Fair valuation of insurance liabilities -
a case study

by

Manabu Sato

Submitted in total fulfilment of the requirements of the degree of
Doctor of Philosophy

December 2007

Centre for Actuarial Studies, Faculty of Economics and Commerce,
The University of Melbourne.
Insurance contracts will be reported at fair values on insurers’ balance sheets from 2010. In this thesis, we will review the conceptual and theoretical backbone of the insurance fair valuation project while providing a summary of the key features of the fair valuation project. Then, we will conduct a case study aimed at finding, under the fair valuation regime, the best asset allocation strategy for a particular business unit that carries a hypothetical annuity portfolio using a single modelling framework for valuation, risk calculation and business appraisal.

Insurers’ profits under fair value accounting will be much more volatile than those under cost accounting that is commonly used today. This is because the fair valuation approach is expected to require all insurance contracts to be revalued using assumptions that are consistent with the current financial market without making any adjustments. The fair value is in principle an exchange value and the expected users of financial reports under fair value accounting are the general public. The characteristics of fair valuation lies in, among others, asset and liability measurement approach, use of market assumptions and independence of liabilities.

The market value of risk premiums of securities is zero. Therefore, under the fair valuation environment, any assumption on risk premiums of securities should not affect expected present values in order for the valuation to be consistent with the market. The financial economics approach provides this consistency. Deflators are a key concept for modern financial economics and the existence of deflators is almost equivalent to the market being arbitrage-free. The market-price-of-risk process solely determines the deflator in a continuous-time setting.

Established interest rate term structure models are classified into several categories. Philosophies behind models in each category are discussed and some important models are illustrated. The positive interest approach, in particular Cairns’ Multifactor Model, is explained in detail.

Parameters of Cairns’ Multifactor Model are obtained so that the resulting model yield curve and price of a Cap at a specific date fit the market. The value of an equity
portfolio is assumed to follow a lognormal distribution, and the volatility is estimated using historical data. Using these parameters, 5,000 market scenarios are generated by simulation.

We investigate the risk-return profile of a hypothetical business unit that carries a hypothetical annuity portfolio. We assume the regulatory capital requirement is based on VaR. We analyze capital flows that arise from initial and interim capital injections and dividends, using the generated market scenarios, to compare investment strategies. Our conclusion is that the investment strategy that minimizes the market risk leads to the most efficient capital usage.
Declaration

This is to certify that

(i) the thesis comprises only my own original work,

(ii) due acknowledgement has been made in the text to all other materials used,

(iii) the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies, appendices and footnotes.

(Manabu Sato)  (Date)
Acknowledgements

This thesis would never have been finished, or even initiated, without the help of my supervisor Professor David Dickson. Since I started my master’s research work in 2001, Professor Dickson offered lots of valuable guidance and assistance both on my academic work and on my personal life in Australia. I am very lucky to have been studying under such an enthusiastic researcher. His support ranged from choosing thesis topics to correcting final drafts. In particular, he gave me, with boundless patience, a huge amount of advice on forming the structure of this thesis and presenting ideas in the right context in correct English. I would like to express my genuine gratitude to him.

I would like to thank staff of the Centre for Actuarial Studies at the University of Melbourne. Professor Daniel Dufresne, Associate Professor Mark Joshi, and Mr. Richard Fitzherbert kindly reviewed my work and gave me much useful advice. In particular, Associate Professor Mark Joshi carefully read a preliminary draft of this thesis and gave me lots of useful suggestions. I hope that this completed thesis has come up to the standard he expects. I would also like to thank Ms. Alison Banford who looked after me in non-academic areas during my PhD candidature.

I also acknowledge the financial support provided to me during my PhD candidature by the Faculty of Economics and Commerce through the Melbourne International Research Scholarship and the Melbourne International Fee Remission Scholarship awarded to me.

Finally, but by no means least, my deepest gratitude is reserved for the encouragement and endurance of my family, Kazuko, Ryo and Yu. The happiest event during my PhD candidature period is that Kazuko, my wife, gave birth to our second son, Yu. I appreciate her great task of bringing up two preschool-aged children in this foreign country. This thesis would not have been completed without her endurance. In addition, I dedicate this thesis to my mother and my father who passed away just recently.
# Table of Contents

Abstract ......................................................................................................................... i  
Declaration ..................................................................................................................... iii 
Acknowledgments .......................................................................................................... iv  
Table of Contents ........................................................................................................... v  
List of Tables .................................................................................................................. viii 
List of Figures ............................................................................................................... ix  

1 Introduction .................................................................................................................. 1  
1.1 Motivation and thesis outline ................................................................................... 1  
1.2 The impact of the change of accounting system on profits ..................................... 3  
1.3 Fundamental arguments on fair value accounting ............................................... 7  

2 Outline of the insurance fair valuation project and principles ................................... 9  
2.1 The status of the fair valuation project ..................................................................... 9  
2.2 Background to the fair valuation of insurance contracts ....................................... 11  
2.3 Definition of fair value ........................................................................................... 14  
2.4 Conceptual difference between fair valuation and traditional valuation ............. 17  
2.5 Major issues under discussion ................................................................................ 18  
2.5.1 Asset and liability measurement approach .......................................................... 18  
2.5.2 Reflection of changes in fair values in the profit and loss account ..................... 19  
2.5.3 Indirect or direct method .................................................................................... 23  
2.5.4 Insurers’ own credit standing .......................................................................... 24  
2.5.5 Inclusion of backing asset returns in the discount rates ................................... 25  
2.6 New risk profile and capital requirement under the fair value principle ............. 27  
2.7 Summary ............................................................................................................... 28  

3 Valuation under fair value principles ......................................................................... 29  
3.1 Risk premium and discount factor ......................................................................... 30  
3.2 Value of the risk premium ..................................................................................... 32  
3.3 Financial leverage effect ....................................................................................... 35  
3.4 Certainty-equivalent cash flow ............................................................................. 36  
3.5 Decomposition of the balance sheet ...................................................................... 38  
3.6 A model of an insurer’s balance sheet .................................................................... 41  
3.7 The scope of our research .................................................................................... 45  
3.8 Summary ............................................................................................................. 45
4 Valuation using deflators ................................................................. 47
  4.1 Introduction ..................................................................................... 48
  4.2 Defining financial markets ............................................................... 50
  4.3 Single-period model......................................................................... 53
    4.3.1 The state price vector ................................................................. 53
    4.3.2 Deflator pricing and risk-neutral pricing ........................................ 57
    4.3.3 Choice of numeraire security....................................................... 59
    4.3.4 Deflator generator and summary .................................................. 62
  4.4 Extension to multi-period model ..................................................... 64
    4.4.1 Information structure .................................................................... 65
    4.4.2 Deflators and deflator generators ................................................. 67
    4.4.3 Trading gain and arbitrage ............................................................. 69
  4.5 Extension to continuous-time setting .............................................. 70
    4.5.1 Model specifications ..................................................................... 71
    4.5.2 Some sufficient conditions for no arbitrage .................................... 74
    4.5.3 The market-price-of-risk ............................................................... 75
    4.5.4 Forms of deflator generator ............................................................ 78
  4.6 An example: pricing a plain call option ......................................... 83
  4.7 Practical issues in simulation ........................................................... 87
  4.8 Summary ......................................................................................... 89

5 Interest rate term structure models .................................................. 90
  5.1 Classification of interest rate term structure models ..................... 91
    5.1.1 Economic hypotheses ................................................................. 92
    5.1.2 Pricing models ............................................................................. 92
    5.1.3 A descriptive model .................................................................... 93
  5.2 Consistency between pricing models and economic hypotheses ....... 94
  5.3 Philosophies behind pricing models .............................................. 96
    5.3.1 Equilibrium models ................................................................. 97
    5.3.2 Evolutionary models ................................................................. 97
    5.3.3 Choosing between equilibrium and evolutionary models ............ 98
  5.4 Models under the classical approach .......................................... 99
    5.4.1 One-factor models ....................................................................... 100
    5.4.2 Two-factor models ...................................................................... 104
    5.4.3 Evolutionary models .................................................................... 104
  5.5 Heath, Jarrow and Morton approach ............................................ 105
  5.6 Positive interest ............................................................................. 108
    5.6.1 Rogers and Rutkowski framework .............................................. 109
    5.6.2 Flesaker and Hughston approach ............................................... 112
  5.7 Cairns’ Integrated Gaussian model .............................................. 116
    5.7.1 Basic characteristic of the model ................................................. 117
    5.7.2 Model description ........................................................................ 119
    5.7.3 Choice of the market-price-of-risk ............................................ 123
List of Tables

Table 1·1 - The net profit under cost accounting and fair value accounting .......... 5
Table 2·1 - The history and the schedule of the insurance fair valuation project ... 9
Table 3·1 - Expected return on capital ............................................................... 35
Table 4·1 - 95% confidence range of the simulated average .............................. 88
Table 5·1 - Commonly-used classical approach models .................................... 100
Table 5·2 - One-factor equilibrium models ..................................................... 100
Table 5·3 - Parameters used for Figure 5·1 ...................................................... 102
Table 6·1 - One-year forward rates and volatility .......................................... 137
Table 6·2 - Yield curve fitting result ................................................................. 146
Table 6·3 - Cap price fitting result .................................................................. 146
Table 6·4 - Cap price dispersion ..................................................................... 148
Table 7·1 - Number of policies in-force of major Japanese insurance products .... 164
Table 7·2 - Summary of the premium calculation .............................................. 166
Table 7·3 - Yearly profits for both accounting methods ..................................... 171
Table 7·4 - Basic statistics of the yearly profits ............................................... 172
Table 7·5 - Legal SMRs .................................................................................. 179
Table 7·6 - Covariance matrix of selected bonds and equity portfolio ................ 183
Table 7·7 - Correlation matrix of selected bonds and equity portfolio ............... 183
Table 7·8 - Capital requirement for the annuity contract .................................... 186
Table 7·9 - Strategy combinations .................................................................... 189
Table 7·10 - Illustration of capital flows under scenario 1 (unit 1,000) .......... 196
Table 7·11 - Statistics of standardized value change ........................................... 202
Table 7·12 - Internal Rates of Return (%) ......................................................... 203
Table 7·13 - EPV and IDV ............................................................................. 207
Table 7·14 - Payback Period and Total Capital Injection ................................. 209
Table 7·15 - Exposure to additional capital injections ........................................ 214
Table 7·16 - EPV and IDV on liquidation option ............................................. 216
Table 7·17 - The total discounted profit for various values of $\alpha_i$ and $\beta$ ...... 226
**List of Figures**

