(\mu_{\eta(t)}, \sigma^2_{\eta(t)}\right).$$

3.4.2.5 Hardy (*op.cit.*) further notes that the modelling complexity increases with increase in the number of regimes. She found no evidence of the need to use more than two regimes.

3.4.2.6 The transition between the different regimes is modelled as a Markov chain. This is governed by a transition matrix $\mathbf{P}$, whose elements have the following interpretation:

$$p_{i,j} = \Pr\left[\eta(t+1) = j \,|\, \eta(t) = i\right], \qquad i, j = 1, 2, ..., n.$$

3.4.2.7 The unconditional regime distribution of this is given as $\boldsymbol{\pi}$, a $n \times 1$ vector, so that $\boldsymbol{\pi}'\mathbf{P} = \mathbf{P}\boldsymbol{\pi} = \boldsymbol{\pi}'$. In the case of a 2 regime model, this simplifies to:

$$\pi_1 = \frac{p_{1,2}}{p_{1,2} + p_{2,1}}, \quad \pi_2 = 1 - \pi_1.$$

3.4.2.8 The six parameters of the model (μ_1, μ_2, σ_1, σ_2, $p_{1,2}$, $p_{2,1}$) can be calibrated by maximum-likelihood methods. Hardy (*op.cit.*) notes that the model provides a better fit to the Toronto Stock Exchange and S&P returns than the standard lognormal model.

3.4.2.9 While this model maintains some of the tractability seen with the lognormal model, it does have some drawbacks:

- Some argue that two regimes do not sufficiently capture the complex dynamics (and serial-autocorrelation) of the economy.
- The model does not extend to more regimes easily. The extension to three regimes increases the number of variables to 12.
- The large number of parameters results in unstable calibrations. This is a sign that the model is a poor fit to historical data.
- Because the model is real-world and calibrated to historical data, it does not reflect the prevailing market conditions. The model cannot be used for the pricing of liabilities.

3.4.3 The Variance–Gamma Model

3.4.3.1 Madan & Seneta (1990) propose an equity model that satisfies four desirable properties:

- long tailedness relative to the normal distribution for daily returns, with returns over longer periods approaching normality;
- finite moments for the lower powers of returns;
- scalable up to any length of time; and

- extendable to multiple processes with elliptical multivariate distributions, thereby maintain validity of the capital asset pricing model.

3.4.3.2 Let $R(t) = \log \frac{S(t+1)}{S(t)}$, where $S(t)$ is the share price at time *t*. The model assumes that $R(t) \sim N(\mu, \sigma V)$, where:

- μ and σ are constants;
- V has a gamma distribution with parameters c and γ, so that it has a density function of:

$$g(v) = \frac{c^{\gamma} v^{\gamma-1} e^{-cv}}{\Gamma(\gamma)}, v > 0,$$

where Γ is the gamma function.

3.4.3.3 $R(t)$ is then said to have the variance–gamma distribution.

3.4.3.4 The density function of $X = R - \mu$ is:

$$f(x) = \int_0^{\infty} \frac{\exp\left(-\frac{x^2}{2\sigma^2 v}\right)}{\sigma\sqrt{2\pi v}} g(v)dv, x \in \mathbb{R}.$$

3.4.3.5 There is no closed-form solution for *f*, but the characteristic function can be expressed as:

$$\begin{aligned}
\phi_X(u) &= \mathrm{E}\left[e^{iuX}\right] \\
&= \int_{-\infty}^{\infty} \left(\int_0^{\infty} \frac{\exp\left(-\frac{x^2}{2\sigma^2 v}\right)}{\sigma\sqrt{2\pi v}} g(v)dv \right) e^{iux} dx \\
&= \int_0^{\infty} \int_{-\infty}^{\infty} \left(e^{iux} \frac{\exp\left(-\frac{x^2}{2\sigma^2 v}\right)}{\sigma\sqrt{2\pi v}} dx \right) g(v)dv \\
&= \int_0^{\infty} \exp\left(-0.5\sigma^2 v u^2\right) \frac{c^{\gamma} v^{\gamma-1} e^{-cv}}{\Gamma(\gamma)} dv \\
&= \int_0^{\infty} \frac{c^{\gamma} v^{\gamma-1} \exp\left(-(c+0.5\sigma^2 u^2)v\right)}{\Gamma(\gamma)} dv \\
&= \frac{c^{\gamma}}{(c+0.5\sigma^2 u^2)^{\gamma}} \int_0^{\infty} \frac{(c+0.5\sigma^2 u^2)^{\gamma} v^{\gamma-1} \exp\left(-(c+0.5\sigma^2 u^2)v\right)}{\Gamma(\gamma)} dv \\
&= \frac{c^{\gamma}}{(c+0.5\sigma^2 u^2)^{\gamma}} \\
&= \left(1 + \frac{\sigma^2 \upsilon}{m} \frac{u^2}{2}\right)^{-\frac{m^2}{\upsilon}},
\end{aligned}$$

where:

- $m = \mathrm{E}[V] = \frac{\gamma}{c}$; and
- $\upsilon = \mathrm{var}[V] = \frac{\gamma}{c^2}$.

3.4.3.6 The authors note that σ^2 is a scaling factor for V, so they picked $\mathrm{E}[V] = 1$, i.e. $m = 1$ or $\gamma = c$. In this case, the characteristic function of X becomes:

$$\phi_X(u) = \left(1 + \sigma^2 \upsilon \frac{u^2}{2}\right)^{-\frac{1}{\upsilon}}.$$

3.4.3.7 The VG model has finite moments of all order. In particular, the kurtosis of X is $3(1+\upsilon)$. The kurtosis of a normal distribution is 3. The parameter υ is therefore the proportionate excess of kurtosis. When $\upsilon > 0$, the VG model has a longer tail than the normal distribution.

3.4.3.8 The process followed by the share price is $S(t+1) = S(t)e^{\mu+X}$, where μ is the drift of the share price.

3.4.4 The Heston Model

3.4.4.1 Heston (1993) notes that while Black & Scholes (1973) are successful in explaining option prices, it does have known biases. He attributed this to the strong assumptions underlying the Black–Scholes model: constant and known mean and volatility. Since the Black–Scholes model depends on the risk-free rate of return as the mean drift, it cannot be generalised to allow the mean to vary. The volatility assumption, however, can be allowed to vary. He notes that various other authors, such as Scott (1987), Hull & White (1987) and Wiggins (1987) have generalised the model to allow for stochastic volatility, but these methods do not have a closed-form solution and rely on numerical techniques to solve two-dimensional PDEs.

3.4.4.2 Heston (*op. cit.*) proposes a model of stochastic volatility that is not based on the Black–Scholes formula, but provides a closed-form solution for the price of European options. Note that the term 'closed-form solution' still requires approximation techniques similar to that used to approximate the cumulative distribution function of the standard normal distribution used in the Black–Scholes formula.

3.4.4.3 Heston (*op. cit.*) allows for correlation between the spot asset return and volatility. He assumes that the spot asset has the following form:

$$dS(t) = \mu S(t)dt + \sqrt{v(t)}S(t)dZ_1,$$

where

- μ is the mean return;
- Z_1 is a Wiener process.

3.4.4.4 The volatility term follows an Ornstein–Uhlenbeck process:

$$d\sqrt{v(t)} = -\beta\sqrt{v(t)}dt + \delta dZ_2,$$

where Z_2 is another Wiener process whose correlation with Z_1 is ρ. This correlation is usually negative and reflects that volatility typically increases when the asset has a negative return.

3.4.4.5 By Ito's Lemma, $v(t)$ follows the following process:

$$dv(t) = \left[\delta^2 - 2\beta v(t)\right]dt + 2\delta\sqrt{v(t)}dZ_2,$$

which can then be rewritten as the more familiar square-root process as described in CIR (*op.cit.*):

$$dv(t) = \kappa\left[\theta - v(t)\right]dt + \sigma\sqrt{v(t)}dZ_2,$$

where:

- κ is the mean reversion speed of volatility of return;
- θ is the long term volatility of return; and
- σ is the volatility of volatility of return.

3.4.4.6 Recall from Section 3.3.4.4 that the $v(t)$ process is strictly positive when $2\kappa\theta > \sigma^2$.

3.4.4.7 A risk-neutral implementation of the Heston (*op. cit.*) model replaces the parameter μ with the risk-free interest rate r.

3.5 Integrated Models

There are many models that describe processes of multiple assets. These include Wilkie (1986), Thomson (1996), Yakoubov, Teeger & Duval (1999) and Whitten & Thomas (1999). These models are time-series based and they are neither arbitrage-free nor equilibrium models. For this reason, they are no longer used by practitioners and so they are not reviewed as a part of this research.

3.5.1 Thomson–Gott

3.5.1.1 Thomson & Gott (2009) proposes an integrated model that purports to be an equilibrium model that is also arbitrage free. A descriptive model is fitted to historical data and the parameters are used to inform inputs to a predictive model that projects the prices of inflation-linked ZCBs, conventional ZCBs, inflation rates and equity returns under a capital asset pricing model (CAPM) framework where the expected return of each asset class is consistent with its risk (defined as the variance). The risk-free asset for the CAPM model is chosen to be the one-year inflation-linked bond. It is assumed that market participants have homogeneous expectations with regard to the forthcoming projection year and make their decisions in mean-variance space.

3.5.1.2 Once the ZCB (nominal and inflation linked) prices are calculated, the corresponding zero-coupon yields are backed out of the prices. This is in contrast with the traditional method of first estimating the yields with term-structure models and then using the term structure to price bonds.

3.5.1.3 The purported arbitrage-free property of the model needs clarification. Within financial-economics literature, no-arbitrage models are so named because they value assets so that the market price is matched, thereby eliminating any opportunities for an immediate risk-free profit. This is done in such a way so that the yield and price

processes evolve in such a way so that they are consistent with the starting input, so that arbitrage opportunities also do not arise in the projections. The arbitrage-free property in this paper refers to the lack of risk-free profits in the projection of asset returns in the sense that higher expected returns are at the cost of higher variance. Despite assumptions of homogenous expectations, to the extent that the model's parameters are calibrated to historical data (as this is a descriptive model), these may be different to the ex-ante expectations of the marginal investor. This difference in definition has implications in the market-consistent valuation of liabilities: the model may not produce asset values that are consistent with the current market values and is therefore unsuitable for the market-consistent valuation of liabilities.

3.5.2 Gott

3.5.2.1 In the later stages of writing of this dissertation, the author became aware of what he believes to be the only paper available in academic literature that describes a single integrated risk-neutral model that models more than one type of asset. Many commercial model providers exist, such as Barrie and Hibbert, Deloitte & Touche's (TSM and TSM+ models), Mathematical Finance Company, DFA Capital Management and Ortec Finance — but access to these models is limited to paid subscribers only.

3.5.2.2 Gott (unpublished) constructed a model that generates an equity index, equity dividend yields, nominal and real term-structures of interest rates (from which bond returns can be obtained) and a consumer-price index. The equity index is constructed using a stochastic volatility model to generate negative skewness and greater-than-normal kurtosis. The dividend yield is modelled using an Ornstein–Uhlenbeck process to allow mean-reverting dividend yields. A three-factor interest-rate model is used to model the zero-coupon yields, where the factors correspond to the principle-component analysis of the correlation matrix of historic deviations from expected changes in zero-coupon yields. The consumer-price index is modelled using geometric Brownian motion where the drift is set equal to the difference between the nominal and real term structures. The correlation between these economic variables is allowed for in the model.

Appendix 3 – Glossary of technical terms

This section provides a high level description of some of the technical terms used in this literature review.

Brownian motion
See Wiener process

Consol
Consol bonds are a form of British government bonds. They are redeemable at par by Act of Parliament. The Consol bonds have a yield of 2.5% payable quarterly. Because of the low yield, there is very little incentive for the government to redeem these. As a result of the uncertain redemption date, Consol bonds are treated as perpetual bonds.

multivariate elliptical distribution
A random vector $\mathbf{X} = (X_1, X_2, ..., X_n)'$ has a multivariate elliptical distribution with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ if its characteristic function can be expressed as:

$$\mathrm{E}\left[e^{i\mathbf{t}'\mathbf{X}}\right] = e^{i\mathbf{t}'\boldsymbol{\mu}}\Psi\left(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right)$$

where

- Ψ is a scalar function,
- $\boldsymbol{\Sigma}$ if a positive semi-definite $n \times n$ matrix, and
- $\boldsymbol{\mu}$ and $\mathbf{t}$ are $n \times 1$ vectors.

This class of distributions include:

- Normal distribution,
- Laplace distribution,
- Student-t distribution,
- Cauchy distribution,
- Logistical distribution, and
- Symmetrical stable laws.

Gauss-Markov process
A Gauss-Markov process is a Markov process whose density function is Gaussian.

Ornstein–Uhlenbeck process
An Ornstein–Uhlenbeck process satisfies the stochastic differential equation:

$$dx(t) = \theta(\mu - x(t))dt + \sigma dZ$$

where

- Z is a Wiener process;
- $\theta > 0, \sigma > 0$; and
- θ, σ and μ are parameters.

stationary process
A stationary process is one whose probability distribution does not change when shifted in time. Consequently, its quantities, such as mean and variance, if they exist, do not change over time. Formally, that is:

Let $\{X_t\}$ be a stochastic process and let $F_X\left(x_{t_1+\tau}, x_{t_2+\tau}, ..., x_{t_n+\tau}\right)$ represent the distribution function of the joint distribution of $\{X_t\}$ at times $t_1+\tau, t_2+\tau, ..., t_n+\tau$. Then $\{X_t\}$ is stationary, if for all n, for all τ, for all $t_1, t_2, ..., t_n$, $F_X\left(x_{t_1+\tau}, x_{t_2+\tau}, ..., x_{t_n+\tau}\right) = F_X\left(x_{t_1}, x_{t_2}, ..., x_{t_n}\right)$.
Since τ does not affect F_X, F_X is not a function of time.

Wiener process

A Wiener process $Z(t), t \geq 0$ is a stochastic process, such that:

- $Z(0) = 0$
- The process is continuous.
- The process has stationary, independent increments.
- $Z(t) - Z(s) \sim N(0, t-s), t > s$

Chapter 4. Stochastic Valuation

4.1 Introduction

4.1.1 Stochastic modelling can be used for different applications, such as pricing of financial options, hedging, stress testing and the simulation of future returns distributions. These different applications require different types of stochastic models. Section 4.2 below describes the two different probability measures in which to model economic scenarios, namely real-world measure (or P-measure) and risk-neutral measure (or Q-measure). Section 4.3 below shows that models may be calibrated in two different ways, namely equilibrium and no-arbitrage. Section 4.4 summarises the findings of these two sections and justifies the use of risk-neutral modelling for this research.

4.1.2 Equilibrium and no-arbitrage models implicitly assume that the market is complete. Section 4.5 argues that this is not the case in real life and proposes a model that can be used to price illiquid assets and liabilities. Section 4.6 below describes the different methods of implementing stochastic models, namely closed-form solutions, lattices and Monte–Carlo simulations.

4.2 Real World vs. Risk-Neutral
(The Estimated Market Probability Measure vs. The Derivative Pricing Probability Measure)

4.2.1 The traditional actuarial valuation reflects the present value of future cashflows. The projections are based on deterministic assumptions that reflect the actuaries' estimates of the future economic outlook. For example, the projections of the future cashflows that are linked to an equity-type asset would require an equity-asset return. This is usually set to be equal to a risk-free rate of return (usually taken as the return of a government bond of some term) plus a risk premium to reflect the higher expected return for holding a more risky asset. This rate is also used for the discounting of cashflows.

4.2.2 In a stochastic actuarial world, the principles have mostly remained. The assumptions, rather than taking on a single value, have been replaced by distributions. An example of this is the model proposed by Wilkie (1986) and extended by Wilkie (1995). Thomson (1996) developed a similar model but from a different vantage point – rather than using an economic approach, it uses a statistical one. In all these models, the future returns of some asset classes are projected according to some distributions. The moments of these distributions reflect the actuary's realistic expectations of future probabilistic distributions. These may be guided by historical experience or actuarial judgement, or both. To the extent that different actuaries can have different views or judgement of future returns distributions, they may each have a different real-world model: each real-world model is a subjective interpretation of each actuary's perceived future reality. The so-called 'real-world' model is only 'real-world' to the extent that it reflects the user's realistic view of the world, but it is no more or less real than another user's real-world model.

4.2.3 Whilst this approach may succeed in producing an expected present value of the cashflows in line with the actuary's view of the world, which may be used for

budgeting purposes, it does not succeed in producing an accurate market-consistent value. This is because of the following reasons:

- The real-world model reflects the actuary's view of future return distributions and is not necessarily the same as that of the marginal investor, who determines the market price of the asset.
- The marginal investor is generally risk-averse. To entice a risk-averse investor to accept a riskier or more volatile investment, a higher expected return, or equivalently, a lower market price, is required. The pricing of the assets using the expected present value under a P-measure therefore requires the estimating of the risk premium. This risk premium differs for different types of assets (since each asset is subject to different types and levels of risk) and for different investors (since each investor has a different level of risk aversion). The assets' risk premium and investors' risk-aversion levels are difficult to quantify.
- The risk premium for each asset should also reflect the risk aversion of the marginal investor and not of the actuary.

4.2.4 For these reasons, it is inappropriate to calculate the market-consistent value of assets and liabilities with scenarios generated under a P-measure without making some assumptions about equilibrium or market price of risk. Thomson (2005) outlines one such method and is described briefly in Section 4.5 below.

4.2.5 The risk-neutral measure, or Q-measure, was originally proposed by Black & Scholes (1973) for valuation of equity options. This has subsequently been extended to other interest-rate-sensitive assets in Black (1976). This involves transforming the P-measure into a measure of a hypothetical risk-neutral investor. In the risk-neutral world, all investors are risk neutral, so no risk premium is needed to entice the investor to invest in risky assets. All assets are expected to earn the risk-free rate. This removes the need to estimate the risk premium of the marginal investor for asset valuation purposes (Fitton & McNatt, 2002). The two wrongs (assumptions that all investors are risk-neutral and all assets earn the risk-free rate) together produce the correct market price for these market instruments.

4.2.6 Fitton & McNatt (*op. cit.*) explain that a risk-averse, non-satiated investor expects to earn a return from his or her risky investments that is higher than the risk-free return. This is consistent with expected utility theory. The risk-free return is usually accepted as the return on a default-free government bond for a very short period of time, and is usually called the short rate. If the investment term on this bond is longer, then it carries a market risk – the risk that the short rate in the next instant is not equal to that initially expected. It is this risk that investors need to be compensated for in the form of a risk premium, even though the asset itself is free of default risks. This risk premium usually varies by term. From this definition it may appear that in order to value a bond, the risk premium for every term is needed in addition to the stochastic short rate. Fitton & McNatt (*op. cit.*) expressed it algebraically:

$$V(t,T) = e^{-s(t,T)(T-t)} = e^{-\phi(t,T)(T-t)} \mathrm{E}_r\left[\exp\left(-\int_t^T r(s)ds\right)\right], \tag{4.1}$$

where

- $V(t,T)$ is the price at time t of a ZCB which matures at time T;

– $s(t,T)$ is the zero (or spot) rate which applies from time t to time T, expressed as a continuously compounded rate;
– $\phi(t,T)$ is the risk premium, expressed as an annual excess rate of return, continuously compounded, which is required by the investor for the term from t to T;
– $r(s)$ is the stochastic short rate at time s; and
– E_r is the expectation at time t across all paths of future risk-free short-rates up to time T, across the realistic probability space.

4.2.7 When using risk-neutral valuation methods, it is sufficient to identify a set of spot rates that value bonds correctly relative to the market – the risk-premium assumptions are not required. Fitton & McNatt (*op. cit.*) show that this is achieved by risk adjusting the short rates $r(s)$. Define:

– $r^*(s) = r(s) + \phi(s,t) + \dfrac{\partial \phi(s,t)}{\partial s}(s-t)$ as the risk-adjusted short-rate; and
– E_{r^*} as the expectation at time t across all paths of future risk-adjusted short-rates up to time T, i.e. across the risk-neutral probability space,

4.2.8 Then

$$V(t,T) = e^{-s(t,T)(T-t)} = \mathrm{E}_r\left[\exp\left(-\int_t^T \phi(t,T) + r(s)ds\right)\right] = \mathrm{E}_{r^*}\left[\exp\left(-\int_t^T r^*(s)ds\right)\right] \qquad (4.2)$$

The last step can be proven by integrating the expansion of $r^*(s)$ by parts.

4.2.9 Note that this process involves changing the probability measure of the term structure of interest rates. This means redefining the term-structure model so that instead of being a random process for the short rate, it is a random process for the short rate plus a function of the term premium. Under this measure, the price of an asset is equal to the expected discounted value of the payoff. Because the price is now the expected discounted value where the variability of the returns is ignored, the probability measure is known as the risk-neutral measure. Other names include equivalent-martingale measure, or risk-adjusted measure.

4.2.10 Equations (4.1) and (4.2) highlight the difference between realistic-probabilities valuation and risk-neutral valuation. Where risk-neutral valuation defines a new probability measure, the realistic probabilities valuation makes no such measure change. Realistic probabilities valuation is more difficult to apply than risk-neutral valuation because a risk-premia function must also be defined.

4.2.11 Incidentally, scenarios that are generated under the P-measure are capable of valuing cashflows and assets on a market-consistent basis, provided an appropriate discount rate is used in conjunction with the projections that are generated under the measure. These discount rates are known as state-price deflators. Jarvis et al. (2001) shows that discounting with deflators under the P-measure is equivalent to discounting with the risk-free rate under the Q-measure. The state-price deflator is a combination of a utility function and the simulated return. Jarvis et al. (*op. cit.*) show that any utility function may be used for this purpose.

4.2.12 State-price deflators are constructed by taking the market price of existing market instruments and splitting this into the prices of a series of state securities. A state security is one that pays out a unit when a future state is achieved, and zero otherwise. The market price of a risky asset (when considering a single period time horizon) is therefore:

$$V^{\mathrm{r}} = \sum_{\forall s} C(s)\psi(s), \tag{4.3}$$

where:

- $C(s)$ is the cash flow from the asset in state s; and
- $\psi(s)$ is the price of an s-state security (i.e. the security pays out 1 in state s, 0 otherwise).

4.2.13 Define deflator for state s:

$$D(s) = \frac{\psi(s)}{p(s)} \tag{4.4}$$

where $p(s)$ is the probability of state s at the end of the single-period time horizon. Note here that the probability measure of $p(s)$ is not specified. It can be shown that this relationship holds for any probability measure P, where P is not necessarily the realistic measure. The deflator $D(s)$ is unique for the probability measure P.

4.2.14 Substituting (4.4) into (4.3):

$$V^{\mathrm{r}} = \sum_{\forall s} p(s)D(s)C(s) = \mathrm{E}_P[DC] \tag{4.5}$$

where E_P is the expectation taken across any probability space of P, provided it is consistent with the deflator.

4.2.15 For the special case of a risk-free asset, we have $C(s) = 1 \forall s$. In this case, (4.5) becomes:

$$V^{\mathrm{f}} = \sum_{\forall s} \psi(s) = \sum_{\forall s} p(s)D(s) = \mathrm{E}_P[D],$$

where V^{f} is the price of the risk-free asset.

4.2.16 In the case of a realistic probability measure, as in (4.1), the deflator is defined as:

$$\exp\left[-\phi(t,T)(T-t) - \int_t^T r(s)ds\right]$$

4.2.17 In the case of a risk-neutral probability measure, as in (4.2),the deflator is defined as:

$$\exp\left(-\int_t^T r^*(s)ds\right)$$

4.2.18 Because the state prices are calculated from the market prices, this approach is necessarily market-consistent. However, the constraint here is that they can be calculated only if there are at least as many market securities as there are future states.

4.2.19 Jarvis et al. (*op. cit.*) give various examples of the use of deflators. They also show that the utility function of a non-satiated and risk-averse investor necessarily satisfies the criterion of a deflator. They also highlight that in certain circumstances, such as a

multiple-currency environment, the use of deflators may be preferable to the use of risk-neutral valuation.

4.2.20 It is also noted by Fitton & McNatt (*op. cit.*) that although a risk-neutral valuation may produce the same results as a realistic-probabilities valuation (with deflators), it is not suitable for all purposes. In particular, for stress testing of asset–liability strategies under adverse movements in interest rates, the relevance of the information provided by the testing depends completely on the realism of the simulated environment. For such purposes, it is imperative that the test environment be as realistic as possible. This is clearly not satisfied by the risk-neutral methodology where all investors are assumed to be risk-neutral. For this reason risk-neutral valuation methodologies are used mainly for pricing, and realistic probabilities used for sensitivity testing.

4.2.21 Any other alternative measure may also be used provided an appropriate discount rate is chosen, including measures that describe worlds the investor is more risk-seeking than risk-neutral Q-measure, or more risk-averse than the real-world P-measure, or somewhere between these two. However, such discount rates are more difficult to calculate than the two approaches described above and are rarely used in practice. The specific change of variables that produce a risk-neutral model simply makes the algebra easier than the others, because one can ignore risk preferences (Fitton & McNatt, *op. cit.*).

4.3 Equilibrium vs. No-arbitrage

4.3.1 The difference between 'real-world' and 'risk-neutral', described in Section 4.2 above, reflects a difference in modelling methodology; the difference between 'equilibrium' and 'no-arbitrage' reflects a difference in assumptions and calibration targets. Cairns (2004) distinguishes between no-arbitrage and equilibrium models by providing the following comparisons:

> "No-arbitrage models... use the observed term structure at the current time as the starting point. Future prices evolve in a way that is consistent with this initial price structure and which is arbitrage-free. Such models are used for the pricing of short-term derivatives.
>
> "Equilibrium models are built on assumptions about how the economy works. They take into account of the varying risk preferences of different investors and aim to achieve a balance between supply of bonds and other securities and the demand for these by investors... This is done in such a way which captures the essential characteristics of the wider economy... Under such a model the theoretical prices evolve in a way which is free from arbitrage."

4.3.2 In other words, equilibrium models are those that make assumptions about the future evolution of the economy and these will then guide the evolution of future prices. No-arbitrage models are those that make an implicit assumption that the current prices reflect fully how the future prices will evolve. Both types of models should produce future prices that are arbitrage free.

4.3.3 Fitton & McNatt (*op. cit.*) also similarly contrast the properties of arbitrage-free and equilibrium models of the term structure of interest rates. They define the former as follows:

> "Despite being called "term structure models", they do not in reality attempt to emulate the dynamics of the term structure. Instead, they assume some computational convenient, but essentially arbitrary, random process underlying the yield curve, and then add time dependent constants to the drift (mean) and volatility (standard deviation) of the process until all market prices are matched. To achieve this exact fit, they require at least one parameter for every market price used as an input to the model."

4.3.4 Because no-arbitrage models attempt to match the current market prices exactly, they need to have the same number of variables as market instruments. They are therefore usually over-parameterised and are susceptible to instruments with outlying prices. As a result, their parameters are usually unstable and rarely do they have any fundamental meaning. Fitton & McNatt (*op. cit.*) observed that while the no-arbitrage models give an exact fit to prices of assets in some asset classes, they do not explain the differences between the behaviours of the models and the actual behaviour of the term structure over time. An arbitrage-free model is hence constructed only with reference to a single point in time.

4.3.5 Fitton & McNatt (*op. cit.*) define equilibrium models as follows:

> "In contrast to arbitrage-free models, equilibrium term structure models are truly models of the term structure process. Rather than interpolating among prices at one particular point in time, they attempt to capture the behaviours of the term structure over time. An equilibrium model employs a statistical approach, assuming that market prices are observed with some statistical error, so that the term structure must be estimated, rather than taken as a given. Equilibrium models do not exactly match market prices at the time of estimation, because they use a small set of state variables (fundamental components of the interest rate process) to describe the term structure. Extant equilibrium models do not contain time dependent parameters; instead they contain a small number of statistically estimated constant parameters, drawn from the historical time series of the yield curve.

4.3.6 In the case of equilibrium models, the theoretical prices at the outset may be different from market prices. To that extent, equilibrium models may produce initial arbitrage opportunities. Cairns (*op. cit.*) also notes that the difference between the theoretical and actual price when using an equilibrium model may be small (say 1%), but when used in the context of pricing derivatives, the gearing effect may lead to an exaggerated difference (say 10%).

4.3.7 Furthermore, arbitrage-free pricing requires the set of market prices to be complete and reliable, which is usually not the case. When there is no freely observable market price (but a price may be quoted by market participants upon request), arbitrage-free pricing requires interpolation between the available data points. In this case, two different arbitrage-free models will produce the same prices only for the instruments that form part of the input data. The form of the model underlying the term-structure process can make a large difference to the valuation of other instruments. Equilibrium models, in contrast, do not aim to match market prices at the time of estimation. Instead they aim to capture the behaviour of the term structure over time. The theoretically true term structure is usually not given and must be estimated. A small set of variables are used to describe this term structure. These variables usually also have intuitive interpretations.

4.3.8 Ahlgrim et al. (unpublished) give similar comparisons between no-arbitrage and equilibrium models.

4.3.9 Horizon pricing (pricing of instruments in some assumed future state of the market), as opposed to current pricing, can only be done using equilibrium (Fitton & McNatt,

op. cit.). This is because no-arbitrage pricing requires market prices as inputs and the future prices are clearly not available for horizon pricing. However, no-arbitrage models are capable of producing future prices in an arbitrage-free manner, albeit these are consistent with the initial prices and not some future state of economy.

4.3.10 One final observation is that equilibrium models are more subjective in their formulation and parameterisation than no-arbitrage models. This is because the modeller makes assumptions about the future evolution of the economy. In contrast, no-arbitrage models rely on observable market prices as inputs, thereby ensuring an objective parameterisation, at least for the assets that form the input data.

4.4 Summary of Model Types

4.4.1 Fitton & McNatt (2002) summarised the differences between applying equilibrium and no-arbitrage models parameterisation to real-world and risk-neutral models as follows:

- Risk-neutral models with no-arbitrage parameterisation are used for current pricing where input market prices are reliable.
- Risk-neutral models with equilibrium parameterisation are used for current pricing where inputs are unreliable or unavailable, or for horizon pricing (defined in Paragraph 4.3.9).
- Real-world models with no-arbitrage parameterisation are unusable because it is difficult to separate the term and risk premium from the model misspecification.
- Real-world models with equilibrium parameterisation are used for stress testing and reserve and asset adequacy test.

4.4.2 Thomson (2005) notes that real-world models with equilibrium parameterisation can also be used to price unmarketable assets in an incomplete market under some conditions. This is discussed in more details in Section 4.5 below.

4.4.3 The desire for objectivity and transparency in the new accounting standards has forced actuaries to adopt market-consistent valuation methodologies. The key here is 'consistent' rather than 'equivalent'. Insurers are required to value their assets and liabilities in a way that is more or less consistent with market prices, but not to necessarily replicate them. In dealing with insurance contracts, which usually have terms of over 30 years, a deep and liquid market does not exist with which to fully calibrate no-arbitrage methods. The focus is therefore to ensure that the model replicates market prices reasonably well but without over-parameterising the model. A no-arbitrage model should be used where the calibration instruments are available and liquid, and the model does not become over-parameterised. An equilibrium model is used where there is insufficient liquidity in the market or where a no-arbitrage model would cause over-parameterisation. In this case, the calibration targets are guided (but not determined) by the market where instruments are available, or by actuarial judgement where these are not.

4.4.4 A risk-neutral model is used for this research, although the author recognises that state-price deflators could similarly be used, as could the Thomson (*op. cit.*) method described in Paragraph 4.4.2 above and Section 4.5 below.

4.5 Valuation in an Incomplete Market

4.5.1 Thomson (*op. cit.*) recognises that the valuation of liabilities through replicating portfolios or risk-neutral and deflator methodologies makes the assumption of a complete market. He argues that since the market is incomplete it cannot be valued using these methodologies. He states:

> "The notion of a price at which an asset or liability would trade if a complete market existed is a fiction. It suggests that a complete market makes no difference to prices. That is clearly untrue: in an incomplete market, extra risks exist, which cannot be hedged. Those risks must affect prices."

4.5.2 He proposes the use of a CAPM-based model that assumes that all investors are rational, non-satiated and risk-averse. CAPM also assumes that all investors make their investment decisions using the mean and the variance of the return on their portfolios over a single-period time horizon.

4.5.3 Thomson (*op. cit.*) argues that the price of a liability in an incomplete market is equal to the sum of the price of a hedge portfolio and the negative price of the residual exposure to unhedgeable risks. The hedge portfolio is the combination of a risk-free asset and constituents of the market portfolio (which carries the same connotation as the CAPM) such that the unhedgeable risk is minimised and the expected return on the hedge portfolio during that time period is equal to that on the liabilities. The price of the residual exposure to unhedgeable risk is equal to that of a portfolio containing the market portfolio and the risk-free rate, whose risk is equal to the unhedgeable risk. 'Risk' is defined here as variance. Note that although the risk and price of the residual exposure are the same as those of the equal-variance portfolio, they are by definition not correlated. The reference to the market portfolio and cash is merely to determine the price of the unhedgeable risk (i.e. through the variance).

4.5.4 The argument against the validity of replicating portfolios or risk-neutral and deflator valuation methodologies on the basis of the market being incomplete first appears dubious. In particular, this is a debate of the purpose of the valuation:

- If the purpose of the valuation is to find a theoretical cost for a hedge for a portfolio of liabilities, should the market for such a hedge exist, then the risk-neutral or deflator methodology should suffice. This hypothetical value may be needed for financial reporting purposes.
- If the purpose of the valuation is to provide an appraisal value of the business for liability transfer, then the Thomson approach may seem more appropriate.

4.5.5 In any case, regulatory and financial-reporting standards of insurers, as well as professional guidance of actuaries, prescribe the method of carrying out market-consistent valuations and these should be followed regardless of their theoretical appropriateness. This makes the arguments less relevant for the purposes of financial and regulatory reporting. Some relevance remains as some insurers argue on the grounds of incompleteness that a valuation using hedging principles is too stringent. The relevance of these arguments will remain as long as insurers and rule-makers disagree on whether the standards are too stringent.

4.5.6 In South Africa, PGN 110.3[15] describes the following methodology as best practice:

[15] *PGN 110.3: Allowance for Embedded Derivatives*, Actuarial Society of South Africa, Version 2, April 2008.

"In order to quantify the reserves, stochastically simulated future economic variables are used to project the liabilities arising from embedded investment derivatives. The additional liabilities (i.e. shortfalls) at claim date are then discounted at an appropriate discount rate to determine the present value of the reserve required at the valuation date.

"... For each policy or group of policies with an applicable minimum contractual maturity value, the market value of the underlying assets (i.e. the asset share) as at the valuation date is used as the starting point. This value is accumulated with future premiums at the stochastically simulated investment returns allowing for charges and taxation to determine the projected maturity value for each policy or group of policies. The projected maturity values are calculated based on best estimates of all future contingencies (e.g. premium increases), other than decrements and the future investment returns.

"... For each policy or group of policies, the projected maturity value is compared with the contractual minimum guaranteed maturity value, where the contractual minimum guaranteed maturity value is also calculated without allowance for decrements... The shortfall or surplus at the maturity date must be discounted to quantify the value of the required reserve at the valuation date.

"... If investment returns are simulated under the risk-neutral probability measure (i.e. if the expected return on each asset class is the risk-free rate), the appropriate discount rate for each projection interval is the simulated risk-free rate of return for that interval. If the expected return is assumed to be different from the risk-free rate, the appropriate discount factor is the state-price deflator for the particular simulation. The risk-free discount factor is a special case of a state-price deflator where the risk premium on each asset class is zero.

"... Market-consistent simulation models can be either risk-neutral or deflator-based."

4.5.7 To address the issue of an incomplete market, PGN 110.3[15] says the following:

"It is recognised that due to the long-term nature of insurance contracts, tradable derivatives in the underlying assets may not be available for certain maturities. Therefore, to determine the assumptions underlying the stochastic model, a combination of historical volatility analysis and solving for volatilities implied by derivative prices may be required.

"To determine the relevant volatility assumptions underlying the stochastic model using historical analyses, a pragmatic approach is required. For example, the realised market price volatility of the underlying asset (over an acceptable historic period) could be calculated for a term consistent with each required volatility parameter. This calculated realised volatility can then be compared with the implied volatility parameter derived from derivative prices for terms where tradable derivatives are available. This relationship can then be extrapolated, together with the calculated realised volatility, to determine the volatility parameter for terms where tradable derivatives are not available."

4.5.8 The method proposed by PGN 110.3[15] should provide an objective yet parsimonious way of extrapolating market prices of options. Extrapolation of the yield curve can be done in many ways, each with its advantages and disadvantages. An outline of these methods is described in Thomas & Maré (unpublished) and Thomas (unpublished).

4.5.9 Towards the end of the writing of this dissertation, the author became aware of a proposal in the IFRS framework that requires an allowance for additional risk margins in the calculation of market-consistent value liabilities to account for the uncertainties of assumptions that are not observable in the market. As at May 2009, this proposal remains open.

4.6 Stochastic Valuation Methods

4.6.1 In general, stochastic models are required in the valuation of investment-related guarantees. This is because deterministic models do not capture the costs of the guarantees. These stochastic models need not be simulations-based – in some cases it is possible to quantify the value of the investment guarantee using closed-form methods, such as Black & Scholes (1973).

4.6.2 There are three main stochastic methods for pricing options. These are:
- closed-form formulae;
- trees (or lattices); and
- Monte-Carlo simulations.

4.6.3 Under some circumstances, a combination of the three methods here can be used together.

4.6.4 These methods all begin with one or more stochastic differential or difference equations that describe the random process. Depending on the complexity of the instrument that needs to be valued and also the complexity of the assumptions, these stochastic equations may be simplified to closed-form equations. Each of these methods has its own advantages and disadvantages and is used under different circumstances.

4.6.5 The closed-form formula is the simplest and fastest way to value simple options. Black & Scholes (*op. cit.*) developed the breakthrough closed-form solution for valuing European options. Black (1976) subsequently extended this model to value vanilla swaptions. A common characteristic of closed-form equations is that they can only be used to price simple options, such as those with single fixed option dates and constant notional units. Black & Scholes (*op. cit.*) can be extended to price American options where the underlying stocks pay constant and continuous dividends. These closed-form formulae also make some extremely stringent assumptions, such as a constant risk-free rate of return and constant volatility. This method, although simple, generally does not allow more complex derivatives to be valued.

4.6.6 Trees allow a term structure to be incorporated into variables such as interest rates and volatility. The extra degrees of freedom provide more flexibility in capturing the dynamics of these variables realistically, but nevertheless they cannot be used to value instruments that have more than one underlying asset class and are computationally more intensive than closed-form solutions. Recombining trees, also known as lattices (see Chapter 5), are used to help improve the efficiency of this method.

4.6.7 Monte–Carlo simulation is the most flexible method for valuing derivatives and can be used to value any type of derivative contracts. The disadvantage of this method is the practical limitations imposed by long runtimes.

4.6.8 Because of the contingent nature of GAOs, it is generally not possible to quantify the cost of GAOs using closed-form formulae. In particular, closed-form methods do not adequately allow for the term structure of interest rates and the volatility surface (where the volatility depends on the term to maturity and the moneyness of the option). In most cases, the guarantee underlying the GAO is also dependent upon the

contract's survival to the option date (i.e. there may be decrements through death and surrender). The dependence of the option value on the two asset types (the term structure of interest rates and equity performance until retirement or vesting date) means that they cannot be valued using trees without introducing multidimensional recombinant lattices. Doing so makes the model more complex. The computational demand will also increase exponentially, perhaps making it even more demanding than Monte-Carlo methods. GAOs are therefore valued with Monte–Carlo simulations.