| Figure 1-1 | The net profit under cost accounting and fair value accounting | 6 |
| Figure 3-1 | Example of the balance sheet of the simple firm | 39 |
| Figure 3-2 | Example of the balance sheet of the simple insurer | 41 |
| Figure 3-3 | Balance sheet of the insurer (Babbel) | 42 |
| Figure 3-4 | Proposed balance sheet of the insurer | 44 |
| Figure 3-5 | The balance sheet assumed in this study | 45 |
| Figure 5-1 | Sample paths of equilibrium models | 102 |
| Figure 6-1 | Historical JPY interest rates | 128 |
| Figure 6-2 | Spot volatility determination process | 134 |
| Figure 6-3 | Historical 30-day volatility and forward rates | 135 |
| Figure 6-4 | Implied volatilities and forward rates | 135 |
| Figure 6-5 | Market and model yield curve | 140 |
| Figure 6-6 | Model forward rate percentage volatilities | 140 |
| Figure 6-7 | Historical 30-year forward rates | 142 |
| Figure 6-8 | Independent forward rate volatility curves | 143 |
| Figure 6-9 | Converged shape of the volatility, yield and risk premium curve | 149 |
| Figure 6-10 | Scenario 836 | 152 |
| Figure 6-11 | Scenario 4,532 | 153 |
| Figure 6-12 | Scenario 917 | 154 |
| Figure 6-13 | Rolling correlation between Nikkei 225 and interest rates | 157 |
| Figure 6-14 | Historical monthly volatility of Nikkei 225 index | 159 |
| Figure 6-15 | Historical daily volatility of Nikkei 225 and TOPIX 500 index | 160 |
| Figure 7-1 | Expected cash flows arising from an annuity contract | 168 |
| Figure 7-2 | Interest rate term structure development (Scenario 5) | 169 |
| Figure 7-3 | Development of liability values under both accounting methods | 169 |
| Figure 7-4 | Yearly profits under both accounting methods | 172 |
| Figure 7-5 | Process to determine capital inflow or outflow amount | 191 |
| Figure 7-6 | Pricing errors of zero coupon bond prices | 193 |
| Figure 7-7 | Illustration of capital flows under three strategies | 198 |
| Figure 7-8 | Q-Q plot of standardized value change against normal distribution | 201 |
| Figure 7-9 | Average capital flows for the strategy with SMR 110%, no equity investment and with hedging | 210 |
| Figure 7-10 | Average capital flows for the strategy with SMR 110%, no equity investment and without hedging | 210 |
| Figure 7-11 | Average capital flows for the strategy with SMR 200%, no equity investment and with hedging | 212 |
| Figure 7-12 | Capital flows for the strategy with SMR 200%, 25% equity investment and with hedging | 213 |
| Figure 7-13 | Mortality improvement ratio | 221 |
| Figure 7-14 | Profit emergence when there is no mortality improvement | 224 |
| Figure 7-15 | Profit emergence when there is mortality improvement | 225 |
1 Introduction

1.1 Motivation and thesis outline

The International Accounting Standards Board (“IASB”, hereafter) is leading a project aimed at implementing standards that require that insurance contracts be reported at fair values on insurers’ balance sheets. IASB completed Phase I of the project in March 2004 by issuing the International Financial Reporting Standard (“IFRS”, hereafter) 4 Insurance Contracts. This IFRS is regarded as a stepping stone to Phase II of the project and provides only limited improvements to, and interim guidelines for, accounting for insurance contracts until the IASB completes Phase II. Phase II is expected to be completed in 2010 and the final standard will be issued then. In spite of strong objections from insurance industry bodies worldwide, in particular from the USA and Japan, hesitation expressed in some academic literature, and bewilderment from some professional bodies, the direction towards fair value reporting has already been established, even if there is the possibility of further delay in its final implementation.

The fair valuation requirement would have an enormous impact on many aspects of insurance businesses as we will illustrate in the next section. Insurers will be required to immediately report the profit or loss caused by the fluctuation of financial markets such as changes in interest rate term structure. This immediate reporting requirement is a significant departure from the accounting methodologies that are commonly used today. Some countries use cost accounting where the original assumptions continue to be used until the expiration of a contract regardless of the current market level. In some countries, assumptions are revised, taking the current market conditions into account. Even in those cases, however, some discretion is usually allowed based on the long term behaviour of the variables, considering the long term nature of the insurance contracts. On the contrary, the fair valuation approach is expected to require all insurance contracts to be revalued using assumptions that are consistent with the current financial market without making any adjustments.

The effects of the introduction of the fair valuation are not just limited to how balance sheets appear and the pattern of emergence of profits of insurers. The changes in the valuation method and reporting system are likely to impact on all aspects of an
insurance business. A change in the method of profit recognition leads to a change in marketing strategy. The reshaped risk profile of an insurance portfolio naturally requires more rigorous asset-liability matching strategy. Fair valuation of insurance liabilities will be one of the greatest challenges all insurers will have to face in the next few years.

The objective of this study is to review the conceptual and theoretical backbone of the insurance fair valuation project, and to find, under the fair valuation regime, the best asset allocation strategy for a particular business unit that carries a hypothetical annuity portfolio backed by bonds and equities, using a single modelling framework for valuation, risk calculation and business appraisal. The business is assumed to be subject to valuation and capital requirements that are expected to be imposed under the fair valuation regime. The business is assumed to be supported by initial and subsequent capital injections, and the business is analyzed from the viewpoint of the capital provider. We seek the asset allocation strategy between bonds and equities that results in the most efficient capital usage. We use modelling techniques based on financial economics. The financial economics approach preserves the consistency of the assumptions with the current financial market, and for that reason we view that this approach will be a dominant methodology under the fair valuation regime. However, the financial economics approach is not widely used for modelling insurance products, in particular products of life insurance companies, and the concept behind this approach is not well understood. Therefore, we not only describe the specific model we use for this study, but also review and attempt to clarify the fundamental principles on which financial economics are based. We use a single modelling framework for pricing, risk calculation and business appraisal. This is a new proposal since at present different approaches tend to be used for pricing, risk calculation and business appraisal for insurance products.

Our thesis, however, starts by introducing the fair value principles and relevant discussion, and moves on to clarifying our scope for this thesis before any quantitative study. Insurance contracts are generally very complex and the fair valuation project covers all aspects of these complicated contracts. Therefore, this preparatory work is necessary for us to clarify what we are studying and on what basis. Traditional valuation methods have been the basis for financial reporting, prudential reporting and risk management for decades. Any departure from the conceptual and technical
framework of these traditional methods for the new fair valuation framework is a task involving great difficulty and significant workload, and the task is still in its infancy. There are still a large number of issues that need to be cleaned up. Some issues are not well-defined yet, some issues are still under discussion, and some issues have been resolved but their contents are not well understood. Among hundreds of issues being or having been discussed, we will select the most important issues. We will fix our standpoints while presenting major discussions on those issues.

1.2 The impact of the change of accounting system on profits
In this section, we will show the magnitude of the change from cost accounting to fair value accounting using a simplified hypothetical endowment assurance policy. We assume that a 20 year endowment assurance is issued to a male aged 30 whose sum assured is 1 million. From the 2004 Japan Male Mortality Table\(^1\), the probability that a male aged 30 survives for 20 years is 0.9673, i.e. the male almost surely survives. Therefore, for simplicity, we treat the contract as a zero coupon liability that expires in 20 years, ignoring any early payment due to death. We also ignore any expenses and margins. The contracts which actual life insurance companies carry are not as simple as this example. However, this example portrays the characteristics of the asset-liability structure of life insurance companies. Life insurance companies have liabilities of very long duration and assets whose duration is short relative to that of liabilities.

We assume that the interest rate was initially 5% per annum on an annual compounding basis. In addition, we assume that the yield curve is always flat and fluctuates only in parallel. Discounting the cash outflow that occurs in 20 years by \(1.05^{-20}\), we obtain the present value of the policy at time 0 is 376,889. We assume that the policyholder paid this amount as a single premium at time 0 and the insurer invests the paid premium in cash all the time.

Under cost\(^2\) accounting, the present value of the policy is calculated using the interest rate 5% regardless of the prevailing market interest rate over the whole policy term. The

---

\(^1\) Published by the Ministry of Health, Labour and Welfare in Japan for the year 2004. Mortality after age 100 was obtained by extrapolation assuming Gompertz’s Law. The table is attached in Appendix I.

\(^2\) In Australia, the discount rates need to be current but rates are not always the market rates. The resulting change of the value of policies is recognized as profit or loss immediately (Australian Accounting Standard Board (“AASB”, hereafter) 1038 (2005) Paragraph 8.5 (c) ).
present values are:

at the end of first year; \(1,000,000 \times 1.05^{-19} = 395,734\),

at the end of second year; \(1,000,000 \times 1.05^{-18} = 415,521\),

and so on.

If the interest rate stays at 5% for the first year, the profit for this term is calculated as follows:

\[
\text{Income: Interest on cash} = 376,889 \times 5\% = 18,844
\]

\[
\text{Expense: Liability increase} = 395,734 - 376,889 = 18,844
\]

Therefore, the net profit is zero.

Similarly, for the second year:

\[
\text{Income: Interest on cash} = 395,734 \times 5\% = 19,787
\]

\[
\text{Expense: Liability increase} = 415,521 - 395,734 = 19,787
\]

Therefore the net profit is again zero. The profit for any period is zero as long as interest rates stay at the same level.

Now suppose the interest rate changes to 4% immediately before the end of the first year, and stays at this level over the second year. Under cost accounting the net profit for the first year is still zero since the liability is measured using 5%, which is the cost. For the second year, the earned interest is calculated using the current interest rate of 4%, but the liability value at the end of the term is still calculated using 5%. Then the net profit is calculated as follows:

\[
\text{Income: Interest on cash} = 395,734 \times 4\% = 15,829
\]

\[
\text{Expense: Liability increase} = 415,521 - 395,734 = 19,787
\]

The net profit is -3,957, which is a 1% interest rate loss over the second year.

However, under fair value accounting, the loss amount is significantly larger than this figure. Under the fair value regime, the value of the policy at the end of the first year is:

\[
1,000,000 \times 1.04^{-19} = 474,642.
\]

Therefore, for the first year:

\[
\text{Income: Interest on cash} = 376,889 \times 5\% = 18,844
\]

\[
\text{Expense: Liability increase} = 474,642 - 376,889 = 97,753
\]

The net profit is -78,908, which is approximately 1% interest rate loss over the whole remaining period of 19 years.
To illustrate the effect of fair value accounting, we now assume that the interest rate fluctuates regularly as 5%, 4%, 5%, 6%, 5%, 4%... year by year. Table 1-1 shows the net profit for each year and accumulated profit to that year under the both accounting methods and Figure 1-1 illustrates the position.

Table 1-1 - The net profit under cost accounting and fair value accounting

<table>
<thead>
<tr>
<th>Year</th>
<th>Cost accounting</th>
<th>Fair value accounting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Net Profit</td>
<td>Accumulated</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-3,957</td>
<td>-3,957</td>
</tr>
<tr>
<td>3</td>
<td>-198</td>
<td>-4,155</td>
</tr>
<tr>
<td>4</td>
<td>4,114</td>
<td>-42</td>
</tr>
<tr>
<td>5</td>
<td>-2</td>
<td>-44</td>
</tr>
<tr>
<td>6</td>
<td>-4,812</td>
<td>-4,856</td>
</tr>
<tr>
<td>7</td>
<td>-243</td>
<td>-5,098</td>
</tr>
<tr>
<td>8</td>
<td>4,997</td>
<td>-101</td>
</tr>
<tr>
<td>9</td>
<td>-5</td>
<td>-106</td>
</tr>
<tr>
<td>10</td>
<td>-5,851</td>
<td>-5,957</td>
</tr>
<tr>
<td>11</td>
<td>-298</td>
<td>-6,255</td>
</tr>
<tr>
<td>12</td>
<td>6,071</td>
<td>-184</td>
</tr>
<tr>
<td>13</td>
<td>-9</td>
<td>-193</td>
</tr>
<tr>
<td>14</td>
<td>-7,115</td>
<td>-7,308</td>
</tr>
<tr>
<td>15</td>
<td>-365</td>
<td>-7,673</td>
</tr>
<tr>
<td>16</td>
<td>7,375</td>
<td>-298</td>
</tr>
<tr>
<td>17</td>
<td>-15</td>
<td>-313</td>
</tr>
<tr>
<td>18</td>
<td>-8,651</td>
<td>-8,964</td>
</tr>
<tr>
<td>19</td>
<td>-448</td>
<td>-9,412</td>
</tr>
<tr>
<td>20</td>
<td>8,959</td>
<td>-453</td>
</tr>
</tbody>
</table>
The final accumulated profits under both methods are identical. An insurer, however, is experiencing enormous swings in its profit under fair value accounting. The yearly profit or loss under fair value accounting is roughly 10 times as much as that under cost accounting.

Suppose that some capital was injected at the issue of this hypothetical endowment assurance business. We assume that the amount of capital is determined so that the capital can cover the accumulated loss as long as the policy is in force using the interest rate scenario we have just used. Under cost accounting, the accumulated loss for this business reaches its maximum of 9,412 in 19 years. Therefore, the initial capital requirement is approximately 10,000. If we take into account the interest income earned on the capital, the initial capital requirement is even smaller. On the other hand, the accumulated loss reaches its maximum of 78,908 in the first year under fair value accounting. That means the initial capital requirement is approximately 80,000 under fair valuation accounting, which is eight times as much as that under cost accounting. This initial capital is released and returned to the capital provider as time passes. Even taking into account this gradual repayment of the initial capital, it is obvious that the
return on capital under fair value accounting is significantly worse that that under cost accounting.

In a real situation, capital requirements are calculated using much more complex formulae. However, in general, a business whose profits are more volatile is required to keep more capital than the business whose profits are less volatile. Therefore, the fair value accounting affects the capital requirements of insurers and hence the return on capital as well.

1.3 Fundamental arguments on fair value accounting
We are of the view that fair value accounting is more appropriate than cost accounting for the general-purpose financial reporting. The primary reason for this view is that no single body can predict future macro economic variables such as interest rates better than the market does. In the above example, when the market interest rates declined to 4%, our stance is that 4% is the best assumption to be used for the valuation. Financial reporting is for the general public, and the prices in the open and liquid markets are the best means to capture the consensus among the general public. Our view is in line with the IASB’s view.

The other view contrary to our view is that the fluctuation of profits due to changes in financial markets does not reflect the true financial conditions of the insurer. The underlying idea of this argument is that the term of an insurance contract is generally very long. Therefore the assumptions on macro economic variables should also be based on the long term equilibrium of macro economics rather than short term factors such as a sudden change in supply and demand. In particular, interest rates tend to move back to their convergence level in the long run. Thus, the interim fluctuation of interest rates does not generate a significant effect on the total profit at the end of the contract term.

This argument, however, is not persuasive even when we admit the strong mean-reversion feature of interest rates, and also admit that the market sometimes makes collective mistakes as seen in bubble economies. The question to be answered is:

“Who can set the assumptions that would be better than those set by the market?”,

7
not:

"Are current market prices absolutely unbiased?".

In the above example, when the market interest rates declined to 4%, nobody could guarantee that the interest rates would move back to 5%. Rather, the market tells that, as a consensus of all market participants, the long term equilibrium level has now become 4%. We will present a full discussion on the above topic in the context of the fair valuation project in detail in the next chapter.

In our example, it would be a rational investment decision to sell bonds at this 4% level anticipating that the level will revert to 5% eventually. Such a specific view of the direction of future market movement is reasonable and essential for the market to be actively functioning, but cannot be used for the purpose of financial reporting whose users are the general public. This is because 4% is the best estimate of the consensus by the general public when the market interest rate is 4%. The forecast that the interest rate will revert to 5% is only one view among a large number of views made by all the market participants. In order for the valuation to be accepted by the general public, it is not correct to use a specific view by a specific entity.
2 Outline of the insurance fair valuation project and principles

In this chapter, we present our standpoint on major issues related to the fair valuation of insurance contracts, introduce the insurance fair valuation project, and provide relevant discussions. The insurance fair value project is a large and complicated project that is expected to take more than ten years to complete from its launch. This is because insurance products are generally very complicated with many covenants, and the underlying concepts behind fair valuation are quite different from those behind traditional methods. Due to the size of the project, it is impossible to look into the details of every single issue. Therefore, it is important for us to understand the whole picture of the project and to select the core arguments out of hundreds of issues. Based on our understanding of the major issues on the project, we will state our standpoint on them.

2.1 The status of the fair valuation project

Table 2-1 shows the history and schedule of the insurance fair valuation project. Information is taken from the web homepage of IASB as of the end of April 2007 and IASB Project Updates Insurance Contracts (Phase II) as of December 2006 (“Project Updates (2006)”, hereafter).

<table>
<thead>
<tr>
<th>Year</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>A Steering Committee was set up</td>
</tr>
<tr>
<td>December 1999</td>
<td>An Issues Paper was published</td>
</tr>
<tr>
<td>2001</td>
<td>Draft Statement of Principles (“DSOP”, hereafter) was issued</td>
</tr>
<tr>
<td>May 2002</td>
<td>IASB split this project into two phases</td>
</tr>
<tr>
<td>March 2004</td>
<td>Phase I completed by issuing IFRS 4 Insurance Contracts</td>
</tr>
<tr>
<td>Mid 2004</td>
<td>Phase II restarted</td>
</tr>
<tr>
<td>May 2007</td>
<td>A discussion paper was issued⁵</td>
</tr>
<tr>
<td>(Plan) 2009</td>
<td>An exposure draft will be issued</td>
</tr>
<tr>
<td>(Plan) 2010</td>
<td>The final standard will be set out</td>
</tr>
</tbody>
</table>

⁵ Since this has just recently issued and still been asking for comments on it, the contents of this document are not reflected in this thesis.
As of 2007, the insurance fair valuation project is still in its infancy. In March 2004, IFRS 4 (2005) was issued as the outcome of Phase I. This standard does not set out any conclusion about the fair valuation of insurance contracts. Instead, this standard explicitly permits insurance contracts to be exempt from the fair valuation requirement even though the standard stipulates some requirements on some specific issues. For example, IFRS 4 (2005, IN4) prohibits catastrophe or equalization provisions, requires a test for the adequacy of recognized insurance liabilities, and requires an insurer to keep insurance liabilities until they are discharged or cancelled. This is, however, all that IFRS 4 (2005) requires for insurance contracts. These items account for a very tiny portion of the full scope of the insurance fair valuation project. On the other hand, IFRS 4 (2005) has a conclusive policy for financial derivatives. IFRS 4 (2005) explicitly requires financial derivatives embedded in the insurance contracts to be measured at the fair value.

The project was forced to split into two phases in 2002. This fact shows that the insurance fair valuation project is a difficult project. IASB needed to declare that it had abandoned past discussions in order to restart the project in Phase II. Paragraph 4 of Project Updates (2006) states:

“Restarting phase II, the Board took a fresh look at financial reporting by insurers. Past work by the Board and by its predecessor was a useful resource, but did not bind the Board.”

In paragraph 9 of Project Updates (2006), there is a list of the preliminary conclusions that are expected to be included in the planned discussion paper. Despite the declaration that IASB is not bound by its predecessor’s work, the conclusions are almost in line with DSOP that was issued in 2001. Therefore, we regard Phase II as an extension of Phase I. Since DSOP is the most comprehensive piece of literature published in the insurance fair valuation project so far, we still use DSOP as our major reference material.

The DSOP article 1.10 describes the objective of setting the accounting standard based on fair value principle as:

“International Financial Reporting Standards are intended to be used in general purpose financial statements directed toward the common
information needs of a wide range of users. These users include present and potential investors, employees, lenders suppliers and other trade creditors, customers (for example, the policyholders of an insurer), governments and their agencies (for example, supervisors and regulators) and the public”.

In other words, the financial reporting is for the general public.

Despite strong opposition from insurance industry bodies worldwide, for example, the American Council of Life Insurance (“ACLI”, hereafter), fair value accounting is now on track. Most professional actuarial bodies, for example, the American Academy of Actuaries (“AAA”, hereafter), the Institute of Actuaries of Australia (“IAAust”, hereafter), and the International Actuarial Association (“IAA”, hereafter), agree with the direction towards fair valuation of insurance liabilities.

### 2.2 Background to the fair valuation of insurance contracts

In this section, we discuss what has brought the insurance fair valuation initiative. According to our view, there are three major issues that are driving the insurance fair valuation project. These are: eliminating accounting mismatches, moving away from the use of any averaging methods, and enforcing more rigorous asset-liability management. All of these issues relate to assets held to back insurance contracts. Therefore, we look into the history of the accounting method for assets first.