Chapter 5. Economic Scenario Generator

5.1 Introduction

5.1.1 In Chapter 4, the concept of Monte-Carlo simulations is introduced. This stochastic valuation method makes it possible to value products with complicated distributions for which no closed-form solutions exist, such as that of the market value of a portfolio of GAOs.

5.1.2 An ESG is an integrated asset model that stochastically simulates future economic variables such as equity returns and term structures of interest rates, whilst allowing for interdependence between the variables in a way that is consistent with economic theory (such as allowing for higher equity returns when interest rates are low). An ESG facilitates Monte–Carlo valuation by providing the stochastic returns projection for cashflows projections and the corresponding discount rates. This chapter describes the theory behind the construction of an ESG and the tools used to calibrate it. This ESG is built in Matlab and contains a model for the term structure of interest rates and another for equity returns. The algorithms of the various components of the ESG are provided. The code of the program is available from the author upon request.

5.1.3 Because the ESG is essential for the aims of this research, extensive research was done on the different model forms and their calibrations. Significant effort was also made to ensure that the different variables produced are market consistent and arbitrage free.

5.2 The ESG at a High Level

5.2.1 With but a few exceptions, most market-consistent ESGs make the assumption that the economy is driven by interest rates. This is because most market-consistent models are risk neutral. In order to perform a risk-neutral valuation of financial options, the risk-free rate of interest is required. This means that the interest rate needs to be projected before the projection of other asset types can be performed. The ESG developed for the purposes of this research makes the same assumption. The model of the term structure of interest rates forms the foundation of the ESG. The model of the equity returns is driven off this interest-rate model.

5.2.2 Figure 5.1 illustrates the different components of the ESG and their hierarchy. Each component is explained in more details below. There is no explicit model for the bond-portfolio returns — these are calculated as the change in bond prices, which are in turn functions of the simulated interest rates. Because the bond portfolio return is not an explicit model, it is not shaded.

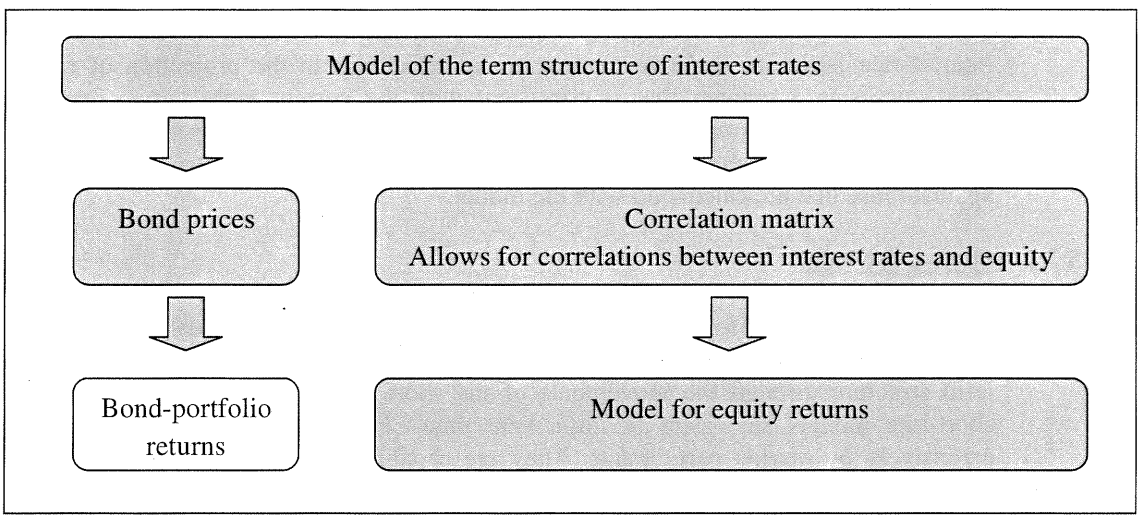


Figure 5.1 High-level structure of the economic scenario generator

5.3 Term Structure of Interest Rates

5.3.1 Introduction

5.3.1.1 Sections 5.3.2 and 5.3.3 gave a brief introduction to two kinds of interest-rate models, *viz.* short-rate models and the LIBOR market models (LMM) (Brace, Gatarek & Musiela, 1997). Both types are capable of modelling the term structure of interest rates. The former assumes that the present term structure and its volatilities may be expressed as a function of future short rates. The latter models a set of forward rates (also called LIBORs).

5.3.1.2 The remainder of Section 5.3 describes the Monte–Carlo implementation of a short-rate model in a market-consistent manner. The model parameters are first calibrated using an interest-rate lattice. These fitted parameters can then be applied to the Monte–Carlo implementation.

5.3.1.3 Section 5.3.4 describes the characteristics of the model that is implemented in the ESG, the Black–Karasinski model. A high level lattice implementation of the Black–Karasinski model to fit a given term structure of interest rates, is also given in this section. This is accomplished by making assumptions that some parameters take on certain values, and fitting the remainder. Because most of the methodology is taken directly from Hull (2006), the details of the implementation are deferred to Appendix 5.1.

5.3.1.4 Section 5.3.5 gives a high level overview of swaps and swaptions, the assets used to calibrate the interest-rate model. Section 5.3.6 describes the methodology of pricing swaptions using an interest-rate lattice, when the parameters of the interest-rate model are assumed to take on certain values. Section 5.3.7 describes the calibration of the remaining parameters of the Black–Karasinski lattice whose values were only assumed as given in Sections 5.3.4 and 5.3.6. Section 5.3.8 describes a process of converting the lattice parameters into another form so that they can be implemented in

a Monte–Carlo method. Section 5.3.9 shows a method of merging Monte–Carlo and interest-rate lattice to improve computational efficiency in the projection of future yield curves in a manner that is consistent with the current yield curve. Sections 5.3.10 to 5.3.12 describe methods of testing the accuracy and appropriateness of the model and its parameterisation by comparing the actual prices of swaps and swaptions against those that are calculated with the model.

5.3.2 Short-rate Models

5.3.2.1 Although named 'short-rate models', these models are capable of modelling the entire term structure — they are so named because they model the movements of the future term structure through the movements of the short rates. Hull (*op. cit.*) notes that short-rate models are easier to implement than LMM type models and are used extensively by market participants. They are capable of producing prices for most nonstandard interest-rate derivatives that are consistent with prices of market-traded instruments. These models have the following general form:

$$df(r(t)) = \kappa(t)\big(\mu(t) - f(r(t))\big) + \sigma(t)dZ, \tag{5.1}$$

where:

- $r(t)$ is the short rate at time t;
- $f(r(t))$ is a one-to-one function of $r(t)$;
- Z is a standard Brownian motion; and
- κ, μ and σ are deterministic parameters and are functions of time t.

5.3.2.2 According to Buetow, Fabozzi & Sochacki (2002), the most popular no-arbitrage, short-rate models are:

- Black & Karasinski (1991);
- Black, Derman & Toy (1990);
- Ho & Lee (1986);
- Heath, Jarrow & Morton (1992); and
- Hull & White (1990, 1993a, 1993b).

They also note that the most popular equilibrium models are:

- Vasicek (1977);
- Cox, Ingersoll & Ross (1985a, 1985b);
- Longstaff (1992);
- Longstaff & Schwartz (1992); and
- Brennan & Schwartz (1979, 1982).

5.3.2.3 Although the above models were originally developed either as equilibrium or no-arbitrage models respectively, they are not mutually exclusive. Buetow, Fabozzi & Sochacki (*op. cit.*) note that no-arbitrage and equilibrium models both start with the same SDEs but apply these under different frameworks to price securities. In particular, equilibrium models describe the basic workings of the macro economy as a function of a given state variable (Audley, Chin & Ramamurthy, 2002), whereas no-arbitrage models use market prices to generate an interest-rate lattice, so that there is no arbitrage relationship between the market and the model (Buetow, Fabozzi & Sochacki, *op. cit.*).

5.3.2.4 For example, Fitton and McNatt (*op. cit.*) presents the Black-Karasinski model in the following risk-neutral, no-arbitrage form:

$$d\log r(t) = \kappa(t)\big(\mu(t) - \log r(t)\big)dt + \sigma(t)dZ, \tag{5.2}$$

where:

- Z is a standard Brownian-motion term;
- $r(t)$ is the short rate at time t;
- $\kappa(t)$ is the rate of reversion to the long-term mean at time t;
- $\mu(t)$ is the mean long rate at time t; and
- $\sigma(t)$ is the volatility level at time t.

5.3.2.5 Although κ(t), $\mu(t)$ and $\sigma(t)$ are expressed as deterministic functions of future times, they are used to describe the current market conditions. As mentioned in Section 5.3.2 above, short-rate models make the implicit assumption that the current term structure of interest rates and its volatilities are characterised by the short rates and their volatilities at the future times. We shall explore the relationship between these future short rates and the current term structure in Sections A 5.1.3 and 5.3.7 below.

5.3.2.6 Fitton and McNatt (*op. cit.*) also presents the Black-Karasinski model in the following real-world, equilibrium form:

$$d\log r(t) = \kappa(\mu - \gamma\log r(t) - \log r(t))dt + \sigma dZ, \tag{5.3}$$

where:

- κ, μ, σ and r have similar definitions to the above, but have scalar values;
- $\log r(0)$ is fitted statistically to bond prices; and
- κ, μ, σ and r are estimated from historical data.

5.3.2.7 The first difference between (5.2) and (5.3) is the presence of the term $\gamma\log r$ in (5.3), representing the drift term to allow for the risk-premium when moving to a real-world measure. This is consistent with the term $\phi(s,t)$ in equations (4.1) and (4.2). The other difference is that κ, μ and σ are assumed to be constant in the equilibrium model. This is consistent with the analysis presented in Section 4.3.

5.3.3 LMM- or BGM-type Models

5.3.3.1 The London interbank offer rate (LIBOR) market model (LMM), or Brace–Gatarek–Musiela (BGM) model, Hull (*op. cit.*) notes, are capable of being fitted to more complex volatility structures than short-rate models because they generally have more degrees of freedom. They also have parameters that are directly observable in the market, avoiding the need to calibrate parameters. However, they require much more computation time for most applications as Monte-Carlo simulation is the only way to implement these models.

5.3.3.2 To illustrate this model, Götsch (2006) used the following notation to describe the bond price process:

$$B(t,T_i) = \big(1 + \alpha_{T_iT_{i+1}} L_{T_iT_{i+1}}(t)\big)B(t,T_{i+1}),$$

where

- $B(t,T_i)$ is the time t price of a bond maturing at time T_i;

- $\alpha_{T_i T_{i+1}}$ is a deterministic parameter; and
- $L_{T_i T_{i+1}}(t)$ is the stochastic process that describes the forward Libor rate that one can enter into at time $t < T_i$, in order to invest in a money-market account for the time period $[T_i, T_{i+1})$.

If $t = T_i$, then $L_{T_i T_{i+1}}(t)$ is the spot Libor rate at time T_i.

5.3.3.3 Rearranging the terms, this becomes:

$$\alpha_{T_i T_{i+1}} L_{T_i T_{i+1}}(t) = \frac{B(t,T_i) - B(t,T_{i+1})}{B(t,T_i)}.$$

5.3.3.4 It can be shown that if $B(t,T_i)$ is taken as the numeraire associated with some unique probability measure Q_{i+1}, then $B(t,T_i) - B(t,T_{i+1})$ and $\frac{B(t,T_i) - B(t,T_{i+1})}{B(t,T_i)}$ are both martingales under Q_{i+1}. This means that $\alpha_{T_i T_{i+1}} L_{T_i T_{i+1}}(t)$ is also a martingale under Q_{i+1}, but because $\alpha_{T_i T_{i+1}}$ is a constant, $L_{T_i T_{i+1}}(t)$ is also a martingale under Q_{i+1}.

5.3.3.5 $L_{T_i T_{i+1}}(t)$ is then modelled as a geometric Brownian motion, given by:

$$dL_{T_i T_{i+1}}(t) = \sigma_{T_i, T_{i+1}}(t) L_{T_i T_{i+1}}(t) dW_{i+1},$$

where

- W_{i+1} represents the Wiener process under the unique probability measure Q_{i+1}; and
- $\sigma_{T_i, T_{i+1}}(t)$ is the instantaneous volatility of $L_{T_i T_{i+1}}(t)$ at time t and is a deterministic function of time.

5.3.3.6 The process followed by $L_{T_i T_{i+1}}(t)$ is lognormal, which is precisely the dynamics underlying the Black's (1976) formula. This means that $\sigma_{T_i, T_{i+1}}(t)$ can be calculated directly from the observed caplet-implied volatility (similar idea to implied volatility on swaptions, in Section 5.3.5), and no other calibration of parameters is necessary. LMM models therefore have parameters that are directly observable in the market.

5.3.3.7 This model relies on a unique probability measure Q_{i+1} for each maturity T_i — each forward rate is modelled by a lognormal process under its own forward measure. The LMM may thus be interpreted as a collection of forward LIBOR dynamics for different forward rates, each forward rate $L_{T_i T_{i+1}}(t)$ being consistent with a Black interest-rate-caplet formula for its maturity. However, $L_{T_i T_{i+1}}(t)$ is only a martingale under its own measure Q_{i+1}. Each process is adequate for calculating the payoffs of a caplet, but not for calculating the payoffs of caplets of other maturities or more complex assets. For these assets, it is necessary to bring the dynamics of the different maturities together under a common measure, such as the forward measure for a preferred single maturity. In this case forward rates will not be lognormal under the general measure, leading to the need of numerical methods (Monte–Carlo simulations) to evaluate the process.

5.3.4 One-factor Black–Karasinski Model

5.3.4.1 The author chose to implement a short-rate model for this research. This is done because short-rate models are simpler to implement. Specifically, the Black–Karasinski (1991) model is used. This model is chosen because it has the desirable property that the logarithm of the short rate is modelled, thus ensuring that the short rate is positive.

5.3.4.2 The Black–Karasinski model is discussed briefly in Section 5.3.2. Equation (5.2) shows the model in its mean-reverting form. To simplify the representation for the remainder of this dissertation, this equation is rewritten as:

$$d\log r(t) = \left(\theta(t) - \kappa(t)\log r(t)\right)dt + \sigma(t)dZ\,. \tag{5.4}$$

This corresponds to equation (5.2) with $\theta(t) = \kappa(t)\mu(t)$.

5.3.4.3 Since we are interested in producing the market-consistent price of a portfolio of GAOs, equation (5.4) must be calibrated to the instruments that are available in the market. This calibration is done with the help of interest-rate trees. Ideally, we should find a combination of $\theta(t)$, $\kappa(t)$ and $\sigma(t)$ such that both the term structure of interest rates and the volatility surface are fitted exactly. Hull & White (1993a) and Hull (2006) note that it is very difficult to achieve this in practice as the parameters often exhibit non-sensible and non-stable results when expressed as functions of time. These time-dependent parameters are useless for extrapolation. They recommend that $\kappa(t)$ and $\sigma(t)$ be reduced to constants κ and σ respectively and the objectives be changed so that the initial term structure of interest rates is matched exactly (i.e. bonds are priced exactly), but the term structure of volatility should be fitted so that the pricing error of interest-rate swaptions is minimised. This is consistent with the principles established in Paragraph 4.4.3 that the emphasis of market-consistent valuation is to be consistent with the market rather than equivalent to the market.

5.3.4.4 Practitioners often believe that a one-factor model, as the one proposed here, does not offer sufficient flexibility for capturing a full range of future term structures movements (such as steepening and changes in curvature), and that a two- or three-factor model offers greater potential in this regard. Whilst this concern is valid for a real-world model, it is not a concern for a risk-neutral and no-arbitrage model. With such a model, the emphasis is on pricing securities at the valuation date. Any model that can replicate the market price readily and implicitly reflects the market's expectations of the future. The modeller's view of the future, which may be projected through his/her choice of model (and the number of factors) in a real-world model, is irrelevant and redundant in calculating the market price of the asset as long as the model is calibrated to fit the market prices. This view is shared by Mossman[16], where he states:

> "The choice of model does not matter as long as the model is calibrated to fit the target. Any recognised model could be used to give sensible answers. Answers are independent of the choice of model."

[16] Mossman, J (2006). *Introduction to Economic Scenario Generators*. Presentation by Towers Perrin Tillinghast. Available online at: http://www.docstoc.com/docs/29487230/Introduction-to-Economic-Scenario-Generators.

5.3.4.5 We see below that with appropriately chosen $\theta(t)$, this one-factor model is capable of replicating the current term structure of interest rates. The advantage of having two factors rather than one in a risk-neutral model is that the model would then be capable of replicating a bigger range of volatility surfaces. This comes at the cost of significantly increased modelling complexity. For the purpose of this research, a one-factor model is considered.

5.3.4.6 The first step of the calibration is to make assumptions for the values of κ and σ. $\theta(t)$ is fitted to ensure that the term structure of interest rates fits exactly to that available in the market. This is done by fitting interest-rate trees using the method proposed by Hull & White (1993a) and Hull (2006). These texts describe a two-stage process to fitting the tree. Because this material is reproduced from these texts, this method is described below in Appendix 5.1. The equations in this appendix are referenced extensively in the remainder of this chapter and carry reference numbers in the form of (A5.1.x). The trees and lattices below model the natural logarithms of the interest rates, but are nevertheless referred to as either interest-rate trees or interest-rate lattices.

5.3.4.7 The second step of the calibration is to price a portfolio of swaptions using the interest-rate tree created in the first step. Swaps and swaptions are discussed in Section 5.3.4.8. The third step of the calibration is to find the values for κ and σ that minimise the sum of square of errors in the swaption prices. This is discussed in Section 5.3.7.

5.3.4.8 Once the initial parameters have been fitted, we can proceed with the Monte–Carlo simulation of the short rate. This will give us, for each simulation and at each timestep, the risk-neutral short rate. The term-structure at the future projection horizons, the short end of which is consistent with this simulated short rate, may be obtained by extending the Hull–White method. This procedure is described in Section 5.3.9. The result of the calibration is presented in Section 5.3.10.

5.3.5 Swaps and Swaptions

5.3.5.1 A swap is an instrument that facilitates an exchange of cashflows. There are two parties under a swap arrangement, the receiver and the payer. The payer agrees to pay a series of predetermined and fixed cashflows to the receiver for a predetermined term (usually referred to as the 'tenor'), who in return pays the payer a series of floating cashflows for the same tenor. A receiver swap is one where the holder receives the fixed cashflows, and the payer swap is one where the holder pays the fixed cashflows. The fixed cashflows are determined upfront as Ns, where:

- s is the swap rate, the flat rate that determines the interest that the payer pays to the receiver; and
- N is some notional principle, to which the interest rate applies, but is itself not exchanged.

5.3.5.2 The floating cashflows are likewise defined upfront as $N\left(e^{r(i)}-1\right)$, where $r(i)$ is the continuously compounded reference rate that will be applicable for the period $[i-1, i)$. The reference rate determines the value of the floating payments in the swap

contract — it is usually related to the LIBOR and is outside the control of the parties to the contract. The market convention for swap quotation is the swap rate.

5.3.5.3 In most instances, the swap rates are quoted such that the swap contract has a net value of zero at the inception of the contract. This contract is then said to be issued at-the-money. The swap rate can be calculated by equating the present value of the cashflows at time 0. This is:

$$\sum_{t=1}^{T}\left(C_{t,R}\prod_{i=1}^{t}e^{-r(i)}\right)=\sum_{t=1}^{T}\left(C_{t,P}\prod_{i=1}^{t}e^{-r(i)}\right). \tag{5.5}$$

where:

- $C_{t,R}$ is the payment made by the receiver at time t;
- $C_{t,P}$ is the payment made by the payer at time t; and
- r_i is the forward risk-free rate that applies for the period $[i-1,\ i)$.

5.3.5.4 Substitute expressions for fixed and floating cashflows into (5.5). This gives:

$$\sum_{t=1}^{T}\left(N(e^{r(t)}-1)\prod_{i=1}^{t}e^{-r(i)}\right)=\sum_{t=1}^{T}\left(Ns\prod_{i=1}^{t}e^{-r(i)}\right).$$

Rearranging the terms gives:

$$s=\frac{\sum_{t=1}^{T}\left((e^{r(t)}-1)\prod_{i=1}^{t}e^{-r(i)}\right)}{\sum_{t=1}^{T}\left(\prod_{i=1}^{t}e^{-r(i)}\right)}. \tag{5.6}$$

5.3.5.5 The numerator can be expanded to give the following telescoping sequence:

$$\begin{aligned}\sum_{t=1}^{T}\left((e^{r(t)}-1)\prod_{i=1}^{t}e^{-r(i)}\right)&=(e^{r(1)}-1)e^{-r(1)}+(e^{r(2)}-1)e^{-r(1)-r(2)}+\\&\qquad(e^{r(3)}-1)e^{-r(1)-r(2)-r(3)}+\ldots+(e^{r(T)}-1)e^{-r(1)-r(2)-r(3)\ldots-r(T)}\\&=(1-e^{-r(t)})+(e^{-r(1)}-e^{-r(1)-r(2)})+(e^{-r(1)-r(2)}-e^{-r(1)-r(2)-r(3)})+\\&\qquad\ldots+(e^{-r(1)-r(2)-r(3)\ldots-r(T-1)}-e^{-r(1)-r(2)-r(3)\ldots-r(T)})\\&=1-e^{-r(1)-r(2)-r(3)\ldots-r(T)}=1-P(T),\end{aligned} \tag{5.7}$$

where $P(T)$ is the current price of a ZCB maturing at time T.

5.3.5.6 The denominator of (5.6) is an annuity factor. Equation (5.7) can be substituted into (5.6) to give the following:

$$s=\frac{1-P(T)}{\underline{a}_{\overline{T}|}}, \tag{5.8}$$

where $\underline{a}_{\overline{T}|}$ is the present value of a certain annuity paying 1 per annum in arrears for T years and the cashflows are discounted at the forward rates r_i.

5.3.5.7 A swaption is an option on a swap. The terms of the swap underlying the swaption are defined upfront and may be of a number of different types. A vanilla swaption refers to one where the notional is fixed. Equity-linked swaptions refer to swaptions where

the notional is linked to an equity index. For the remainder of this chapter, 'swaption' refers to the vanilla interest-rate type. A receiver swaption is one where the holder has the right to enter into a receiver swap. In a payer swaption the holder has the right to enter into a payer swap. An option term and a tenor need to be defined in a swaption contract.

5.3.5.8 The holder of a receiver swaption should exercise the option only if the prevailing market swap rate is lower than that defined in the swaption contract (the strike rate). The opposite applies for a payer swaption.

5.3.5.9 The at-the-money strike rate is calculated similarly to that of the swap rate of a swap. Using similar arguments as those used for swaps in ¶5.3.5 to ¶5.3.5.6 and same derivation as equations (5.5) to (5.8), we get:

$$s(t,T-t)=\frac{P(t)-P(T)}{{}_{t|}\underline{a}_{\overline{T-t}|}}, \tag{5.9}$$

where $s_{t,d}$ is the at-the-money strike rate, at time 0, for a swap with a tenor of d starting at time t and ${}_{t|}\underline{a}_{\overline{T-t}|}=\sum_{j=t+1}^{T}\exp\left[-\sum_{i=1}^{j}r(i)\right]$ is the present value of a deferred term-certain annuity payable in arrears from time t (first payment is at time $t+1$) to time T and the cashflows are discounted at the forward rates $r(i)$.

5.3.5.10 Having established the swap rate $s(t,d)$ at time 0, the payoff of a receiver swaption at time t is equal to:

$$\max(s(t,T-t)-s^{*}(t,T-t),0)\underline{a}^{*}_{\overline{T-t}|}(t),$$

where:

- $s^{*}_{t,d}$ is the prevailing d-year swap rate at time t; and
- $\underline{a}^{*}_{\overline{d}|}(t)$ is the prevailing d-year annuity factor at time t.

5.3.5.11 The payoff of a payer swaption at time t is equal to:

$$\max(s^{*}(t,T-t)-s(t,T-t),0)\underline{a}^{*}_{\overline{T-t}|}(t).$$

5.3.5.12 Black (1976) derived a closed-form formula for the pricing of swaptions using assumptions and arguments similar to those made in Black & Scholes (*op. cit.*) for the pricing of equity options. The model had the following forms:

$$p=e^{-rT}\left[S\Phi(-d_2)-F\Phi(-d_1)\right], \text{ and} \tag{5.10}$$

$$c=e^{-rT}\left[F\Phi(d_1)-S\Phi(d_2)\right], \tag{5.11}$$

where:

- $d_1=\dfrac{\ln(F/S)+(\sigma^2/2)t}{\sigma\sqrt{t}}$;
- $d_2=\dfrac{\ln(F/S)-(\sigma^2/2)t}{\sigma\sqrt{t}}=d_1-\sigma\sqrt{t}$;
- p is the price of a receiver swaption;
- c is the price of a payer swaption;

- r is the risk-free discount rate, continuously compounded;
- $S = {}_{t|}\underline{a}_{\overline{T-t}|}s$ is the strike price;
- s is the agreed strike rate;
- $F = {}_{t|}\underline{a}_{\overline{T-t}|}s_{t,T-t}$ is the forward price; and
- σ is the implied volatility on swaptions.

5.3.5.13 Like the Black–Scholes formula, some of the assumptions made in the derivation of the Black formula proved to be too stringent and unrealistic. As a result, (5.10) and (5.11) are generally not used in the market for pricing. The only unobservable variable in the equation is the volatility parameter σ and there is a one-to-one relationship between the price and σ. This has led to the market convention of using σ to quote for the price of swaptions. This is known as the implied volatility on swaptions. The parameter σ is therefore no more than a proxy for the price of swaptions and reflects the market's ex-ante expectations of future volatilities.

5.3.6 Pricing Swaptions Using the Interest-rate Tree

5.3.6.1 We apply the Hull–White model to price swaptions. Consider a swaption with option term t and tenor d so that the last payment, if swaption is exercised, is at time $T = t + d$. To price this swaption, we must build an interest-rate lattice that is of length T. Define (i, j) as the node after i timesteps, $i \in \mathbb{N}$ and j steps above (or below, if negative) the central path, $j \in \mathbb{Z}$.

5.3.6.2 We first calculate the price, at time t, of a ZCB that starts at time t and matures at time T for each node (t, j) in the interest-rate lattice. We do this by discounting the payoff at time T back to t using the path-specific rates from the interest-rate lattice. The payoff and price of the zero-coupon risk-free bond at time T is one unit notional for all nodes. Define $P^*(i, j, d)$ as the prevailing price at node (i, j) of a ZCB of duration d years. By definition, $P^*(T, j, 0) = 1, \forall j$. To discount the payoff, we use the following equation iteratively:

$$P^*(i, j, T-i) = e^{-r(i,j)} \sum_k P^*(i+1, k, T-i-1) q(i, k, j), \tag{5.12}$$

where the summation is across all nodes from which (i, j) may be accessed, $q(i, k, j)$ is the probability of moving from (i, k) to $(i+1, j)$.

5.3.6.3 We also calculate the price of an annuity that begins at time t and ends at time T. Without loss of generality, we assume that the annuity is payable annually in arrears, so that the first payment occurs just before time $t + 1$. The following equation holds:

$$\underline{a}^*_{\overline{T-i}|}(i, j) = \begin{cases} e^{-r(i,j)} \left(1 + \sum_k \underline{a}^*_{\overline{T-i-1}|}(i+1, k) q(k, j) \right) & \text{for } i < T, \forall j \\ 0 & \text{for } i = T, \forall j \end{cases}. \tag{5.13}$$

where $\underline{a}^*_{\overline{d}|}(i, j)$ is the prevailing value of the annuity at node (i, j), where the annuity pays 1 per annum in arrears from time i (i.e. first payment at time $i + 1$) for a period of d years.

5.3.6.4 $P^*(i, j, T-i)$ and $\underline{a}^*_{\overline{T-i}|}(i, j)$ are calculated using equations (5.12) and (5.13) by backward induction from $i = T$ until $i = t$. We calculate the prevailing swap rate for each node (t, j) as:

$$s^*(t, j, T-t) = \frac{1 - P^*(t, j, T-t)}{a^*_{\overline{T-t}|}(t, j)}.$$

5.3.6.5 The payoff of a receiver swaption at this node is as follows:

$$V_R(t, j) = \max(s(t, T-t) - s^*(t, j, T-t), 0)\underline{a}^*_{\overline{T-t}|}(t, j). \quad (5.14)$$

5.3.6.6 The payoff of a payer swaption at this node is

$$V_P(t, j) = \max(s^*(t, j, T-t) - s(t, T-t), 0)\underline{a}^*_{\overline{T-t}|}(t, j). \quad (5.15)$$

5.3.6.7 Having determined the payoff at time t, we continue to discount this back from time t to time 0 iteratively using the following equations:

$$V_R(i, j) = e^{-r(i,j)} \sum_k V_R(i+1, k) q(i, k, j); \text{ and}$$

$$V_P(i, j) = e^{-r(i,j)} \sum_k V_P(i+1, k) q(i, k, j).$$

where the summations are across all nodes from which (i, j) may be accessed. The theoretical prices of the receiver and payer swaptions are the time-zero discounted values $V_R(0,0)$ and $V_P(0,0)$ respectively.

5.3.7 Calibration of κ and σ

5.3.7.1 The steps in Section 5.3.6 are repeated for every market-traded swaption. The tree will give a theoretical price for each swaption. Because there are only two degrees of freedom (κ and σ), it is not possible to fit the tree to all market prices. The two-stage tree-building (steps described in Appendix 5.1) and Section 5.3.6 are repeated for different combinations of κ and σ until some weighted minimum sum of squares of errors (SSE) is minimised. Because of the subjectivity involved in the choosing of the weights, getting the absolute optimal values of κ and σ is not as important. In any case, in the author's experience, the valuation results are not overly sensitive to the values of these two parameters. For these two reasons, any number of optimisation techniques can be used.

5.3.7.2 In this research, the author uses a simple method where a large grid is defined. The two dimensions of the grid are κ and σ respectively. Five horizontal and five vertical lines are drawn in this grid, each equidistant to the next, so that there are 25 intersecting points in total. The weighted SSE is calculated and compared for each of these 25 points. The point with the smallest weighted SSE is identified and becomes an estimate of the optimal solution. The square formed by the 4 lines immediately surrounding this point becomes the new grid and the procedure is repeated. In this way, the solution is refined iteratively until the desired level of accuracy is achieved. This method is usable because the objective function (i.e. the weighted SSE) is a convex function of κ and σ.

5.3.7.3 The weights that are applied are subjective but can be guided by the amount of business in each tenor and their option date. The SSE can be defined either by the

errors of the prices or by the errors of the implied volatilities. The advantage of minimising the SSE on prices is that the swaption prices have a more linear relationship to the value of the GAO liabilities. For this reason, the weights have a more direct and intuitive meaning. If higher weights are applied to the tenor and option-term combination with the most business should ensure that that business is proportionately more accurately modelled at the expense of having the smaller terms and tenors modelled proportionately less accurately. On the other hand, the advantage of minimising the weighted SSE on implied-volatilities is that the implied volatilities tend to be more similar to each other in value than the corresponding swaption prices. For example, under normal market conditions, the implied volatilities on swaptions in a developed economy should be in the range of 10% to 14%, regardless of the term and tenor. Swaption prices, in contrast, vary more greatly by moneyness, term and tenor — the longer the term and tenor, the more expensive the swaption. The moneyness of the swaption describes the relationship between the strike rate in the swap agreement and the forward-swap rate $s(t, T-t)$. It determines if intrinsic value exists in the swaption and the degree to which the swaption is likely to have a positive value at its maturity. In-the-money swaptions have higher prices than out-the-money swaptions.

5.3.7.4 This means that SSE on implied volatilities should ensure a more even contribution to the optimisation. This is in contrast to SSE on prices, where the more expensive in-the-money swaptions with longer terms and tenor are by construction more influential in the SSE and needs to be compensated for by the weights. In the extreme, an out-the-money swaption with a very short term has a near-zero price and virtually no contribution to the SSE. For the purpose of this research, the weighted SSE on implied-volatilities is used, rather than the weighted SSE on prices.

5.3.7.5 Hull & White (*op. cit.*) also provides a method for the development of trees that allow either κ or σ to be a function of time such that all market swaption prices may be matched. They however note that when this happens, $\kappa(t)$ and $\sigma(t)$ both exhibit values that vary significantly with *t*. This causes unjustifiable prices if the functions $\kappa(t)$ and $\sigma(t)$ were interpolated or extrapolated to price assets of other terms. They recommend that the two variables remain as constants rather than functions of *t*. Even though this means that the modelled prices of swaptions do not exactly replicate the market prices, they should represent a reasonable approximation. Based on the rationale put forward in Paragraph 4.4.3 above, constant κ or σ should be acceptable for the purpose of market-consistent valuations.

5.3.8 Simulating the Short Rate

5.3.8.1 So far we have discussed the calibration of the interest-rate tree. The mean-reversion target, in the recombining lattice, is characterised by the drift term $\alpha(t)$. In equation (A5.1.3), this mean-reversion target is equal to $\frac{\theta(t)}{\kappa}$. In order to simulate the short rate, the relationship between $\alpha(t)$ and $\theta(t)$ is needed.

5.3.8.2 From (A5.1.13) we have

$$\alpha(t) = x(t, j) - x^*(t, j), \forall j . \tag{5.16}$$

5.3.8.3 From the definition in ¶A 5.1.2.5 of Appendix 5.1, we have:

$$x^*(t,0)=0. \tag{5.17}$$

5.3.8.4 Combining (5.16) and (5.17), we have:

$$\alpha(t)=x(t,0)-x^*(t,0)=x(t,0). \tag{5.18}$$

5.3.8.5 Without loss of generality, the drift of the midpoint of the lattice is:

$$\frac{x(t+1,0)-x(t,0)}{\Delta t}=\frac{\alpha(t+1)-\alpha(t)}{\Delta t}. \tag{5.19}$$

5.3.8.6 From (A5.1.3), we also know that the drift of the lattice at the same point is:

$$\theta(t)-\kappa x(t,0)=\theta(t)-\kappa(x^*(t,0)+\alpha(t))=\theta(t)-\kappa\alpha(t). \tag{5.20}$$

5.3.8.7 Equating the right-hand side of (5.19) and (5.20), we get:

$$\frac{\alpha(t+1)-\alpha(t)}{\Delta t}=\theta(t)-\kappa\alpha(t), \text{ and} \tag{5.21}$$

$$\theta(t)=\frac{\alpha(t+1)-\alpha(t)}{\Delta t}+\kappa\alpha(t). \tag{5.22}$$

5.3.8.8 $\theta(t)$ is the mean reversion level of the short rate at time t. This is true regardless of the level of the simulated short rate. By substituting the values of κ, σ and $\theta(t)$ into the discretised version of equation (A5.1.1) and setting $r(0)$ as the starting short-rate, we can simulate the future short rates.

5.3.9 Simulating the Future Bond Prices and Yield Curves

5.3.9.1 It is not sufficient to simulate the short rate at each future time. The valuation of a GAO that vests at a future date t requires a full term structure of interest rates at that time.

5.3.9.2 For short-rate models, there are three ways to project the future term structures that are consistent with the initial term structure. The first method uses closed-form solutions. These solutions do not always exist – their availability depends on the model form. Some models such as the one-factor Hull–White model, the closed-form solution is available. For the Black–Karasinski model, the closed-form solution does not exist.

5.3.9.3 The second method uses nested simulations. For each simulation s at time t, we simulate say, 1000 nested paths of short rates using $r(s,t)$ as the starting point (see Figure 5.2). We can use these nested simulations to price bonds of various terms and from these back out the term structure of interest rates. This method is highly inefficient.

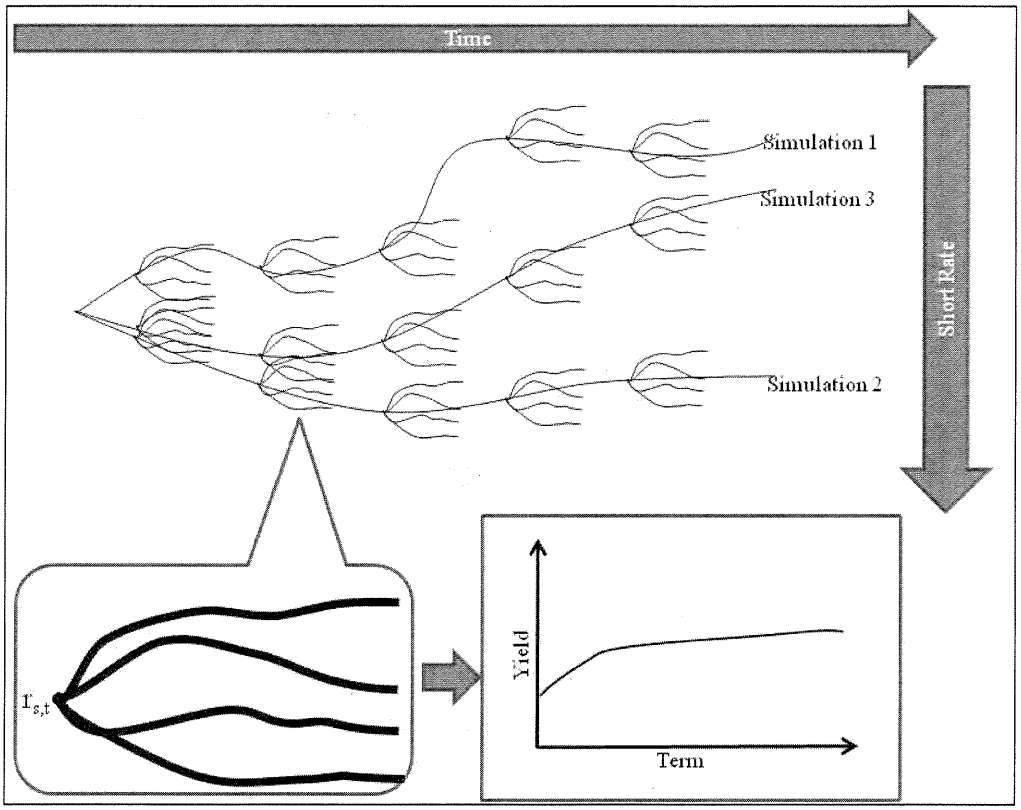


Figure 5.2: Simulating future term-structures by nested simulations

5.3.9.4 The last method is to nest interest-rate lattices at time t in the place of nested simulations, using the simulated $r(s,t)$ as the starting value. Here, we have a lattice nested within each projection horizon t and simulation s (see Figure 5.3). Because bonds are simple instruments, they can be priced using interest-rate lattices, which are much more efficient than nested simulations.