As of 2006, the International Accounting Standard (“IAS”, hereafter) 39 is already in force. Under the IAS 39, financial instruments held for trading purposes need to be measured at fair value. IFRS 4 (2005) explicitly exempts insurance contracts from being measured at fair value. Therefore, current accounting standards create an accounting mismatch⁶ between insurance liabilities and the assets backing them. We agree with IASB’s view, described in detail in paragraph 10 of IFRS 4 Insurance Contracts Frequently Asked Questions (2004)⁷ (“FAQ (2004)”, hereafter), that measuring both assets and liabilities at fair value is the best way to eliminate the accounting mismatch. Eliminating the accounting mismatch is the first driving force for insurance fair

---

⁶ Paragraph 7 in FAQ (2004) states: “Accounting mismatch arises if changes in economic conditions affect assets and liabilities to the same extent, but the carrying amounts of those assets and liabilities do not respond equally to those economic changes. Specifically, accounting mismatch occurs if an entity uses different measurement bases for assets and liabilities”.

Before IAS 39 was set out, cost accounting was commonly used, and is still being used in areas where IAS 39 are not yet been adopted, for assets as well as for liabilities. Cost accounting on the financial assets has been a source of misleading reporting as Vanderhoof (1998) shows using the example of the United States.

In the United States, Financial Accounting Standard ("FAS", hereafter) 115, where debt securities and equities are required to be evaluated at market values, had been in force before IAS 39 was set out. The primary objective of FAS 115 is to eliminate “gains trading”, where appreciating assets are sold to realize a booked capital gain while the assets that had depreciated are carried at book value. According to Vanderhoof (1998), “gains trading” was common among banks and the Savings and Loans companies as well as insurance companies. Under FAS 115, while the debt securities that were classified as “hold to maturity” are still allowed to be accounted for on a cost basis, the debt securities for trading purposes are required to be held at market value, and changes in market value are carried through to earnings. FAS 115 aims at increasing the transparency of financial reports. The introduction of FAS 115 created the accounting mismatch since some debt securities are measured at market value while the liabilities are still measured on cost basis. The most rational way to eliminate this accounting mismatch is that we measure the liabilities in a consistent way with the assets since we cannot return to the world where assets are measured at cost.

Moving away from the use of any “averaging” methods is the second driving force for insurance fair valuation. The nature of FAS 115, marking assets only to market, has created many problems. Apart from conceptual issues related to the accounting mismatch, earnings under Generally Accepted Accounting Principles ("GAAP", hereafter) and capital are highly volatile, even if the company operation is smooth (see, for example, Carr (1998)). Some efforts have been made to eliminate this inappropriate volatility. One example is the interest maintenance reserve (see, for example, Doll et al. (1998)), which allows the realized gains and losses on investments supporting long-term liabilities to be amortized over the life of the investment sold. Besides these newly created reserves, traditionally insurers have been using various reserves such as catastrophe or equalization provisions.
While insurers insist that such “averaging” or “smooth emergence of profits” are appropriate considering the long term nature of insurance contracts, there have been strong objections to them. Traas (2003) believes that the annual reporting of insurance companies is completely wrong because of averaging. Averaging over a long period would be allowed if an insurance company, which in fact does not last forever, were to be immortal. According to Traas (2003), annual reporting should provide the state of the business by presenting a picture of the company’s ability to generate profits and cash flow.

The third driving force is the pressure for more rigorous asset-liability management. Even though insurance companies tend to use sophisticated asset-liability matching techniques, the current reporting system does not provide enough incentives to reduce asset-liability mismatch risks. The existence of protection funds even encourages holding risky assets, as we will show in Chapter 3. Stakeholders have paid more and more attention to asset-liability mismatch positions taken by insurance companies after experiencing asset devaluation in the 1990s. Sheldon and Smith (2004) argue that one motivation for examining market consistent valuation is the notion of hedging or closing out of risks. Sheldon and Smith (2004) state in paragraph 2.3.1:

“If an entity is solvent as measured on a market consistent basis, then it might, in theory, be able to transact at those prices and close out its assets and liabilities. Therefore, ensuring market consistent solvency offers some protection to policyholders or creditors”.

The realistic balance sheet devised by the FSA in the UK is designed, according to Sheldon and Smith (2004), to capture the cost of guarantees and smoothing on a market consistent basis. Sheldon and Smith (2004) state in paragraph 2.3.3:

“The introduction of the realistic balance sheet is, in part, a response to the difficulties that un-hedged guarantees have caused the life industry in recent years. Reliance on long-term solvency tests runs the risk that we overlook more imminent problems, compounded by the use of over-optimistic assumptions and models used to determine capital needs”.

Similarly, the failure of several Japanese life insurance companies in the 1990s that had
not shown warning solvency margin levels in advance might have driven the need for fair value accounting.

### 2.3 Definition of fair value

Fair value is a hypothetical value. DSOP Principle 3.1 defines fair value as:

“*The amount for which an asset could be exchanged or a liability settled between knowledgeable, willing parties in an arm’s length transaction. In particular, the fair value of a liability is the amount that the enterprise would have to pay a third party at the balance sheet date to take over the liability*”.

We understand the concept of fair value more clearly when fair value is compared with economic value. The IAAust Economic Value Taskforce (2002, p.808) defines the economic value as:

“*An Economic Value of an Economic Asset is an estimate of the current cash equivalent of all future cash flow benefits (or costs) that are expected to be derived from ownership or use of the Economic Asset by a specified client in the specified circumstances*”.

IAAust Economic Valuations Taskforce (2002, page 809) characterize fair value as a hypothetical market value in a perfect market, stating:

“*Fair Value appears to contemplate a perfectly competitive and rational market where market ‘sentiment’ is muted, or even suppressed, leaving the rational component of trade in the Economic Asset to determine exchange prices*”.

Note that the economic value for a particular client is very likely to differ from that for other clients. In other words, there are various economic values perceived by various clients for the same asset or liability; and a client does not, in principle, have to consider economic values that are perceived by other clients. On the other hand, fair value is an exchange value determined uniquely in the market.

We support using fair value for measuring insurance contracts even though fair value does not exist in reality and it is difficult to estimate it. The market value would be best for this purpose if it were to exist. The fair value is the proxy for that. Since there is no liquid market for insurance contracts from which we would obtain some indication of fair value, the fair value of insurance contracts is obtained only by calculations.
IASB is now proposing to use a new concept “current exit value” in Phase II for the time being. Project Updates (2006) paragraph 24 defines current exit value as:

“The amount the insurer would expect to have to pay now if it transferred all its remaining contractual rights and obligations to another entity”.

Our view is that there is no substantial difference between fair value and current exit value; these are in principle exchange prices.

A similar but different concept is the “entity specific value”. DSOP Principle 3.1 defines the entity specific value as:

“the value of an asset or liability to the enterprise that holds it, and may reflect factors that are not available (or not relevant) to other market participants. In particular, the entity-specific value of an insurance liability is the present value of the costs that the enterprise will incur in settling the liability with policyholders or other beneficiaries in accordance with its contractual terms over the life of the liability”.

The entity specific value can be regarded as an Economic Value for the insurer. Despite the conceptual difference, the actual difference between the fair value and the entity-specific value is likely to be small (UK Working Party on fair valuation (2001, p.4)). One reason for the similarity lies in the shared assumptions. The DSOP article 3.25 states:

“Except where otherwise stated, measurement principles in this DSOP apply to both entity-specific value and fair value”.

Next we discuss the basics of calculating fair value of insurance contracts. We agree with the principle that is set out in Project Updates (2006). This principle is asking an insurer to use three building blocks in measuring its insurance liabilities. These building blocks are: current unbiased probability-weighted estimates of future cash flows, current market discount rates, and an explicit and unbiased estimate of the margin that market participants require. This method is largely in line with the method proposed in DSOP. DSOP Principle 4.1 states:

“The starting point for measuring insurance assets and insurance liabilities should be the expected present value of all future pre-income-tax cash flows arising from the contractual rights and contractual obligations associated
with the closed book of insurance contracts”.

The cash flows should be adjusted to reflect risks and uncertainties. DSOP Principle 5.1 states:

“The entity-specific value and fair value of insurance liabilities and insurance assets should always reflect risk and uncertainty”.

Also DSOP Principle 5.2 states:

“Adjustments for risk and uncertainty should be reflected preferably in the cash flows, or alternatively in the discount rate(s), without any double counting”.

The risk-adjusted cash flows are discounted by risk free rates. The DSOP Principle 6.1 states:

“The starting point for determining the discount rate for insurance liabilities and insurance assets should be the pre-tax market yield at the balance sheet date on risk free assets….Risk free assets are those assets with readily observable market prices whose cash flows are least variable for a given maturity and currency”.

Project Updates (2006) more explicitly requires incorporating liquidity in setting the discount rates. Project Updates (2006) paragraph 18 states:

“be consistent with observable market prices for cash flows whose characteristics match those of the insurance liability. Those characteristics include timing, currency and liquidity.”, and,

“exclude any factors that influence observed interest rates but are not relevant to the liability (for example, risks that affect the observed rate but are not present in the liability)”.

The IFRS 4 (2005) already allows insurers to use market interest rates in measuring insurance contracts. IFRS 4 (2005) article 24 states:

“An insurer is permitted, but not required, to change its accounting policies so that it remeasures designated insurance liabilities to reflect current market interest rates and recognises changes in those liabilities in profit or loss”.

16
2.4 Conceptual difference between fair valuation and traditional valuation

Fair valuation is significantly different from traditional valuation. This difference is essentially the difference between the market value and insurer’s own economic value. This conceptual difference leads to the difference in methodology. DSOP article 4.10 (b) states:

“it differs from the traditional approaches by focusing on direct analysis of cash flows and on more explicit statements of the assumptions used in the measurement”.

We conceive that there are three important differences. We will discuss each of them now, aiming at highlighting the characteristics of fair valuation.

(1) Asset and liability measurement approach

We are of the view that the asset and liability measurement approach is more appropriate than the deferral and matching approach for financial reporting of insurance contracts. DSOP article 2.9 states that income and expenses are defined in terms of changes in measurement of insurance assets and insurance liabilities. This is a fundamental departure from the currently prevailing deferral and matching approach where the revenue and expenses from an insurance contract are recognised progressively over time as services are provided. The asset and liability measurement approach also results in a closed book approach that accounts only for contracts in force at the reporting date, since an insurance contract without contractual rights and contractual obligations is not recognized as a liability. The concept has already been incorporated in IFRS 4 (2005) as a prohibition of catastrophe and equalization provisions. IFRS 4 (2005) article 14 (a) states:

“an insurer shall not recognise as a liability any provisions for possible future claims, if those claims arise under insurance contracts that are not in existence at the reporting date (such as catastrophe provisions and equalisation provisions)”.

(2) Use of market assumptions

Our view is that assumptions should be consistent with market prices as much as practically possible. Our view is in line with IASB’s proposals. DSOP article 4.124 states:
“Market assumptions are assumptions about variables, such as interest rates and asset prices, that are readily observed in the financial markets. Fair value will clearly reflect such market data”.