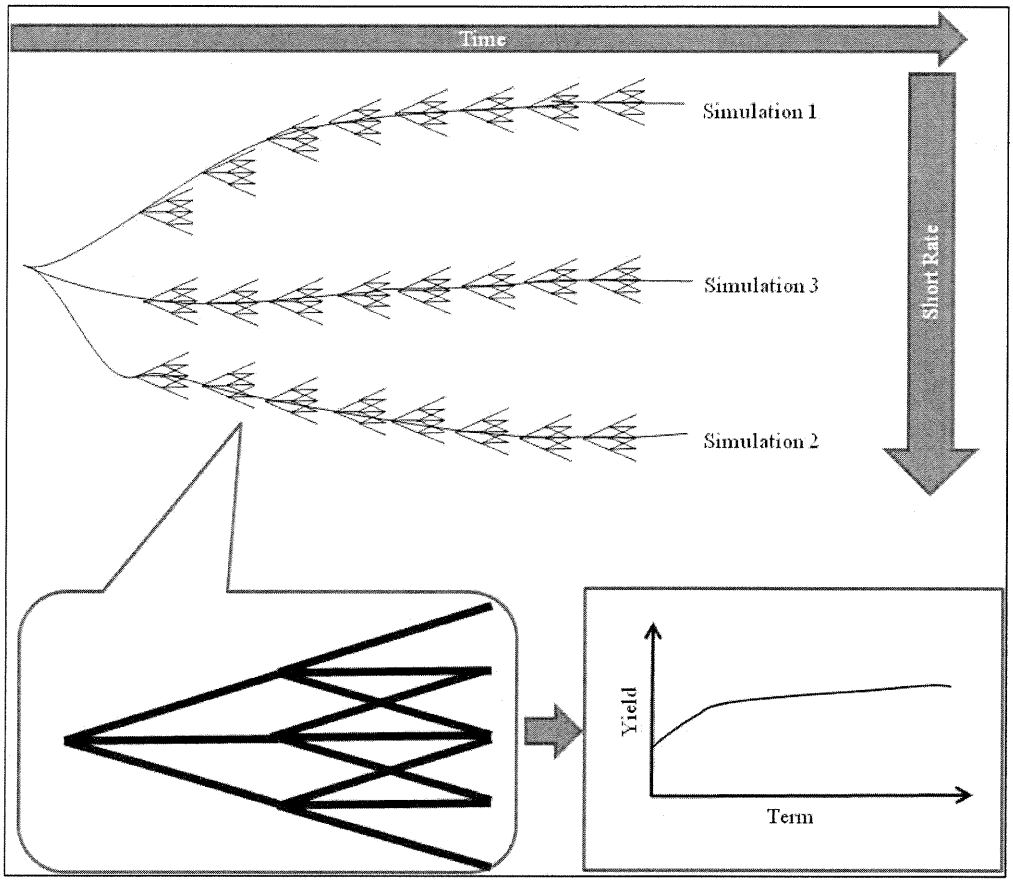


Figure 5.3: Simulating future term-structures by nesting lattices within simulations

5.3.9.5 To demonstrate the nesting of interest-rate lattice described in Paragraph 5.3.9.4, consider a situation where we need a term structure of $T-t$ years for each simulation s at each projection horizon t. The simulated $r(s,t)$ serves as a starting point for each lattice. We refit the interest-rate lattice using the two-stage Hull–White method as described in Appendix 5.1. The first-stage tree that was constructed for calibration purposes may be reused. This is because this tree is not time-dependent and is thus the same irrespective of the starting time t. Because we make the assumption that $x^*(0,0)=0$, this tree is also independent of the starting short rate.

5.3.9.6 To fit the second-stage tree, we note that the mean-reversion level of the tree is defined by $\theta(t+d)$ for $d=0,1,2,...,T-t$, calculated from (5.22) above. This is true because the mean-reversion level of the interest-rate tree $\theta(t+d)$ for $d=0,1,2,...,T-t$ remains the same irrespective of the simulated starting short rate $r(s,t)$. The mean reversion parameter $\alpha(t+d)$ for $\theta(t+d), d=0,1,2,...,T-t$ may be calculated iteratively using the following equation (from (5.21)):

$$\alpha(s,t+d+1)=\left(\theta(t+d)-\kappa\alpha(s,t+d)\right)\Delta t+\alpha(s,t+d).$$

5.3.9.7 The parameters κ and σ are the same as those calibrated in Section 5.3.7. Δt and Δx are the same as those calculated in the first stage of the tree-building process

described in Appendix 5.1 below. The starting condition for the lattice for simulation s at time t can be set as:

$$\alpha(s,t) = f(r(s,t)),$$

where the subscript s and t refer to the simulation number and projection time respectively.

5.3.9.8 Set the payoff at projection time t of a ZCB of a term of $T-t$ years (i.e. ZCB redemption occurs at time T) to equal to 1 for all nodes. The following equation may be used iteratively, working back from time T, to derive the price, at time t, of the ZCB with a term of $T-t$ years:

$$P(s,t+d,j,T-t-d) = \exp\left(-r(s,t+d,j)\Delta t\right)\sum_{k} P(s,t+d+1,k,T-t-d-1)q(k,j),$$

where

- the summation is across all nodes from which $(i+1,j)$ may be accessed;
- $q(k,j)$ is the probability of moving from (i,k) to $(i+1,j)$ as defined in Appendix 5.1; and
- $r_{s,t,j}$ is as defined in Appendix 5.1, i.e.

$$\begin{aligned} r(s,t+d,j) &= f^{-1}\left(x(s,d,j)\right) \\ &= f^{-1}\left(\alpha(s,t+d)+x^{*}(s,d,j)\right) \\ &= \exp\left(\alpha(s,t+d)+x^{*}(s,d,j)\right). \end{aligned}$$

5.3.10 Testing the ESG's Fit to the Initial Yield Curve

5.3.10.1 Here we test the fit of the ESG's outputs to the initial yield curve. We note that there are two aspects of the model that need to be tested:

- the short rates, as described in Section 5.3.8 above; and
- the future yield curves, as described in Section 5.3.9 above.

5.3.10.2 These two components need to be consistent with each other and also consistent with the starting yield curve. We ensure this consistency by pricing ZCBs with both approaches and then comparing the modelled prices to the actual market prices. To price the ZCBs using only the short rates, we use the following equation:

$$\hat{P}(t) = \frac{1}{N}\sum_{s=1}^{N}\prod_{j=0}^{t-1} \mathrm{e}^{-r(s,j)}, \tag{5.23}$$

where

- $\hat{P}(t)$ is the simulated price of a t-year ZCB;
- N is the number of simulations; and
- $r(s,j)$ is the simulated short rate from simulation s at time j.

5.3.10.3 $\hat{P}_t$ is calculated for $t = 1,2,...,T$. T is the length of the initial yield curve.

5.3.10.4 The payoff of a ZCB is discounted at the simulated short rates and averaged across all simulations. We must ensure that $P(t) \approx \hat{P}(t), t = 1,2,...,T$. The approximation is due to numerical inaccuracies that arise from Monte-Carlo simulations.

5.3.10.5 To test that the model constructs future yield curves correctly, we consider, for each s, pricing a ZCB of duration $t - t_1$ at time t_1 using the simulated term structure then. This price can then be discounted from time t_1 to time 0 at the corresponding short rates of that simulation. This is illustrated in Figure 5.4. The price of the t-year ZCB can be calculated by taking an average of the discounted values across all simulations.

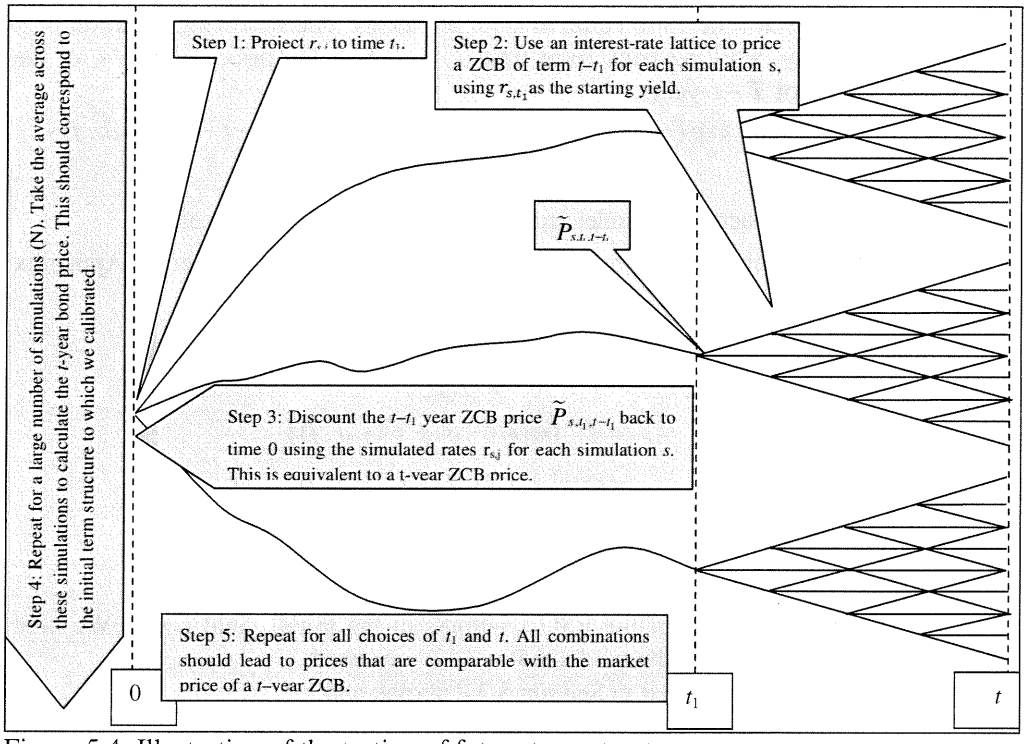


Figure 5.4: Illustration of the testing of future term structures

5.3.10.6 Using notation from (5.23), the price of a t-year ZCB can be calculated as:

$$\hat{P}(t) = \frac{1}{N}\sum_{s=1}^{N} \tilde{P}(s, t_1, t-t_1) \prod_{j=0}^{t_1-1} e^{-r(s,j)}, \qquad (5.24)$$

where $\tilde{P}(s, t_1, t-t_1)$ is the price, at time t_1, of a $t - t_1$ year ZCB for simulation s, calculated using the interest-rate lattice with the initial rate of $r(s, j)$. Again we should ensure that $P(t) \approx \hat{P}(t)$ for $t = 1, 2, ..., T$, and also for $t_1 = 0, 1, 2, ..., t$.

5.3.10.7 The interest-rate tree is calibrated to the swap curve of South Africa as at the end of December 2007. The test results attached in Appendix 5.2 show that prices and yields are matched when these are calculated using both the short rates and future term structures approaches.

5.3.11 Simulating Swaptions Prices and Implied Volatilities

5.3.11.1 The pricing of swaptions with trees was discussed in Section 5.3.6. The pricing of swaption using simulations follows the same method. Using the simulated bond prices from Section 5.3.9, we calculate the annuity factors as follows:

$$\tilde{a}_{\overline{t-t_1}|}(s,t_1) = \sum_{j=1}^{t} \tilde{P}(s,t_1,t-j),$$

where:

- $\tilde{a}_{\overline{d}|}(s,t_1)$ is the value of a certain annuity at time t_1 payable annually in arrears for d years for simulation s; and
- $\tilde{P}(s,t,d)$ has the same definition as in (5.24).

5.3.11.2 Let $\tilde{s}(t,j,T) = \frac{1-\tilde{P}(s,t,T)}{\tilde{a}_{\overline{T}|}(s,t)}$. The receiver and payer swaption prices are simulated as:

$$\frac{1}{N}\sum_{s=1}^{N}\left[\max(s(t,T)-\tilde{s}(t,j,T),0)\tilde{a}_{\overline{T}|}(s,t)\prod_{j=0}^{t-1}e^{-r(s,j)}\right] \tag{5.25}$$

and

$$\frac{1}{N}\sum_{s=1}^{N}\left[\max(\tilde{s}(t,j,T)-s(t,T),0)\tilde{a}_{\overline{T}|}(s,t)\prod_{j=0}^{t-1}e^{-r(s,j)}\right] \tag{5.26}$$

respectively.

5.3.11.3 Once the swaption prices are calculated, the implied volatility is backed out with (5.10) or (5.11) by substituting in the simulated price as the equation price in the appropriate equation and solving for the value of σ.

5.3.12 Testing the ESG's Fit to Market Yield Curves and Implied Volatility on Swaptions

5.3.12.1 In this section, we test the model's fit to a market term structure of interest rates and the implied volatility surface. The 31 December 2007 South-African swap-curve is used as the calibration target of the initial term structure. This is provided by Barrie and Hibbert (2008) and is replicated in Appendix 5.3. The term structure of interest rates is only market-observable for the first 30 years. For purposes that will become apparent later in Chapter 6, we extend this curve to 60 years by assuming a constant continuously compounded forward rate at a level that is equal to the forward rate at 30 years, the longest observable point. The forward and spot curves are shown in Figure 5.5.

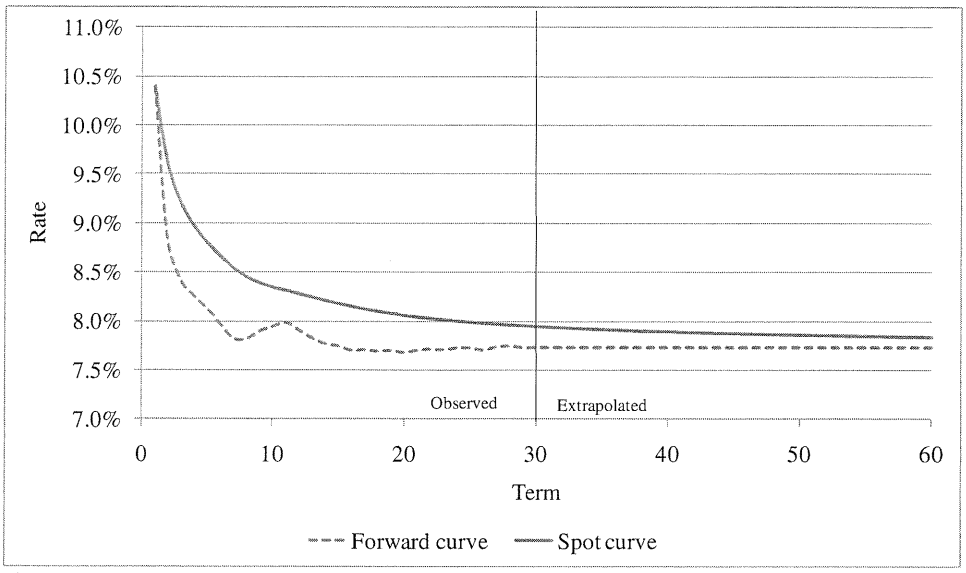


Figure 5.5: Extrapolation of spot curve by assuming constant forward curve

5.3.12.2 Figure 5.6 compares the observed term structure of interest rates to that simulated with Monte–Carlo simulations and interest-rate tree respectively. The simulated curve at time 0 is based on 2000 simulations.

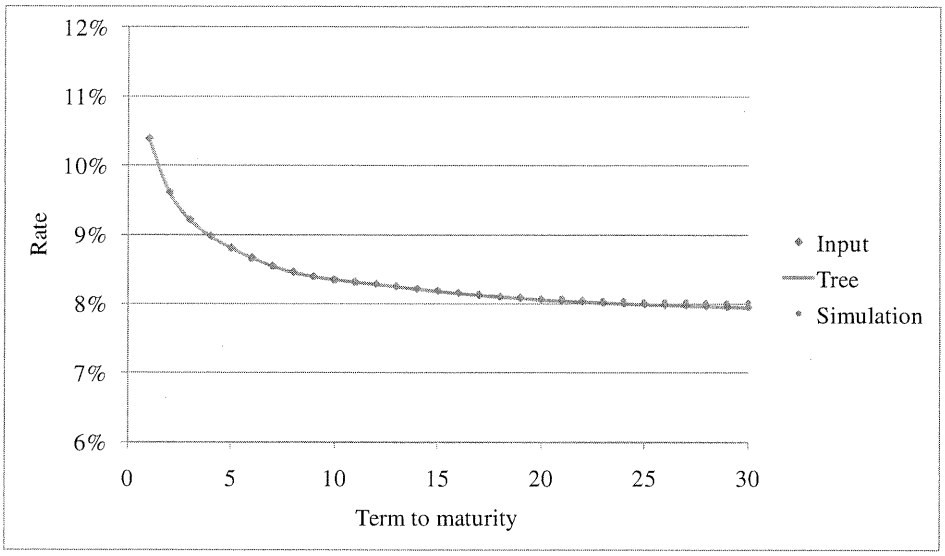


Figure 5.6: A comparison of yield curve generated by the interest-rate tree and Monte–Carlo simulations to the input

5.3.12.3 Three types of information are presented in Figure 5.6. 'Input' refers to the market yield curve that formed the calibration target. 'Tree' refers to the term structure of interest rates that was calibrated with the interest-rate lattice at time 0, as described by Appendix 5.1 and Section 5.3.7. 'Simulation' refers to the term structure that was backed out of the simulated prices of ZCB that were calculated with the procedure described in the first part of Section 5.3.10. By construction, the input and the lattice rates are exactly the same so they lie on the same line. The simulated rates

for the longer terms to maturity seem slightly higher than the input and tree rates — these were caused by sampling noise — but the differences should not be material for the purposes of market-consistent valuation.

5.3.12.4 As discussed in Section 5.3.7, κ and σ are calibrated to minimise the sum of squares of errors of the volatilities. It is difficult for the model to replicate the entire volatility surface with only two degrees of freedom. As a result, only one tenor is fitted. This is chosen so that it best replicates the liabilities — the swaption that has a tenor similar to the expected payment term of the annuity is used — the 20-year receiver swaption is used in this example. Figure 5.7 shows the market implied volatility on swaptions, the implied volatility as calculated by the interest-rate lattice and the implied volatility as calculated by the simulations. The simulated implied volatility is given here with a 95% confidence band (whose calculation allowed for the use of antithetic variables — see Appendix 5.4 for a more detailed description). We note that despite our targeting of only one tenor, there is still a statistically significant difference between the simulation-implied volatilities and the market-implied volatilities. This is still due to insufficient degrees of freedom.

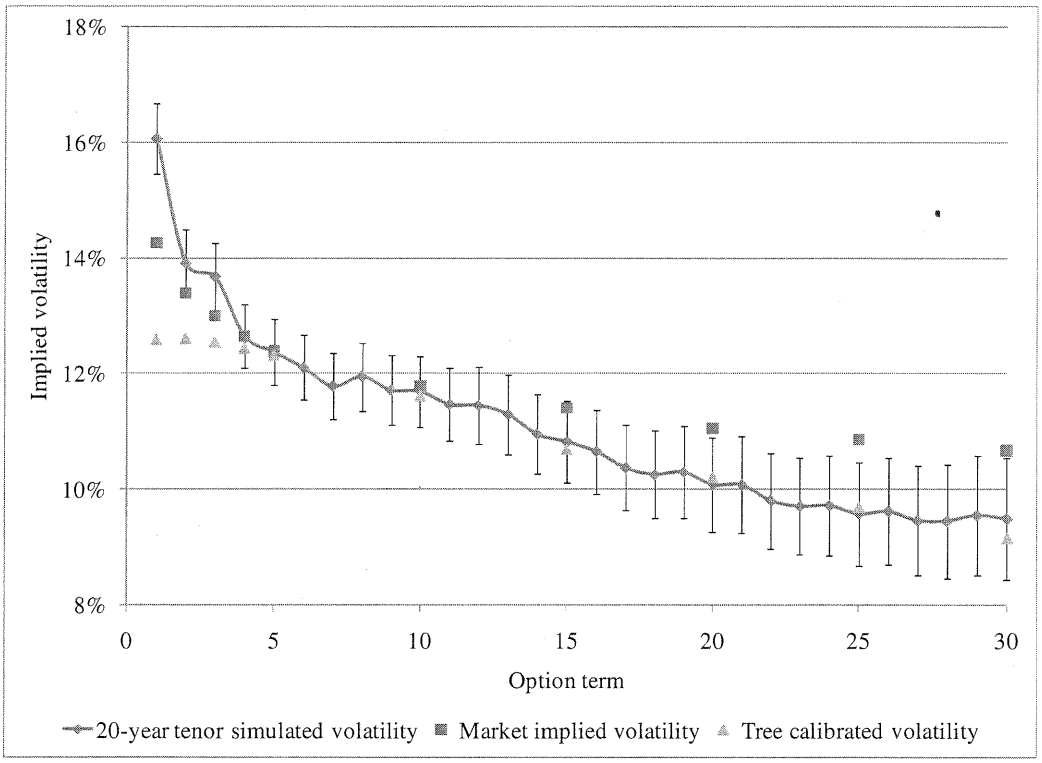


Figure 5.7: A comparison of implied volatilities of receiver swaptions with 20-year tenor: market, simulated and interest-rate lattice

5.3.12.5 The final parameters are given in Appendix 5.5.

5.4 Martingale Test

5.4.1 The concept of martingale test is introduced in this section. The concept will be used in Sections 5.5 and 5.6 of this chapter. All risk-neutral models should pass the

martingale test, sometimes also known as the '1 = 1' or '100 = 100' test. The purpose of this test is to ensure that the ESG does not create any bias in the projection of returns of simple assets, such as bonds and equities.

5.4.2 In the risk-neutral framework, the value of the asset is projected at the simulated asset return and is discounted at the simulated risk-free rate. The test tests the hypothesis that:

$$E[L]=1, t=1,2,...,T, \tag{5.27}$$

where L is the time-0 value of the asset calculated by the model.

5.4.3 Define

$$l(s,t)=\frac{\prod_{i=1}^{t}\exp(R(s,i))}{\prod_{i=1}^{t}\exp(r(s,i))}, \tag{5.28}$$

where

- $R(s,i)$ is the return from the risky asset at time i in simulation s, and
- $r(s,i)$ is the return from the risk-free asset at time i in simulation s.

5.4.4 We simulate n such values. Define $\widehat{l(t)}=\frac{1}{n}\sum_{s=1}^{n}l(s,t)$. We test that $\hat{l}$ is an unbiased estimator of L in a risk-neutral framework. This expression has a variance of $\frac{1}{n-1}\sum_{s=1}^{n}\left(l(s,t)-\widehat{l(t)}\right)^2$. If antithetic variables are used, the variance has a different expression (see Appendix 5.4).

5.4.5 The mean and the variance can be used to test the null hypothesis that $\hat{l}$ is an unbiased estimator of L by assuming the normal distribution.

5.5 Bond-portfolio Returns

5.5.1 Introduction

5.5.1.1 This section tests a portfolio of bonds for compliance with the martingale tests as outlined in Section 5.4. This requires the calculation of the return on a bond portfolio.

5.5.1.2 For simplicity, we assume that only T-year ZCBs are in the bond portfolio. At the end of a one-year period, the bond only has a remaining term of T–1 years. At this time, the portfolio rebalanced to T-year ZCBs of the same market value. In practice, this portfolio may be extended to any number and combinations of bonds.

5.5.2 Returns from the Bond Portfolio

5.5.2.1 Bond-portfolio returns are modelled as the change in the ZCB prices over the year. Define $B(s,t,d)$ as the return on the ZCB of duration d in the period $(t,t+1)$ for simulation s, then

$$B(s,t,d) = \log\left(\frac{P(s,t+1,d-1)}{P(s,t,d)}\right). \tag{5.29}$$

where $P(s,t,d)$ is the price of a d-year ZCB at time t in simulation s.

5.5.2.2 It is possible to construct a more complex portfolio than the one described above simply by including more ZCBs in the portfolio. This is done by replacing the bond prices in (5.29) with those of the portfolio.

5.5.3 Martingale Test on the Bond-portfolio Returns

5.5.3.1 This section presents the results of the martingale test for bond prices. The martingale test should be passed by all bond portfolios. Figure 5.8 illustrates the result of the martingale test on a portfolio of 15-year ZCBs. $B(s,t,15)$ is used as the return term in the numerator of (5.28).

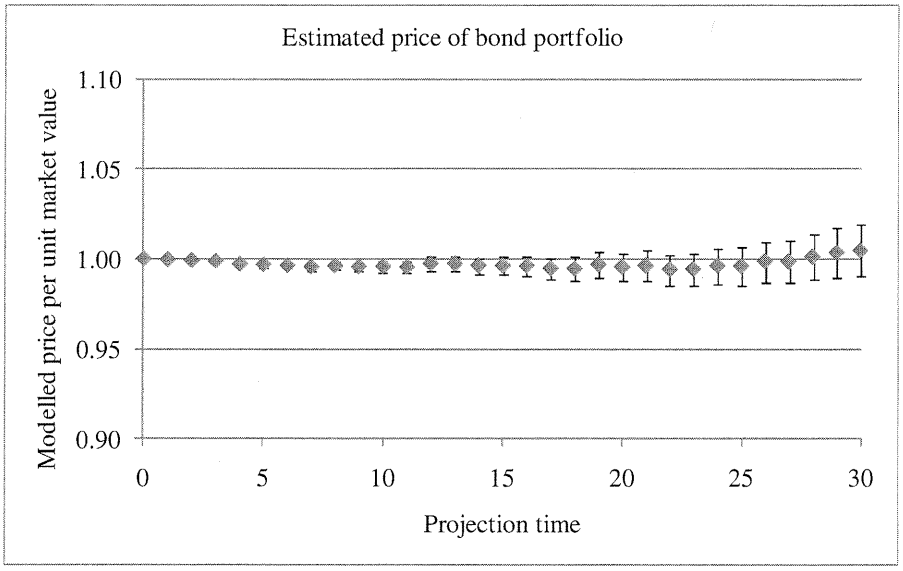


Figure 5.8: Martingale test for 15-year ZCBs

5.5.3.2 Figure 5.9 shows the martingale tests of a bond portfolio with six ZCBs in equal weightings of one-sixth each: one sixth of the portfolio is invested in each and dynamically rebalanced annually to this ratio. The six bonds have maturities of 5 years, 10 years, 15 years, 20 years, 25 years and 30 years. It is clear in both figures that the martingale test is passed — the modelled price is very close to the unit-line for all projection time t. 95% confidence bands are also shown on the two charts. We fail to reject the martingale hypothesis.

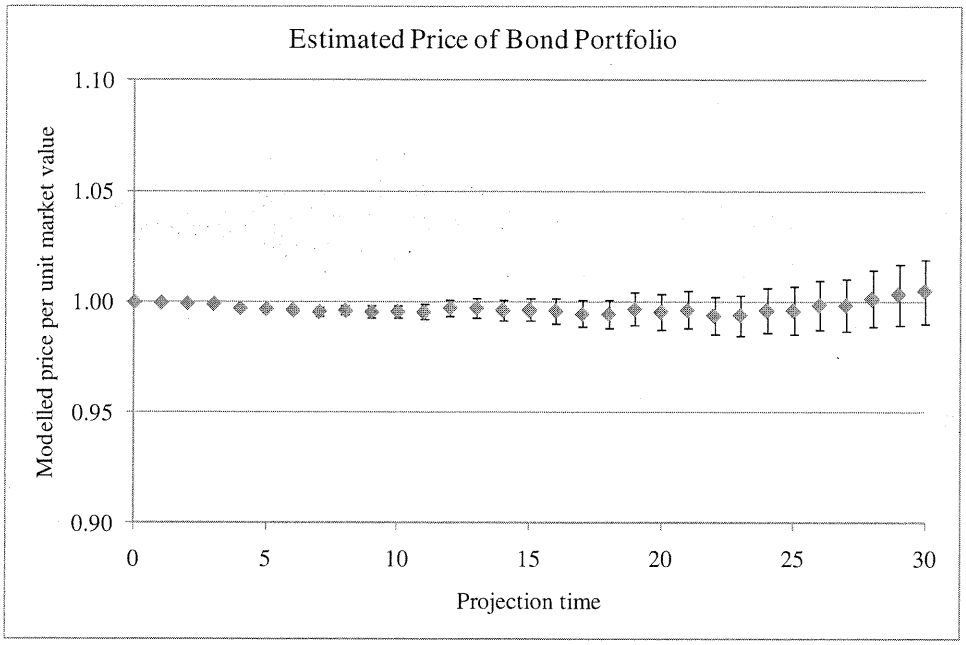


Figure 5.9: Martingale test for a bond portfolio with six ZCBs in equal weightings

5.6 Equity-portfolio Returns

5.6.1 Introduction

5.6.1.1 The unit funds to which GAOs apply are usually balanced funds that are focused on growth. A large equity component can usually be found in these portfolios. An equity model is needed in order to project the fund values at retirement. We consider using the Black–Scholes model to model a total-returns equity index. Assume without loss of generality that no dividend is paid, so that all returns are in the form of capital gains. The Black–Scholes model may be applied using a Monte–Carlo approach. The model has the following stochastic differential equation:

$$d\log S(t) = \left[r(t) - \frac{1}{2}\sigma^2(t) \right] dt + \sigma(t) dZ ,$$

where S_t is the value of the total-returns index at time t.

5.6.1.2 This model has the following discrete form:

$$\log S(t+\Delta t) - \log S(t) = \left[r(t) - \frac{1}{2}\sigma^2(t) \right] \Delta t + \sigma(t)\sqrt{\Delta t} Z , \qquad (5.30)$$

where Z is random and has a $N(0,1)$ distribution.

5.6.1.3 Simplifying (5.30), we get:

$$S(t+\Delta t) = S(t) \exp\left[\left(r(t) - \frac{1}{2}\sigma^2(t) \right) \Delta t + \sigma(t)\sqrt{\Delta t} Z \right] \qquad (5.31)$$

5.6.1.4 Taking expectations of both sides of (5.31), we get:

$$\mathrm{E}\left[S(t+\Delta t)\right] = S(t)\exp\left[\left(r(t)-\frac{1}{2}\sigma^2(t)\right)\Delta t\right]\exp\left[\frac{1}{2}\sigma^2(t)\Delta t\right] = S(t)\exp\left[r(t)\Delta t\right].$$

5.6.1.5 Because the expected return from the share is the risk-free return r, this method is risk-neutral. We also note that the left-hand side of (5.30) is the logarithmic return $\log\left(\frac{S(t+\Delta t)}{S(t)}\right)$. The drift term in the simulation of the logarithmic return is $\left(r(t)-\frac{1}{2}\sigma^2(t)\right)\Delta t$, its standard deviation is $\sigma(t)\sqrt{\Delta t}$.

5.6.1.6 Note the subtle difference between the $\sigma(t)$ here and the $\sigma(t)$ in the classic Black–Scholes equation: the former is the ex-ante forward instantaneous volatility that occurs from time t to time $t + \Delta t$, whereas the latter refers to the implied volatility of a t-year option, which may be interpreted as the ex-ante volatility that applies from time 0 to time t. In the classic Black–Scholes world where a constant σ is assumed, the two definitions are equivalent. Throughout the rest of this text, the former will be denoted as $\sigma^*(t)$ to distinguish between the two. The following describes their relationship.

$$\sigma(t) = \sqrt{\frac{1}{t}\int_0^t \sigma^{*2}(u)du}\,. \tag{5.32}$$

5.6.1.7 Analogous to the same parameter in the Black equation for swaptions, $\sigma(t)$ may be interpreted as a proxy for the price of an equity option. This section describes the modelling of $\sigma^*(t)$ and Z.

5.6.2 Modelling Equity Returns

5.6.2.1 In the modelling of the equity returns, it is assumed that equities are subject to two types of risks:

- systematic risk, which is the volatility related to the drivers of the economy; and
- specific risk, which is the noise generated by equities only and is unaffected by other factors.

5.6.2.2 The reason for distinguishing between these two risks will be made clearer below. Note that the definitions of systematic risk and specific risk presented here differ from their traditional definitions, where the former refers to the volatility of the market portfolio of equity that is undiversifiable and the latter refers to the volatility of individual shares that is diversifiable. In our model, it is assumed that the short rate is the main driver of the economy. The systematic risk is therefore a function of the level of the simulated short rate.

5.6.2.3 The two components together should reflect the forward volatility, which in turn should reflect the implied volatility. Using the terminology defined in the previous section, this is:

$$\sigma^{*2}(t) = \sigma^{*2}_{sys} + \sigma^{*2}_{spe}(t),$$

where

- $\sigma^*(t)$ is the forward implied volatility at time t;
- σ^*_{sys} is the systematic volatility, assumed to be constant in all future times; and
- $\sigma^*_{spe}(t)$ is the forward specific-risk component.

5.6.2.4 Systematic volatility, representing the volatility of the equity market portfolio that results from the volatility of interest rates, is difficult to observe in a market-consistent way. This is because there is no deep and liquid market of derivatives that span across multiple asset classes. As a result, this component has been left as a constant and was fitted by calculating the historical correlation between the change of the Johannesburg Interbank Agreed Rage (JIBAR) and the monthly equity returns. The change in JIBAR is used rather than the level of JIBAR because the dZ in the short-rate model describes the change in the short rate. The choice of the historical period and the frequency of observation can affect the correlation and so this method can introduce subjectivity into the process. In the author's experience, GAO valuation is fortuitously insensitive to the correlation structure.

5.6.2.5 The specific volatility is modelled as a deterministic function of t. This allows the market-observed term structure of implied volatility to be approximated by the model.

5.6.3 Correlation between Equity Returns and the Short Rate

5.6.3.1 Equation (5.30) specifies an equation for simulating equity returns as a function of the short rate. This equation does not allow for the correlation between the short rate and the equity returns. We modify this equation as follows:

5.6.3.2 Define

$$\sigma^*_{sys} = \beta\sigma; \tag{5.33}$$

where:

- β is the sensitivity of the equity return to the short rate;
- σ is the volatility parameter in the Black-Karasinski model.

5.6.3.3 Define:

$$\begin{aligned} E(t,t+\Delta t) &= \log S(t+\Delta t) - \log S(t) \\ &= r(t)\Delta t - \frac{1}{2}\left[\sigma^{*2}_{sys} + \sigma^{*2}_{spe}(t)\right]\Delta t + \left[\sigma^*_{sys} Z_1 + \sigma^*_{spe}(t) Z_2\right]\sqrt{\Delta t} \end{aligned} \tag{5.34}$$

where:

- $E(t,t+\Delta t)$ is the logarithmic return of the equity portfolio in the period $(t,\Delta t)$;
- Z_1 and Z_2 are simulations from the $N(0,1)$ distribution; and
- $\text{cov}(Z_1,Z_2)=0$.

5.6.3.4 By construction, we have:

$$\text{cov}\left(E(t,t+\Delta t), \log r(t+\Delta t)\right)$$

$$
\begin{aligned}
&= \operatorname{cov}\begin{pmatrix} r(t)\Delta t - \frac{1}{2}\left[\sigma_{sys}^{*2} + \sigma_{spe}^{*2}(t)\right]\Delta t + \left[\beta\sigma Z_1 + \sigma_{spe}^{*}(t)Z_2\right]\sqrt{\Delta t}, \\ \log r(t) + \left[\theta(t) - a\log r(t)\right]\Delta t + \sigma\sqrt{\Delta t}Z \end{pmatrix} \\
&= \operatorname{cov}\left(\left[\beta\sigma Z_1 + \sigma_{spe}^{*}(t)Z_2\right]\sqrt{\Delta t}, \sigma\sqrt{\Delta t}Z\right) \\
&= \operatorname{cov}\left(\beta\sigma Z_1\sqrt{\Delta t}, \sigma\sqrt{\Delta t}Z\right), \text{ since } Z_2 \text{ is independent of } Z_1 \text{ and } Z \text{ by definition} \\
&= \Delta t\beta\sigma^2 \operatorname{cov}(Z_1, Z)
\end{aligned}
$$

$$
\begin{aligned}
&\operatorname{correlation}\left(E_{t,t+\Delta t}, \log r(t+\Delta t)\right) \\
&= \frac{\operatorname{cov}\left(E(t,t+\Delta t), \log r(t+\Delta t)\right)}{\sqrt{\operatorname{var}\left(E(t,t+\Delta t)\right)\operatorname{var}\left(\log r(t+\Delta t)\right)}} \\
&= \frac{\Delta t\beta\sigma^2 \operatorname{cov}(Z_1, Z)}{\sqrt{\operatorname{var}\left(\left(\beta\sigma Z_1 + \sigma_{spe}^{*}(t)Z_2\right)\sqrt{\Delta t}\right)\operatorname{var}\left(\sigma Z\sqrt{\Delta t}\right)}} \\
&= \frac{\Delta t\beta\sigma^2 \operatorname{cov}(Z_1, Z)}{\sqrt{\beta^2\sigma^2 + \sigma_{spe}^{*2}(t)}\Delta t\sigma} \\
&= \frac{\beta\sigma\rho_{Z_1,Z}}{\sqrt{\beta^2\sigma^2 + \sigma_{spe}^{*2}(t)}} \text{ since } \operatorname{var}(Z_1) = 1, \operatorname{var}(Z) = 1 \\
&\approx \rho_{Z_1,Z}
\end{aligned} \tag{5.35}
$$

The approximation in the last step holds if $\sigma_{spe}^{*2}(t)$ is small relative to $\beta^2\sigma^2$.

5.6.3.5 Equation (5.35) shows that simulating correlated normal variates leads to similar correlations in the simulated equity returns and the logarithm of yields.

5.6.4 Simulating Correlated Normal Variates

5.6.4.1 Section 5.6.3 makes reference to the simulation of multivariate standard normal variates that are correlated with each other. Cholesky decomposition is used for this simulation.

5.6.4.2 Let

- $Z_i^* \sim N(0,1)$, $i = 1,2,...,n$;
- $\operatorname{cov}(Z_i^*, Z_j^*) = 0, i \neq j$; and
- $\mathbf{z}^* = \left(Z_1^* \quad Z_2^* \quad ... \quad Z_n^*\right)$.

5.6.4.3 Let **W** be the upper triangular matrix in the Cholesky decomposition of the desired correlation matrix **P**, such that:

$$\mathbf{W}^T\mathbf{W} = \mathbf{P}.$$

5.6.4.4 If we define **z** as:

$$\mathbf{z} = \left(Z_1 \quad Z_2 \quad ... \quad Z_n\right) = \mathbf{z}^*\mathbf{W},$$

we ensure that:

- $Z_i \sim N(0,1), i = 1,2,...,n$; and
- $\text{cov}(Z_i, Z_j) = P_{i,j}$.

5.6.5 Specifying the Term Structure of Specific Volatility

5.6.5.1 To value GAO liabilities, we need to project the fund values at the vesting dates of the option at retirement. To do this, we need to project equity returns. Given that GAOs were still being sold in the 1990s to people who were employed at that time, some policyholders are potentially still in the mid 40s and are potentially another 30 years from retirement.

5.6.5.2 In a risk-neutral world, equity-returns are calibrated to option-implied volatilities. In a market such as South Africa, equity options of terms greater than three years are rarely quoted in the market, and options with terms greater than five years are never quoted. Although life insurers no longer sell GAOs, many of them are still offering other forms of investment guarantees such as guaranteed maturity benefits. Because these guarantees typically have terms in excess of thirty years, assumptions must be made for implied volatilities for terms longer than those available in the market. The common practice is to make an assumption for the implied volatilities at a very long term. A separate assumption is then made to interpolate between the market volatilities at the short term and this long-term assumption.

5.6.5.3 For the purposes of illustration, we assume that the implied volatility is 27% at 30 years. We define the following equation to characterise the term structure of implied volatilities:

$$\sigma^*_{spe}(t) = e^{-\alpha^* t}\sigma^*_0 + (1 - e^{-\alpha^* t})\sigma^*_\infty . \tag{5.36}$$

5.6.5.4 This is an exponential interpolation of forward-specific volatility from σ^*_0, the short-term specific volatility and σ^*_∞, a long-term specific volatility. α^* here represents the speed of the reversion. The value of σ^*_0 is affected mainly by the options traded in the market, whereas σ^*_∞ is determined largely by our long-term volatility assumption. σ^*_0, α^* and σ^*_∞ are determined by minimising the sum of squares of errors between observed and modelled implied volatilities.

5.6.5.5 Substituting (5.33) and (5.36) into (5.32) gives the term structure of implied volatility:

$$\begin{aligned}\sigma(t) &= \sqrt{\frac{1}{t}\int_0^t \left(\beta^2\sigma^2 + \left(e^{-\alpha^* u}\sigma^*_0 + (1 - e^{-\alpha^* u})\sigma^*_\infty\right)^2\right) du} \\ &= \sqrt{\beta^2\sigma^2 + \frac{1}{t}\left(\sigma^{*2}_0 f_1(t) + \sigma^{*2}_\infty f_2(t) + 2\sigma^*_0\sigma^*_\infty f_3(t)\right)}\end{aligned}$$

where:

- $f_1(t) = \int_0^t e^{-2\alpha^* u} du = \left(\frac{e^{-2\alpha^* t} - 1}{-2\alpha^*}\right)$,

- $f_2(t) = \int_0^t \left(1 - e^{-\alpha^* u}\right)^2 du = t - \frac{2}{-\alpha^*} - \frac{2e^{-\alpha^* t}}{-\alpha^*} - \frac{1}{-2\alpha^*} + \frac{e^{-2\alpha^* t}}{-2\alpha^*}$, and
- $f_3(t) = \int_0^t e^{-\alpha^* u}\left(1 - e^{-\alpha^* u}\right) du = \frac{e^{-\alpha^* t}}{-\alpha} - \frac{e^{-2\alpha^* t}}{-2\alpha^*} - \frac{1}{-\alpha} + \frac{1}{-2\alpha^*}$.

5.6.5.6 This simplifies to:

$$\sigma(t) = \sqrt{\beta^2\sigma^2 + \sigma_\infty^{*2} + 2\left(\sigma_0^* - \sigma_\infty^*\right)\sigma_\infty^*\left(\frac{1 - e^{-\alpha^* t}}{\alpha^* t}\right) + \left(\sigma_0^* - \sigma_\infty^*\right)^2\left(\frac{1 - e^{-2\alpha^* t}}{2\alpha^* t}\right)}. \qquad (5.37)$$

5.6.5.7 This is the analytic, implied volatility of the model. This function may be used to calibrate the term structure of volatility by minimising the sum of squares of errors between the analytic model-implied volatility and the market-implied volatility.

5.6.6 Simulating Equity Option Prices and Implied Volatilities

5.6.6.1 Prices of equity options can be simulated using the simulated risk-free and equity returns and applying the following equations:

$$\hat{c} = \frac{1}{N}\sum_{s=1}^{N}\left[\max\left(S(0)\left[\prod_{i=0}^{t-1} e^{E(s,i)}\right] - K(t), 0\right)\prod_{i=0}^{t-1} e^{-r(s,i)}\right]; \text{ and}$$

$$\hat{p} = \frac{1}{N}\sum_{s=1}^{N}\left[\max\left(K(t) - S(0)\left[\prod_{i=0}^{t-1} e^{E(s,i)}\right], 0\right)\prod_{i=0}^{t-1} e^{-r(s,i)}\right].$$

where:

- $K(0)$ is the strike price at time;
- $S(0)$ is the spot price of an equity index;
- N is the number of simulations;
- $E(s,t)$ is the simulated equity return for the period $[i, i+1)$ of simulation s;
- $r(s,i)$ is the simulated short rate for period $[i, i+1)$ of simulation s;
- $\hat{c}$ is the price of the t-year call option with strike price $K(t)$ and
- $\hat{p}$ is the price of the t-year put option with strike price$K(t)$.

5.6.6.2 Like swaptions, the implied volatility σ is backed out using the Black–Scholes equation by setting the price equal to the simulated price and r equal to the spot-rate of duration t.

5.6.7 Calibrating the ESG to Market Equity-option Implied-volatility Structure

5.6.7.1 The short end of the implied volatility curve is calibrated to mid-market at-the-money put-option prices as at the end of December 2007. The data are provided by Barrie and Hibbert (*op. cit.*). These are in Appendix 5.3. As mentioned in Section 5.6.5, the 30-year implied volatility has been fixed at 27%. Equation (5.37) is used to calibrate the term structure of volatility by minimising the sum of squares of errors between the analytic model-implied volatility and the market-implied volatility.

5.6.7.2 However, the implied volatilities that are backed out from the simulated options prices usually differ from the analytic model-implied volatilities calculated here: the implied volatilities that are backed out of the simulated options prices are usually higher than that suggested by the analytic formula, particularly at the long end of the term structure. This is because our calibration, as in the Black–Scholes formulaic framework, makes an implicit assumption that the only sources of volatility are captured in the projection of equity returns in equation (5.31). In a Monte–Carlo framework, there is an additional source of volatility in the form of the stochastic risk-free rate, used in the discounting of cashflows. The simulation of equity option prices is affected by the volatility of the short rate, whereas the Black–Scholes equation used to back out the implied volatility assumes a constant risk-free rate of return. The difference between the analytic and the simulated volatility is higher when the σ parameter and the level of the term structure are both high. A practical solution to this problem is to artificially decrease σ_∞ in the calibration so that the implied volatilities that are backed out from the simulated option prices are lower. The downside of this approach is that it may result in a non-sensible implied-volatility term structure in the medium term.

5.6.7.3 Figure 5.10 shows the market-implied volatilities, the calibrated implied volatility and the simulated implied volatility. For the last item, a 95% confidence interval is also shown. These confidence bands show that the simulated implied-volatility term-structure satisfies the calibration target. Note that the calibrated implied-volatility curve has been forced lower than 27% so that the simulation output fits the 27% calibration target. This is to compensate for the over-estimation caused by stochastic interest rates described in Paragraph 5.6.7.2.

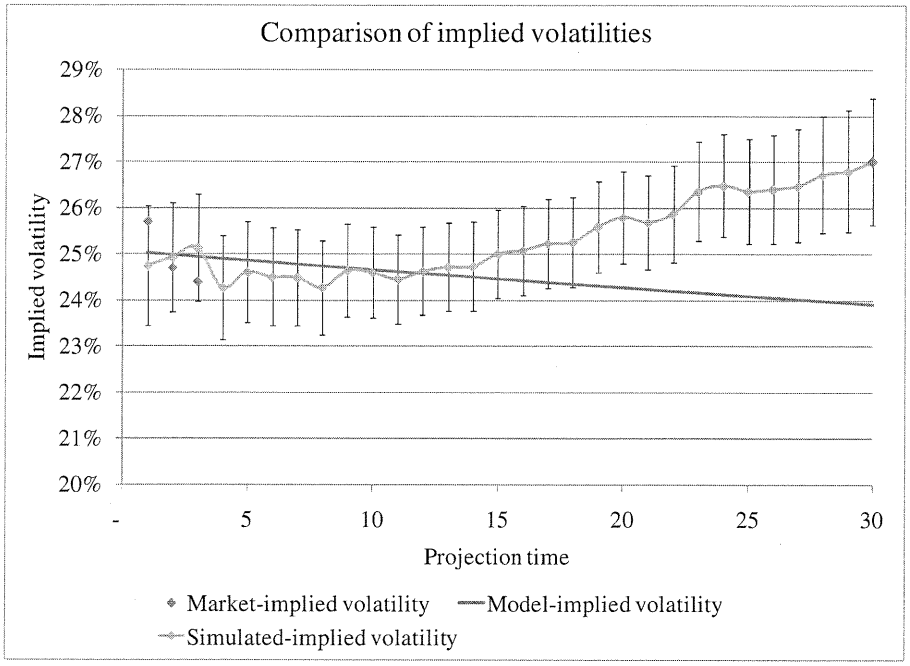


Figure 5.10: Market implied-volatility and its comparison against those calculated using the model and simulations

5.6.7.4 Figure 5.11 shows the martingale test for the equity portfolio. As before, the confidence bands reflect a 95% confidence band. This shows that the mispricing error is larger than for the bond portfolios illustrated in Section 5.5.3.2. This is reasonable as equities exhibit higher volatility than bond portfolios. Nevertheless, we fail to reject the martingale hypothesis with 95% confidence.

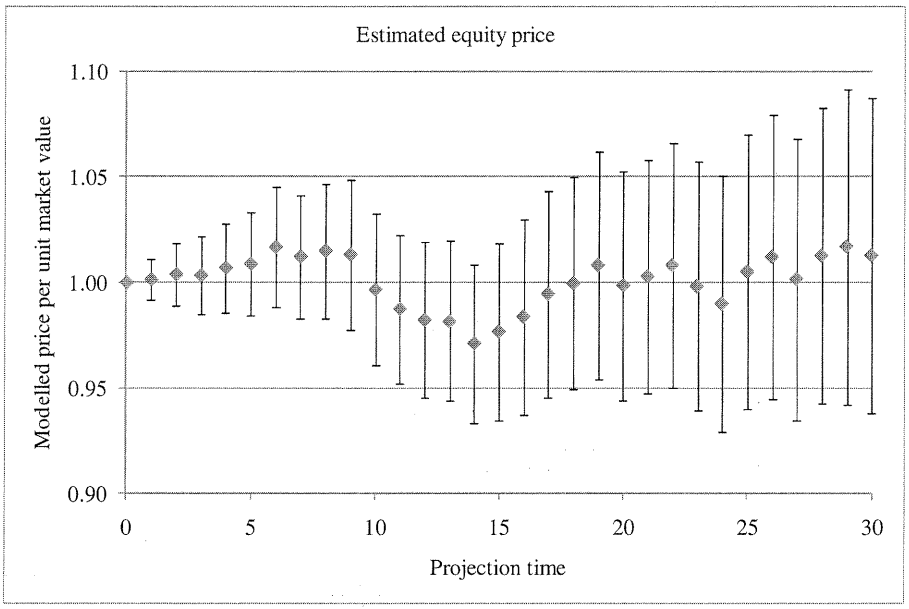


Figure 5.11: Martingale test for the equity portfolio

Appendix 5.1 – Building an Interest-rate Tree

A 5.1.1 Introduction

A 5.1.1.1 This section describes a methodology for fitting interest-rate trees. The method was proposed by Hull & White (1993a, 1993b) and Hull (2006) and may be used for all one-factor models of the short rate described in ¶5.3.2.1, where the model can be expressed in the form of (5.1). Here we apply this method to the Black–Karasinski model, where $f(r) = \log r$.

A 5.1.1.2 Consider the model given by equation (5.4) with constant κ and σ:

$$d\log r(t) = \left[\theta(t) - \kappa \log r(t)\right]dt + \sigma dZ\,. \tag{A5.1.1}$$

A 5.1.1.3 To simplify the notation, define:

$$x = f(r(t)) = \log r(t)\,. \tag{A5.1.2}$$

This gives:

$$dx = \left[\theta(t) - \kappa x\right]dt + \sigma dZ\,. \tag{A5.1.3}$$

A 5.1.1.4 Hull & White (1993a) note that although binomial lattices can be fitted to this model, it does provide some practical challenges. The biggest obstacle is that the tree cannot be recombining whilst the timesteps are constant. Recombination is important to keep the complexity and runtime low. They suggest using a trinomial tree, which offers an additional degree of freedom, so that the recombination obstacle can be overcome. The tree-building procedure is split into two stages.

A 5.1.2 Building an Interest-rate Tree – the First Stage

A 5.1.2.1 The first stage involves building a tree for x_t^* where $dx^* = -\kappa x^* dt + \sigma dZ$ and $x_0^* = 0$. Discretising this we have:

$$x_{t+\Delta t}^* - x_t^* = -\kappa \Delta t x_t^* + \sigma\sqrt{\Delta t}Z\,, \tag{A5.1.4}$$

where $Z \sim N(0,1)$, so that:

$$x_{t+\Delta t}^* - x_t^* \mid x_t^* \sim N\left(-\kappa x_t^* \Delta t, \sigma^2 \Delta t\right), \tag{A5.1.5}$$

A 5.1.2.2 The expected change in x_t^* is therefore:

$$x_t^*(-\kappa\Delta t) = x_t^* M\,,$$

where $M = -\kappa\Delta t$.

A 5.1.2.3 Similarly, the variance of the change in x_t^* is:

$$V = \sigma^2 \Delta t\,.$$

A 5.1.2.4 Like Hull & White (*op. cit.*), we set the size of the interest-rate step in the tree as:

$$\Delta x^* = \sqrt{3V}\,. \tag{A5.1.6}$$

A 5.1.2.5 Consider the model given by equation (5.4) with constant κ and σ: We build a tree that is evenly spaced in both x^* and t. Consider a node after i timesteps (i.e. where $t = i\Delta t$) and has experienced a net movement of j upward drifts, (i.e. so that it attains a value of $j\Delta x^*$). Denote this as node (i, j), $i \in \mathbb{N}$ and $j \in \mathbb{Z}$, the set of integers. Define $x^*(i, j)$ as the value of node (i, j).

A 5.1.2.6 Note that $x^*(i, j) = j\Delta x^*, \forall i \in \mathbb{N}$. The inclusion of the first dimension i is to simplify the interpretation of equations that appear below.

A 5.1.2.7 Define p_u, p_m and p_d as the probabilities of the highest, middle and lowest branches that emanate from a node respectively. Figure 5.12 shows an example of the interest-rate tree.

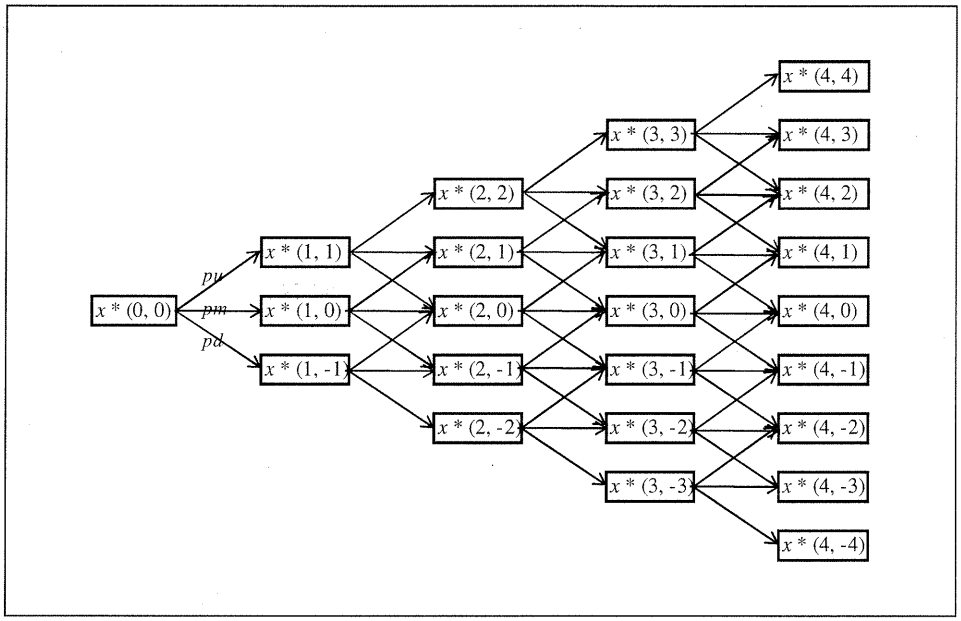


Figure 5.12: Stage 1 interest-rate tree

A 5.1.2.8 We want to calculate p_u, p_m and p_d such that the moments are matched, subject to the boundary condition that they sum to unity.

A 5.1.2.9 To match the first moment of the change, we have:

$$E\left[x^*_{t+\Delta t} - x^*_t \middle| x^*_t = j\Delta x^*\right] = p_u\Delta x^* - p_d\Delta x^* = -\kappa j\Delta x^*\Delta t\,. \qquad \text{(A5.1.7)}$$

A 5.1.2.10 To match the second moment of the change, we have:

$$E\left[\left(x^*_{t+\Delta t}-x^*_t\right)^2\middle| x^*_t = j\Delta x^*\right] = p_u(\Delta x^*)^2 + p_d(\Delta x^*)^2$$
$$= \text{var}\left[x^*_{t+\Delta t}-x^*_t\middle| x^*_t = j\Delta x^*\right] + E\left[x^*_{t+\Delta t}-x^*_t\middle| x^*_t = j\Delta x^*\right]^2 \quad \text{(A5.1.8)}$$
$$= \text{var}\left[x^*_{t+\Delta t}\middle| x^*_t = j\Delta x^*\right] + E\left[x^*_{t+\Delta t}-x^*_t\middle| x^*_t = j\Delta x^*\right]^2$$
$$= \sigma^2\Delta t + \kappa^2 j^2(\Delta x^*)^2\Delta t^2.$$

A 5.1.2.11 The probabilities must also sum to unity:

$$p_u + p_m + p_d = 1. \quad \text{(A5.1.9)}$$

A 5.1.2.12 Solving for p_u, p_m and p_d simultaneously using equations (A5.1.7) to (A5.1.9), we get:

$$p_u = \frac{1}{6} + \frac{1}{2}(\kappa^2 j^2\Delta t^2 - \kappa j\Delta t), \quad \text{(A5.1.10)}$$

$$p_m = \frac{2}{3} - \kappa^2 j^2\Delta t^2, \text{ and} \quad \text{(A5.1.11)}$$

$$p_d = \frac{1}{6} + \frac{1}{2}(\kappa^2 j^2\Delta t^2 + \kappa j\Delta t). \quad \text{(A5.1.12)}$$

A 5.1.2.13 We note here that using these equations, it is possible, depending on the values of κ, j and Δt, that any one or more of these probabilities becomes less than 0 or greater than 1. Hull & White (*op. cit.*) suggest that 'trimming' be used to prevent this from happening. This means that at the uppermost node, the highest branch will be horizontal, the middle branch will point one node down, and the lowest branch will point two nodes down. This has the effect of limiting the value of j in equations (A5.1.10) to (A5.1.12), thereby ensuring that the probabilities remain within [0, 1]. These are illustrated in Figure 5.13.

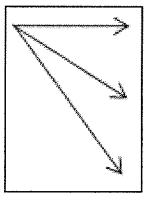
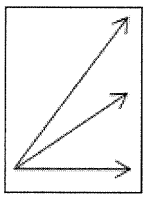
(a)  (b) 

Figure 5.13: Trimming of the tree at the (a) uppermost and (b) lowermost nodes respectively

A 5.1.2.14 The tree will then look like Figure 5.14.

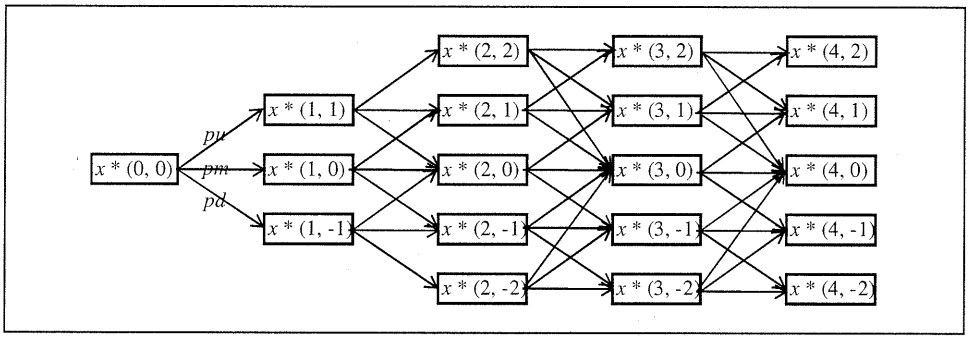


Figure 5.14: Stage 1 trimmed interest-rate tree.

A 5.1.2.15 Hull & White (*op. cit.*) recommend that trimming be done when j reaches the following upper bound:

$$j > j_{max} = \frac{0.184}{\kappa \Delta t},$$

and lower bound:

$$j < j_{min} = -\frac{0.184}{\kappa \Delta t}.$$

A 5.1.2.16 When trimming is needed, equations (A5.1.7) and (A5.1.8) need to be altered accordingly. If we are at the uppermost node, then tree branching as in Figure 5.13 (a) above applies. In this case:

$$E\left[x^*_{t+\Delta t} - x^*_t \middle| x^*_t = j\Delta x^*\right] = -p_m \Delta x^* - 2p_u \Delta x^* = -\kappa j \Delta x^* \Delta t,$$

$$E\left[\left(x^*_{t+\Delta t} - x^*_t\right)^2 \middle| x^*_t = j\Delta x^*\right] = p_m (\Delta x^*)^2 + 4p_d (\Delta x^*)^2 = \sigma^2 \Delta t + \kappa^2 j^2 (\Delta x^*)^2 \Delta t^2, \text{ and}$$

$$p_u + p_m + p_d = 1.$$

These equations solve to give the following probabilities:

$$p_u = \frac{7}{6} + \frac{1}{2}(\kappa^2 j^2 \Delta t^2 - 3\kappa j \Delta t),$$

$$p_m = -\frac{2}{3} - \kappa^2 j^2 \Delta t^2 + 2\kappa j \Delta t, \text{ and}$$

$$p_d = \frac{1}{6} + \frac{1}{2}(\kappa^2 j^2 \Delta t^2 - \kappa j \Delta t).$$

A 5.1.2.17 If we are at the lowermost node, then tree branching as in Figure 5.13 (b) above applies. In this case:

$$E\left[x^*_{t+\Delta t} - x^*_t \middle| x^*_t = j\Delta x^*\right] = 2p_u \Delta x^* + p_m \Delta x^* = -\kappa j \Delta x^* \Delta t,$$

$$E\left[\left(x^*_{t+\Delta t} - x^*_t\right)^2 \middle| x^*_t = j\Delta x^*\right] = 4p_u (\Delta x^*)^2 + p_m (\Delta x^*)^2 = \sigma^2 \Delta t + \kappa^2 j^2 (\Delta x^*)^2 \Delta t^2, \text{ and}$$

$$p_u + p_m + p_d = 1.$$

A 5.1.2.18 These equations solve to give the following probabilities:

$$p_u = \frac{1}{6} + \frac{1}{2}(\kappa^2 j^2 \Delta t^2 + \kappa j \Delta t),$$

$$p_m = -\frac{2}{3} - \kappa^2 j^2 \Delta t^2 - 2\kappa j \Delta t, \text{ and}$$

$$p_d = \frac{7}{6} + \frac{1}{2}(\kappa^2 j^2 \Delta t^2 + 3\kappa j \Delta t).$$

A 5.1.2.19 Note that every one of these probabilities relies only on j and that the probabilities are symmetrical about $j = 0$.

A 5.1.3 Building an Interest-rate Tree – the Second Stage

A 5.1.3.1 In the first stage, Hull & White (*op. cit.*) constructed a tree that is recombining but has zero drift. This tree is useful in determining the risk-neutral probabilities of an increase, a decrease and no change of rates in a trinomial tree. In the second stage, a drift term is added to the tree to ensure that the starting term structure of interest rates is replicated. They convert the tree of x_t^* into one of x_t by defining:

$$\alpha(t) = x_t - x_t^*. \qquad \text{(A5.1.13)}$$

A 5.1.3.2 The drift parameters α_t are calculated iteratively so that the initial term structure is matched exactly. Define $Q_{i,j}$ as the value at time 0 of a state security that pays off 1 if node (i, j) is reached and zero otherwise. It follows that:

$$Q_{0,0} = 1, \text{ and}$$

$$\alpha_0 = x_0 - x_0^* = x_0.$$

A 5.1.3.3 The subsequent values of $Q_{i,j}$ and $\alpha(t)$ can be calculated using forward induction so that the initial term structure is matched exactly. Figure 5.15 shows the new tree.

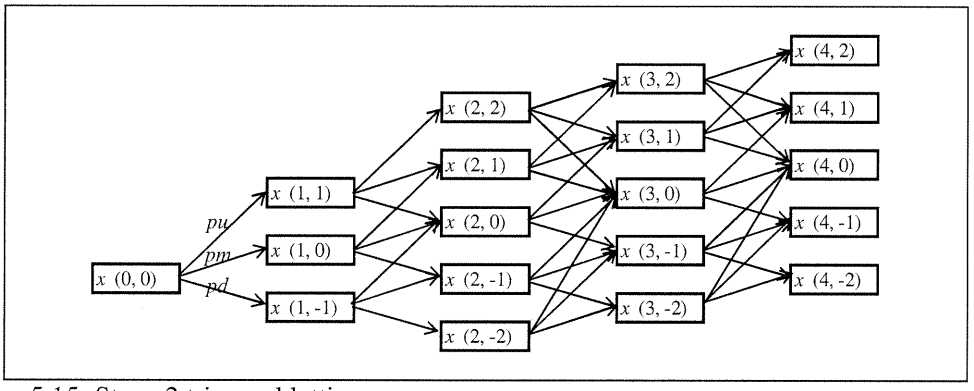


Figure 5.15: Stage 2 trimmed lattice

A 5.1.3.4 $Q_{1,1}$ is the price at time 0 of a state-security that pays off 1 if node (1,1) of Figure 5.15 is reached. Node (1,1) can be reached only from node (0,0) with an upward branch. $Q_{1,1}$ has the value of :

$$Q_{1,1} = Q_{0,0} p_{u,0} e^{-r(1,1)} = Q_{0,0} p_{u,0} e^{-\exp(x(1,1))} = Q_{0,0} p_{u,0} e^{-\exp\left(x^*(1,1)+\alpha(0)\right)}.$$

Similarly:

$$Q_{1,0} = Q_{0,0}p_{m,0}e^{-r(1,0)} = Q_{0,0}p_{m,0}e^{-\exp(x(1,0))} = Q_{0,0}p_{m,0}e^{-\exp\left(x^*(1,0)+\alpha(0)\right)}, \text{ and}$$
$$Q_{1,-1} = Q_{0,0}p_{d,0}e^{-r(1,-1)} = Q_{0,0}p_{d,0}e^{-\exp(x(1,-1))} = Q_{0,0}p_{d,0}e^{-\exp\left(x^*(1,-1)+\alpha(0)\right)}.$$

A 5.1.3.5 The values of all the variables on the right-hand side are known, so $Q_{1,1}$, $Q_{1,0}$ and $Q_{1,-1}$ can be calculated explicitly. We can check that $Q_{1,1}+Q_{1,0}+Q_{1,-1}=P(1)$, where $P(1)$ is the price of a one-year ZCB.

A 5.1.3.6 Similarly, the values of the second-year state-security prices are as follows:

$$Q_{2,2} = Q_{1,1}p_{u,0}e^{-r(2,2)} = Q_{1,1}p_{u,1}e^{-\exp(x(2,2))} = Q_{1,1}p_{u,1}e^{-\exp\left(x^*(2,2)+\alpha(1)\right)}; \quad \text{(A5.1.14)}$$

$$\begin{aligned} Q_{2,1} &= Q_{1,1}p_{m,1}e^{-r(2,1)} + Q_{1,0}p_{u,0}e^{-r(2,1)} \\ &= Q_{1,1}p_{m,1}e^{-\exp(x(2,1))} + Q_{1,0}p_{u,0}e^{-\exp(x(2,1))} \\ &= Q_{1,1}p_{m,1}e^{-\exp\left(x^*(2,1)+\alpha(1)\right)} + Q_{1,0}p_{u,0}e^{-\exp\left(x^*(2,1)+\alpha(1)\right)}; \end{aligned} \quad \text{(A5.1.15)}$$

$$\begin{aligned} Q_{2,0} &= Q_{1,1}p_{d,1}e^{-r(2,0)} + Q_{1,0}p_{m,0}e^{-r(2,0)} + Q_{1,-1}p_{u,0}e^{-r(2,0)} \\ &= Q_{1,1}p_{d,1}e^{-\exp(x(2,0))} + Q_{1,0}p_{m,0}e^{-\exp(x(2,0))} + Q_{1,-1}p_{u,0}e^{-\exp(x(2,0))} \\ &= Q_{1,1}p_{d,1}e^{-\exp\left(x^*(2,0)+\alpha(1)\right)} + Q_{1,0}p_{m,0}e^{-\exp\left(x^*(2,0)+\alpha(1)\right)} + Q_{1,-1}p_{u,0}e^{-\exp\left(x^*(2,0)+\alpha(1)\right)}; \end{aligned} \quad \text{(A5.1.16)}$$

$$\begin{aligned} Q_{2,-1} &= Q_{1,-1}p_{m,-1}e^{-r(2,-1)} + Q_{1,0}p_{d,0}e^{-r(2,-1)} \\ &= Q_{1,-1}p_{m,-1}e^{-\exp(x(2,-1))} + Q_{1,0}p_{d,0}e^{-\exp(x(2,-1))} \\ &= Q_{1,-1}p_{m,-1}e^{-\exp\left(x^*(2,-1)+\alpha(1)\right)} + Q_{1,0}p_{d,0}e^{-\exp\left(x^*(2,-1)+\alpha(1)\right)}; \text{ and} \end{aligned} \quad \text{(A5.1.17)}$$

$$Q_{2,-2} = Q_{1,-1}p_{d,-1}e^{-r(2,-2)} = Q_{1,-1}p_{d,-1}e^{-\exp(x(2,-2))} = Q_{1,-1}p_{d,-1}e^{-\exp\left(x^*(2,-2)+\alpha(1)\right)}. \quad \text{(A5.1.18)}$$

A 5.1.3.7 We do not know the value of $\alpha(1)$, but it can be solved for iteratively by using the equation $\sum_{i=-2}^{2} Q_{2,i} = P(2)$ and substituting equations (A5.1.14) to (A5.1.18) for $Q_{2,i}$. Similarly, subsequent values of $Q_{i,j}$ and $\alpha(t)$ can be calculated using forward induction. In general, the following equation holds for the price $P(m)$:

$$P(m) = \sum_{j=-n_m}^{n_m} Q_{m,j}, \quad \text{(A5.1.19)}$$

where:

- $g(x) = f^{-1}(x)$; and
- f is defined as is in (A5.1.2).

A 5.1.3.8 The following equation holds for the state price $Q_{m,j}$:

$$Q_{m+1,j} = \sum_{k} Q_{m,k}\, p_{k,j} e^{-g(\alpha(m)+k\Delta x)\Delta t}. \quad \text{(A5.1.20)}$$

where:

- $\alpha(m)$ is solved for numerically;
- the summation is across all nodes from which $(m+1, j)$ may be accessed; and
- $p_{k,j}$ is the probability of moving from (m,k) to $(m+1, j)$.

A 5.1.3.9 Equations (A5.1.19) and (A5.1.20) may be used to solve for the values of α_t. Care must be taken in the construction of $Q_{i,j}$ where the end nodes have been trimmed. At the uppermost node, the following equation applies:

$$Q_{m+1,j} = Q_{m,j-2}p_{j-2,j}e^{-g(\alpha(m)+(j-2)\Delta x)\Delta t} + Q_{m,j-1}p_{j-1,j}e^{-g(\alpha(m)+(j-1)\Delta x)\Delta t} + Q_{m,j}p_{j,j}e^{-g(\alpha(m)+j\Delta x)\Delta t}.$$

A 5.1.3.10 At the non-extreme nodes, the following equation applies:

$$Q_{m+1,j} = Q_{m,j-1}p_{j-1,j}e^{-g(\alpha(m)+(j-1)\Delta x)\Delta t} + Q_{m,j}p_{j,j}e^{-g(\alpha(m)+j\Delta x)\Delta t} + Q_{m,j+1}p_{j+1,j}e^{-g(\alpha(m)+(j+1)\Delta x)\Delta t}.$$

A 5.1.3.11 At the lower-most nodes, the following equation applies:

$$Q_{m+1,j} = Q_{m,j}p_{j,j}e^{-g(\alpha(m)+j\Delta x)\Delta t} + Q_{m,j+1}p_{j+1,j}e^{-g(\alpha(m)+(j+1)\Delta x)\Delta t} + Q_{m,j+2}p_{j+2,j}e^{-g(\alpha(m)+(j+2)\Delta x)\Delta t}.$$

Appendix 5.2 – Model Tests

A 5.2.1 Test of Model Fit to the Initial Term Structure Using Monte-Carlo Method

A 5.2.1.1 This section compares the term-structure calibration target to the simulated results. The latter are calculated using the method specified in Section 5.3.10.2 (i.e. generating 2000 simulations of the short rate over 30 years).

Table 5.1 Comparison of market and simulated bond prices and yield curves

T	Market ZCB price	Simulated ZCB price	Input spot yield	Simulated spot yield
0	1.0000	1.0000		
1	0.9013	0.9013	10.40%	10.40%
2	0.8251	0.8252	9.61%	9.61%
3	0.7585	0.7586	9.21%	9.21%
4	0.6984	0.6986	8.98%	8.97%
5	0.6438	0.6442	8.81%	8.79%
6	0.5945	0.5951	8.67%	8.65%
7	0.5498	0.5504	8.55%	8.53%
8	0.5084	0.5091	8.46%	8.44%
9	0.4698	0.4705	8.39%	8.38%
10	0.4339	0.4346	8.35%	8.33%
11	0.4006	0.4012	8.32%	8.30%
12	0.3702	0.3706	8.28%	8.27%
13	0.3423	0.3424	8.25%	8.24%
14	0.3167	0.3166	8.21%	8.22%
15	0.2931	0.2928	8.18%	8.19%
16	0.2714	0.2709	8.15%	8.16%
17	0.2512	0.2506	8.13%	8.14%
18	0.2326	0.2318	8.10%	8.12%
19	0.2154	0.2144	8.08%	8.10%
20	0.1994	0.1984	8.06%	8.09%
21	0.1847	0.1835	8.04%	8.07%
22	0.1710	0.1697	8.03%	8.06%
23	0.1583	0.1569	8.02%	8.05%
24	0.1465	0.1451	8.00%	8.04%
25	0.1356	0.1341	7.99%	8.04%
26	0.1255	0.1240	7.98%	8.03%
27	0.1162	0.1146	7.97%	8.02%
28	0.1075	0.1059	7.96%	8.02%
29	0.0995	0.0979	7.96%	8.01%
30	0.0921	0.0905	7.95%	8.01%

A 5.2.1.2 The simulated prices of ZCB are similar to their input prices. Similarly, the simulated yields are similar to the input yields. This demonstrates that the model is parameterised appropriately to the initial term structure.

A 5.2.2 Test of Model Fit to the Initial Term Structure Using Future Yield Curves

A 5.2.2.1 Table 5.2 and Table 5.3 compare the term-structure calibration target to the simulated results. The simulated results are calculated using the method specified in Section 5.3.10. Two thousand simulations of the short rate over t years are generated and a 30-year term structure is constructed at the end of the t years for each of the 2000 simulations. For each path, the prices of ZCBs are calculated using these term structures and then discounted at the short rate back to time 0 from time t. These discounted values are averaged across the 2000 simulations to calculate the simulated ZCB prices. Figure 5.16 compares the term structure using ZCB prices. Figure 5.17 compares the term structure using spot rates.

Table 5.2: Comparison of simulated and actual prices of ZCBs

T	Input ZCB price	Simulated ZCB price (tree)		
		t=5	t=15	t=25
0	1.0000			
1	0.9013			
2	0.8251			
3	0.7585			
4	0.6984			
5	0.6438	0.6444		
6	0.5945	0.5952		
7	0.5498	0.5505		
8	0.5084	0.5091		
9	0.4698	0.4705		
10	0.4339	0.4345		
11	0.4006	0.4010		
12	0.3702	0.3704		
13	0.3423	0.3424		
14	0.3167	0.3166		
15	0.2931	0.2929	0.2927	
16	0.2714	0.2710	0.2707	
17	0.2512	0.2507	0.2504	
18	0.2326	0.2320	0.2315	
19	0.2154	0.2146	0.2140	
20	0.1994	0.1985	0.1979	
21	0.1847	0.1837	0.1829	
22	0.1710	0.1699	0.1690	
23	0.1583	0.1571	0.1561	
24	0.1465	0.1453	0.1442	
25	0.1356	0.1343	0.1332	0.1332
26	0.1255	0.1242	0.1230	0.1230
27	0.1162	0.1149	0.1136	0.1136

T	Input ZCB price	Simulated ZCB price (tree)		
		t=5	t=15	t=25
28	0.1075	0.1062	0.1049	0.1048
29	0.0995	0.0982	0.0968	0.0967
30	0.0921	0.0908	0.0894	0.0893
31	0.0853	0.0839	0.0825	0.0824
32	0.0789	0.0776	0.0762	0.0761
33	0.0731	0.0717	0.0703	0.0702
34	0.0676	0.0663	0.0649	0.0648
35	0.0626	0.0613	0.0599	0.0598
36	0.0579		0.0553	0.0552
37	0.0536		0.0511	0.0510
38	0.0496		0.0472	0.0470
39	0.0459		0.0435	0.0434
40	0.0425		0.0402	0.0401
41	0.0394		0.0371	0.0370
42	0.0364		0.0342	0.0342
43	0.0337		0.0316	0.0315
44	0.0312		0.0292	0.0291
45	0.0289		0.0269	0.0269
46	0.0267			0.0248
47	0.0247			0.0229
48	0.0229			0.0212
49	0.0212			0.0196
50	0.0196			0.0181
51	0.0182			0.0167
52	0.0168			0.0154
53	0.0156			0.0142
54	0.0144			0.0132
55	0.0133			0.0122

Table 5.3: Comparison of simulated and actual spot rates

T	Input spot rate	Simulated spot rates (tree)		
		t=5	t=15	t=25
1	10.40%			
2	9.61%			
3	9.21%			
4	8.98%			
5	8.81%	8.79%		
6	8.67%	8.65%		
7	8.55%	8.53%		
8	8.46%	8.44%		
9	8.39%	8.38%		
10	8.35%	8.34%		
11	8.32%	8.31%		
12	8.28%	8.28%		
13	8.25%	8.24%		
14	8.21%	8.21%		
15	8.18%	8.19%	8.19%	
16	8.15%	8.16%	8.17%	
17	8.13%	8.14%	8.15%	
18	8.10%	8.12%	8.13%	
19	8.08%	8.10%	8.11%	
20	8.06%	8.08%	8.10%	
21	8.04%	8.07%	8.09%	
22	8.03%	8.06%	8.08%	
23	8.02%	8.05%	8.07%	
24	8.00%	8.04%	8.07%	
25	7.99%	8.03%	8.06%	8.06%
26	7.98%	8.02%	8.06%	8.06%
27	7.97%	8.01%	8.06%	8.06%
28	7.96%	8.01%	8.05%	8.06%

T	Input spot rate	Simulated spot rates (tree)		
		t=5	t=15	t=25
29	7.96%	8.00%	8.05%	8.05%
30	7.95%	8.00%	8.05%	8.05%
31	7.94%	7.99%	8.05%	8.05%
32	7.94%	7.99%	8.04%	8.05%
33	7.93%	7.98%	8.04%	8.05%
34	7.92%	7.98%	8.04%	8.05%
35	7.92%	7.98%	8.04%	8.05%
36	7.91%		8.04%	8.05%
37	7.91%		8.04%	8.05%
38	7.90%		8.04%	8.04%
39	7.90%		8.04%	8.04%
40	7.89%		8.04%	8.04%
41	7.89%		8.03%	8.04%
42	7.89%		8.03%	8.04%
43	7.88%		8.03%	8.04%
44	7.88%		8.03%	8.04%
45	7.88%		8.03%	8.04%
46	7.87%			8.03%
47	7.87%			8.03%
48	7.87%			8.03%
49	7.86%			8.03%
50	7.86%			8.03%
51	7.86%			8.03%
52	7.86%			8.02%
53	7.85%			8.02%
54	7.85%			8.02%
55	7.85%			8.02%

A 5.2.2.2 The simulated zero-coupon-bond prices and yields are similar to the input yields for all three choices of t. This demonstrates that the model projects future term structures of interest rates consistently with the initial term structure. Figure 5.16 to Figure 5.21 demonstrate this point.