DSOP article 4.133 proposes that the market assumptions should be consistent with current market data, not those projected based on long-term economic relationships. As we have already shown, IFRS 4 (2005) article 24 already permits, but does not require, the use of current market interest rates in measuring insurance contracts.

(3) Independence of liabilities
We strongly support the idea that the insurance liabilities should be valued independently of their backing assets. The cash flows need to be estimated excluding any effects from the backing assets, and discount rates should not be affected by the constituents of backing assets. IASB is also pursuing this method. DSOP article 6.11 states:

“...the discount rate for a liability should reflect the characteristic of that liability, not the characteristics of some other instrument with different features. Accordingly, this DSOP article does not permit a discount rate for insurance liabilities based on any of the following: a) an insurer’s incremental borrowing rate. ...b) an insurer’s cost of capital. ...c) returns on assets held,...”.

2.5 Major issues under discussion
We will discuss some issues, selected from hundreds of issues that have been or are still being discussed, that we think are most important in the insurance fair valuation project. These issues are the backbone of the new valuation paradigm and in most cases cause significant changes in valuation methods that are used at present. Therefore, they are not easily accepted and there is usually resentment and opposition to these issues. While introducing these discussions, we try to point to the essence of the issues, and then clarify our standpoint on them.

2.5.1 Asset and liability measurement approach
We support the asset and liability measurement approach as we have already mentioned in the previous section. Life insurance industry bodies were opposed to, and still seem
uncomfortable with, the asset and liability measurement approach. The American Council of Life Insurers (“ACLI”, hereafter) et al. (2003, p.2), which represents the life industries in the United States, Japan, Germany, and Austria, states:

“We continue to believe that there are viable approaches for insurance contracts. National GAAP in many countries follows a model that tends to be consistent with a Deferral and Matching approach that effectively captures the nature of the business in each respective country”.


“We accept the preference of the Steering Committee for the Asset and Liability Measurement approach under the IASC framework. ... Beyond these observations, we note the importance of developing relevant and appropriate disclosure under an asset-and-liability measurement basis. This will be critical to understandability of the financial statements and “profits” results emerging.”

Nobody would oppose eliminating an accounting mismatch even though there are disagreements on the methods to achieve it. In other words, there would not be any objection to consistent measurements between assets and liabilities. Now IAS 39 is already in force and requires\(^8\) that financial assets on trading books be measured at fair value. In order to eliminate an accounting mismatch, it will be inevitable that the liabilities are measured at fair value.

2.5.2 Reflection of changes in fair values in the profit and loss account
Our view is that the changes in fair values need to be recognized as profit or loss for the period when the changes have occurred. This way of recognition is consistent with

\(^8\) IAS 39 permits a financial asset is categorised in available-for-sale category. Financial assets in this category are measured at fair value, but the unrealised movements in fair value are recognised directly in equity without going through profit or loss.
international accounting standards. Under the asset and liability measurement approach, changes in fair values are changes in liabilities (or assets); therefore, the changes should be recognized as income or expenses. Otherwise, a portion of liability that is not included in the fair value of the liability would be created. In addition to this requirement for income statements to be consistent with balance sheets, practical implications need to be considered. If the changes in fair values are not recognized as profit or loss, insurers would not pay any serious attention to the mismatch between assets and liabilities.

DSOP article 2.9 prohibits the recognition of assets or liabilities other than insurance assets or liabilities. This principle requires that any change in the assets and liabilities directly flows into the profit and loss account in that financial year. IAA (2000a, p.40) is supportive to the DSOP’s view:

“The IAA believes that any deferral of recognition of changes in experience assumptions, whether through a corridor or any other approach, would be inappropriate. Efforts to smooth earnings as a result of the use of average long-term historical experience would not be appropriate”.

The UK Working Party (2001) emphasizes, however, the necessity of sound development of fair valuation, insisting that financial markets sometimes behave irrationally. It states that even in the most liquid and deepest of available markets, market values can sometimes reflect elements of exuberance or pessimism rather than financial views and probabilities. It insists that a reporting framework within which the underlying realities are not obscured by overall earnings volatility should be found. Sheldon and Smith (2004) insist that market consistent valuation requires using market prices in any circumstances saying in paragraph 2.2.1:

“They may have argued that prices were distorted by supply and demand. Fortunately, supply and demand are what modern asset theory is all about. It is when prices are not set by supply and demand that economic theory falters”.

Also in paragraph 3.3.5 they state:

9 The IASB Framework for Preparation and Presentation of Financial Statements (“Framework”, hereafter) (2005) articles 92 and 94 set out that income is recognised when a liability has decreased, and expense is recognized when a liability has increased.
“market consistency dictates the use of actual market prices, even when those prices are affected by unusual supply or demand effects. A believer in market inefficiency might try to correct market prices to where they ’truly’ should be, but the result is not then market consistent”

We agree with the above idea of Sheldon and Smith (2004) because the purpose of financial reports is to provide the general public with neutral and comparable information10. We admit that the market is not perfect, but no single body would present assumptions that the general public regards as better than those set by the market. The financial reports should not be based on managements’ specific views of the future.

In Australia, an advanced accounting system has already been in place known as “Margin On Services (MOS) accounting”. Australian Accounting Standards Board ("AASB", hereafter) 1038 (2005) article 8.5 requires that the effects of a change to adopted discount rates and related economic assumptions be recognised as income or expense immediately while effects of changes to other sources are spread over the remaining period of the contract. This redistribution of the effects to other sources is made by adjusting the margins of revenues that are expected over the periods during which services are provided to policyholders. This treatment is based on the notion that an insurance contract has an aspect of a service contract relationship between the insurer and its retail customers while it has an aspect of a wholesale component that directly deals with financial markets.

In line with this MOS accounting principle, IAAust (2003) proposed the wholesale-retail accounting model. This model is aiming at suppressing the earnings volatility caused by relatively subjective assumptions while allowing the earnings volatility caused by the fluctuation of financial assumptions such as interest rates. IAAust (2003, p.38) states:

“any change in the valuation assumptions relating to the wholesale component (interest rates, claim rates, claims amounts, claims delays, economic margins for value of risk/uncertainty, etc) are reflected immediately in the value of liability. Changes in other assumptions

10 Framework (2005) lists; understandability, materiality, faithful representation, substance over form, neutrality, prudence, and completeness as qualitative characteristics of financial statements.
(voluntary lapses, expense levels, interest rate effects on DAC, etc) are not reflected in the value of the liability, but are absorbed in the value of the profit margins, subject to loss recognition (i.e. profit margins being no less than the minimum above)

Although this approach seems to conform with the practical situation and seems to be working well, there is a theoretical deficiency in light of the asset and liability measurement approach. Under MOS accounting, the planned margin can be decomposed into the current margin, which is the margin level for new contracts, and a de facto reserve. The real margin portion can be regarded as a constituent of the liability, but the reserve portion does not constitute any liability. Therefore, the strict enforcement of the asset and liability measurement approach would require that the reserve be taken outside of the liability explicitly if it is admitted. Project Updates (2006) article 11 proposes a margin model where the margin is composed of a risk margin and a service margin. Both margins should be explicit and unbiased estimates. The function of this proposed service margin is likely to be different from the function of the margin on service that is used in MOS accounting.

Industry bodies insist (for example, ACLI (2000)) that the swings of financial results due to changes in factors such as risk free interest rates damage the reliability and comparability of financial reports. However, any objections whose rationalization is based simply on the proposed method causing excessive volatility are not persuasive. The volatile profit or loss may show the real financial condition of insurers. One objective of fair valuation is to eliminate the practice of averaging out losses. Babbel (1998, p.125) states:

“..., what you are really worried about is volatility of economic surplus, and you remove volatility of economic surplus if you are reasonably well matched. ... If unwanted volatility is revealed when marking-to-market both sides of the balance sheet, you had better take some remedial economic (as opposed to window dressing) actions, because the effect of a change in value is real”.

Fair valuation of liabilities should help to reduce volatility of earnings if assets and liabilities are matched.
2.5.3 Indirect or direct method

Indirect methods are still widely used for life insurance contracts. DSOP article 3.32 states:

“...Direct methods measure the liability by discounting future cash flows arising from a book of insurance contracts. Indirect methods measure the liability by discounting all cash flows arising from both the book of insurance contracts and the assets supporting the book, to arrive at a net measurement for the contracts and supporting assets. The measurement of the assets is then deducted to arrive at a measurement of the book of contracts”.

According to our view, indirect methods should not be used in calculating the fair value of insurance contracts for financial reporting purposes. The reason is that it is extremely difficult for indirect methods to create results that are consistent with current financial markets. The indirect methods, also called embedded value methods, are used in the block acquisition market; hence, they could be characterized as a business appraisal rather than asset or liability valuation. Both methods can result in the same value if assumptions are set carefully, as, for example, Girard (2000) states. They are, however, unlikely to result in the same value in complicated scenarios. It is practically impossible to determine the cost of capital used for discounting net cash flows in the indirect method so that all the assumptions regarding leverage, tax, and risk premium on backing assets, etc, can be incorporated precisely. Babbel (1998) points out that discounting the cash flows net of assets and liabilities using the combined discount rate may lead to errors, and argues that the expected cash flows should be adjusted for their riskiness first. We will show how difficult it is to find correct combined discount rates, using a hypothetical simple example in Chapter 3. Doll et al. (1998) conclude that the direct methods provide a valuation of liabilities which is more consistent with asset valuation, after comparing eight fair valuation methods. The UK Working Party (2001, p.15) also supports the direct method and states:

“an embedded value approach would not be appropriate for the new accounting standard. We understand that the main reason for rejecting embedded value methods, at least as currently applied, is to avoid discounting liability and asset cash flows on bases inconsistent with market prices”.

23
DSOP article 3.33 shows a strong preference for direct methods:

“this DSOP takes the view that direct methods are more transparent and, hence, preferable”.

2.5.4 Insurers’ own credit standing

We believe that the discount rate should not reflect the insurer’s own credit standing contrary to the IASB’s view in DSOP. DSOP Principle 4.8 states:

“The entity-specific value of an insurance liability should not reflect the insurer’s own credit standing. Conceptually fair value should reflect the insurer’s own credit standing, but this would have practical implications that need further investigation”.

The primary reason for our objection against including the insurer’s own credit standing is that the fair value of a specific contract should be determined solely by the nature of the contract, and therefore should not be tied to a specific party even when the party is the insurer of the contract. The UK Working Party (2001, p.20) states:

“if accounting standards permit allowance to be made for the credit standing of the reporting entity, then this should apply only to debt issued by the company and not to its insurance liabilities”.