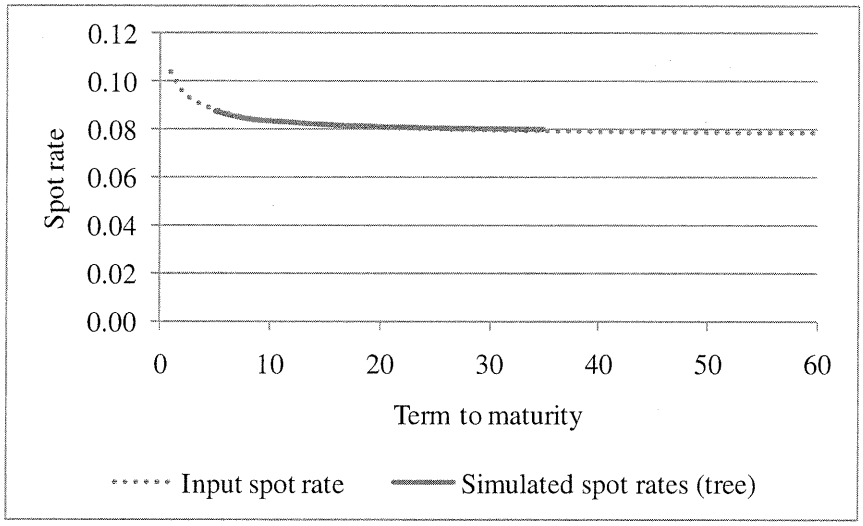


Figure 5.16: Comparison of simulated ZCB prices (t=5) to the calibration targets

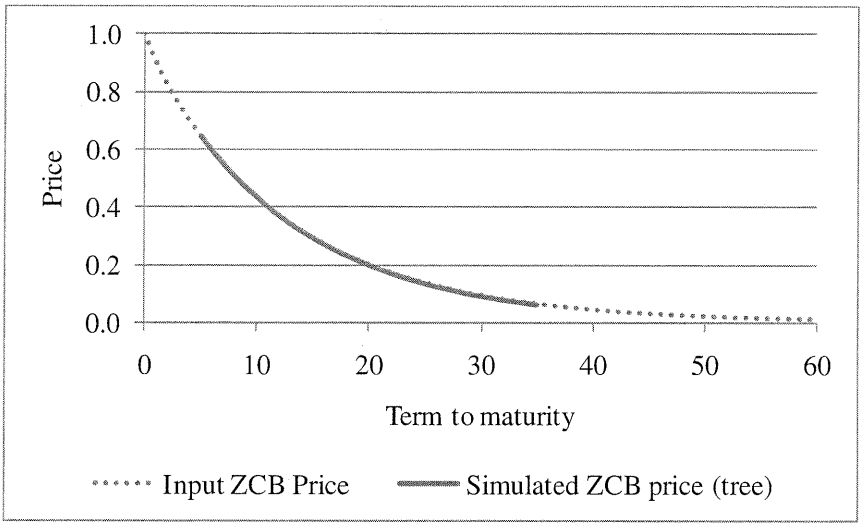


Figure 5.17: Comparison of simulated spot yields (t=5) to the calibration targets

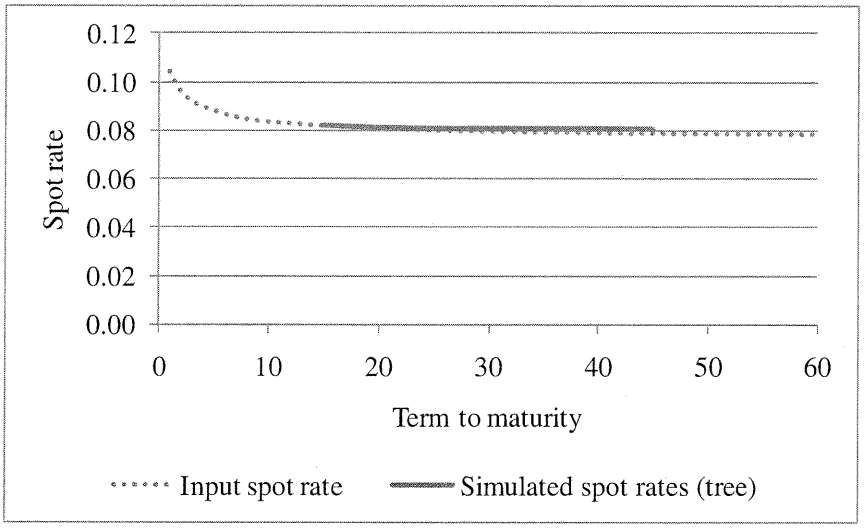


Figure 5.18: Comparison of simulated ZCB prices (t=15) to the calibration targets

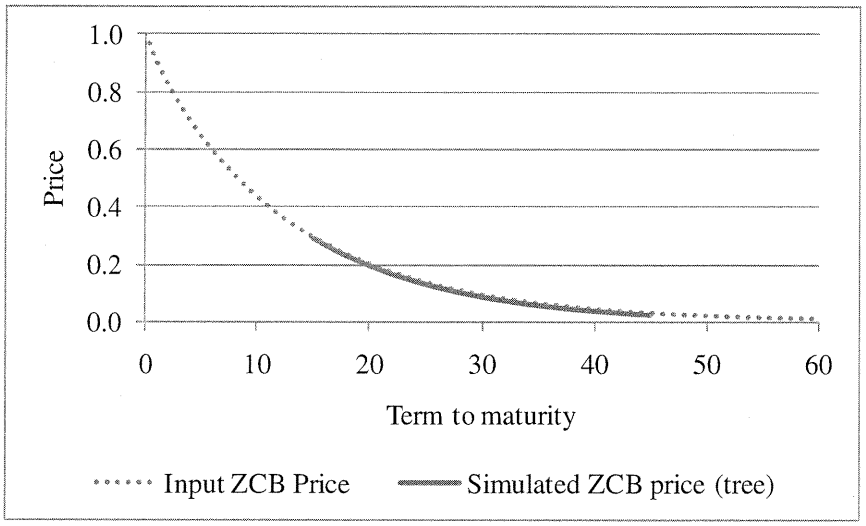


Figure 5.19: Comparison of simulated spot yields (t=15) to the calibration targets

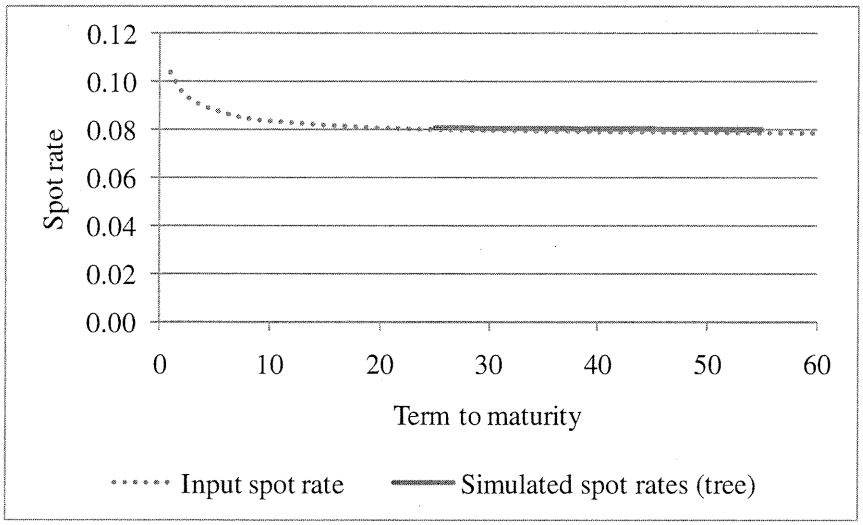


Figure 5.20: Comparison of simulated ZCB prices (t=25) to the calibration targets

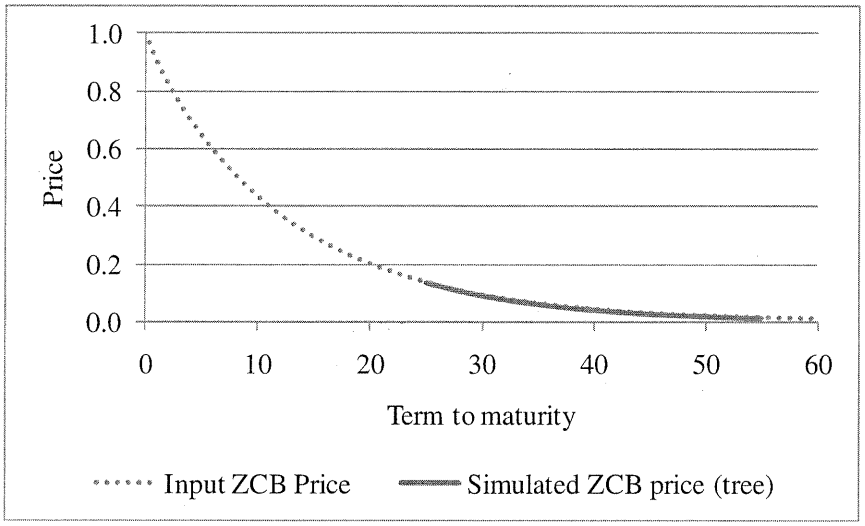


Figure 5.21: Comparison of simulated spot yields (t=25) to the calibration targets

Appendix 5.3 – Calibration Data

This section contains the market information that forms the calibrating targets in this chapter.

Term Structure of Interest Rates

Term (years)	Annual spot yield
1	10.40%
2	9.61%
3	9.21%
4	8.98%
5	8.81%
6	8.67%
7	8.55%
8	8.46%
9	8.39%
10	8.35%
11	8.32%

Term (years)	Annual spot yield
12	8.28%
13	8.25%
14	8.21%
15	8.18%
16	8.15%
17	8.13%
18	8.10%
19	8.08%
20	8.06%
21	8.04%
22	8.03%

Term (years)	Annual spot yield
23	8.02%
24	8.00%
25	7.99%
26	7.98%
27	7.97%
28	7.96%
29	7.96%
30	7.95%

Implied Volatilities of Equities Options

Term	Implied volatility
1	25.7%
2	24.7%
3	24.4%

Implied Volatilities of Swaptions

Implied volatilities		Tenor (years)					
		1	5	10	15	20	25
Option term (years)	1	13.4%	14.6%	14.3%	14.3%	14.3%	13.7%
	2	13.5%	14.5%	14.3%	13.7%	13.4%	13.3%
	3	13.5%	14.6%	14.2%	13.1%	13.0%	12.9%
	4	13.5%	14.4%	14.1%	12.8%	12.6%	12.6%
	5	13.5%	14.5%	14.3%	12.5%	12.4%	12.3%
	10	13.6%	14.5%	14.3%	11.9%	11.8%	11.7%
	15	12.2%	11.8%	11.6%	11.5%	11.4%	11.4%
	20	11.8%	11.5%	11.2%	11.1%	11.1%	11.0%
	25	11.6%	11.2%	11.0%	10.9%	10.9%	10.8%
	30	11.5%	11.0%	10.8%	10.7%	10.7%	10.6%

The shaded column represents the tenor that was targeted in the calibration.

Appendix 5.4 – Antithetic Variables

A 5.4.1 Let $x_1, x_2, x_3, ..., x_n$ be n independently and identically distributed normal random variables with expectation 0 and variance σ_X^2. Define $y_i = f(x_i)$ so that $y_1, y_2, y_3, ..., y_n$ are also independently and identically distributed random variables, each having finite values of expectation μ_Y and variance σ_Y^2. Define $\bar{y}$ as the sample mean of these n random observations from y_i as follows:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}.$$

A 5.4.2 The sample mean $\bar{y}$ is an unbiased estimator of the population mean $\bar{Y}$. From the central limit theorem,

$$\bar{Y} \sim N\left(\mu_Y, \frac{\sigma_Y^2}{n}\right). \tag{5.21}$$

A 5.4.3 Equation (5.21) shows that when the simulation size n doubles, the standard deviation of $\bar{y}$ decreases only by a factor of $\sqrt{2}$. This is computationally very inefficient.

A 5.4.4 The idea of antithetic variables was first proposed by Hammersley & Morton (1956). This method is implemented by simulating scenarios in pairs so that the second trial uses the negative of the first trial's randomly generated number.

A 5.4.5 Consider another sample of $\frac{n}{2}$ simulations of x_i^a. Consider another series of $\frac{n}{2}$ values of x_i^b, where $x_i^b = -x_i^a$. This way, we have $\frac{n}{2}$ pairs of values x_i^a and x_i^b. Denote a pair of antithetic variables $y_i^a = f(x_i^a)$ and $y_i^b = f(x_i^b) = f(-x_i^a)$. Since x_i^a is from a normal distribution with an expectation of 0, it is identically (but not independently) distributed as x_i^b. This means that y_i^a and y_i^a are also identically (but not independently) distributed with mean μ_Y and variance σ_Y^2. We define $\overline{y^m}$ as the sample mean of these n observations. This is calculated as follows:

$$\overline{y^m} = \frac{2}{n}\sum_{i=1}^{n/2}\left(\frac{y_i^a + y_i^b}{2}\right) = \frac{2}{n}\sum_{i=1}^{n/2} y_i^m, \tag{5.22}$$

where $y_i^m = \frac{y_i^a + y_i^b}{2}$.

A 5.4.6 We note that $\overline{y^m}$ is still an unbiased estimator of $\bar{Y}$:

$$E\left[\overline{Y^m}\right] = E\left[\frac{2}{n}\sum_{i=1}^{n/2}\left(\frac{Y_i^a + Y_i^b}{2}\right)\right] = \mu_Y.$$

A 5.4.7 We further note that y_i^m is identically and independently distributed. Because of this independence, an unbiased estimator of the variance of Y as:

$$\hat{\sigma}_{Y^m}^2 = \frac{2}{n-2}\left[\sum_{i=1}^{n/2}\left(y_i^m\right)^2 - \frac{\left(\sum_{i=1}^{n/2}\left(y_i^m\right)\right)^2}{\frac{n}{2}}\right]. \tag{5.23}$$

A 5.4.8 From (5.22) we also have:

$$\overline{y^m} = \frac{2}{n}\sum_{i=1}^{n/2} y_i^m . \tag{5.24}$$

A 5.4.9 It follows that the following variance is an unbiased estimator of the variance of $\overline{Y^m}$:

$$\text{var}\left[\overline{y^m}\right] = \text{var}\left[\frac{2}{n}\sum_{i=1}^{n/2} y_i^m\right] = \left(\frac{2}{n}\right)^2 \text{var}\left[\sum_{i=1}^{n/2} y_i^m\right] = \left(\frac{2}{n}\right)^2 \sum_{i=1}^{n/2} \text{var}\left[y_i^m\right] = \frac{2}{n}\text{var}\left[y_i^m\right],$$

so that:

$$\hat{\sigma}_{\overline{Y^m}}^2 = \frac{\hat{\sigma}_{Y^m}^2}{n/2}. \tag{5.25}$$

Appendix 5.5 – ESG Parameters

This section lists the variables used in the ESG and their values in the calibration as at the end of December 2007.

A 5.5.1 Interest-rate Parameters for the South African Swap Curve and Implied Volatility on Swaptions

A 5.5.1.1 The parameters for the Black–Karasinski model used for the term structure of interest rate are as follows:

- $a = 0.008$,
- $\sigma = 0.135$,
- $\Delta t = 1$,
- $\Delta x = \sigma(3dt)^{\frac{1}{2}} = \sqrt{3}\sigma = 0.2338$,
- $j_{\max} = \left\lceil \frac{0.184}{a\Delta t} \right\rceil = 23$, where $\lceil x \rceil$ is the smallest integer not less than x, and
- $j_{min} = -j_{max} = -23$.

A 5.5.1.2 The time-dependent parameters $a(t)$ and $\theta(t)$ are listed in Table 5.4. The values of x^* in the first stage interest-rate lattice, and their associated probabilities of upward, middle and downward movements p_u, p_m and p_d are listed in Table 5.5.

Table 5.4: Time dependent parameters $a(t)$ and $\theta(t)$ used in the modelling of the term structure of interest rate

t	$a(t)$	$\theta(t)$	$a(t)$	$\theta(t)$	$a(t)$	$\theta(t)$	$a(t)$	$\theta(t)$
0	-2.2638	-0.0169	-2.5256	-0.0323	-2.2522	0.0047	-1.9024	0.0073
1	-2.4357	-0.0173	-2.5157	-0.0336	-2.2316	0.0049	-1.8799	0.0075
2	-2.4895	-0.0184	-2.5016	-0.0342	-2.2107	0.0050	-1.8574	0.0077
3	-2.5140	-0.0192	-2.4909	-0.0355	-2.1895	0.0051	-1.8349	0.0078
4	-2.5331	-0.0200	-2.4735	-0.0363	-2.1682	0.0052	-1.8123	0.0080
5	-2.5548	-0.0215	-2.4570	-0.0377	-2.1466	0.0054	-1.7899	0.0081
6	-2.5738	-0.0220	-2.4421	-0.0385	-2.1249	0.0054	-1.7674	0.0083
7	-2.5742	-0.0234	-2.4230	-0.0392	-2.1031	0.0055	-1.7450	0.0084
8	-2.5621	-0.0241	-2.4058	-0.0405	-2.0811	0.0056	-1.7226	0.0086
9	-2.5527	-0.0252	-2.3911	0.0411	-2.0590	0.0057	-1.7002	0.0087
10	-2.5425	-0.0260	-2.3688	0.0420	-2.0368	0.0058	-1.6779	0.0088
11	-2.5461	-0.0274	-2.3489	-0.0438	-2.0146	0.0059	-1.6557	0.0089
12	-2.5492	-0.0284	-2.3319	0.0448	-1.9922	0.0060	-1.6335	
13	-2.5482	-0.0291	-2.3125	0.0453	-1.9698	0.0067		
14	-2.5419	-0.0301	-2.2927	0.0468	-1.9474	0.0069		
15	-2.5377	-0.0312	-2.2726	0.0472	-1.9249	0.0071		

Table 5.5: Values of x^* for each j and their associated up, middle and down probabilities

j	x^*	p_u	p_m	p_d
-23	-5.37802	0.09159	0.00081	0.90759
-22	-5.14419	0.27015	0.63569	0.09415
-21	-4.91036	0.26478	0.63844	0.09678
-20	-4.67654	0.25947	0.64107	0.09947
-19	-4.44271	0.25422	0.64356	0.10222
-18	-4.20888	0.24903	0.64593	0.10503
-17	-3.97506	0.24391	0.64817	0.10791
-16	-3.74123	0.23886	0.65028	0.11086
-15	-3.50740	0.23387	0.65227	0.11387
-14	-3.27358	0.22894	0.65412	0.11694
-13	-3.03975	0.22407	0.65585	0.12007
-12	-2.80592	0.21927	0.65745	0.12327
-11	-2.57210	0.21454	0.65892	0.12654
-10	-2.33827	0.20987	0.66027	0.12987
-9	-2.10444	0.20526	0.66148	0.13326
-8	-1.87061	0.20071	0.66257	0.13671
-7	-1.63679	0.19623	0.66353	0.14023
-6	-1.40296	0.19182	0.66436	0.14382
-5	-1.16913	0.18747	0.66507	0.14747
-4	-0.93531	0.18318	0.66564	0.15118
-3	-0.70148	0.17895	0.66609	0.15495
-2	-0.46765	0.17479	0.66641	0.15879
-1	-0.23383	0.17070	0.66660	0.16270
0	0.00000	0.16667	0.66667	0.16667

j	x^*	p_u	p_m	p_d
1	0.23383	0.16270	0.66660	0.17070
2	0.46765	0.15879	0.66641	0.17479
3	0.70148	0.15495	0.66609	0.17895
4	0.93531	0.15118	0.66564	0.18318
5	1.16913	0.14747	0.66507	0.18747
6	1.40296	0.14382	0.66436	0.19182
7	1.63679	0.14023	0.66353	0.19623
8	1.87061	0.13671	0.66257	0.20071
9	2.10444	0.13326	0.66148	0.20526
10	2.33827	0.12987	0.66027	0.20987
11	2.57210	0.12654	0.65892	0.21454
12	2.80592	0.12327	0.65745	0.21927
13	3.03975	0.12007	0.65585	0.22407
14	3.27358	0.11694	0.65412	0.22894
15	3.50740	0.11387	0.65227	0.23387
16	3.74123	0.11086	0.65028	0.23886
17	3.97506	0.10791	0.64817	0.24391
18	4.20888	0.10503	0.64593	0.24903
19	4.44271	0.10222	0.64356	0.25422
20	4.67654	0.09947	0.64107	0.25947
21	4.91036	0.09678	0.63844	0.26478
22	5.14419	0.09415	0.63569	0.27015
23	5.37802	0.90759	0.00081	0.09159

A 5.5.2 Equity Parameters for the South African Implied Volatilities for Equity Options (31 December 2007)

A 5.5.2.1 The systematic-risk parameters for the equity model are as follows:

- $\beta = 0.6$,
- $\sigma = 0.2$,
- $\beta^*_{sys} = \beta\sigma = 0.12$.

A 5.5.2.2 The specific-risk parameters for the equity model are as follows:

- $\alpha = 0.006236$,
- $\sigma_0 = 0.22$,
- $\sigma_\infty = 0.06715$.

A 5.5.2.3 The term structure of implied volatilities suggested by these parameters is presented in Table 5.6.

Table 5.6: The term structure of implied volatilities suggested by the model

t	Model-implied volatility
1	25.02%
2	24.98%
3	24.94%
4	24.89%
5	24.85%
6	24.81%
7	24.77%
8	24.73%
9	24.69%
10	24.65%
11	24.61%
12	24.57%
13	24.53%
14	24.50%
15	24.46%

t	Model-implied volatility
16	24.42%
17	24.38%
18	24.34%
19	24.30%
20	24.27%
21	24.23%
22	24.19%
23	24.15%
24	24.12%
25	24.08%
26	24.04%
27	24.01%
28	23.97%
29	23.94%
30	23.90%

Chapter 6. Valuation of Guaranteed Annuity Options

6.1 Introduction

This chapter describes the market-consistent valuation of a hypothetical portfolio of GAOs. The hypothetical portfolio and the modelling assumptions are described in Section 6.2. The model is described in Section 6.3. The results are presented and interpreted in Section 6.4.

6.2 The Hypothetical Portfolio and Modelling Assumptions

6.2.1 Assume that we have a simple GAO portfolio consisting of identical policies but issued to many different lives, so that the portfolio is well diversified. The policyholders and the policies have the following characteristics:

- The expected retirement age (age at conversion) is 65.
- The policyholder is currently 40 years of age.
- The policyholder is male.
- The unit fund has current value of R100,000.
- The guarantee on the fund is currently at the money (i.e. guaranteed value is equal to the fund value of R100,000).
- The guaranteed return on the unit fund is 5%.
- The future recurring premium is R5000 a year, payable in advance.
- The future premiums are allocated 100% to the unit fund.
- The annuity after the conversion is payable annually in arrears for life, subject to a maximum of 30 years.
- The guaranteed annuity-conversion factor is 9.

6.2.2 The annuities are limited to a maximum of 30 years because the nested term-structure model is limited to 30 years. This limitation is used because of the computational limitations imposed by the calibration of the model, in which the number of nodes increases exponentially as the modelling term increases. We shall see below, in Chapter 9, that this assumption does not make a material difference. Without any loss of generality, only one such policy is modelled in this research.

6.2.3 The current fund value and the future premiums will be invested in a unit fund. At maturity in 25 years' time, the unit-fund value is guaranteed to be at least:

$$K_{25} = 100\ 000 \times 1.05^{25} + 5000 \sum_{i=1}^{25} 1.05^{i} = 589\ 203.$$

6.2.4 This is the guaranteed maturity value (GMV). Depending on the value of the fund at that time, S_{25}, this represents a guarantee cost to the insurer of:

$$F_{25} = \max\left(K_{25} - S_{25}, 0\right).$$

6.2.5 Furthermore, the guaranteed-annuity-conversion factor is 9. The policyholder will be guaranteed an annual payment of at least:

$$\frac{F_{25}}{9}.$$

6.2.6 This represents a cost to the insurer of:

$$\frac{F_{25}}{9}\max\left(a_{65:\overline{30}|}-9,0\right),$$

where $a_{65:\overline{30}|}$ is the actual annuity factor for an annual annuity payable to a life aged 65 until his death or 30 years, whichever is the earlier. The factor $a_{65:\overline{30}|}$ should be calculated using the interest rates applicable at the time of retirement.

6.2.7 The portfolio underlying the unit fund (to which the GMV applies) is a typical balanced portfolio consisting of cash, equities and ZCBs. The asset split is given in Table 6.1.

Table 6.1: Unit-fund asset allocation

Asset type	Composition
Cash	5%
Equities	60%
Fixed income	35%
Total	100%

6.2.8 The fixed-income assets are invested in ZCBs of 5 years, 10 years, 15 years and 20 years to maturity. Each of these maturities has an equal allocation of 25% of the fixed-income market value. The portfolio is rebalanced to these allocations at the end of each year.

6.2.9 We make other assumptions as follows:

- There is no tax payable on the investment income, gains or losses.
- There is no explicit deduction for asset management fees.
- Mortality follows SAIML98 after vesting date and SA (85–90) before vesting.
- Lapse rate is 3% each year before the vesting date.
- The investment-guarantee option will always be exercised if it is in the money at conversion date.
- The annuity-conversion option will always be exercised if it is in the money at conversion date.

6.3 The Model

6.3.1 We are interested in the value of the guarantees that are provided in the contract. We first calculate the value of the GMV and then that of the GAO.

6.3.2 The market-consistent value of the GMV benefit is calculated by projecting the unit-fund value and the guaranteed value. The shortfall, where the fund value is less than the guaranteed value, is discounted back to the current date using the simulated short rate, allowing for decrements from deaths and lapses until the vesting date. This is repeated for each simulation s and the results are averaged. That is:

$$V_U = \frac{1}{N}\,{}_{25}p^{*}_{40}\sum_{s=1}^{N}\left[\max\left(K_{25}-S_{s,25},0\right)\prod_{i=0}^{24}e^{-r_{s,i}}\right], \tag{6.1}$$

where:

- V_U is the market-consistent value for the GMV benefit.

- $S_{s,25} = 100000\prod_{t=0}^{25-1} e^{R_{s,t}} + 5000\sum_{i=0}^{25-1}\prod_{j=i}^{25-1} e^{R_{s,t}}$ is the simulated value of the unit fund at retirement in 25 years' time, for simulation s.
- $r_{s,i}$ is the short rate at time i for simulation s.
- $R_{s,i}$ is the unit-fund portfolio return in simulation s in period $[i, i+1)$.
- ${}_{25}p^{*}_{40}$ is the probability that the contract is in force in 25 years' time given that the policyholder is currently aged 40. The policy is subject to both mortality and lapse decrements.
- N is the number of simulations.

6.3.3 The value of the guaranteed annuity option is dependent on the value of the fund and the level and shape of the yield curve at the vesting date, and the expected mortality after vesting. This is:

$$V_A = \frac{1}{N}\,{}_{25}p^{*}_{40}\sum_{s=1}^{N}\left[\frac{S_{s,25}}{9}\max\left(\ddot{a}^{s,25}_{65:\overline{30}|} - 9, 0\right)\prod_{i=0}^{T-1} e^{-r_{s,i}}\right]. \tag{6.2}$$

where:

- $\ddot{a}^{s,25}_{65:\overline{30}|} = \sum_{t=0}^{30} {}_{t}p_{65}P_{s,25,t}$ is the value under simulation s, at time 25, of the annuity benefit payable to a life who will be aged 65 then;
- ${}_{t}p_{65}$ is the probability that a life aged 65 survives for t years, given that mortality is the only decrement; and
- $P_{s,25,t}$ is the value of a t-year ZCB at time 25 in simulation s.

6.3.4 Equation (6.2) does not recognise that if the simulated fund value at time 25, $S_{s,25}$, is less than the guaranteed maturity value K_{25}, the fund value to which the GAO applies will be increased by the GMV benefit. This causes the GAO cost for that particular scenario to increase proportionately. This is a secondary cost (referred to as the 'cross term' below). To capture this cost, the following equation is used:

$$V_C = \frac{1}{N}\,{}_{25}p^{*}_{40}\sum_{s=1}^{N}\left[\frac{\max(K_{25} - S_{s,25}, 0)}{9}\max\left(\ddot{a}^{s,25}_{65:\overline{30}|} - 9, 0\right)\prod_{i=0}^{T-1} e^{-r_{s,i}}\right]. \tag{6.3}$$

6.3.5 The total value of market-consistent liabilities for all of the guarantees is:

$$V = V_U + V_A + V_C. \tag{6.4}$$

6.4 Results and Interpretation

6.4.1 We calculate the market-consistent value of the policy with the model and the assumptions described above. The simulations are calibrated to the South African market conditions as at 31 December 2007 (as described in Chapter 5). The results are in Table 6.2.

Table 6.2: Market-consistent value of the policy liabilities

Component	**Value (R)**
GMV reserve (V_U)	2,785
GAO reserve (V_A)	1,406
Cross-term reserve (V_C)	157
Total	4,349

6.4.2 The GMV reserve and GAO reserve are the market-consistent reserves that need to be held in respect of the GMV and GAO benefits respectively. The cross-term reserve is held in respect of the secondary-order cost as identified in ¶6.3.4. These reserves, at over 4% of the current unit value, are not immaterial. It should also be noted that these guarantees are currently out the money — the intrinsic value of the guarantees is zero. Therefore, this reserve represents only the time value of the options. The cost of the options may elevate to much higher levels if the options move into the money. This was the case of Equitable Life and other life insurers in the UK. This high cost was especially burdensome given that no charges were imposed on the policyholders for a guarantee that was deemed worthless at the point of sale.

Chapter 7. Risk Measures

7.1 Introduction

In Section 2.3, we argued that risk management is intimately related to economic capital management. In Section 7.2 of this chapter, the differences between the two main methods that are used for calculating economic capital are discussed. The full implementation of economic capital is computationally intensive. A simplified implementation of one of these methods, utilising shocks that are calibrated to specific confidence levels, is proposed. This simplification requires calibration of a series of shocks. Section 7.3 discusses the calibration of these shocks.

7.2 Two Different Methods of Determining Economic Capital

7.2.1 In practice, there is no standard method to calculate economic capital. There are two different approaches used by typical North American and European companies. Turnbull[17] outlines the difference:

> *"The first defines capital as the amount required to fund all future liability cashflows from existing business as they fall due, at some specified level of confidence. This is perhaps the most natural probabilistic implementation of traditional actuarial thinking on the purpose of prudential capital.*
>
> *"The second approach takes a different perspective: instead of asking how much capital is required to fund the run-off of all existing liabilities and their embedded risks, this approach looks at how much capital is required to fund the short-term transfer of liabilities and their risks to a willing third-party, again at some specified level of confidence. This amount is assessed by calculating market-consistent values for liabilities, and projecting the market value balance sheet (usually over a one-year horizon). Required capital is then defined as the amount needed to ensure sufficient assets are available to meet the end-year market-consistent liability value at the specified confidence level (this is usually referred to as the Value-At-Risk (sic))."*

7.2.2 The first method is used mainly in North America whereas the second in Europe through the Solvency II framework. The United Kingdom, through its ICA regulatory framework (introduced in Section 2.2), also uses the second method. This difference in capital approach is related to the difference between the reporting requirements of the respective geographical regions. North American accounting bodies have not yet embraced the market-consistent methods like the rest of the world through the International Financial Reporting Standards (IFRS). Canadian life insurers are planning to adopt IFRS by 2012, whereas those in the USA are currently undecided.

7.2.3 The author focuses on the value-at-risk (VaR) approach to economic capital since the focus is on South Africa, whose accounting rules prescribe market-consistency. The VaR approach requires the distribution of the balance sheet on a market-consistent basis at the end of one year. The balance sheet at that time is a function of the then current assets and liabilities. This problem of projecting the distribution of the balance sheet then becomes one of projecting the market-consistent value of assets and liabilities to the end of the first year.

[17] Turnbull, C (2008). Capital to do what? Barrie and Hibbert Limited.

7.2.4 In a perfect world where there are no time, money and technological constraints, this would be done using a nested-simulation approach: in the first year, the state of the economy is projected forward using a real-world economic scenario generator. At the end of the first year, risk-neutral scenarios are then simulated in order to value the options-based assets and liabilities, which cannot be valued using simpler and more efficient methods. These risk-neutral scenarios are calibrated to each of the real-world scenarios at the end of the first year. The results will provide us with the balance-sheet positions at the end of the first year. The real-world scenarios provide a realistic distribution of these balance sheets.

7.2.5 The nested-simulations approach usually involves significant computing run time that makes this approach impractical even with today's processing power. The process consists mainly of three phases:

- calibration of the ESG;
- scenarios generation; and
- asset and liability cashflow projections and discounting.

7.2.6 In the first phase, run time increases with the number of calibrations. A separate calibration of the risk-neutral generator is needed for each real-world simulation at the end of the first year. In the second phase, run time increases with the number of calibrations, the number of scenarios generated for each calibration, the number of years projected and the length of the term structure projected. In the last phase, run time increases with the number of calibrations, number of scenarios, number of policies and the complexity of the policies. The asset projections and discounting depend on the complexity and number of assets – some (such as option-type derivatives) may require nested simulations to be valued, whereas others (such as fixed income) can be valued using more efficient methods (such as discounting fixed cashflows at the real-world yield curve).

7.2.7 Advances in technology have helped shorten the run times in three ways:

- Processing speed is improving.
- High speed processors are becoming cheaper per unit. This allows more processors to be purchased for the same cost.
- New software is allowing distributed processing for some applications so that independent scenarios may be processed concurrently on separate computers and the results recombined.

7.2.8 Policies with similar characteristics such as age, gender and moneyness may be grouped together into a single data point. This will also significantly reduce run times. Other scenario reduction techniques have also been developed to significantly reduce the number of scenarios required. A full review of these scenario reduction techniques is beyond the scope of this research. Interested readers may consult Dardis *et al* (2008) for a bibliography of papers on scenario reduction techniques.

7.2.9 The calibration of the ESG developed in Chapter 5 required approximately one day on a computer [18]. The simulations of the scenarios in the same chapter required approximate 40 minutes, for 2000 scenarios projecting 30-year term-structures for 30 years, on a single processor. Lastly, the cashflow projections and discounting required

[18] AMD Athlon 64 X2 Dual Core Processor 5400+, 2.8GHz, 6 GB dual channel RAM @ 800MHz

approximately 10 minutes for one contract, with 2000 scenarios and projecting 30 years. The arbitrary policy that we projected is a simple contract with no bells and whistles that tend to increase the complexity of the model and therefore run time.

7.2.10 It is desirable to use at least 10,000 real-world scenarios for the first year so that the tails of the distribution can be adequately captured. Approximately 2000 risk-neutral scenarios are needed for convergence under a normal real-world scenario with reasonable volatility — convergence is usually slower when volatility is high, typically when the economy is under stress. Assume that the company has at least 100 sample data points after grouping its portfolio of GAO policies. This would require 10,000 days for 10,000 calibrations, 28 days to generate 10,000 sets of 2000 risk-neutral scenarios and 14 million days to project and discount cashflows. Assume further that the author's programs are inefficient and that efficiency can be improved by 90% and that the company has 1000 processors at its disposal. Notwithstanding the expenses of maintaining 1000 processors and that 100 sample points is extremely ambitious, this would still require over 1400 days to be completed. This would render the information out-of-date for most dynamic hedging strategies.

7.2.11 Some companies have therefore opted to use a shock-scenario approach. This means that rather than simulating real-world scenarios for the first year, some important underlying variables are identified and shocked to reflect some pre-specified confidence level. These shocks are usually calculated independently for individual variables and are not very different to the traditional actuarial sensitivity tests that are used in some regulatory regimes — the key differences are:

- The shocks are calibrated to reflect a certain confidence level for each risk factor.
- The shocks are applied to both sides of the balance sheet, not just to the liabilities as is the case with most traditional capital regimes.

7.2.12 The capital of each individual risk factor is then aggregated to approximate the total capital. This aggregation is non-trivial. The use of the sum of the capital held in respect of the individual risk factors as the total capital implicitly assumes that all extreme scenarios under all risks would happen at the same time. Diversification benefits are ignored. Some assumptions can be made about the correlation of the various risk factors when aggregating the capital of individual risks. Even so, the total capital estimated this way should not be as accurate as if the nested simulations approach were used. Despite this shortcoming, this method does quantify the primary risks associated with GAOs in a timely manner so that they can be managed appropriately.

7.2.13 The shocks should reflect the real-world distribution of the variables. An historical analysis is done to determine the appropriate shocks. The next section describes the determination of the shocks that are to be applied.

7.3 Determining the Shocks to Apply

7.3.1 Introduction

7.3.1.1 Many factors affect the balance sheet. The main economic factors that affect the balance sheet of a GAO portfolio with GMV are:

- the maturity value to which the GAO benefits can apply;
- the value of the unit fund relative to the GMV;
- the term structure of interest rates;
- the implied volatilities on swaptions; and
- the implied volatilities on equity options.

7.3.1.2 This section describes the methods used to determine the shocks that can be applied to these factors so that the risks can be quantified and that a suitable risk-mitigating strategy can be developed.

7.3.2 Term Structure of Interest Rate

7.3.2.1 The yields at different durations on a yield curve have a tendency to move together simultaneously. It is insufficient to simply apply shocks to a single point on the term structure without effecting a corresponding change to others, yet the different durations do not necessarily shift by the same amount. A method of shocking the different points of the term structure is needed so that they are consistent with each other.

7.3.2.2 The most popular method in practice is the use of principle-components analysis (PCA). Shlens (unpublished) provides a method of analysing complex correlation structures with PCA using an example based in the physics realm. A simple description of PCA is that any point in an m-dimensional system where the axes are $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, ..., \mathbf{X}_m)$, say with co-ordinates $\mathbf{x} = (x_1, x_2, x_3, ..., x_m)$, may be linearly transformed onto a new set of axes $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3, ..., \mathbf{Y}_m)$, say with co-ordinates $\mathbf{y} = (y_1, y_2, y_3, ..., y_m)$, such that the axes of $\mathbf{Y}$ are orthogonal (and hence independent). In other words, $\mathbf{Y}_i \bullet \mathbf{Y}_j = 0,\ i \neq j$.

7.3.2.3 Suppose there are n points. The selection of axes in $\mathbf{Y}$ is done so that $\mathbf{Y}_1$ is the axis with the greatest deviations in the values. $\mathbf{Y}_2$ is then selected as the axis with the next greatest deviations in the values and also orthogonal to $\mathbf{Y}_1$, and so on.

7.3.2.4 Maitland (unpublished, 2002) shows that when PCA is applied to the covariance matrix of the spot yields at different durations of the curve, the first axis explains around 77.1% of all South African yield curve volatility, the first two together explain 98.4%, and the first three together explain 99.4%. When PCA is applied to the covariance matrix of the absolute change of the spot yields, the first, first two and first three components explain 92.8%, 97.3% and 98.4% respectively. Either way, the first three components together explain most of the yield-curve movements. A correlation matrix for the different durations in a 30-year yield curve has 30 correlated dimensions. PCA has effectively reduced these thirty dimensions to three independent dimensions.

7.3.2.5 Novosyolov & Satchkov (2008) made similar comparisons between the PCA on level changes (absolute) and percent changes (relative) of the yield curve. The US Treasury and US LIBOR curves were used in their investigation. Their results are summarised in Table 7.1.