This statement is on the grounds that it is unlikely that terms of transfer of insurance liabilities to a third party would reflect that entity’s credit standing. There are also concerns that inclusion of insurers’ credit standing may cause wrong incentives for insurers. AAA (2000, pp.11-12) states:

“We believe that the insurer’s credit standing should not be reflected in the estimated fair value of liabilities associated with insurance contracts. ... We are concerned that reflecting an insurer’s credit standing in the fair value of its liabilities could be misleading to equity investors. A deterioration in an insurer’s credit standing would cause a decrease in its liabilities and thus an increase in its earnings. Conversely, an insurer with an improving credit standing would find its earnings penalized”.

Babbel (1998) proposes that the insurer’s own credit risk should be displayed as the value of a put option to default on the asset side of the balance sheet, leaving the insurance liabilities calculated on the non-default basis. We will look into the insurer’s
balance sheet in Chapter 3.

There has been a lot of opposition to including insurers’ credit standings in the fair value of insurance liabilities. In addition, it should be noted that the credit spread would be small anyway. DSOP article 1.161 states:

“In practice, for many regulated insurers, the impact of their own credit standing may be negligible, given supervisory procedures that aim to minimize the possibility of any losses to policyholders”.

This statement implicitly suggests that the insurer’s own credit standing would be ignored in a practical situation when the fair value of insurance liabilities is calculated.

2.5.5 Inclusion of backing asset returns in the discount rates

We strongly support the IASB’s view that backing asset returns should not be included in the interest rates for discounting liabilities. The primary reason for our insistence is that an insurance contract is in principle a contract independent of its backing assets. Thus, the value of the insurance contract cannot be affected by the nature of the backing assets. DSOP article 3.52 recognises that some argue the insurance liabilities should be discounted by the investment return that the enterprise expects to obtain over the lifetime of the liability. Nevertheless, DSOP article 3.53 insists that identical policies provided by insurers with the same credit standing should have the same value regardless of each insurer’s investment strategy; DSOP article 3.55 states that the characteristics of liabilities should not affect the measurement of assets, thus the opposite should be true. Also DSOP article 3.56 states that the asset-liability mismatch risk should not affect the fair value. IFRS 4 (2005) article 27 represents the intention of IASB, saying:

“An insurer need not change its accounting policies for insurance contracts to eliminate future investment margins\textsuperscript{11}. However, there is a rebuttable presumption that an insurer’s financial statements will become less relevant and reliable if it introduces an accounting policy that reflects future investment margins in the measurement of insurance contracts, unless those margins affect the contractual payments”.

\textsuperscript{11} There is no explicit definition found for the term “future investment margins” in IFRS 4. We interpret this term as a synonym for the risk premium, i.e. a portion of an expected return over the risk free rate.
McLaughlin (1998) supports independence of liabilities on the grounds that the assets supporting the liabilities are being traded and are changing their features every day. He also mentions that, in company acquisition situations, the purchaser feels free to determine a price in the knowledge that the portfolio can be turned over, if desired, in a reasonable length of time. The IAA (2000b) proposes to use the replicating portfolio, which is independent of backing assets. The IAA (2000b, p.5) states:

“A portfolio of insurance liabilities can be marked-to-model by creating a matching portfolio and discounting the liability cash flows at the asset yield of the matched portfolio... The matching portfolio is defined here as the “replicating” portfolio and is the set of assets whose cash flows match the liability cash flows as closely as possible, given the assets available in the market”.

The UK Working Party (2001) also follows this approach. The replicating portfolio is in principle composed of risk free assets. On the contrary, the IAA (2000d) proposes that the investment margin, which is the difference between a set of investment earnings and any corresponding interest payments, should also be considered as well as the replicating portfolio. It (p.7) concludes:

“It is thus appropriate to use a market-based discount rate incorporating investment margins to value liabilities”.

We do not agree with this view.

The holding of risky assets may lead to a lower liability value through a higher default probability. Babbel (2002) argues that the increase in the value of a put option is caused by an increase in the volatility of the assets, which is not explicitly related to the expected return of assets. He also insists that the value of the put option to default should be excluded from the values of insurance liabilities, and it should be recognized as an asset separately. We think this is a reasonable approach.

In addition to the discussion about inclusion of backing asset returns, it still needs to be decided what interest rate is used for the discount rates. The AAA (2000) asserts that the discount rates should be chosen carefully. The AAA (2000, p.10) states:

“...However, many experienced pricing and appraisal actuaries consider it
inappropriate to use a risk free rate (where the risk free rate is defined as the government bond rate) as the discount rate even if risks are explicitly provided for in the discount model. An example is the market for Guaranteed Investment Contracts (GICs) ... as these contracts are priced and sold using an interest rate assumption more in line with rates for investment-grade corporate bonds rather than for government bonds”.

One reason why the interest rates in line with corporate bonds are used for pricing GICs is that the discount rates take into account liquidity, as well as credit standings, of GICs. Insurance products are much less liquid than government bonds; therefore investors demand some liquidity premiums over the yield of government bonds. Project Updates (2006) article 18 requires consideration of liquidity in determining the discount rates, and we agree with this principle.

2.6 New risk profile and capital requirement under the fair value principle
A change in the accounting system affects the way of risk recognition and probably the regulatory capital requirements. The counter argument would be that an accounting system is just an interpretation, or a reporting, of what the company has done as hindsight; therefore a change in an accounting system should not have changed any future cash flows arising from the business. However, in reality, the accounting system affects the way of conducting a business significantly. This is because the performance of the company and the rewards for managers of a company for a period is usually measured on the reported profit for the period. Therefore, in many cases, managers set high priorities to items that cause significant effects on reported profits for the current period. These items are commonly called “risk factors”.

Once a loss is officially reported, the nominal capital decreases. The capital deterioration happens even when the financial market is collectively making mistakes and it is almost certain that the market will revert to the previous level sooner or later. The regulators need to take some corrective action once the nominal capital, not the market capitalization, reaches some critical level. In other words, the regulatory capital regime is based on nominal capital. Thus it is influenced significantly by accounting systems.
The fair value of long term insurance contracts is very sensitive to changes in market and non-market assumptions, especially to changes in the yield curve. Unless the assets and liabilities are matched, insurers will experience enormous fluctuation of reported profit or loss due to changes in the yield curve and hence will be subject to high capital requirement.

The risk exposure against the yield curve has in fact been recognized by actuaries, but it was regarded as a form of short-term fluctuation that will be eventually offset by opposite movements anyway, and the impact of it on earnings was averaged out in financial reporting. Thus, there has been little focus on it.

In our study, however, we will investigate the risk-return profile of a hypothetical business placing the change in the yield curve as the major risk factor.

2.7 Summary

In this chapter, we have reviewed the outline of the insurance fair valuation project and explained major issues and showed related discussions. Our view is that the fair value approach is a rational approach for measuring insurance liabilities and this approach will replace the traditional approach sooner or later. We agree with IASB’s proposals, in particular those described in DSOP, for most of the issues under discussion. Since fair valuation is significantly different from traditional valuation methods, it is critical to clarify what conceptual and technical frameworks we are using before we launch actual quantitative work. We will describe the frameworks in Chapters 3, 4, and 5.
3 Valuation under fair value principles

This chapter explains why the financial economics approach is expected to be the dominant methodology for valuing insurance contracts under the fair valuation regime. We highlight the most important concepts underpinning the financial economics approach while comparing them to those in the traditional approach. In addition, we analyse an insurers’ balance sheet structure and decompose the balance sheet into fundamental components. Then we clarify what sort of balance sheet we will assume in our study.

There are many valuation models for financial products, including insurance contracts. We categorize those models into two classes: models under the traditional approach and models under the financial economics approach. We characterize models in the traditional approach, which actuaries have been using for decades, to be return-oriented and projection-based. In the traditional approach, models try to forecast the future development of economic variables as realistically as possible, bearing the long term equilibrium of the variables in mind. In most cases, the most important variable to be forecast is the return on assets. Based on the future projection, present values are calculated using some discount factors. On the other hand, we characterize models in the financial economics approach, which is commonly used for pricing derivatives, to be pricing-oriented and market-based. In the financial economics approach, models are calibrated so that the expected present value of cash flows arising from a financial product fits to the current market price. As we will show in Chapter 4, thanks to the assumption that the market is arbitrage-free, the risk premiums vanish through the process called “probability measure conversion”. Therefore, there is no need to forecast risk premiums. The probability distribution of the price of securities is chosen, on a case-by-case basis, so that the final results become simple as long as the chosen distribution is reasonably realistic.

Under the fair valuation regime, valuation models should guarantee consistency with financial markets. This is the reason why the financial economics approach is expected to be dominant under the fair valuation regime.
3.1 Risk premium and discount factor

We now illustrate that the most notable difference between the traditional approach and the financial economics approach lies in the treatment of risk premiums and discount factors.

Let \( S(t) (t = 0,1) \) be the price of a with-dividend security at time \( t \). \( S(1) \) is a random variable while \( S(0) \) is constant. We assume there is a risk free security and the risk free rate is \( r \) per annum effective. Consider the pricing model:

\[
S(0) = D_s \cdot E[S(1)], \tag{3.1}
\]

where \( D_s \) is the discount factor for this security and \( E[\ ] \) is the expectation operator.

In the financial economics approach, we start from the prerequisite that \( S(0) \) is known. Now, we know neither \( D_s \) nor \( E[S(1)] \). These two are unknown values but they are not free quantities; these are constrained by the formula (3.1).

In a risk-averse economy, any security except risk free bonds incurs a risk premium, by which the expected return of the security exceeds the risk free rate. According to the Capital Asset Pricing Model (“CAPM”, hereafter), the excess return is positive as long as the security price is positively correlated to the market wealth. Now we set:

\[
E\left[\frac{S(1)}{S(0)}\right] = 1 + r + \alpha_s, \tag{3.2}
\]

where \( \alpha_s \) is the risk premium for this security.

Then, \( D_s \) is solved for as:

\[
D_s = \left[1 + r + \alpha_s\right]^{-1}. \tag{3.3}
\]

Note that \( \alpha_s \) and \( D_s \) depend on the choice of the security, i.e. they differ from one security to another.

In the traditional approach, \( \alpha_s \) is one of the most important quantities; therefore \( \alpha_s \) is chosen so that it is as close to reality as possible. In theory, once \( \alpha_s \) is obtained, the correct discount factor can be obtained using (3.3). However, in a practical situation where a portfolio that is composed of a number of assets and liabilities needs to be valued, it is very difficult to get a correct single discount factor as we will show in the
next section.