Table 7.1: Cumulative variations explained by principle components, expressed as a percentage of total variations

	Absolute Change		Relative Change	
Factors	Treasury	LIBOR	Treasury	LIBOR
1	57%	66%	53%	48%
2	70%	86%	73%	75%
3	78%	91%	88%	88%
4	83%	93%	90%	92%

7.3.2.6 Based on this table alone, there is very little to distinguish between using the relative change and absolute change. Novosyolov & Satchkov (*op. cit.*) decided to use the percentage changes because economies that have high levels of interest rates would naturally also have larger absolute changes.

7.3.2.7 The PCA calculated here to derive the interest-rate shocks is based on the correlation matrix of the logarithmic changes in the yield curve. Shlens (*op. cit.*) provides an algorithm for extracting principle components from a data set using singular value decomposition.

7.3.2.8 The amount of variability explained by the first three factors of the South African term structure of interest rate is much more than those of the USA. This difference could be partly explained by choice of yield curve, data period and frequency. Maitland's (*op. cit.*) investigation is done on monthly spot yields from January 1986 to December 1998. Novosyolov & Satchkov's (*op. cit.*) investigation is done on bond-equivalent yields (par, compounded semi-annually), and was based on daily data from 1 August 2006 to 8 August 2007. The shorter time period in the latter consists of data in very similar economic environment, resulting in less data variability and higher explanatory power of the first few factors.

7.3.2.9 The data used for this research are the daily logarithmic changes to the JIBAR spot curve, from 4 August 2003 to 17 October 2008. The data series does not include a full economic cycle, but at the time of writing this dissertation, the author could not obtain any longer data series.

7.3.2.10 Figure 7.1 illustrates the coefficients for mapping observations from the old axes to the new axes. The coefficients represent one standard deviation annualised shock to the three components of the logarithmic changes of the yield curves. If we need to apply a shock to the yield curve that incorporates the first component at 95% confidence, then the shocked yield of duration t is equal to:

$$\tilde{i}_{1,t} = i_t \left[1 + f_{1,t} \Phi^{-1}(0.95) \sigma_1 \right],$$

where:

- i_t is the actual yield of duration t,
- $f_{1,t}$ is the coefficients of duration t for factor 1 in **Y** space; and
- σ_1 is the standard deviation of factor 1 in **Y** space, which is equivalent to the first eigenvalue from the singular-value decomposition.

Similarly, the shocked curve incorporating the second component is represented by:

$$\tilde{i}_{2,t} = i_t \left[1 + f_{2,t} \Phi^{-1}(0.95) \sigma_2 \right].$$

7.3.2.11 Because the PCA is done on daily changes, the shocks that are determined need to be annualised. Assume independence between observations, then:

$$\sigma_{Y,i} = \sigma_{D,i}\sqrt{365\times 5/7},$$

where

- $\sigma_{Y,i}$ is the annualised standard deviation of factor i; and
- and $\sigma_{D,i}$ is the daily standard deviation of factor i.

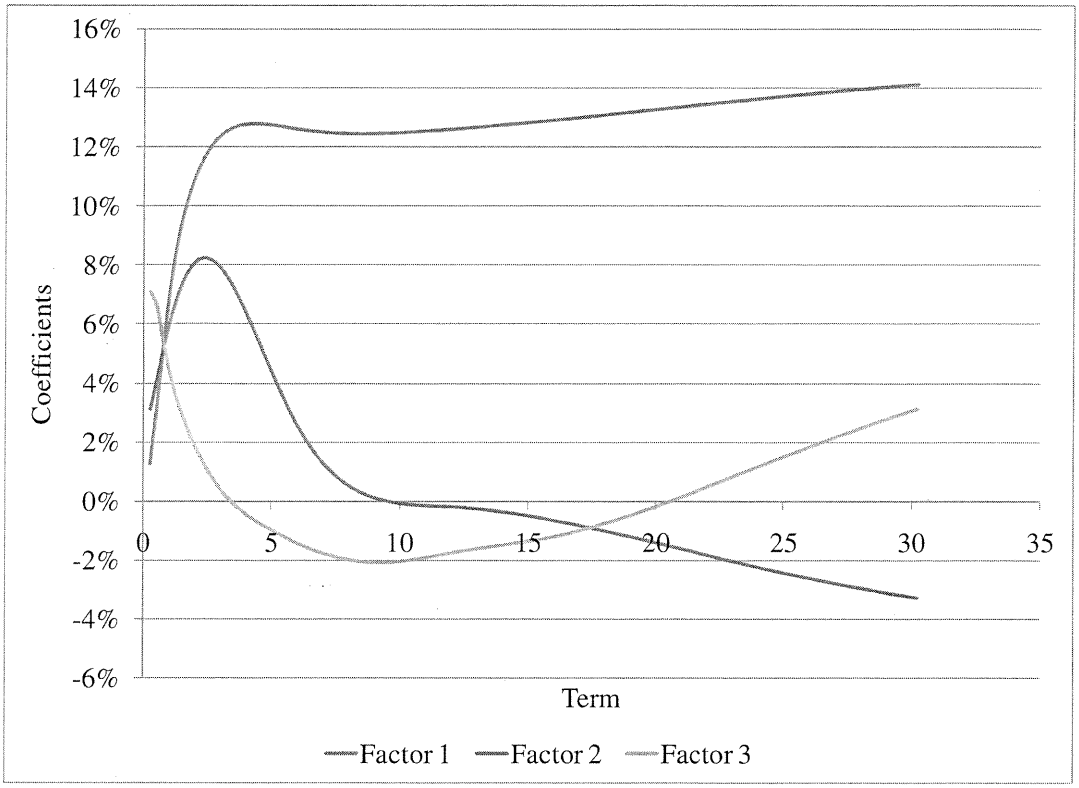


Figure 7.1: Coefficients for mapping observations from the old axes to the new axes

7.3.2.12 Factor 1 (with the exception of the anomaly at the short durations) is almost a parallel shift. Practitioners usually refer to this as the level or parallel shock. Factor 2 (again with the exception of the anomaly at the short durations) shows an increase in the short end of the term structure and a decrease to the long end. This is usually referred to as the inversion of the yield curve. Factor 3 represents an increase to both the short end and the long end of the term structure, but a decrease to the medium durations. This is often referred to as the butterfly effect in the yield curve.

7.3.2.13 Figure 7.2 shows the coordinates of the historical observations in **Y** space. One obvious observation is that the variability associated with factor 1 (observed by the magnitude of the deviations from the x-axis) is much larger than that associated with factor 2. This is similarly true for factor 2 when compared with factor 3. In fact, factors 1, 2 and 3 explain 90.16%, 5.55% and 2.06% of all variability respectively. The remaining 27 factors explaining the remaining 2.23%. This means that term-structure movements are parallel approximately 90% of the time.

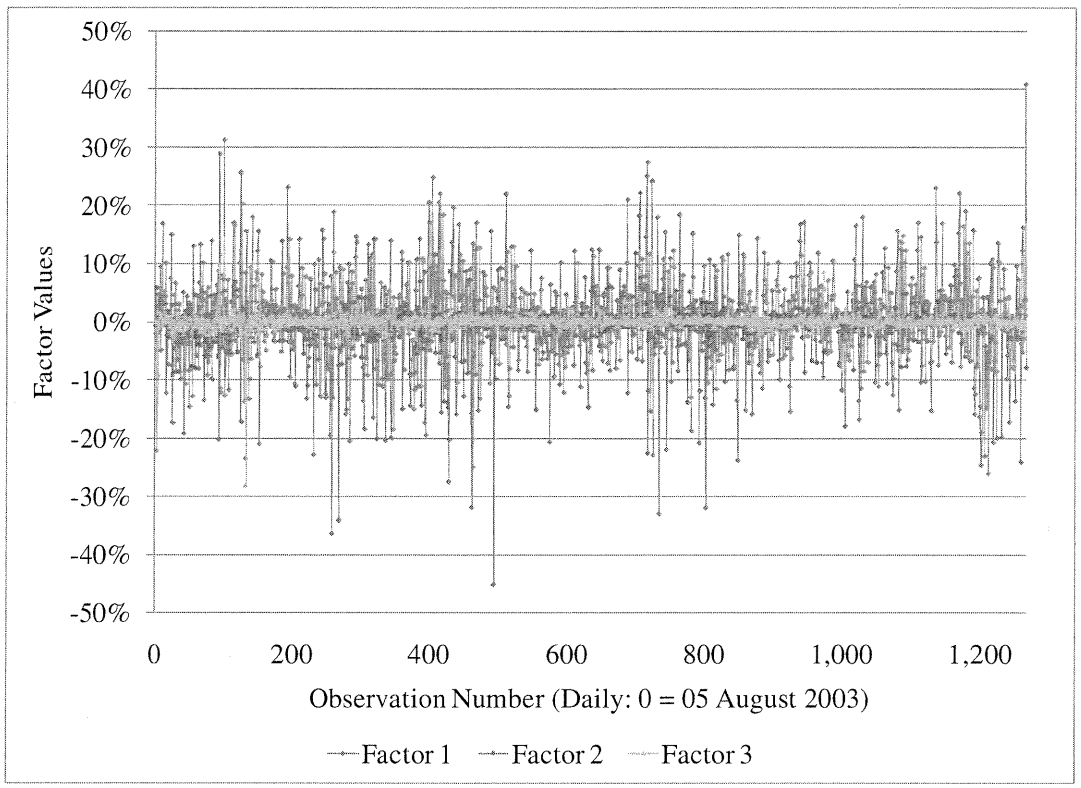


Figure 7.2: Historic observed log changes mapped onto the new axes

7.3.2.14 Figure 7.3 to Figure 7.5 illustrate the original yield curve and the shocked yield curves. The shocked yield curves reflect the worst possible position at the end of one year at the 90th and 99.5th percentile confidence levels in both directions. These form the calibration targets for the shocked yield curves.

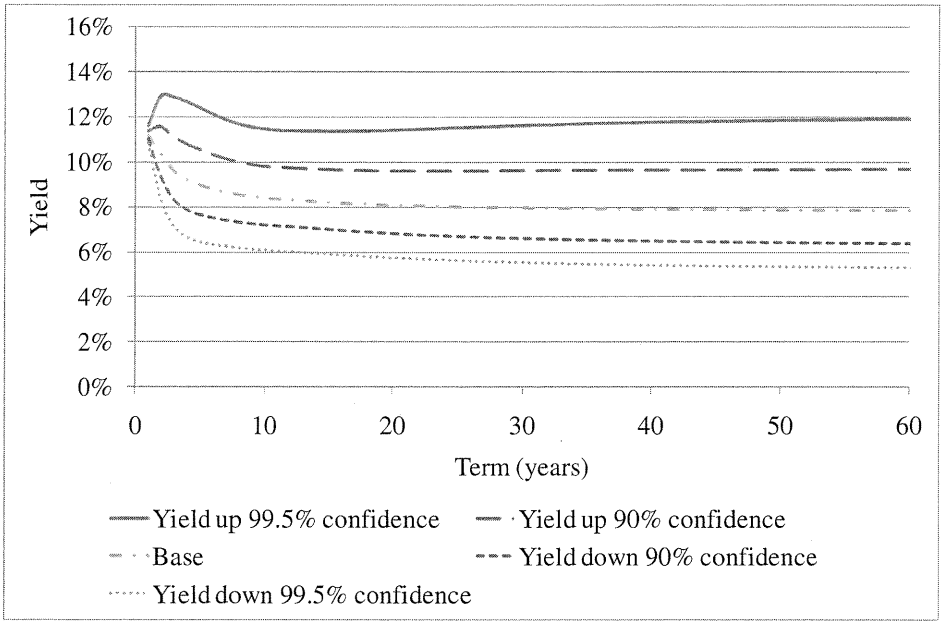


Figure 7.3: Shocked term structures of interest rates when using the first factor

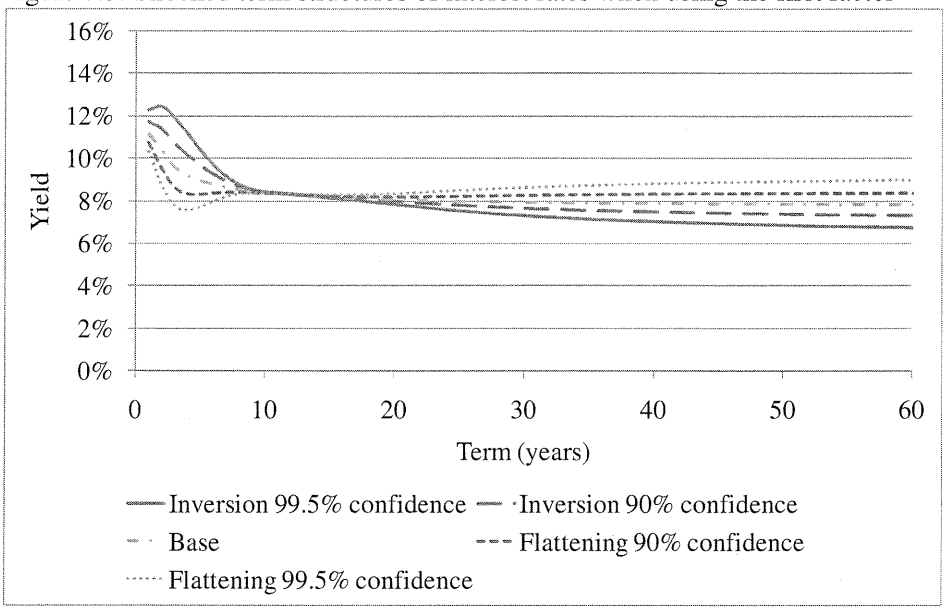


Figure 7.4: Shocked term structures of interest rates when using the second factor

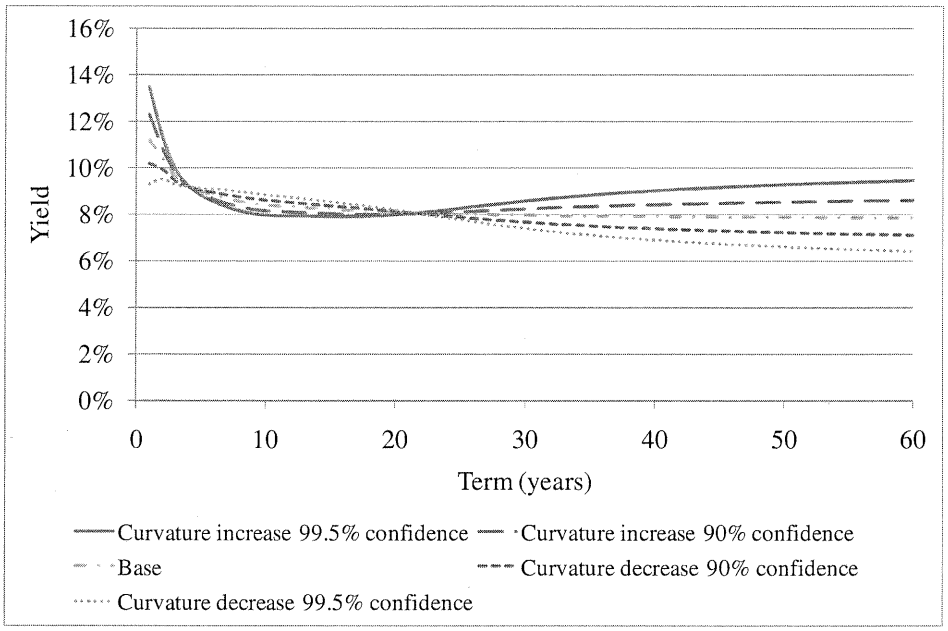


Figure 7.5: Shocked term structures of interest rates when using the third factor

7.3.3 Asset Values

7.3.3.1 The logarithm of the total return on daily JSE All Share Index between 30 June 1995 and 25 January 2008 is used for this investigation. The standard deviation of the daily logarithmic returns is calculated. The logarithmic returns from non-overlapping periods are assumed to be independent of each other. The standard deviation of the annual logarithmic return is then calculated just as in the case of the yield curve. That is:

$$\sigma_Y = \sigma_D \sqrt{365 \times 5/7}\,. \tag{7.1}$$

7.3.3.2 Similarly, the expected logarithmic return over a year as approximated as:

$$\mu_Y = \mu_D 365 \times 5/7\,,$$

where the term μ_D is calculated as the arithmetic mean of the historical daily logarithmic returns. The potential equity return (E) shock over a single year is:

$$\log(E) = \mu_Y + \sigma_Y \Phi^{-1}(p)\,, \tag{7.2}$$

where p is the desired confidence level. Table 7.2 lists the values of the fitted parameters and the shocks to the equity values over the first year.

Table 7.2: Parameters values and shocks to equity assets

Variable	**Value**
μ_D	0.054%
μ_Y	14.05%
σ_D	1.20%
σ_Y	19.44%
Annual upward shock (99.5% confidence)	64.14%
Annual downward shock (99.5% confidence)	-36.03%

7.3.3.3 The annual shock is applied at time 0 to the policyholder assets to determine the increase in the market-consistent reserves caused by a sudden rise or drop in the equity market. The 'policyholder assets' refer to the unit-linked investment portfolios. The shocks that are applied to the policyholder assets affect the reserves because they affect the value of the investment funds and therefore the value of the GMV guarantee. They also affect the amount at maturity to which the GAO guarantees will apply and therefore the value of the GAO guarantees.

7.3.3.4 If equities are held in the guarantees reserves, these too will be affected by the shock.

7.3.3.5 Changes to fixed-income assets such as bond values have a similar effect to those of equities. The shocks to be applied to the bond portfolios are directly linked to the shocks to the level and shape of the yield curve — the values of the bond portfolio are calculated, using the term structure, before and after the shocks are applied to the term structure. The price appreciation and depreciation reflect the shocks that should be applied to the bond portfolio.

7.3.3.6 The overall shock to the policyholder assets is determined as a weighted average of the amounts allocated in each asset category. Assume that the asset allocation follows the unit-fund benchmark, in this case the same as the portfolio described in ¶6.2.7.

7.3.4 Implied Volatility on Swaptions

7.3.4.1 The shocks for implied volatility on swaptions, like those of the term structure of interest rates, are determined using PCA. This is because the different points on the volatility surfaces also tend to move together. The logarithmic changes for each point on the volatility surface are used to calculate the correlation matrix on which PCA is performed. Unlike the term structure of interest rates, which only has one dimension

(term to maturity), the implied volatility on swaptions have two dimensions (option term and swap tenor). We use mid-market volatilities from at-the-strike payer swaptions issued in the United States to determine the correlation matrix. Similar credible South African data are not available so the United States data is used as a reasonable approximation.

7.3.4.2 The result from the investigation shows that five factors are required to explain 97% of movements in the implied volatility surface for swaptions. The cumulative volatility of the five factors is listed in Table 7.3.

Table 7.3: Cumulative volatility of the first five factors in the principle component analysis of changes in the implied volatility surface of swaptions

Factors	Cumulative volatility
1	85.1%
2	92.7%
3	94.9%
4	96.7%
5	97.3%

7.3.4.3 Because our interest-rate model cannot model all points of the volatility surface accurately (due to limited degrees of freedom), it is difficult for it to capture the twists represented by factors higher than the first order. We therefore focus on just the first factor. Figure 7.6 to Figure 7.8 illustrate the implied volatility on swaptions before and after applying a 99.5% shock to the first principle component.

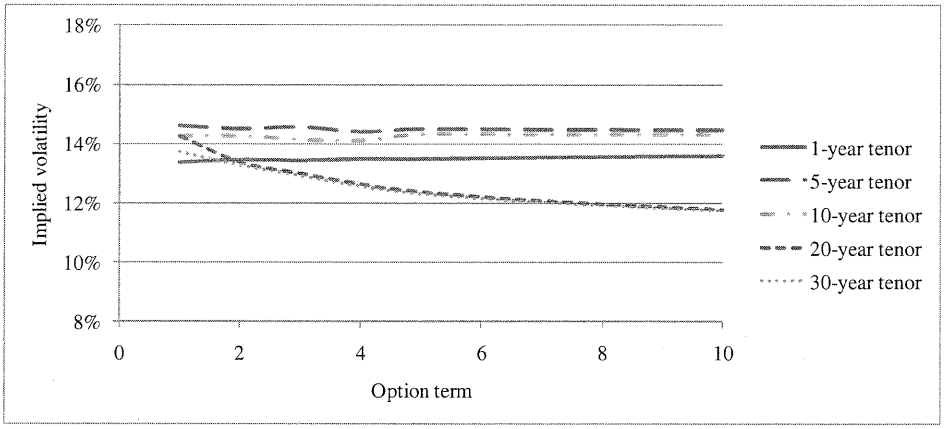


Figure 7.6: Implied volatility on swaptions before any shocks

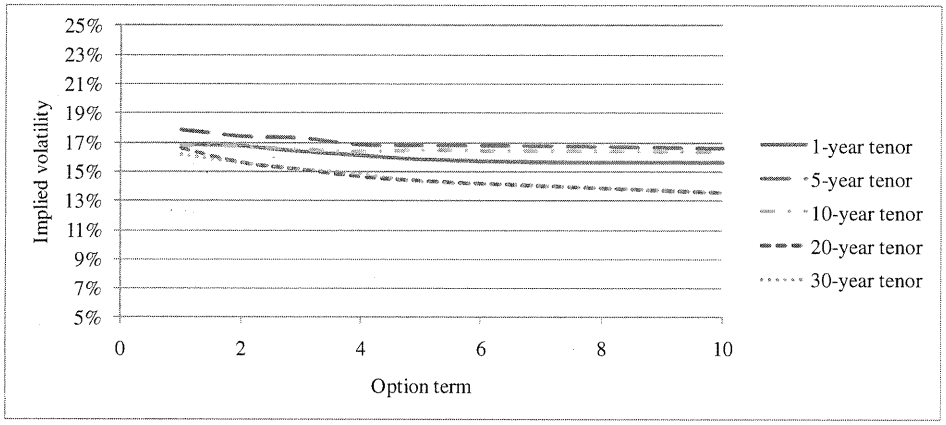


Figure 7.7: Implied volatility on swaptions where an upward shock corresponding to a 99.5% confidence is applied to the first factor

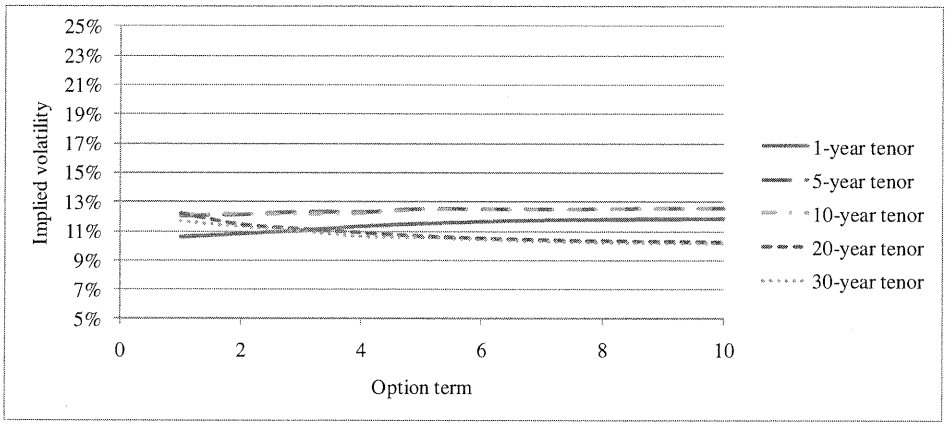


Figure 7.8: Implied volatility on swaptions where a downward shock corresponding to a 99.5% confidence is applied to the first factor

7.3.4.4 Once the shocked implied volatility on swaptions is calculated, the interest-rate model is recalibrated, using the approach described in Section 5.3.7, to the new implied volatilities. The new ESG is used to generate scenarios that can then be used to evaluate the post-shock market-consistent value of the liabilities.

7.3.5 Implied Volatility on Equity

7.3.5.1 Like implied volatilities of swaptions very limited data are available on the implied volatilities of options with JSE equities as underlying instruments. The shocks for implied volatilities of equity options are calculated using the relative changes in the VIX index. The VIX index reflects a weighted average of prices for a range of options on the S&P 500 index. The method used is similar to that used in Section 7.3.3 for equity values — the logarithm of the change in implied volatilities of equities is assumed to be normally distributed. We assume that the entire term structure of implied volatility suffers the same relative shock as the index. Table 7.4 shows the term structure of the at-the-money volatility before and after applying a 99.5% shock.

Table 7.4: Parameters values and shocks to the equity volatility

Variable	Value
μ_D	0.004%
μ_Y	1.107%
σ_D	5.8%
σ_Y	93.0%
Multiplicative upward shock (99.5% confidence)	240%
Multiplicative downward shock (99.5% confidence)	41.7%

7.3.5.2 Again, the ESG is recalibrated to the shocked scenarios to test the effects on the liabilities. We note that in this case, the upward shock to the volatility structure is sufficiently large to cause a very slow convergence of results. The 2000 scenarios are not enough to pass the martingale test. The ESG was rerun using increments of 1000 scenarios in order to investigate the convergence of scenarios, until 15,000 scenarios at which time the program stopped responding because of memory insufficiency. Even at this simulation count the convergence was unsatisfactory. The 99.5th percentile is replaced by the 90th percentile is investigated for the upward shock.

Table 7.5: Parameters values and adjusted shocks to the equity volatility

Variable	Value
μ_D	0.004%
μ_Y	1.107%
σ_D	5.8%
σ_Y	93.0%
Multiplicative upward shock (90% confidence)	119%
Multiplicative downward shock (99.5% confidence)	41.7%

7.4 Summary

This chapter introduced two main methods of calculating economic capital. The first one of these calculates the amount required to fund all future liability cashflows with a specified level of confidence and is used predominantly in North America. The other calculates the amount required to transfer the liability and risks to a third party at the end of a one-year horizon and is used mainly in Europe. The research focuses on the second method. The implementation of this method requires the use of nested simulations: risk-neutral simulations are nested at the end of real-world simulations at the end of one year. Because of the intensive computational requirements, a simplified approach that uses calibrated shocks is proposed as an approximation. The methodology for the calibration of these shocks is discussed. The application of these shocks in the calculation of economic capital and the market-risk management of GAOs is discussed in the next chapter.

Chapter 8. Managing Market Risks

8.1 Introduction

8.1.1 In this chapter, the sensitivities of market-consistent liabilities to changes in market conditions are calculated. Various hedging strategies are proposed. These utilise simple listed assets and derivative instruments. Investigations are conducted on their effectiveness as hedges for the market risks underlying the liabilities under each of the stressed scenarios developed in the previous chapter.

8.1.2 The following hedging strategies are investigated:
- interest-rate duration hedging using a single ZCB;
- interest-rate duration hedging using geared ZCBs;
- a basket of vanilla swaptions; and
- swaptions and equity-put options.

8.2 Sensitivity of Market-consistent Liabilities to Changes in Market Conditions

8.2.1 Sensitivity to Changes in the Term Structure of Interest Rates

8.2.1.1 Table 8.1 to Table 8.3 illustrate the sensitivity of the market-consistent value of the hypothetical GAO portfolio to the changes in the term structure of interest rates when allowing for movements of the first three factors. The shocks that are applied are the same as those derived in Section 7.3.2. The liabilities are affected in two ways:
- The starting fund values are affected by the changes to the values of the ZCBs as the yield curves are shifted.
- The option values changed as the risk-neutral discount rates changed.

Table 8.1: Sensitivity of market-consistent liabilities to changes in the term structure of interest rates when allowing only for movement in the first factor

	Increase in term structure		Base	Decrease in term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	131	872	2,785	6,418	11,409
GAO reserve	134	487	1,406	3,854	7,010
Cross-term reserve	1	28	157	663	1,668
Total guarantee value	266	1,387	4,349	10,935	20,086

Table 8.2: Sensitivity of market-consistent liabilities to changes in the term structure of interest rates when allowing only for movement in the second factor

	Inversion of term structure		Base	Flattening of term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	4,526	3,705	2,785	1,492	1,145
GAO reserve	3,015	2,086	1,406	967	648
Cross-term reserve	442	272	157	69	39
Total guarantee value	7,983	6,063	4,349	2,529	1,831

Table 8.3: Sensitivity of market-consistent liabilities to changes in the term structure of interest rates when allowing only for movement in the third factor

	Increase in curvature		Base	Decrease in curvature	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	2,258	2,841	2,785	2,017	2,116
GAO reserve	359	718	1,406	2,671	4,772
Cross-term reserve	51	104	157	186	300
Total guarantee value	2,668	3,663	4,349	4,874	7,188

8.2.1.2 It is evident that an increase in the term structure of interest rates results in lower liability for GMV and GAO. This results in lower total liabilities. This is consistent with option-pricing theory that an increase in the risk-free rate results in a lower option price. However, the extent to which the value of the liabilities can decrease is limited as the value of the liabilities has a floor of 0. Conversely, the increase in liabilities from a decrease in the term structure of interest rates can potentially be much greater. Figure 8.1 to Figure 8.3 illustrate this non-linearity clearly.

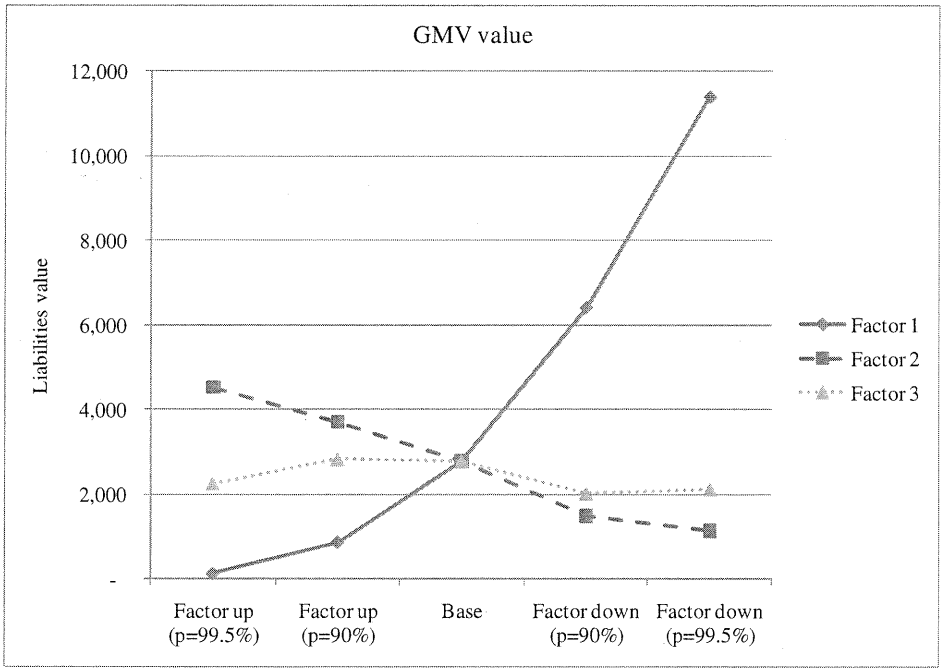


Figure 8.1: Sensitivity of GMV reserves to the changes in term structure of interest rates

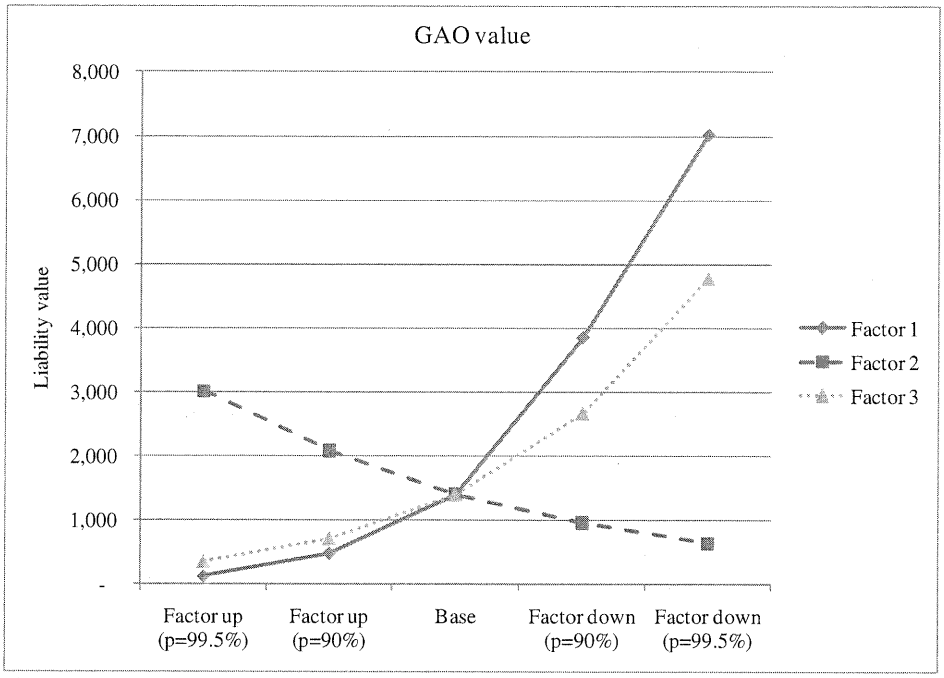


Figure 8.2: Sensitivity of GAO reserves to the changes in term structure of interest rates

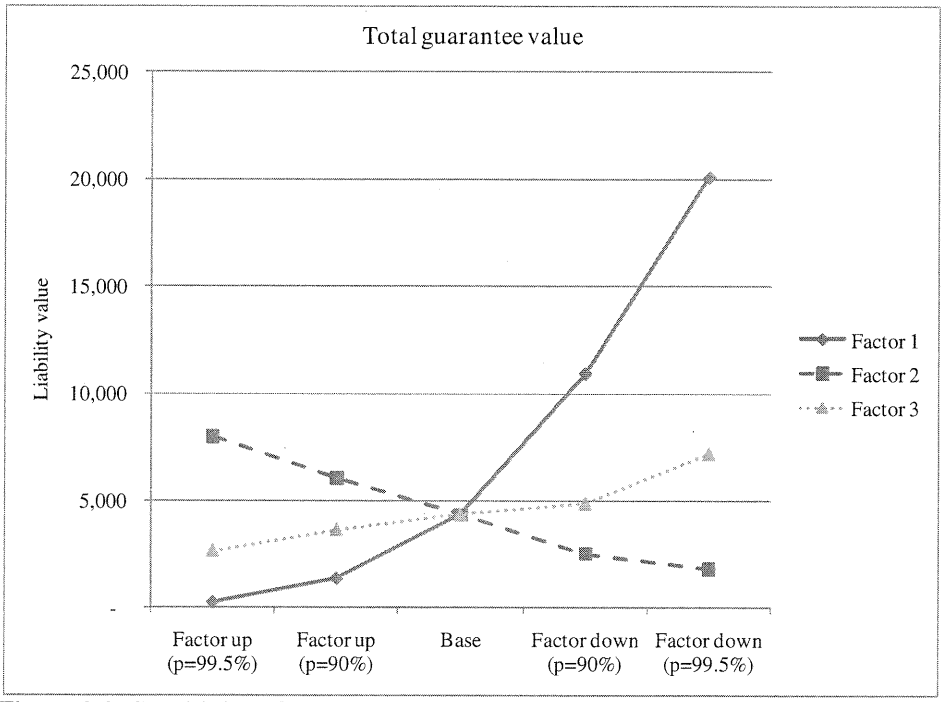


Figure 8.3: Sensitivity of total reserves to the changes in term structure of interest rates

8.2.1.3 Furthermore, it is worth noting that the change to the values of the liabilities when the first factor is shocked is by far the biggest. The 99.5th percentile second-factor shock results in a significant decline at the long end of the yield curve. Because the GAO and GMV liabilities are long-term, this fall in long-term interest rates resulted in an increase in liabilities that cannot be offset by the increase in interest rates at the short

end. Lastly, the third factor has very little impact on GMV liabilities because the level of the term structure at time 25 (when the GMV matures) has not changed significantly. However, the GAO liabilities have increased significantly – this is because the GAO liabilities are affected by the term structure beyond 25 years and at these durations, the 99.5th percentile shock results in a much lower term structure.

8.2.2 Sensitivity to Changes in Starting Fund Values

8.2.2.1 Table 8.4 illustrates the sensitivity of the market-consistent liability value of the hypothetical GAO portfolio to changes in the starting fund values. This change is made only to the fund value allocated to equities — the bond component remains unaffected. The shocks that are applied are the same as those derived in Section 7.3.3.

Table 8.4: Sensitivity of market-consistent liabilities of the hypothetical liability portfolio to changes in the starting fund values

Description	Increase in starting fund values	Base	Decrease in starting fund values
Confidence level	99.5%		99.5%
Shock applied	41.71%		-24.62%
GMV reserve	1,697	2,785	3,824
GAO reserve	1,782	1,406	1,184
Cross-term reserve	95	157	209
Total guarantee reserve	2,747	3,803	4,870

8.2.2.2 Table 8.4 shows that as the starting fund value falls, the GMV liability increases. This is because the GMV acts like a put-option on the underlying fund. Conversely, the GAO liability decreases as the GAO guarantees a conversion factor that converts the maturity value into a life annuity at the retirement date — the lower the fund value at maturity, the less onerous the guarantee becomes.

8.2.3 Sensitivity to Changes in Implied Volatility of Swaptions

8.2.3.1 Table 8.5 shows the sensitivity of the market-consistent liability value of the hypothetical GAO portfolio to the changes in the implied volatility on swaptions. The shocks that are applied are the same as those derived in Section 7.3.4.

Table 8.5: Sensitivity of market-consistent liabilities of the hypothetical liability portfolio to changes in implied volatility on swaptions

Description	Increase in the implied volatility of swaptions	Base	Decrease in the implied volatility of swaptions
Confidence level	99.5%		99.5%
GMV reserve	4,426	2,785	2,551
GAO reserve	3,994	1,406	1,234
Cross-term reserve	531	157	123
Total guarantee reserve	8,952	4,349	3,908

8.2.3.2 This table shows that the liability increases as the implied volatility on swaptions increases. Intuitively, this should only affect the GAO liabilities and not the GMV liabilities, because the former is swaption-related whereas the latter is not. However, in this case of a risk-neutral, Monte-Carlo valuation, a change to the volatility of the interest-rates affects the discount rates of the model and creates knock-on effects on

the equities options. In practice it is unlikely that the implied volatility on equity would remain stable if there are large movements in the interest-rate volatility.

8.2.4 Sensitivity to Changes in Implied Volatility of Equity Options

8.2.4.1 Table 8.6 shows the sensitivity of the market-consistent liability value of the hypothetical GAO portfolio to changes in the equity-option-implied volatility. The shocks that are applied are the same as those derived in Section 7.3.5.

Table 8.6: Sensitivity of market-consistent liabilities to changes in equity-option-implied volatilities

Description	Increase in implied volatility of equity	Base	Decrease in implied volatility of equity
Confidence level	90.0%		99.5%
GMV reserve	4,106	2,785	496
GAO reserve	1,417	1,406	1,362
Cross-term reserve	211	157	42
Total guarantee reserve	5,733	4,349	1,900

8.2.4.2 Table 8.6 shows that the GMV liability increases as the implied volatility of equity options increases. This is consistent with option-pricing theory.

8.2.5 Summary of Sensitivities

8.2.5.1 Figure 8.4 compares the changes to the market-consistent value of total guarantee liabilities after the various shocks to market conditions have been applied. A positive number indicates an increase in the value of the liabilities and vice versa. The 99.5th percentile downward shock for factor 1 of the term structure results in an increase in the liabilities whereas the 99.5th percentile downward shock for the equity volatility results in a decrease.

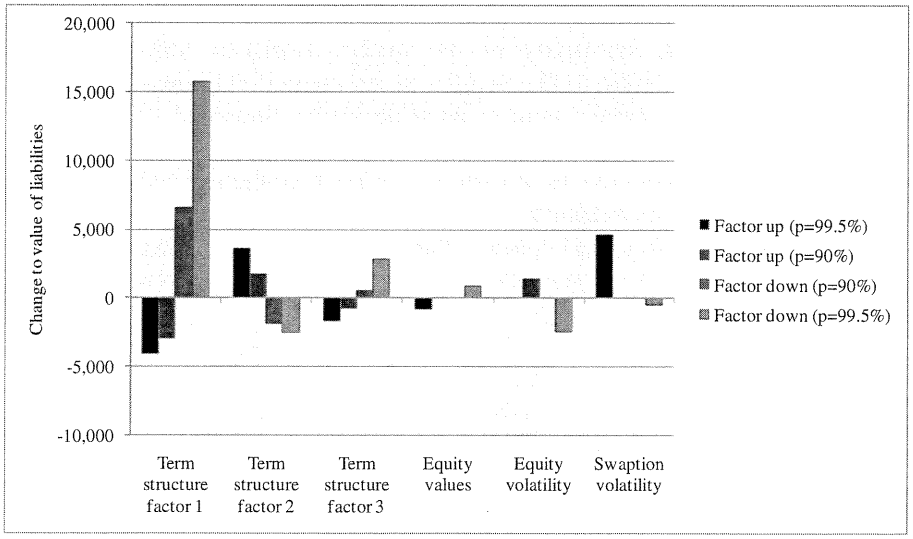


Figure 8.4: The effects of the various shocks at various confidence levels on the total guarantee liabilities

8.2.5.2 This chart demonstrates clearly that the most significant risk is that of the first-factor term-structure. For shocks of the same confidence level and have the same likelihood of occurring, the first-factor term-structure has a much larger financial impact than the other market variables.

8.2.5.3 The explanation of the relative sizes of the impacts for the three factors of the term structure was already done in Section 8.2.1 above. Shocks to the equity values do not make a large impact to the guarantee liabilities because the starting asset values only affect the intrinsic value of the guarantee. As the option still has a long period until expiry and as a result the intrinsic value is small relative to the time value of the option. Equity volatility and swaption volatilities also have a relatively small impact on the liability values because the option values are less sensitive to changes in volatilities than term-structures.

8.2.5.4 Because interest rates, in particular shift in the first factor, represents the biggest risk, our hedging proposals below shall focus on addressing this risk.

8.3 Interest-rate Duration Hedge Using a Single Zero-coupon Bond

8.3.1 A simple hedge that attempts to duration-match the liabilities is investigated. The effective duration of the liabilities is calculated using the following equation:

$$D = \frac{V_{-\Delta y} - V_{+\Delta y}}{2V_0 \Delta y}, \tag{8.1}$$

where:

- D is the stochastic duration;
- V_0 is the value of the liabilities under the base calibration;
- $V_{-\Delta y}$ and $V_{+\Delta y}$ are the values of the liabilities after a parallel decrease and increase in the term structure of interest rates of Δy, respectively; and
- Δy is chosen to be 50 basis points (bps).

This represents the sensitivity of the value of the liabilities to the change in the level of the yield curve.

8.3.2 This calculation of the stochastic duration requires that the ESG be recalibrated with the new term structure of interest rates and the liabilities recalculated. The results are in Table 8.7.

Table 8.7: The value of guarantees liabilities after a reduction and increase of 50 bps to the term structure of interest rates

	Liabilities after 50 bps reduction in term structure	Original liabilities	Liabilities after 50 bps increase in term structure
GMV reserve	2,091	2,785	3,706
GAO reserve	1,017	1,406	1,940
Cross-term reserve	101	157	245
Total guarantee reserve	3,209	4,349	5,891

8.3.3 This resulted in a stochastic duration of 58 years for GMV, 74 years for the GAO in isolation and 61.7 years for the total guarantees liability.

8.3.4 If a 61.7-year ZCB were to exist, this would have a value of 0.01061 per unit nominal. This value is calculated using a spot rate of 7.8385% (assuming the same value as the 60-year spot rate in Appendix 5.2). Equating the value of the liabilities (3,803) to that of the assets, one would need to invest in 358,430 units nominal of these ZCBs.

8.3.5 Table 8.8 to Table 8.13 compare the value of the liabilities under various shocked scenarios against the value of this ZCB. A negative value-at-risk (VaR) means that a mismatch loss would result under the scenario, whereas a positive VaR means a mismatch profit under the scenario. Figure 8.5 then compares the post-shock positions if a hedge is held and if no hedge is held.

Table 8.8: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor

Description	Increase in term structure		Base	Decrease in term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	131	872	2,785	6,418	11,409
GAO reserve	134	487	1,406	3,854	7,010
Cross-term reserve	1	28	157	663	1,668
Total guarantee reserve	266	1,387	4,349	10,935	20,086
Value of assets	414	1,523	4,349	10,214	18,907
Value at risk	-148	-136	0	721	1,179

Table 8.9: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor

Description	Inversion of term structure		Base	Flattening of term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	4,526	3,705	2,785	1,492	1,145
GAO reserve	3,015	2,086	1,406	967	648
Cross-term reserve	442	272	157	69	39
Total guarantee reserve	7,983	6,063	4,349	2,529	1,831
Value of assets	8,198	6,017	4,349	3,142	2,221
Value at risk	-215	46	0	-614	-390

Table 8.10: Value at risk: changes in the term structure of interest rates when allowing only for the changes in of the third factor

Description	Increase in curvature		Base	Decrease in curvature	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	2,258	2,841	2,785	2,017	2,116
GAO reserve	359	718	1,406	2,671	4,772
Cross-term reserve	51	104	157	186	300
Total guarantee reserve	2,668	3,663	4,349	4,874	7,188
Value of assets	1,679	2,745	4,349	6,714	10,157
Value at risk	989	918	0	-1,840	-2,970

Table 8.11: Value at risk: changes in the starting fund values

Description	Increase in equity values (41.71%)	Base	Decrease in equity values (–24.62%)
Confidence level	99.5%		99.5%
GMV reserve	1,697	2,785	3,824
GAO reserve	1,782	1,406	1,184
Cross-term reserve	95	157	209
Total guarantee reserve	3,574	4,349	5,217
Value of assets	4,349	4,349	4,349
Value at risk	-775	0	869

Table 8.12: Value at risk: changes in implied volatility on swaptions

Description	Increase in implied volatility on swaptions	Base	Decrease in implied volatility on swaptions
Confidence level	99.5%		99.5%
GMV reserve	4,426	2,785	2,551
GAO reserve	3,994	1,406	1,234
Cross-term reserve	531	157	123
Total guarantee reserve	8,952	4,349	3,908
Value of assets	4,349	4,349	4,349
Value at risk	4,603	0	-440

Table 8.13: Value at risk: changes in implied volatilities of equity options

Description	Increase in implied volatility of equities	Base	Decrease in implied volatility of equities
Confidence level	90%		99.5%
GMV reserve	4,106	2,785	496
GAO reserve	1,417	1,406	1,362
Cross-term reserve	211	157	42
Total guarantee reserve	5,733	4,349	1,900
Value of assets	4,349	4,349	4,349
Value at risk	1,385	0	-2,448

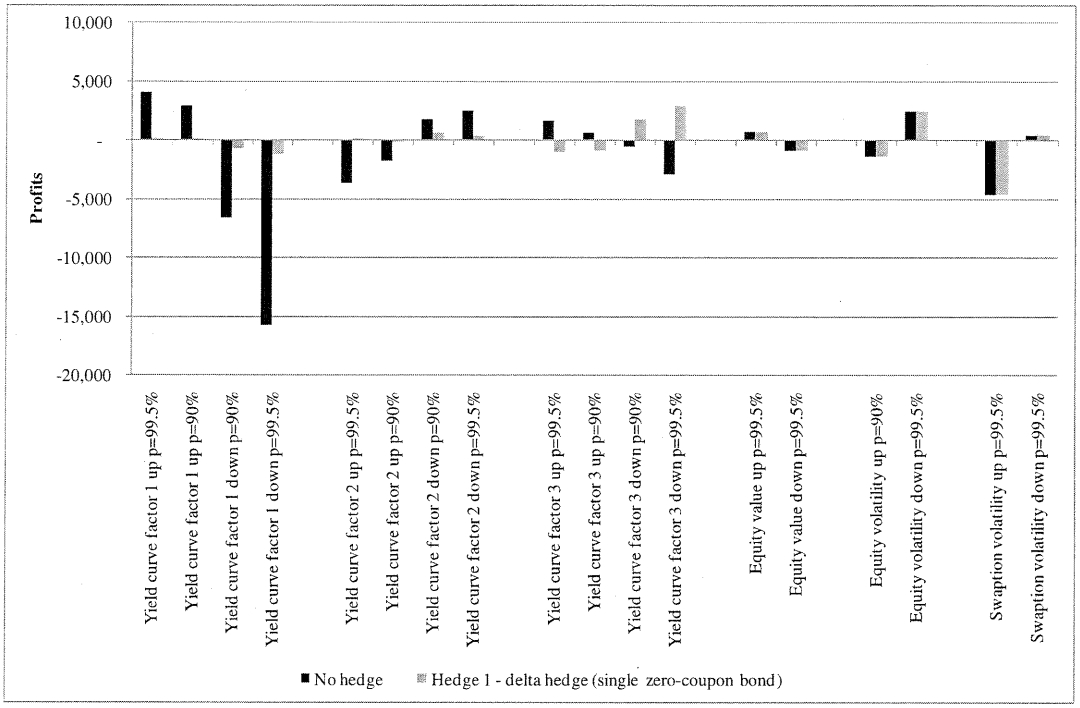


Figure 8.5: Comparison of the value at risk between (a) no hedge, and (b) duration hedge with a ZCB of 61.7 years

8.3.6 This duration hedge has removed most of the downside risk associated with the decrease in the term structure of interest rates but it also removed the upsides associated with the increase in the term structure. However, the potential losses that have been reduced are significantly more than the potential profits that are sacrificed. At the second factor level, again the downside risk has been removed but the upside is also decreased. At the third factor level, it has created some small losses but also some potential gains. As expected, the exposures to equity values, equity volatility and swaption volatility are not affected.

8.3.7 This hedge has successfully reduced most of the downside financial risks (particularly interest-rate risks) that are associated with this product. However, we note that bonds longer than 30 years do not exist in today's market so the investigation of this hedge is purely academic. Asset duration of 61.7 years must be achieved through other means. In the next section, we investigate a possible methodology of artificially extending the yield curve.

8.4 Interest-rate Hedge Using Bond Gearing

8.4.1 In the previous section, we noted that bond duration of greater than 30 years can be very difficult to achieve because the South African and the UK yield curves are limited to 30 years. Even so, asset duration of 30 years can only be achieved if the bond does not have any coupons — the presence of coupons will put some weight on earlier cashflows and decrease the duration of the asset.

8.4.2 To artificially extend the duration beyond the longest available term, one could invest in the longest bond available and simultaneously hold a short position in a bond of a

shorter term. We assume here for illustrative purposes that the long bond is a 30-year ZCB whereas the shorter bond is a 29-year old bond. The 29-year bond is chosen because its yield should be very highly correlated to that of the 30-year bond.

8.4.3 The nominal holdings in the two bonds provide two variables. We also have two constraints: the market value and the duration of the hedge portfolio should be the same as that of the liability. Under the market condition at the valuation date, the spot yield on the 30-year bond is 7.9485% and has a price of 9.21295 per 100 nominal. The spot yield on the 29 year bond has a spot yield of 7.9560% and has a price of 9.95355 per 100 nominal. Holding a long position of R1,521,071 nominal of the 30-year bond and a short position of R1,364,206 nominal of the 29-year bond will give us total market value equal to that of the liabilities, R4 349. At the same time it also gives the duration of 61.7 years. We investigate the effectiveness of this hedge portfolio under the various market stresses. The results are presented in Table 8.14 to Table 8.16. Only the interest-rate scenarios are attached as the other scenarios remain the same as those in the previous section. Figure 8.6 then compares the post-shock positions if this hedge is held and if no hedge is held.

Table 8.14: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor.

Description	Increase in term structure		Base	Decrease in term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	131	872	2,785	6,418	11,409
GAO reserve	134	487	1,406	3,854	7,010
Cross-term reserve	1	28	157	663	1,668
Total guarantee reserve	266	1,387	4,349	10,935	20,086
Value of assets	-601	1,010	4,349	9,749	16,317
Value at risk	868	377	0	1,186	3,769

Table 8.15: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor.

Description	Inversion of term structure		Base	Flattening of term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	4,526	3,705	2,785	1,492	1,145
GAO reserve	3,015	2,086	1,406	967	648
Cross-term reserve	442	272	157	69	39
Total guarantee reserve	7,983	6,063	4,349	2,529	1,831
Value of assets	7,785	5,972	4,349	2,976	1,748
Value at risk	198	91	0	-447	84

Table 8.16: Value at risk: changes in the term structure of interest rates when allowing only for the changes in of the third factor.

Description	Increase in curvature		Base	Decrease in curvature	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	2,258	2,841	2,785	2,017	2,116
GAO reserve	359	718	1,406	2,671	4,772
Cross-term reserve	51	104	157	186	300
Total guarantee reserve	2,668	3,663	4,349	4,874	7,188
Value of assets	658	2,402	4,349	6,515	8,926
Value at risk	2,010	1,261	0	-1,641	-1,738

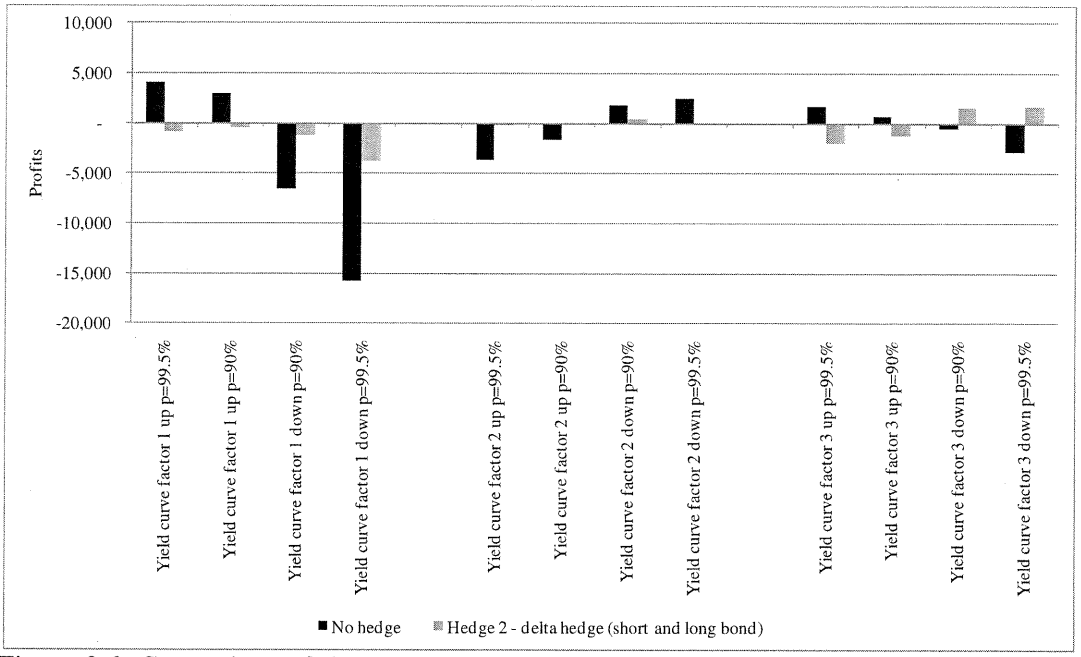


Figure 8.6: Comparison of the value at risk between (a) no hedge, and (b) duration hedge with a long and short positions in ZCBs of 30 and 29 years respectively

8.4.4 The different points at the long end of the yield curve are usually very highly correlated with each other, so in theory the value of the hedge portfolio should be similar to that of the previous hedge under the various scenarios. Like the previous hedge, this hedge has significantly reduced the potential losses that are associated with interest-rate movements of the first factor.

8.4.5 However, this hedge requires a long and a short position of very large notional amounts, making this hedge difficult to implement in the absence of a perfectly liquid market. Even if implemented, the constant dynamic rebalancing of the large notional will incur significant trading costs. Also, as discussed in Section 3.2.2, this dynamic hedging strategy can potentially cause a feedback loop. These factors cause this hedge to be impractical to implement. More practical hedges are investigated below.

8.5 Interest-rate Hedge Using a Basket of Vanilla Swaptions

8.5.1 Determining the Guaranteed Cashflows and Swaption Strike Rate

8.5.1.1 In this section a way to hedge only the GAO reserve using a portfolio of vanilla swaptions is proposed. We noted earlier that GAOs, given the value of the fund value at the maturity of the policy, behave similarly to a vanilla swaption. However, whereas the cash inflows that a receiver vanilla swaption provides are constant, the cash outflows that a GAO would command, if the option is exercised, are not. This is because the expected annuity payments under a well diversified portfolio of life annuities decrease over time as a result of mortality. This means that a single swaption cannot be used to completely match the cashflows of a GAO.

8.5.1.2 A method of matching the minimum cashflows of the GAO (i.e. the guaranteed payments) is to use a portfolio of vanilla swaptions. The notional underlying each swaption is calibrated so that the projected minimum GAO liability cashflow will always be met.

8.5.1.3 Because we are investigating the effectiveness of a swaption hedge for just the GAO component, the fund value at maturity is assumed to be fixed for this investigation. The assets in the unit fund are assumed to grow at the forward risk-free rate until the vesting date so that it has a maturity value of:

$$S_{25} = 100{,}000 \exp\left(\sum_{t=1}^{25} r_t\right) + \sum_{t=1}^{25} 5000 \exp\left(\sum_{j=t}^{25} r_t\right) = 1{,}140{,}682,$$

where r_t is the forward risk-free rate that applies for the period $[t-1,\ t)$.

8.5.1.4 The guaranteed annuity conversion factor is 9, so that the guaranteed minimum annuity payment, conditional on the policy reaching maturity, is:

$$\frac{S_{25}}{9} = 126{,}742.$$

8.5.1.5 The expected guaranteed minimum payment for the first year of annuity payment is:

$$126{,}742\ {}_{25}p^{*}_{40}\ p_{65} = 45{,}823.$$

8.5.1.6 The expected minimum payments for the subsequent years decrease further with mortality so that the expected payment t years after vesting is as follows:

$$126{,}742\ {}_{25}p^{*}_{40}\ {}_{t}p_{65}$$

These values are shown in Table 8.17 and Figure 8.7.

Table 8.17: Decrease in the expected guaranteed minimum payments.

Age	Expected minimum payment conditional on retirement	Expected minimum payment	Age	Expected minimum payment conditional on retirement	Expected minimum payment	Age	Expected minimum payment conditional on retirement	Expected minimum payment
65	123,777	45,823	76	82,200	30,431	87	29,117	10,779
66	120,660	44,669	77	77,651	28,747	88	24,790	9,177
67	117,397	43,461	78	72,981	27,018	89	20,769	7,689
68	113,994	42,201	79	68,200	25,248	90	17,100	6,331
69	110,456	40,891	80	63,325	23,443	91	13,816	5,115
70	106,790	39,534	81	58,375	21,611	92	10,936	4,048
71	103,004	38,133	82	53,379	19,761	93	8,467	3,134
72	99,099	36,687	83	48,370	17,907	94	6,400	2,369
73	95,070	35,195	84	43,387	16,062			
74	90,911	33,656	85	38,476	14,244			
75	86,621	32,068	86	33,699	12,475			

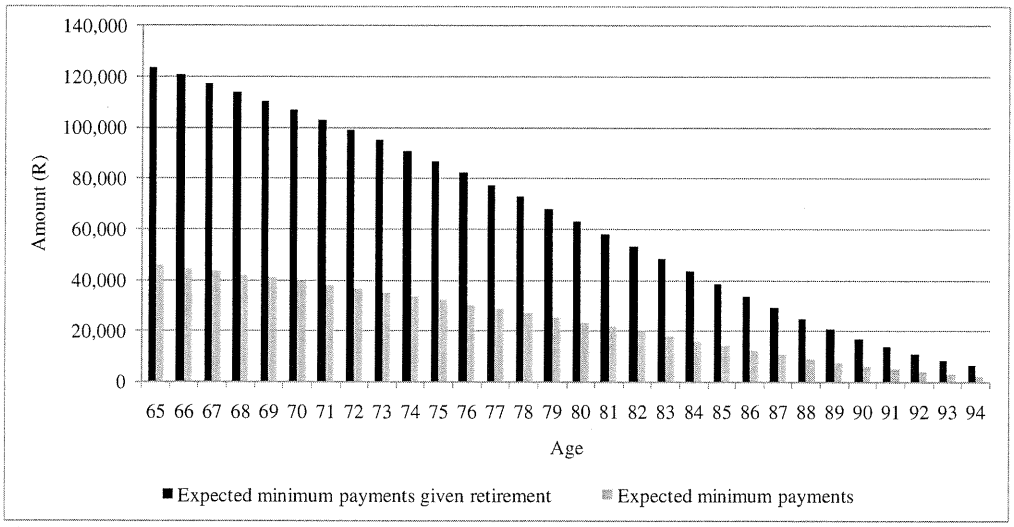


Figure 8.7: Projected cashflow profiles

8.5.1.7 To calculate the average guaranteed interest rate, we solve:

$$a^{i}_{65:\overline{30}|} = 9$$

for *i*. In this case, the guaranteed interest rate is 5.97%. This implies that should the average interest rate drop below this rate, the guarantee would be in the money.

8.5.1.8 The forward yield curve implies a market annuity factor of:

$$\underline{a}_{65:\overline{30}|} = \sum_{t=1}^{30} {}_{t}p_{65} \exp\left(-\sum_{j=1}^{t} r_j\right) = 7.66$$

is expected at the retirement date.

8.5.1.9 The average market yield is calculated by solving for *i* in:

$$\underline{a}^{i}_{65:\overline{30}|} = \sum_{t=1}^{30} {}_{t}p_{65}e^{-ti} = 7.66.$$

8.5.1.10 In this case, the average market yield is 8.24%. This implies that the strike rate underlying the GAO, 5.97%, is 72.5% of market rate, 8.24%. The strike rate that should be used for the swaption hedge is 72.5% of the current market interest rate.

8.5.2 Calculating the Notional Amounts for the Swaption Portfolio

8.5.2.1 Starting from the last annuity payment at age 94, we work backwards towards retirement at age 65. To hedge this last payment, we need to have a swaption that provides a 25-year option to enter into a 30-year receiver swap, which provides the payment at age 95. The at-the-money strike rate is calculate with (5.9) of Chapter 5:

$$s_{25,30} = \frac{P_{25} - P_{55}}{{}_{25|}\underline{a}_{\overline{30}|}}.$$

8.5.2.2 This implies at-the-money strike rate of 8.04%. The swaption strike rate that we target is 72.5% of this, or 5.83%. To generate the expected minimum payment of R2,369 at age 94, at a swap rate of 5.83% per year, requires a notional of

$$\frac{2,369}{0.0583} = 40,634.$$

8.5.2.3 This 30-year swap will also provide fixed cash inflows for the first 29 years of R2,369 each year. The expected minimum payment for year 29 is R3,134. This means that only the shortfall of R765 needs to be provided for by a swaption with a 25-year option term entering into a 29-year swap. A similar method can be used to calculate the notional that is needed to generate this fixed inflow. This process can be applied iteratively until the first guaranteed payment is matched. Table 8.18 lists the notional of each 25-year swaption that should be held in our hedge portfolio.

Table 8.18: Notional amounts of 25-year swaptions of various tenors to be held in the hedge portfolio

Tenor (years)	Notional
1	19,856
2	20,743
3	21,614
4	22,469
5	23,279
6	24,038
7	24,794
8	25,585
9	26,403
10	27,240

Tenor (years)	Notional
11	28,071
12	28,883
13	29,651
14	30,356
15	30,956
16	31,425
17	31,719
18	31,804
19	31,642
20	31,180

Tenor (years)	Notional
21	30,331
22	29,089
23	27,475
24	25,526
25	23,295
26	20,854
27	18,284
28	15,676
29	13,121
30	40,634

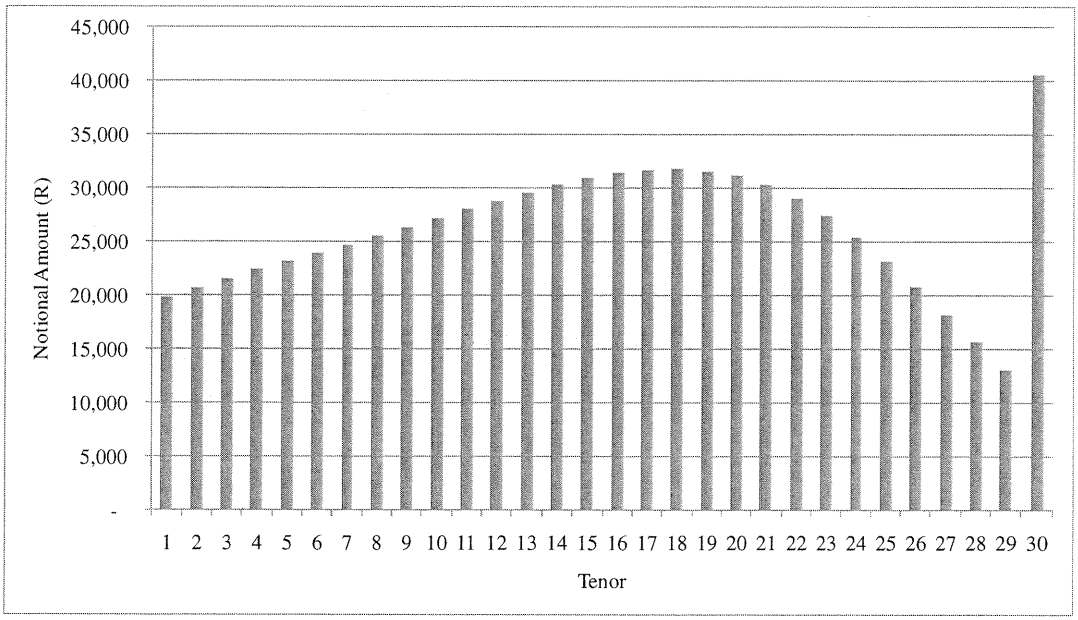


Figure 8.8: Notional amounts of 25-year swaptions of various tenors to be held in the hedge portfolio

8.5.3 Calculating the Value of the Swaption Portfolio

8.5.3.1 Having calculated the notional for each of the vanilla swaptions in our portfolio, the price of this portfolio is calculated with simulations, using the same method introduced in (5.25). Table 8.19 lists the price of each swaption under the base scenario.

Table 8.19: Prices of 25-year swaptions of various tenors to be held in the hedge portfolio

Tenor (years)	Value
1	17
2	33
3	48
4	64
5	78
6	91
7	104
8	116
9	128
10	139
11	150
12	160
13	169
14	177
15	184
16	190
17	194
18	197
19	197
20	196
21	191
22	184
23	174
24	161
25	147
26	131
27	115
28	98
29	82
30	253
Total	4,169

8.5.3.2 The market-consistent value of the GAO liabilities can be calculated using equation (6.2). The result of this calculation is R2,002. The value of this hedge portfolio is R4,169 (see Table 8.19). This shows that the value of the hedge costs more than the amount of money we have in reserve for the GAO liabilities. This is because the GAO liabilities consist of a single option whereas the hedge portfolio consists of 30 different swaptions and it is possible that one or more of these swaptions be in the

money whilst the others are not. Under these circumstances it is possible and beneficial to the option holder to exercise only the swaptions that are in the money but not others. This additional option in the hedge portfolio increases the cost of the hedge portfolio. To analyse the effectiveness of this hedge, assume that cash is borrowed to make up the difference between the reserve and the cost of the swaptions.

8.5.4 Value at Risk

8.5.4.1 Table 8.20 to Table 8.24 show the VaR of this hedging strategy under the various shocked scenarios. The value of liabilities represents the market-consistent value of the GAO benefit of the respective scenarios when the underlying fund is assumed to grow at the forward rate of the base scenario.

Table 8.20: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor

Description	Increase in term structure		Base	Decrease in term structure	
Confidence level	99.5%	90%		90%	99.5%
Value of liabilities	105	500	2,002	4,952	10,372
Value of swaptions	281	1,313	4,169	13,212	27,913
Cash	-2,167	-2,167	-2,167	-2,167	-2,167
Swaption + Cash	-1,886	-854	2,002	11,045	25,746
VaR	1,990	1,354	0	-6,093	-15,374

Table 8.21: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor

Description	Inversion of term structure		Base	Flattening of term structure	
Confidence level	99.5%	90%		90%	99.5%
Value of liabilities	336	813	2,002	3,268	5,881
Value of swaptions	4,300	4,256	4,169	4,080	3,957
Cash	-2,167	-2,167	-2,167	-2,167	-2,167
Swaption + Cash	2,133	2,089	2,002	1,913	1,791
VaR	-1,797	-1,276	0	1,355	4,091

Table 8.22: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the third factor

Description	Increase in curvature		Base	Decrease in curvature	
Confidence level	99.5%	90%		90%	99.5%
Value of liabilities	88	428	2,002	5,241	12,654
Value of swaptions	4,340	4,282	4,169	4,041	3,833
Cash	-2,167	-2,167	-2,167	-2,167	-2,167
Swaption + Cash	2,173	2,115	2,002	1,874	1,666
VaR	-2,085	-1,687	0	3,367	10,989

Table 8.23: Value at risk: changes in the starting fund values

Description	Increase in starting fund values (41.71%)	Base	Decrease in starting fund values (–24.62%)
Confidence level	99.5%		99.5%
Value of liabilities	2,837	2,002	1,509
Value of swaptions	4,169	4,169	4,169
Cash	-2,167	-2,167	-2,167
Swaption + Cash	2,002	2,002	2,002
VaR	835	0	-493

Table 8.24: Value at risk: changes in implied volatilities of swaptions

Description	Increase in implied volatility of swaptions	Base	Decrease in implied volatility on swaptions
Confidence level	99.5%		99.5%
Value of liabilities	4,722	2,002	1,297
Value of swaptions	12,569	4,169	3,564
Cash	-2,167	-2,167	-2,167
Swaption + Cash	10,402	2,002	1,398
VaR	-6,831	0	-101

8.5.4.2 Figure 8.9 shows the VaR under the various scenarios graphically.

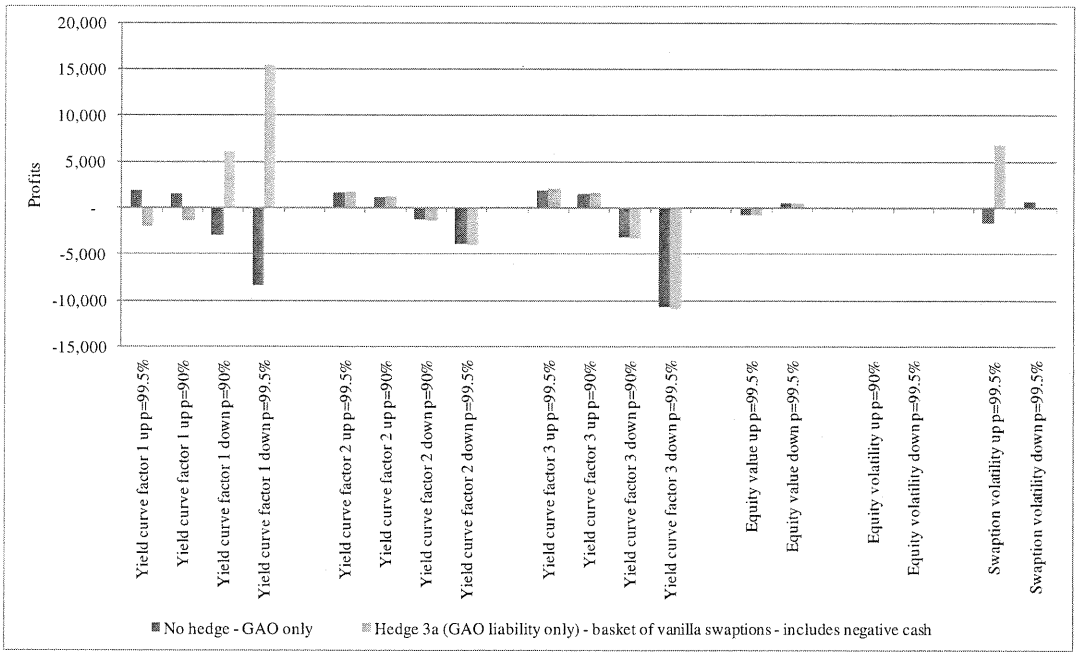


Figure 8.9: Comparison of the VaR for GAO liabilities only when (a) no hedge is held in respect of the risks; and (b) a hedge consisting of receiver swaptions and cash is held

8.5.4.3 Figure 8.9 shows a seemingly effective hedge for term structure movements of the first factor. By holding this portfolio of swaptions as a hedge, we have reduced most of the downside risks of a GAO, even if the equity risk component is ignored, and also created some very large upside potential. The large upside potential is created because this hedge consists of a series of swaptions that can be exercised independently of each other, whereas the liability is a single option. This extra

optionality on the hedge assets overhedges the liability and allows for higher increase to the asset values than the liabilities in the case of a decrease in interest rates. The swaption portfolio's value is marginally reduced when the value of the liabilities increased for some second- and third-factor movements of the term structure. This marginally increases the VaR.

8.5.4.4 The analysis in the paragraph above ignores the loss of time value over time. In the bond hedges that were previously investigated, no or little value is lost over time as the bond portfolios unwound through coupon and redemption payments. With the swaption hedge, however, we note that a significant proportion of the initial hedge cost of R4,169 is the time value of the option. The option itself has an intrinsic value of R0. If the interest rates do not change in value until the redemption of the policy, then this R4,169 is lost. This is partially offset by the liability value that would decrease by R2,002. It nevertheless represents an overall loss of R2167, the amount that was borrowed to fund the purchase of the swaption portfolio. This cost should be taken into account in determining whether or not the hedge should be implemented.

8.5.4.5 Furthermore, the conclusions drawn above are based on the assumption of deterministic equity returns, which is not a realistic assumption. Figure 8.10 and Table 8.25 illustrate the financial effects of implementing this hedge if equities are not forced to be deterministic. The liabilities values here include the GMV and cross-term reserves. These add up to R4,349. The swaption portfolio has a value of R4,169. The difference of R180 is assumed to be invested in cash.

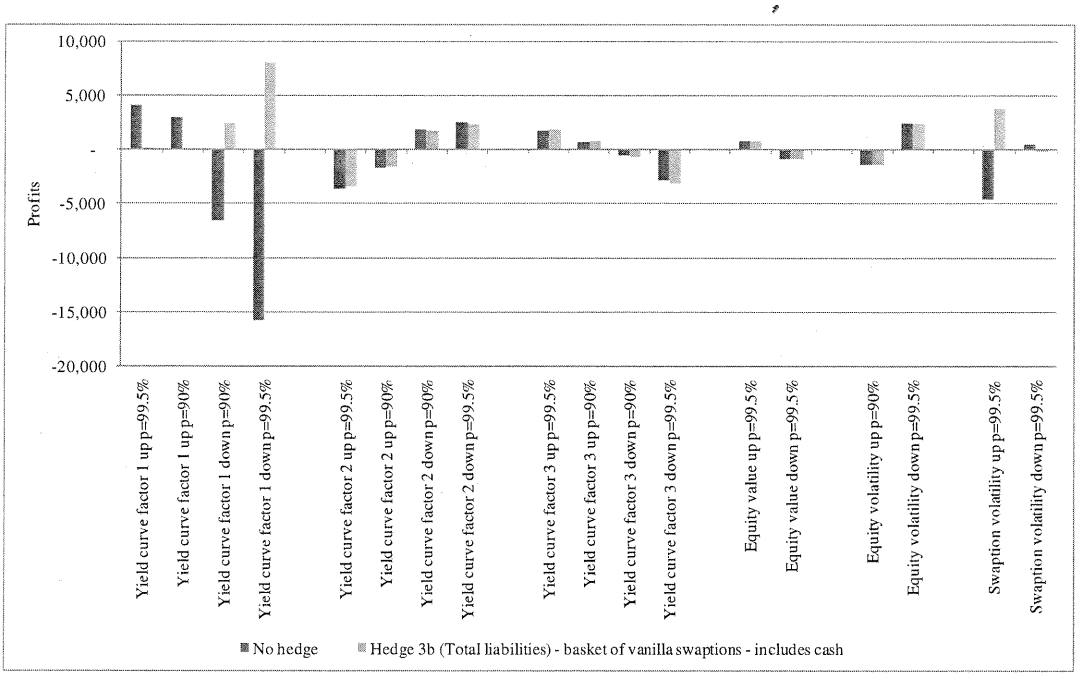


Figure 8.10: Comparison of VaR of all guarantee liabilities when (a) no hedge is held in respect of the risks; and (b) a hedge consisting of receiver swaptions and cash is held

Table 8.25: Comparison of VaR of all guarantee liabilities when (a) no hedge is held in respect of the risks; and (b) a portfolio of swaptions and cash

Description	Confidence level	Value of liabilities	Value of swaptions	Swaption + cash	VaR
Base		4,349	4,169	4,349	-
Increase in term structure	99.5	266	281	461	195
Increase in term structure	90	1,387	1,313	1,493	105
Decrease in term structure	90	10,935	13,212	13,392	2,457
Decrease in term structure	99.5	20,086	27,913	28,093	8,007
Inversion of term structure	99.5	7,983	4,300	4,480	-3,504
Inversion of term structure	90	6,063	4,256	4,436	-1,627
Flattening of term structure	90	2,529	4,080	4,260	1,731
Flattening of term structure	99.5	1,831	3,957	4,137	2,306
Increase in curvature	99.5	2,668	4,340	4,520	1,852
Increase in curvature	90	3,663	4,282	4,462	799
Decrease in curvature	90	4,874	4,041	4,221	-653
Decrease in curvature	99.5	7,188	3,833	4,012	-3,175
Increase in starting fund values	99.5	3,574	4,169	4,349	775
Decrease in starting fund values	99.5	5,217	4,169	4,349	-869
Increase in implied volatility of equities	90	5,733	4,169	4,349	-1,385
Decrease in implied volatility of equities	99.5	1,900	4,169	4,349	2,448
Increase in implied volatility of swaptions	99.5	8,952	12,569	12,749	3,797
Decrease in implied volatility of swaptions	99.5	3,908	3,564	3,744	-164

8.5.4.6 The implementation of this hedge, aimed at only hedging the interest-rate risks underlying the GAO benefit, has unwittingly also hedged the interest-rate risks of the GMV benefit (up to the first factor of the term-structure movements). The GAO and GMV benefits are both at risk to a decrease in interest rates. The overhedging described in Paragraph 8.5.4.3 above has fortuitously also reduced the interest-rate risk of the GMV benefit. Term-structure movements of factors 2 and 3 are not hedged because the swaptions are constructed to hedge the post-retirement interest-rate risks associated with the annuitisation. The majority of the interest-rate risks in the guaranteed liabilities are associated with the GMV benefit, which is affected by the short-term interest rates. This means that the most of the guarantee liabilities are not protected from inversion and curvature changes of the term structure of interest rates, associated with the second and third factors. The hedge also does not provide any protection against changes in equity values and implied volatilities. The comment on time value of option in ¶8.5.4.4 still applies.