On the other hand, the financial economics approach starts by assuming that the market is arbitrage-free. Then $\alpha_s$ is set to zero\(^{12}\). Formally, we convert the measure into the risk-neutral measure $Q$ that satisfies:

$$E^Q[S(1)] = 1 + r,$$

where $E^Q[\quad]$ is the expectation operator using the measure $Q$.

Now let $D_B$ denote the risk-free discount factor, which is $[1 + r]^{-1}$. Then we have

$$S(0) = D_B E^Q[S(1)],$$

i.e. $D_B$ is the discount factor for this security under the measure $Q$. It is obvious that $D_B$ is also the discount factor for any security under the measure $Q$. This method is called risk-neutral pricing and is widely used for pricing financial derivatives. The method is applicable for financial products whose payoff can be replicated by other existing securities. Note that the risk-neutral pricing is not assuming that the real world is risk-neutral. This method suggests that risk premiums disappear after the technical procedure; and expected present values (“EPVs”, hereafter) can be correctly calculated in this hypothetical world. We will justify this method mathematically in Chapter 4.

One notable feature of this method is that the probability distribution under the measure $Q$ is not exactly describing what will happen to the security at time 1. This apparent deficiency does not matter when the purpose of the modelling is just to calculate EPVs. This characteristic leads us to state that the financial economics approach is pricing-oriented and market-based. In our example, using the risk-neutral measure, the EPV of $S(1)$ is correctly calculated to be $S(0)$, which is the market price of the security.

In the financial economics approach, there is another pricing method. One drawback of the specification in (3.1) is that $D_j$ is assumed to be deterministic while $S(1)$ is a random variable. Instead of (3.1), we consider the model:

$$S(0) = E[D(1) \cdot S(1)],$$

(3.4)

\(^{12}\) Or some other number depending on the choice of the numeraire. See Chapter 4.
where $D(1)$ is a random discount factor that does not depend on the choice of the security.

This random discount factor is called a deflator. In Chapter 4, we give a formal definition of a deflator and we show its properties mathematically. Using a deflator, the EPVs are calculated consistently with the market while using the real-world probability distribution without using measure conversion technique. The EPV of any security, as long as its payoff can be replicated by existing securities, can be obtained using the deflator.

3.2 Value of the risk premium

Consider the model (3.1). Suppose $S(0)$ is already given as a market price. Then whatever the risk premium we may assume for this security, the EPV of $S(1)$ is, and should be, $S(0)$. That means the assumed risk premium should not affect the result of calculation for EPV of $S(1)$. This principle holds in general; given the market price for a security, the assumed risk premium for this security should not affect the EPV. Now let us assume that the risk premium is zero for this security. Clearly, the price of the risk premium is zero in this case. Then, we change the risk premium assumption for this security to, for example, 3%. The price of this security is still $S(0)$, which means the price of the risk premium is still zero. In fact, the “risk premium” is so-called because the economic benefit on it is given without paying any monetary compensation. Risk premiums are given in exchange for taking risks, not for any cash payments.

Even though the price of risk premiums is zero and risk premiums are given without any monetary compensation, that does not mean the risk premiums do not bring any economic benefits. On the contrary, risk premiums are in fact expected revenue; and therefore significant efforts have been made to precisely quantify risk premiums. Realistic risk premiums need to be included in the expected cash flows for revenue projection purposes. However, the resulting EPV of these cash flows may not be consistent with market prices unless the discount rates are carefully chosen. If the discount rates are determined without taking into account consistency with the market, the calculated EPV may well include some risk premiums. This EPV is still useful and sometimes is even desirable as an economic value that is specific to the valuer, but we
need to bear in mind that this EPV does not reflect the consensus in the market made by
the general public. For reporting to the general public, risk premiums should not be
included in EPVs. Under the fair valuation regime, the prices should follow the market
and the market does not award any price on risk premiums.

In summary, we emphasizes the following points:
1. The market regards the price of risk premiums as zero.
2. Hence, the EPV of cash flows from a security should not be affected by their risk
   premium assumptions.
3. In other words, the discount factor should be set so that the risk premiums vanish in
   an EPV calculation.
4. An attempt to discount net future expected cash flows by a single deterministic
discount rate is prone to incur errors.

The first two points are easily seen by the definition of the risk premium.
Following the notation in (3.2), let \( \alpha_s \) denote the risk premium for this security. We
want to calculate the present value of \( \alpha_s \). We do not know what discount factor should
be applied to \( \alpha_s \) yet. However, we know \( D_s \) that satisfies \( D_s E\left[\frac{S(1)}{S(0)}\right] = 1 \), and
\( D_B \) that satisfies \( D_B (1 + r) = 1 \).

Then, we have:
\[
\text{Present value of } \alpha_s = \text{Present value of } (1 + r + \alpha_s) - (1 + r) \\
= \text{Present value of } E\left[\frac{S(1)}{S(0)}\right] - (1 + r) \\
= \text{Present value of } E\left[\frac{S(1)}{S(0)}\right] - \text{Present value of } (1 + r) \\
= D_s E\left[\frac{S(1)}{S(0)}\right] - D_B (1 + r) \\
= 1 - 1 = 0.
\]

For example, suppose the risk free rate is 5%. Consider two non-dividend paying
securities \( S_1 \) and \( S_2 \), both worth $100, whose risk premiums are 2% and 3% respectively
(total expected returns are 7% and 8% respectively). We construct two portfolios \( A \) and
\( B \) with a 1-year investment horizon:
$A$: hold one $S_1$ share funded by selling the risk free asset, and,

$B$: hold one $S_2$ share funded by selling the risk free asset.

The market values of portfolios $A$ and $B$ are both zero. The expected future value of portfolio $A$ 1 year later is $2$, and that of portfolio $B$ $3$. These values represent risk premiums for each security, assuming that the market values are equal to the expected present values. This fact supports the first point above.

It is also obvious that no discount rate for the net expected future values of portfolio $A$ or $B$ can produce the correct present value, which is zero. This supports the fourth point above. Babbel (1998) shows a similar but more realistic example.

We emphasize that we need to apply the discount factor that depends on the choice of security, and that the discount factor needs to be also multiplied to the principal portion as well as the interest portion when present values are calculated. When a future cash flow arises from a portfolio of securities or insurance contracts, it is difficult to find the correct discount factor which can be applied to the net single expected cash flow. For example, suppose we have the random cash flows $C_1, C_2, \ldots, C_n$ at time 1, and we know the net single expected cash flows $C$, i.e.:

$$C = E[C_1] + E[C_2] + \cdots + E[C_n].$$

Then, the present value of this net single expected cash flow $C$ is obtained by:

$$D_1 \cdot E[C_1] + D_2 \cdot E[C_2] + \cdots + D_n \cdot E[C_n],$$

where $D_i$ is the discount factor for the cash flow $C_i$.

The simplified method might be that we apply a single discount factor $D$ to the net single expected cash flows. Then the present value is calculated as:

$$D \cdot E[C].$$

However, there is no simple way to find the correct $D$.

The difficulty with finding the correct discount factors leads to point 4. It is very likely to cause a result that is not consistent with market prices if we apply a single deterministic discount factor to the net cash flow arising from different securities or contracts.
3.3 Financial leverage effect

The example shown in the previous section did not consider any capital injection. In reality, any business has capital injected as well as some funding arrangement at its inception. The ratio between the capital and the funding is called financial leverage. Now, let us consider the following simple structure of a business:

Asset: Some equity portfolio whose risk premium is \( a \)%,

Funding: 
\((1 - x\%)\) of selling the risk free security, and,
\(x\%\) of capital injection.

For example, suppose we set \( x \) to 20\% and the total portfolio size to 100 million. Then 20 million is raised as the capital of this business, and 80 million is funded by issuing a bond or borrowing from banks. We assume that the interest rate on the bond or borrowing is the same as the risk free rate, which is \( r \)% per annum. The total fund raised amounts to 100 million, and this whole amount is invested in an equity portfolio. The expected return of this equity portfolio is \((r + a)\)% per annum.

We consider the expected return for the capital provider for a one-year time horizon. For simplicity, we assume that default of the issued bond or failure to repay the borrowed money never happens. By a simple calculation:

\[
\text{Expected return on capital} = \text{Risk Free Rate} + \frac{a}{x}.
\]

When the risk free rate is 5\%, the expected return on capital, which depends on the risk premium of the equity and the financial leverage, is shown in Table 3-1.

<table>
<thead>
<tr>
<th>Risk Premium ((a))</th>
<th>Capital Portion ((x%))</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0%</td>
<td>25.0%</td>
<td>15.0%</td>
<td>9.0%</td>
<td>7.0%</td>
<td></td>
</tr>
<tr>
<td>2.0%</td>
<td>45.0%</td>
<td>25.0%</td>
<td>13.0%</td>
<td>9.0%</td>
<td></td>
</tr>
<tr>
<td>3.0%</td>
<td>65.0%</td>
<td>35.0%</td>
<td>17.0%</td>
<td>11.0%</td>
<td></td>
</tr>
</tbody>
</table>

The expected return on capital is very sensitive both to the risk premium of the equity and to the financial leverage. The expected return on capital is the discount rate that is applied to the capital of this business. Therefore this result shows the difficulty in setting correct discount rates for the capital. This is the reason why indirect methods in
insurance valuation, in which the cash flows arising from liabilities are mixed with those arising from assets, are likely to produce results that are not consistent with market-based assumptions, even though these two results can be theoretically reconciled as Girard (2000) suggests. Note, however, that risk-neutral valuation, where the risk premium is set to zero and any cash flows are discounted by the risk free rate, always gives the correct present value of the capital regardless of the financial leverage.

3.4 Certainty-equivalent cash flow

We now choose the model (3.4) where discount factors are random. Consider a life insurance contract which simply provides protection for death or survival without any participating features. Let \( V(t,T) \) denote the expected present value at time \( t \) of the benefit that is payable at time \( T (t<T) \). Then \( V(t,T) \) is given by:

\[
V(t,T) = E_{r,y} [D_r(t,y,r) \cdot X_r(t,y,r)],
\]

where \( r \) represents the risk free rate that is random, \( y \) represents mortality, \( X_r(t,y,r) \) denotes the random benefit amount observed at \( t \) and payable at time \( T \), and \( D_r(t,y,r) \) is the risk discount factor for the cash flow \( X_r(t,y,r) \).

In this study, we assume that:

- \( X_r(t,y,r) \) is not a function of \( r \), so that this term is written as \( X_r(t,y) \),
- \( D_r(t,y,r) \) can be written as the product of independent two parts, one related to interest rates and the other to mortality, as \( D_r^{(i)}(t,y) \cdot D_r^{(m)}(t,r) \).