8.6 Swaptions and Equity Options

8.6.1 From the previous section, we see that although swaptions have largely mitigated some of the interest-rate risks underlying GAO contracts, it nevertheless leaves the life insurer unprotected against equity implied volatilities. In this section, we investigate the effectiveness of a hedge portfolio that consists of swaptions, equity puts and cash. The equity puts provide protection against the vega risks associated

with the GMV benefit. The swaptions provide protection against the rho risks associated with GMV and GAO.

8.6.2 It should be noted that the equity puts that are freely traded in today's markets cannot hedge the GMV exposure adequately. This is because equity options in the South African market rarely have maturity dates that are longer than three years. GMV liabilities, on the other hand, have terms that are over 30 years (25 years in our hypothetical portfolio). The vega that is available on the equity puts is therefore much lower than that available on the liabilities. For this reason, short-dated equity puts do not provide an adequate hedge against changes in equity implied volatility, but merely serve to dampen the effect of an increase in the implied volatility. More than a unit of equity put needs to be held for each unit of equity in the fund if the full vega risk is to be fully hedged.

8.6.3 The longest available equity put option is used to hedge the implied volatility on equity options. In South Africa, there is reasonable liquidity for equity options out to approximately 3 years. The stochastic vega for liabilities can be calculated using a similar method to that of the stochastic duration in (8.1):

$$\upsilon_L = \frac{V_{-\Delta\sigma} - V_{+\Delta\sigma}}{2(V_0)\Delta\sigma},$$

where:

- υ_L is the vega of the liabilities;
- V_0 is the value of the liabilities under the base calibration;
- $V_{-\Delta\sigma}$ and $V_{+\Delta\sigma}$ are the values of the liabilities under a parallel decrease and increase in the equities implied volatilities of $\Delta\sigma$ respectively; and
- $\Delta\sigma$ is chosen to be 200 bps in this instance.

This represents the sensitivity of the value of the liabilities to the change in the level of the implied volatilities.

8.6.4 This calculation of vega requires the recalibration of the ESG with the new volatility term structures and the corresponding re-evaluation of liabilities. The results are in Table 8.26.

Table 8.26: The value of guarantees liabilities for a reduction and increase of 200 bps to the term structure of equity implied volatilities

	200 bps increase in implied volatilities	Original liabilities	200 bps reduction in implied volatilities
GMV reserve	3,188	2,785	2,411
GAO reserve	1,402	1,406	1,410
Cross-term reserve	175	157	141
Total guarantee reserve	4,765	4,349	3,962

8.6.5 This resulted in a vega of 4.6 for the total guarantees liability. The vega for the equity put can be calculated using a closed-form solution derived from the Black–Scholes formula:

$$\upsilon_A = S_0\left(\frac{t}{2\pi}\right)^{0.5} \mathrm{e}^{-\frac{d_1^2}{2}},$$

where

- υ_A is the vega of a put option;
- $d_1 = \dfrac{\ln(K/S)+(\sigma^2/2)t}{\sigma\sqrt{t}}$;
- $t = 3$;
- $\sigma = 0.251$;
- $S_0 = 1$; and
- $K = 1.318$ is the accumulated value of S_0 at the forward rate.

8.6.6 This results in a vega of 0.6748 for a three-year put with unit stock value and the strike is at-the-forward. To hedge the vega risk, we need to have

$$4349\frac{\upsilon_L}{\upsilon_A} = 29,750$$

units of equity put. The price of the option, determined using the Black–Scholes formula, is R0.172 per unit value at time 0. The value of the equity put portfolio is therefore R5,126.

8.6.7 The rho of this equity put can likewise be calculated with closed-form solution:

$$\rho_A = -Kt\,\mathrm{e}^{-rt}\,N(-d_2) = -1.76.$$

8.6.8 The rho of the liabilities is analogous to the stochastic duration calculated in Section 8.3.1, which is 61.7 years. The value of the liabilities is expected to increase by

$$\frac{61.7}{100}4,349 = 2,682$$

for every 100 bps parallel reduction in the term structure of interest rates. For the same change to the yield curve, the value of the equity put options is expected to increase by

$$\frac{1.76}{100}29,750 = 523$$

for the same reduction in the term structure of interest rates. This difference represents the rho risk that is not hedged. This amount is approximately R2159 for every 100 bps. Swaptions are used to eliminate this. Equation (8.1) is used to calculate the duration of the swaption portfolio in Section 8.5. This swaption portfolio has duration of 19.68 years, or an increase in price of R0.1968 for each 100 bps reduction in the term structure of interest rate per unit price swaption portfolio. To hedge the remaining rho risk, a swaption portfolio to the value of

$$\frac{2159}{0.1968} = 10,971$$

must be held.

8.6.9 The swaption portfolio and the put options have a combined value of R16,097. The guarantees reserve is only R4349. To hedge this guarantees liabilities, the insurer needs to borrow R11,748. Table 8.27 to Table 8.32 and Figure 8.11 illustrate the effectiveness of this hedge against the various market shocks.

Table 8.27: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the first factor

Description	Increase in term structure		Base	Decrease in term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	131	872	2,785	6,418	11,409
GAO reserve	134	487	1,406	3,854	7,010
Cross-term reserve	1	28	157	663	1,668
Total guarantee reserve	266	1,387	4,349	10,935	20,086
Cash value	-11,748	-11,748	-11,748	-11,748	-11,748
Swaption value	740	3,455	10,971	34,770	73,460
Equity option value	4,351	4,841	5,126	5,279	5,367
Value of assets	-6,658	-3,452	4,349	28,301	67,078
Value at risk	-6,924	-4,839	-	17,366	46,992

Table 8.28: Value at risk: changes in the term structure of interest rates when allowing only for the changes in the second factor

Description	Inversion of term structure		Base	Flattening of term structure	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	4,526	3,705	2,785	1,492	1,145
GAO reserve	3,015	2,086	1,406	967	648
Cross-term reserve	442	272	157	69	39
Total guarantee reserve	7,983	6,063	4,349	2,529	1,831
Cash value	-11,748	-11,748	-11,748	-11,748	-11,748
Swaption value	11,316	11,201	10,971	10,737	10,415
Equity option value	4,221	4,697	5,126	5,424	5,856
Value of assets	3,788	4,150	4,349	4,413	4,523
Value at risk	-4,195	-1,913	-	1,885	2,692

Table 8.29: Value at risk: changes in the term structure of interest rates when allowing only for the changes in of the third factor

Description	Increase in curvature		Base	Decrease in curvature	
Confidence level	99.5%	90%		90%	99.5%
GMV reserve	2,258	2,841	2,785	2,017	2,116
GAO reserve	359	718	1,406	2,671	4,772
Cross-term reserve	51	104	157	186	300
Total guarantee reserve	2,668	3,663	4,349	4,874	7,188
Cash value	-11,748	-11,748	-11,748	-11,748	-11,748
Swaption value	11,422	11,270	10,971	10,634	10,086
Equity option value	4,943	5,086	5,126	5,045	5,113
Value of assets	4,617	4,607	4,349	3,931	3,451
Value at risk	1,949	945	-	-942	-3,736

Table 8.30: Value at risk: changes in the starting equity values

Description	Increase in starting equity values (41.71%)	Base	Decrease in starting equity values (–24.62%)
Confidence level	99.5%		99.5%
GMV reserve	1,697	2,785	3,824
GAO reserve	1,782	1,406	1,184
Cross-term reserve	95	157	209
Total guarantee reserve	3,574	4,349	5,217
Cash value	-11,748	-11,748	-11,748
Swaption value	10,971	10,971	10,971
Equity option value	1,825	5,126	9,052
Value of assets	1,047	4,349	8,274
Value at risk	-2,527	-	3,057

Table 8.31: Value at risk: changes in implied volatilities on swaptions

Description	Increase in implied volatility of swaptions	Base	Decrease in implied volatility of swaptions
Confidence level	99.5%		99.5%
GMV reserve	4,426	2,785	2,551
GAO reserve	3,994	1,406	1,234
Cross-term reserve	531	157	123
Total guarantee reserve	8,952	4,349	3,908
Cash value	-11,748	-11,748	-11,748
Swaption value	33,078	10,971	9,380
Equity option value	5,408	5,126	5,407
Value of assets	26,738	4,349	3,039
Value at risk	17,786	-	-869

Table 8.32: Value at risk: changes in implied volatilities on equity options.

Description	Increase in implied volatility of equities	Base	Decrease in implied volatility of equities
Confidence level	90%		99.5%
GMV reserve	4,106	2,785	496
GAO reserve	1,417	1,406	1,362
Cross-term reserve	211	157	42
Total guarantee reserve	5,733	4,349	1,900
Cash value	-11,748	-11,748	-11,748
Swaption value	10,971	10,971	10,971
Equity option value	6,410	5,126	2,151
Value of assets	5,632	4,349	1,373
Value at risk	-102	-	-527

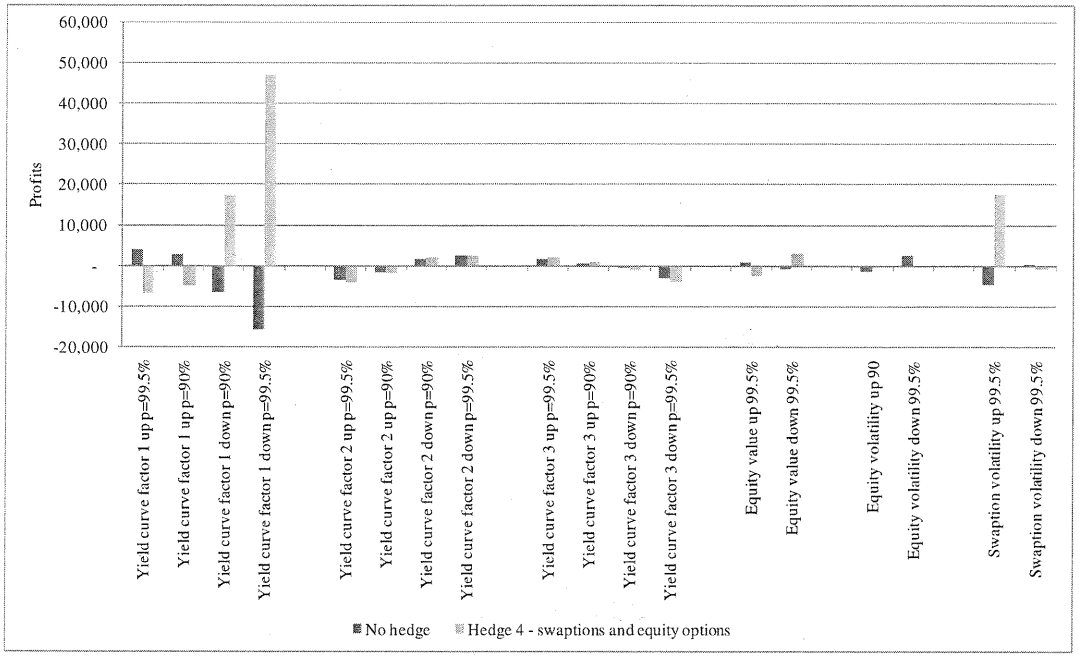


Figure 8.11: Comparison of VaR of all guarantee liabilities when (a) no hedge is held in respect of the risks and (b) rho and vega hedge consisting of receiver swaptions, equity puts and cash

8.6.10 The implementation of this hedge is effective in reducing large potential losses. Because this hedge is calibrated to eliminate the rho and vega risks, one expects that interest rate movements of the first factor to be significantly reduced. This did not happen. The rho hedge only eliminates small parallel shifts in the term structure, but not bigger movements such as those of the 90th and 99.5th percentiles. By holding this portfolio of swaptions and short-dated equity puts as a hedge, the downside risks associated with the downward shifts of the yield curve has been converted into large upside potential. However, the previous small mismatch profits associated with the increase in the yield curve has been replaced with a larger loss. This increased mismatch is created by the existence of swaptions that can be exercised independently of each other, whereas the liability is a single option. This increased optionality increases the volatility of the price.

8.6.11 The hedge introduced small additional risks to interest-rate movements of second and third factors. The hedge does offer significant protection against increases in equity volatility, but introduced higher VaR against equity values.

8.6.12 The comments on time value of options in ¶8.5.4.4 still apply. Because the swaption is currently out of the money, the time value of the swaption is equal to the option price of R10,971. This would be unwound over the 25 years until the maturity date of the policy. The equity put option is also out of the money, so the time value of the put option is equal to its price of R5126. This unwinds over only three years, after which the position needs to be rebalanced again at a significant cost. This is a very expensive hedge that is unsustainable. In the author's experience, life insurance companies rarely hedge their vega risks – there are no equity puts with suitable option term available in the market.

8.7 Summary of Findings

8.7.1 In this chapter, the sensitivities of the guaranteed liabilities to various risk factors are calculated. It's discovered that the guarantees are most sensitive to interest rate term structure movements of the first factor.

8.7.2 Four hedges are investigated to address the risks associated with these guarantees:
- duration hedge using a ZCB;
- duration hedge by shorting a 29 year ZCB and longing a 30 year ZCB;
- a portfolio of vanilla swaptions and short position in cash; and
- a portfolio of equity puts, vanilla swaptions and a short position in cash.

8.7.3 The first hedge is dynamic and successfully reduced the VaR of the interest rate term structure change of the first and second factors. Unfortunately this hedge is not implementable because it requires a ZCB that is significantly longer than what is available in the market.

8.7.4 The second hedge, like the first, is dynamic and also reduces the VaR associated with the same risks. This hedge is also impractical to implement because it requires a very large long position in a long-dated ZCB and simultaneously a large short position in a shorter-dated ZCB.

8.7.5 The third hedge is aimed only at hedging the GAOs but not the GMVs. It can be implemented statically and succeeded in converting the potential losses, associated with the first factor reduction of the term structure, into potential gains. Conversely, it also converted the potential gains, associated with the first factor reduction of the term structure, into small losses. Fortuitously, when this hedge is considered taking GMVs into account, the downside risk associated with the first factor movements of the term structure is removed. This is by accident and not by design. The VaR associated with the implied volatility of swaptions is also reduced. The cost of the swaptions is also higher than the value of the GAO liabilities. Therefore, the purchasing of this portfolio of swaptions necessitates a loan. Because the swaptions are out the money, their time values are equal to the price of the swaption portfolio. If the swaptions remains out the money at maturity, a net loss equal to the value of the loan will be realised.

8.7.6 The last hedge is aimed at hedging the rho and equity-vega risks associated with the guarantees. Short-date equity puts are used to address the equity-vega risks. These puts also reduced the VaR associated with interest rate movements of the first factor. Swaptions are used to reduce the remainder of this risk. This hedge, like the previous one, requires a loan to purchase. It is also more expensive than the previous hedge and the equity puts require regular rebalancing as they expire. The cost associated with this hedge makes it unsustainable.

8.8 Limitations of the Hedges

8.8.1 The hedges proposed above are tested in the absence of non-market risks, such as mortality and lapses. The presence of other market risks, such as credit-default risk and credit-spread risks, require additional risk-mitigating strategies. Also, to the extent that the actual experience deviates from the assumptions, some of the hedges

may need to be rebalanced regularly. Such rebalancing is subject to bid–offer spread. Over-the-counter instruments, such as swaptions, are subject to larger bid–offer spreads than standard instruments like bonds and equities. However, swaptions may be implemented statically — the only rebalancing needed is in respect of changes in demographic experience and fund values — but the duration hedges require more frequent rebalancing as market conditions change. An investigation into these spreads is outside of the scope of this research.

8.8.2 Furthermore, the hedges are investigated under the settings of sensitivity testing — each risk testing independently of other risks. In real life, independent shocks are rare — when extreme events happen in the market, they tend to affect all variables at the same time, as witnessed in the credit crisis in the second half of 2008. The calibration of such a scenario to a specific probability would necessitate the construction of a real-world scenario generator, which falls outside of the scope of this research.

Chapter 9. Longevity Risk

9.1 Sensitivity to Longevity

9.1.1 The hypothetical portfolio of liabilities introduced in Section 6.2 assumes that the policy that's modelled in this research is one of many in a well diversified portfolio so that the random fluctuations associated with mortality are reduced in aggregate. The actuary should be confident that the year-to-year fluctuation in mortality rates is minimal, and that years with good mortality experience relative to expectations are in most cases cancelled out by years with bad experience.

9.1.2 Longevity risk, on the other hand, is not diversifiable. This is the systematic reduction in mortality rates caused by improvements in medical care, resulting in longer lifespan of the population. This increase in lifespan prolongs annuity payments to the annuitants, so that the insurer needs to set aside more provisions now to meet its payments obligations as they become due.

9.1.3 The sensitivity of GAO liability to improvements in mortality rates was tested by Boyle & Hardy (2003). In that paper, the authors note that using a(55), the life expectancy of a life aged 65 is 14.3 years. Life expectancy increased by 2.6 years from a(55) to PMA80(C10) and over 5 years from a(55) to PMA92(C20) tables. At these mortality levels, the break-even interest rates using these tables are 5.61%, 7% and 8.2% per year respectively. The improvements in mortality required higher yields to keep the guarantee from moving into the money.

9.1.4 In this section, we test the sensitivity of the liabilities to longevity. We previously assumed in Chapter 6 that the maximum payment period is 30 years, due to a limitation on the ESG. This assumption clearly limits our ability to investigate the impact of longevity. To get around this limitation, we assume that the projected term structure has a constant spot rate after 30 years. Although this assumption is not strictly arbitrage free, it should not be too problematic as we are more interested in the impact of longevity than interest rates. This extension of the payment period increased our liabilities. The changes are listed in Table 9.1.

Table 9.1: Guarantees reserves with and without a limit of 30 years on annuities

	GMV	GAO	Cross term	Total
Limited to 30 years	2,785	1,406	157	4,349
Paid to end of mortality table	2,785	1,432	160	4,377

9.1.5 The changes to both the GAO reserves and the cross-term reserves are attributed to the improvement in mortality. This represents an increase in GAO liability from R1563 to R1592, or 2%.

9.1.6 We then apply a mortality improvement factor of 2.5% each year to the post-vesting mortality rates. This is approximately consistent with the factors calculated by the Continuous Mortality Investigation Bureau (1999) for the UK annuitant population. These UK factors are used because no South African equivalent is available. The new mortality rate is calculated as follows:

$$\tilde{q}_{x,t} = q_x(1-0.025)^t,$$

where:

- $\tilde{q}_{x,t}$ is the mortality rate at time t for a life then aged x;
- q_x is the mortality rate in SAIML98.

9.1.7 Towards the end of the writing of this dissertation, Continuous Mortality Investigation (2010) recommends an annual long term mortality improvement factor of between 1.0% and 1.5%. This revision of improvement factor is not reflected in this research.

9.1.8 The inclusion of the 2.5% annual improvement in mortality increases the life expectancy at retirement from 15.4 years to 20.6 years. The effects on the liability values are summarised in Table 9.2.

Table 9.2: The effects of mortality improvements on the value of guaranteed liabilities

	GMV	GAO	Cross Term	Total
No mortality improvement	2,785	1,432	160	4,377
With mortality improvement	2,785	5,634	510	8,930

9.1.9 This is an increase in total GAO liability from R1592 to R6144, or an increase of 286%. This demonstrates that the GAO liability is extremely sensitive to longevity.

9.2 Hedging Longevity Risk

9.2.1 Mortality and longevity bonds can be purchased from the market, but these are mostly based on a cohort of lives living in the UK — to use these to hedge South African lives would introduce significant basis risk. There is also limited depth in this instrument, so hedging mortality risks with this instrument is highly unrealistic.

9.2.2 As noted in Section 3.2.6, Milevsky & Promislow (2001) proposed the natural hedging of GAOs by selling life insurance. Boyle & Hardy (*op. cit.*) note that there are limitations to this method for various reasons.

9.2.3 This indicates that there is currently limited ability to reduce longevity risks. Additional capital should be held to protect the company against uncertain improvements in longevity.

Chapter 10. Conclusions and Future Research

10.1 Conclusions

10.1.1 This dissertation provides a method of calculating the market-consistent value of a simple portfolio of GAOs. A risk-neutral ESG was constructed in Matlab to generate stochastic economic scenarios to facilitate this calculation. A one-factor Black–Karasinski model was used for the projection of interest rates. A Black–Scholes model, extended to incorporate a term structure of volatility, was used to project future equity returns. The model allows for correlation between interest rates and equity returns. A simple Excel model was constructed to facilitate the projection of the guarantees-related cashflows under each scenario. The guarantees-related payments are discounted at the short rates simulated under each scenario. The market-consistent value is calculated as the arithmetic average of the discounted values across all of the scenarios.

10.1.2 Various authors have identified the main risks of GAOs. These are:
- decrease in the term structure of interest rates;
- increase in the fund returns until the time of retirement; and
- increase in life expectancy, or longevity.

Other risks include:
- increase in implied volatility of swaptions; and
- increase in implied volatility of equities (although this only affects the GMV benefits).

10.1.3 The author devised methods to calculate the potential variability of each of these risk factors. The sensitivity of the market-consistent liability of the GAO portfolio to these risk factors is calculated. It was found that parallel decrease in interest rates is the largest market risk for the portfolio of GAOs, followed by the implied volatility of interest rates.

10.1.4 Four simple hedge portfolios were proposed to hedge the market-risk exposures of the GAO portfolio:
- duration hedge using a ZCB;
- duration hedge by shorting a 29 year ZCB and longing a 30 year ZCB;
- a portfolio of vanilla swaptions; and
- a portfolio of equity puts and vanilla swaptions.

10.1.5 The first hedge is impossible to implement in current environments because the duration of the liabilities is higher than that of the longest available bond. The second hedge is difficult to implement because it requires the insurer to simultaneous hold very large long and short position in ZCB. Both of these hedges, even if implemented, are at risk of the feedback loop (Pelsser, 2003). The last hedge requires the rolling forward of equity puts because the equity options that are traded in the market are too short when compared with liabilities. This makes the last hedge prohibitively expensive.

10.1.6 For this portfolio of GAOs, at the end of December 2007, it was found that the portfolio of vanilla swaptions is capable of reducing most of the risk associated with a parallel reduction of interest rates. It did very little to reduce the inversion and curvature risks of interest rates, nor is it effective in reducing the risk associated with implied volatility of equities. This reduction in interest-rate risk comes at the cost of purchasing the swaptions, which is completely time value at the outset and is unwound over time. It is irrecoverable should the swaption remain out the money. This is the most successful hedge of the four.

10.1.7 The sensitivity of GAOs to changes in future life expectancy is also investigated. For a modest 2.5% relative improvement in mortality each year, the cost of the GAO is expected to increase almost 300%. There is currently no deep and liquid market instrument that can be used for hedging this risk.

10.2 Comparison with Prior Research

10.2.1 Boyle and Hardy (2003) identify three main factors that cause the embedded guarantee of the GAO to become onerous. These are falling interest rates, longevity and strong equity performance. They conclude that falling interest rates are by far the most risky of the three. These findings are based on the attribution of the actual cost of the payoff of a GAO contract into the three components. This investigation is deterministic and does not account for the cost of the option.

10.2.2 Boyle and Hardy (*op. cit.*) also show that interest-rate volatility has minimal effect on the value of the guarantee except where the guarantee is very close to at the money. They do sensitivity tests of the parameters in the Heath–Jarrow–Morton (1992) model. The base case is not marked-to-market. It is not clear how severe the sensitivity tests are relative to the real-world dynamics of the swaption-implied volatility.

10.2.3 Their findings are not consistent with the findings of this research. This research found that falling interest rates are the biggest risk for a block of GAOs, followed by an increase in interest-rate volatility. This is true with and without the GMV components. These differences can be attributed to the use of fair-valuation (as opposed to deterministic) and the calibration of the stresses to historical data (as opposed to stress testing the parameters of a model).

10.2.4 The results of this research showed an increase in cost of approximately 300% with the inclusion of a 2.5% p.a. mortality improvement factor. Ballotta & Haberman (2003) show an increase in GAO liabilities of approximately 100% to 200% when switching the mortality table from PA90 to PMA92, depending on the interest-rate environment. This only captures the improvement in mortality via the revision to mortality tables over a period of time, but does not capture future improvements. Ballotta & Haberman (2006) note that the incorporation of stochastic mortality into the valuation of GAOs further increases their costs by approximately 26%. The results of this research are not directly comparable with the results from Ballotta & Haberman (*op. cit.*).

10.3 Gaps for Future Research

10.3.1 The interest-rate model in the ESG may be extended to include more factors, so that the implied volatility on swaptions may be better matched. The LIBOR model may also be implemented as a more flexible alternative for interest-rate modelling. Similarly, the equities model may be extended to allow the volatility surface to be matched. The correlation of the different models may be modelled using copulas, which allows the modelling of non-stationary correlation structures.

10.3.2 The choice of the risk-free curve in the modelling of assets and liabilities is a contentious one. Some practitioners argue that because derivative products are priced on the swap curve, liabilities with embedded options should similarly be priced and hedged on the swap curve. Conversely, some argue that the reserving should be based on the true risk-free curve, so that the government bond curve is better suited and more conservative for this purpose. Some argue that the choice of curve should be determined by the purpose: swap curve for pricing and hedging and government curve for reserving. However, if the choice of the yield curve is determined by the purpose of the calculation, it goes against the idea of a single and objective value that is fundamental to market consistency. The author makes no further judgement on the appropriateness of each approach. The International Accounting Standards Board and the Groupe Consultatif Actuariel Europeen (the actuarial body representing the European Union which is also the main instigator for the drafting of the Solvency II standards) have done a significant amount of research on this topic and more guidance is expected to be provided in the near future.

10.3.3 Four hedging strategies are investigated, but three of them are impractical. The last strategy, involving a portfolio of swaptions and a short position in cash, is only successful in reducing interest-rate risk but not equity- and equity-option related risks. As more complex structured assets (such as equity-linked swaption or varying notional swaption) gain liquidity in the open market in future, these may become more viable as hedging instruments for GAOs. The suitability of these instruments should be investigated.

10.3.4 This research focused mainly on the market-consistent valuation of the guarantees and some simple methods of hedging the market risks underlying this liability. Demographic risks, such as those of mortality and lapse, have not been investigated.

10.3.5 There is currently a lack of guidance on the definition of market consistency in the absence of a deep and liquid market for demographic risks, but some research has been published, notably Thomson (2005) and Thomson (unpublished). More guidance is expected from the International Accounting Standards Board with the release of phase 2 of IFRS4. This lack of a deep and liquid market for demographic risks is also the reason for a lack of hedging instruments. As more market instruments become available, more research should be done on the mitigation of these risks.

10.3.6 At the time of writing, Solvency II is still in the development stages. At the final stages of the writing of this dissertation, some insurers have provided some very convincing arguments to allow some credit for liquidity premium in the market-consistent valuation. Instead of valuing the liabilities on a risk-free yield curve, they argue that a liquidity premium should be included. While this is against the principles of risk-neutral pricing, these insurers argue that the risk-free curve is derived from

very liquid government bonds or swaps. Insurance contracts such as life annuities and GAOs, are very illiquid liabilities and allow the insurers to invest in other less liquid assets (such as corporate bonds) to earn higher yields. Whilst these higher yielding assets, by their very nature, include an allowance for expected defaults and default risk premium, they also include a liquidity premium because their issues are usually much smaller. The debate of whether or not liquidity premium should be allowed for is beyond the scope of this research, but if extra allowance is to be had, then the method of splitting of the extra yield into these three categories should be researched.

Acknowledgement

The author would like to thank Professor Robert Thomson for his knowledge and support through what has easily been the longest single undertaking in my life. His help with the review of this dissertation is also acknowledged. Any errors in this dissertation are my responsibilities and mine alone.

In no particular order of importance, the author would also like to thank:

- Linda McRae-Samuels, for making the administrative tasks a breeze;
- the National Research Foundation, for their generous financial support;
- Paul Lancaster for the assistance in reviewing parts of this dissertation and the interesting discussions;
- David Jewell for providing such an interesting thesis topic and the wonderful insights, without which this dissertation may not have begun at all;
- Philip Harrison, for the interesting discussions;
- Coldplay and U2, whose music kept me company through the late nights and early mornings;
- William Blake, for the meaning of life;
- Andries Groenewald, Angela Rademeyer and Simona Levet, for their friendship, for their listening to my woeful moans, for the messages of support and understanding, for letting me see the light at the end of the tunnel; and
- Colin Wakefield, for making me believe that the light may belong to a train.

Last, but not least, I am truly thankful for the daily assistance that my parents have given me. Your support and love know no boundaries and have shaped me into the person that I am today.

References

Ahlgrim, KC, D'Arcy, SP & Gorvett, RW (unpublished). Modelling of economic series coordinated with interest rate scenarios. *Casualty Actuarial Society* and *Society of Actuaries*

Audley, D, Chin R & Ramamurthy, S (2002). Term structure modelling. In Fabozzi (2002)

Ballotta, L & Haberman, S (2003). Valuation of guaranteed annuity conversion options. *Insurance: Mathematics and Economics* **33**, 87–108

Ballotta, L & Haberman, S (2006). The fair valuation problem of guaranteed annuity options: The stochastic mortality environment case. *Insurance: Mathematics and Economics* **38**, 195–214

Ballotta, L & Haberman, S (unpublished). Guaranteed annuity conversion options and their valuation

Barrie and Hibbert (2008), Global market-consistent asset model calibrations – end-December 2007, Barrie and Hibbert Limited

Benhamou, E (unpublished). Brennan and Schwartz Model (1982). Goldman Sachs Internatioal

Bierwag, GO, Kaufman, GC & Toevs, A (eds.) (1983). *Innovations in bond portfolio management duration analysis and immunisation.* JAI Press, Greenwich

Biffis, E & Millossovich, P (2006), The fair value of guaranteed annuity options, *Scandinavian Actuarial Journal*, **1**, 23–41

Black, F (1976). The pricing of commodity contracts, *Journal of Financial Economics* **3**, 167–179

Black, F, Derman, E & Toy, W (1990). A one factor model of interest rates and its application to the Treasury bond options. *Financial Analyst Journal* (January–February), 33–39

Black, F & Karasinski, P (1991). Bond and option pricing when short rates are log-normal. *Financial Analyst Journal* (July–August), 52–59

Black, F & Scholes, M (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* **81(3)**, 637–54

Bolton, MJ, Carr, DH, Collis, PA, George, CM, Knowles, VP & Whitehouse, AJ (1997). Reserving for annuity guarantees. *The Report of the Annuity Guarantees Working Party*. Institute of Actuaries, London, UK

Boyle, P & Hardy, M (2003). Guaranteed annuity options. *ASTIN Bulletin* **33 (2)**, 125–52

Brace, A, Gatarek, D & Musiela, M (1997). The market model of interest rate dynamics. *Mathematical Finance* **7 (2)**, 127–47

Brennan, M & Schwartz, E (1977). Savings bonds, retractable bonds and callable bonds. *Journal of Financial Economics* **5**, 67–88

Brennan, M & Schwartz, E (1979). A continuous time apporach to the pricing of bonds. *Journal of Banking and Finance* 3, 133–55

Brennan, M & Schwartz, E (1980). Conditional predictions of bond prices and returns. *Journal of Finance* **35**, 405–17

Brennan, M & Schwartz, E (1982). An equilibrium model of bond pricing and a test of market efficiency. *Journal of Financial and Quantitative Analysis* **17 (3)**, 301–29

Brennan, M & Schwartz, E (1983). Duration, bond pricing and portfolio management. In Bierwag, Kaufman & Toevs (1983)

Buetow, GW (Junior), Fabozzi, FJ & Sochacki, J (2002). A review of no arbitrage interest rate models. In Fabozzi (2002)

Cairns, AJG (unpublished). A family of models or the term-structure of interest rates: an application to guaranteed annuity options. Presentation at the Insurance, Mathematics and Economics 2002 Conference, Lisbon

Cairns, AJG (2004). *Interest Rate Models.* Princeton University Press, New Jersey

Continuous Mortality Investigation Bureau (1999). *Mortality Improvement Model for Use with the '92' Series of Mortality Tables.* Faculty and Institute of Actuaries

Continuous Mortality Investigation Bureau (2010). *CMI Working Paper 49. CMO Mortality Projections Model, 'CMI_2010'.* Faculty and Institute of Actuaries

Cox, J, Ingersoll, J & Ross, S (1985). A theory of the term structure of interest rates. *Econometrica* **53 (2)**, 385–408

Culbertson, JM (1957). The term structure of interest rates. *Quarterly Journal of Economics* **71**, 485–517

Dardis, T, Beeson, M, Blaske, S, Campbell, N, Craighead, S, Davlin, M, Howes, T, Ipsen, R, Lefkowitz, L & Lenius, B (2008). *Modeling Efficiency Bibliography for Practicing Acturies.* American Academy of Actuaries

Duffie, D (2002). *A Short Course on Credit Risk Modeling with Affine Processes.* Associazione Amici della Scuola Normale Superiore

Fabozzi, FJ (eds.) (2002). *Interest Rate, Term Structure, and Valuation Modelling.* John Wiley & Sons, New Jersey

Fitton, P & McNatt JF (2002). The four faces of an interest rate model. In Fabozzi (2002)

Ford, A, Benjamin, S, Gillespie, RG, Hager, DP, Loades, DH, Rowe, BN, Ryan, JP, Smith, P & Wilkie, AD (1980). Report of the maturity guarantees working party. *Journal of the Institute of Actuaries* 107, 103–212

Götsch, I (2006). *Libor Market Model: Theory and Implementation.* VDM Verlag Dr Müller. E. K. Und Lizenzgeber, Saarbrücken

Gott, DV (unpublished). Stochastic modelling of asset returns with applications in pricing life-insurance embedded derivatives. A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa, in fulfilment of the requirements for the degree of Master of Science. University of the Witwatersrand, Johannesburg, 2009

Hammersley, JM & Morton KW (1956). A new Market-Carlo technique: antithetic variates. *Proceedings of the Cambridge Philosophical Society* **52 (3)**, 449-75

Hardy, M (2004). A regime-switching model of long-term stock returns. *North American Actuarial Journal* **5(2)**, 41–53

Heath, D, Jarrow, RA & Morton, A (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* **60**, 77–105

Heston, SL (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* **6 (2)**, 327–43

Hicks, JR (1946). *Value and Capital*, 2nd ed. Oxford University Press, London

Ho, T & Lee, S (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance* **41 (5)**, 1011–29

Hull, J (2006). *Options, Futures, and Other Derivatives*, 6th ed. Prentice Hall, New Jersey

Hull, J & White, A (1987). The pricing of options on assets with stochastic volatilities, *Journal of Finance* **42 (2)**, 281-300

Hull, J & White, A (1990). Pricing interest rate derivative securities. *Review of Financial Studies* **3 (4)**, 573–92

Hull, J & White, A (1993a). One factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis* **28 (2)**, 235–54

Hull, J & White, A (1993b). Numerical procedures for implementing term structure models I: single-factor models. *Journal of Derivatives.* **2 (1)**, 7–16

Jarvis, S, Southall, F & Varnell, E (2001). Modern valuation techniques. Presented to The Staple Inn Actuarial Society, 6 February 2001

Lee, P (unpublished). Pricing and reserving for guaranteed annuity options. Presentation at the Den Danske Aktuarforening, April 2001

Longstaff, F (1992). A non-linear general equilibrium model of the term structure of interest rates. *Journal of Financial Economics* **23 (2)**, 333–44

Longstaff, F & Schwartz, E (1992). Interest rate volatility and the term structure: a two-factor general equilibrium model. *Journal of Finance* **47 (4)**, 1259–82

Luciano, E & Vigna, E (2005). A note on stochastic survival probabilities and their calibration. *Applied Mathematics Working Paper Series.* International Centre for Economic Research

Madan, DB & Seneta, E (1990). The Variance Gamma (VG) model for share market returns. *The Journal of Business* **63 (4)**, 511–24

Maitland, AJ (unpublished). Interpolating the South African yield curve. Presentation at the AFIR Colloquium, Tromsø, June 20–23, 2000

Maitland, AJ (2002). Interpolating the South African yield curve using principle-component analysis: a descriptive approach. *South African Actuarial Journal* **2**, 129–45

Milevsky, M & Promislow, D (2001). Mortality derivatives and the option to annuitize. *Insurance: Mathematics and Economics* **29 (3)**, 299–316

Novosyolov, A & Satchkov, D (2008). Global term structure modelling using principle component analysis. *Journal of Asset Management* **9 (1)**, 49–60

Pelsser, A (2003). Pricing and hedging guaranteed annuity options via static option replication. *Insurance: Mathematics and Economics* **33**, 283–96

Renshaw, A & Haberman, S (2003). Lee–Carter mortality forecasting: a parallel generalised linear modelling approach for England and Wales mortality projections. Applied Statistics **52**, 1–19

Scott, LO (1987). Option pricing when the variance changes randomly: theory, estimation, and an application. *Journal of Financial and Quantitative Analysis* **22 (4)**, 419–38

Smith, AD (1996). How actuaries can use financial economics. *British Actuarial Journal* **2**, 1057–194

Shlens, J (unpublished). A tutorial on principle component analysis. Systems Neurobiology Laboratory, Salk Institute for Biological Studies and Institute for Nonlinear Science, University of California, San Diego

Thomas, MP (unpublished). Long term extrapolation and hedging of the South African yield curve. A dissertation submitted to the Faculty of Natural and Agricultural Science, University of Pretoria, in partial fulfilment of the requirements for the degree of Master of Science. University of Pretoria, Pretoria, South Africa, 2008

Thomas, MP & Maré, E (unpublished). Long term forecasting and hedging of the South African yield curve. Convention of the Actuarial Society of South Africa, 2007

Thomson, RJ (1996). Stochastic investment modelling: the case of South Africa. *British Actuarial Journal* **2 (III)**, 765–801

Thomson, RJ (2005). The pricing of liabilities in an incomplete market using mean-variance hedging. *Insurance: Mathematics and Economics* **36**, 441–55

Thomson, RJ (unpublished). The arbitrage-free equilibrium pricing of liabilities in an incomplete market: application to a South African retirement fund, 19th International AFIR Colloquium, 2009

Thomson, RJ & Gott, DV (2009). Stochastic models for actuarial use: the equilibrium modelling of local markets. *ASTIN* **39 (1)**, 339–70

Vasicek, O (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* **5**, 177–88

Whitten, S & Thomas, G (1999). A non-linear stochastic asset model for actuarial use. *British Actuarial Journal* **5**, 919–53

Wiggins, JB (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics* **19 (2)**, 351–72

Wilkie, AD (1986). A stochastic investment model for actuarial use. *Transactions of the Faculty of Actuaries* **39**, 341–403

Wilkie, AD (1995). More on stochastic investment model for actuarial use. *British Actuarial Journal* **1**, 777–964

Wilkie, AD, Waters, HR & Yang, S (2003). Reserving, pricing and hedging for policies with guaranteed annuity options. *British Actuarial Journal* **9**, 263–425

Yakoubov, Y, Teeger, M & Duval, DB (1999). The TY model: a stochastic investment model for asset and liability management. Proceedings of the 9th International AFIR Colloquium, Tokyo. 237–266

Yang, S (unpublished). Reserving, pricing and hedging for policies with guaranteed annuity options. A PhD thesis submitted to the Department of Actuarial Mathematics and Statistics, Heriot Watt University, Edinburgh, Scotland, 2001

Zhou, R (unpublished). Actuarial and financial valuation of guaranteed annuity options. Masters of Science project. Department of Statistics and Actuarial Science, Simon Frasier University, 2007