These assumptions are not far from reality because any mortality experience, i.e. whether people survive or die, is largely independent of interest rates. There may be some dependency between the lapse or surrender rates and interest rates, but we do not assume any dependency in our study since:

- Our primary focus is the effect of market risks;
- The policyholders hold insurance policies as protection for their mortality, not as financial assets. Therefore, the correlation is expected to be low, if any; and
- There is not enough data available for a statistical study for this purpose.

Under these assumptions,
\[ V(t, T) = E_{t,y} \left[ D_t^{(y)}(t,y) \cdot D_T^{(y)}(t,y) \cdot X_T(t,y) \right] \]
\[ = E_{t} \left[ D_t^{(y)}(t,y) \right] \cdot E_{y} \left[ D_T^{(y)}(t,y) \cdot X_T(t,y) \right]. \]  

(3.5)

Since \( E_{t} \left[ D_t^{(y)}(t,y) \right] = E_{t} \left[ D_t^{(y)}(t,y) \times 1 \right] \), this is the price of the risk free zero coupon bond at time \( t \) that expires at time \( T \), which is denoted by \( P(t,T) \). Formula (3.5) states that the EPV of the benefit is the risk-adjusted cash flow of the benefit, \( E_{y} \left[ D_T^{(y)}(t,y) \cdot X_T(t,y) \right] \), discounted by the price of the risk free zero coupon bond. In other words, \( E_{y} \left[ D_T^{(y)}(t,y) \cdot X_T(t,y) \right] \) defines a certainty-equivalent cash flow.

We further assume that the expected benefit is not a function of \( t \), i.e. the cash flows occur as expected. Then \( E_{y} \left[ X_T(t,y) \right] \) is solely a function of \( T \) and we let \( C(T) \) denote it, i.e. \( C(T) = E_{y} \left[ X_T(y) \right] \).

We define \( RA(t,T) \) to be:
\[ RA(t,T) = \frac{E_{y} \left[ D_T^{(y)}(t,y) \cdot X_T(t,y) \right]}{E_{y} \left[ X_T(y) \right]} - 1. \]

Then:
\[ E_{y} \left[ D_T^{(y)}(t,y) \cdot X_T(t,y) \right] = C(T) \cdot \left[ 1 + RA(t,T) \right]. \]

We call \( RA(t,T) \) a mortality risk adjustment factor.

Using this notation, (3.5) can be expressed as:
\[ V(t, T) = P(t,T) \cdot C(T) \cdot \left[ 1 + RA(t,T) \right] \]
\[ = P(t,T) \cdot C(T) + P(t,T) \cdot C(T) \cdot RA(t,T). \]  

(3.6)

The first term in the formula (3.6) is the expected present value of the benefit without considering the compensation for the mortality risk taken by the insurer. We call this portion “the core protection value”. The term \( C(T) \cdot RA(t,T) \) is the risk premium the policyholder pays at time \( t \) in exchange for transferring the mortality risk to the insurer. IAA (2000c) call this term a market value margin ("MVM", hereafter)\(^{13} \), and we follow

\(^{13}\) Project Updates (2006) article 11 calls it a risk margin. While the risk margin is required for bearing risk, the article 11 proposes another margin called a service margin for providing other services. In our study, we let MVM include the service margin and mean the total margin.
this naming. The CAPM asserts that the second term should be zero since the insurance risk is almost uncorrelated with the risk on the market portfolio. However, in real insurance markets, the MVM still exists probably because insurers cannot distribute insurance risks efficiently over the entire market portfolio (see, Babbel (2002)).

The form of \( D_y(t, y) \) and that of \( RA(t, T) \) in the real world would be complex and would be changing over time with the development of insurance and capital markets. These are not observable and hence it would not be practical to look for the exact form of \( RA(t, T) \). In our study, we assume that \( RA(t, T) \) is constant regardless of \( t \). Then the MVM is proportional to the core protection value and is exposed to interest rate risk.

3.5 Decomposition of the balance sheet

Insurance contracts are generally very complicated and they often include covenants where some part of the economic benefit arises from the performance of the backing assets or even the performance of the business. Therefore, in order to fully understand insurance contracts, we should look into the structure of insurers’ balance sheets, not only insurance liabilities themselves. Here we analyze insurers’ balance sheets using a simplified setting.

Consider the simple business structure we have used in the Section 3.3 where the risk free security (we simply call this “a bond” hereafter) was issued by the firm for funding. We did not consider the situation where the firm becomes insolvent. Now we discuss the value of the bond and capital when default by the issuer of the bond is allowed. In this section from now on, we will use the term “equity” for the capital of the firm, not for the equity as the assets held by the firm.

The real option argument says that, in general, the value of any corporate bonds can be decomposed into a risk free bond and a (short) put option on the assets of the firm, under a simple setting of the firm with a single zero coupon bond issuance (see, for example, Babbel (2002)). As the put-call parity suggests, we can view equity as a call option on the assets of a firm with a strike price equal to the face value of the bond. This simple structure is illustrated in Figure 3-1.
When the financial condition of the firm is sound, the put option is deeply out-of-the-money; equivalently the call option is deeply in-the-money. Therefore the value of the put option is too small to be noticed. The put option value, however, increases significantly when the firm comes close to insolvency.

![Figure 3-1 - Example of the balance sheet of the simple firm](image)

As shown, the value of the bond, from bond holders’ point of view, is the value of the risk free bond with the value of the put option subtracted. As the assets of the firm deteriorate, the value of the put option increases; and then the value of the bond decreases. Note that the put option does not create any value for the firm. The option just shifts a part of the bond value to equity value.

The key observation that Babbel (2002) makes is that, given a risk free rate, the value of the bond is determined by the volatility of the assets while the expected return on assets is not an explicit component of the value of the bond. This is true since the value of the risk free bond is determined solely by the risk free rate, and the value of the put option is determined by the volatility of the underlying asset given the risk free rate. Adding the risk premium onto the risk free rate for discounting the bond results in the wrong put option value, since there is no direct\(^\text{14}\) relationship between the option value and risk premiums. This observation is also consistent with the point we have already made that risk premiums should not affect any present value calculations.

Note that the volatility of the assets does not affect the value of the assets itself; it changes only the proportion of ownership of the firm’s assets between the bondholders

\(^{14}\) However, there is an indirect relationship. High risk premiums lead to high volatilities, and hence high option value.
and equity holders of the firm. The more volatile assets the firm holds, the more value shifts to equity holders through the higher put option value caused by the higher volatility of the assets. Clearly, the credit standing of the bond decreases when the firm holds more volatile assets. Note that this is caused by the higher volatility not by the higher risk premiums.

We agree with the view by Babbel (2002) that the decomposition of the bond into the risk free bond and the put option increases transparency of reporting and is valuable for analysts, regulators, investors and management. Even though insurance liabilities are not bonds, an insurance liability works as a funding instrument for an insurer. The fund raised by the insurance liabilities is invested into backing assets. Therefore, the structure of the balance sheet is in essence the same as what we have analyzed so far. The redemption amount of insurance liabilities is exposed to mortality risk, i.e. the amount varies depending on mortality. We, however, put aside mortality risk in order to focus on the financial aspect of insurers. Under the assumption that the insurance payments are made as initially expected as we formulated in formula (3.6), the issuing insurance liabilities are economically the same as issuing a portfolio of zero coupon bonds. We will proceed with our discussion on this basis.

Insurance liabilities are usually protected by policyholder protection schemes. In other words, the put option to default is provided by the protection funds, giving ‘default free’ status to policyholders. The put option provided by the protection funds constitutes an intangible asset. The provision of the protection inflates both sides of the balance sheet, the intangible assets and the liabilities for policyholders. The intangible asset belongs to policyholders since only the policyholders benefit from the protection. Equity holders of the insurer do not benefit from the protection; thus equity holders have no share of the intangible asset. Figure 3.2 shows the balance sheet of an insurer with the protection fund under this simple setting.
Due to the protection fund, the insurance liabilities are always solvent. Thus the value of the liabilities is always measured on a no-default basis. In this simple case, the insurer would be tempted to hold as many volatile assets as possible. Without protection funds, the total asset value is unaffected while holding more volatile assets leads to a higher equity value and a lower liability value. With the protection fund, however, the liability value is protected; and only the equity value increases while the liability value is unchanged. In this sense, the protection scheme practically encourages insurers to hold riskier assets since insurers are always required to achieve a higher return on equity.

### 3.6 A model of an insurer’s balance sheet

In reality, another intangible asset is created through running the business, namely a franchise value. Babbel (1998, p.116) states:

“The franchise value stems from what economists call economic rents. This is the present value of rents that an insurer is expected to garner because it has scarce resources, scarce capital, charter value, licenses, a distribution network, personnel, reputation, and so forth. It includes renewal business and that sort of thing”.

Babbel (1998) argues that the net value of any financial institution is comprised of three elements, the franchise value, the liquidation value and the put option value. He states (p.116):

“Liquidation value is simply the market value of tangible assets, less the
Now the balance sheet of an insurer is shown in Figure 3-3, adding the franchise value to the balance sheet of Figure 3-2.

Compared to Figure 3-2, both sides of the balance sheet have increased by the franchise value. Note that the franchise value has increased only the equity value with the liability value unaffected. The franchise value is extremely difficult to quantify. Therefore, it is not shown on financial statements except in very limited circumstances.

One of the reasons why the valuation of insurance liabilities is so difficult and conceptually confusing is that, according to our view, a policyholder may enjoy some economic benefits, even when a contract is not explicitly stated as participating, which would belong to equity holders of the firm if the fund were raised by bonds not by insurance liabilities. A policyholder tends to have rights similar to those of equity holders in that the policyholder receives payments that are linked to the asset (or firm) performance in the form of dividends or variable benefit amounts. This participation takes away some portion of equity value that would otherwise belong to shareholders wholly.

Therefore, life insurance valuers are tempted to apply the method that is used for equity
valuation to insurance liability valuation since the life insurance liabilities have some equity component. Such a valuation method is, however, not applicable to general purpose reporting that requires consistency with markets.

The equity valuation usually involves forecasting expected returns on the assets that include the risk premiums, and other benefits such as franchise value. As a result, calculated expected cash flows tend to include the risk premiums and franchise value. The present value is calculated by discounting these expected cash flows using some discount rate that is the cost of the capital of the firm in most cases. This is correct or sometimes even essential if the objective of the valuation is to forecast the expected revenues and to calculate an economic value that is perceived solely by the valuer. However, as we have already shown, the resulting present value is not consistent with markets unless the discount rate is intentionally and carefully chosen. Therefore this method is not applicable for reporting purposes in a fair valuation environment where the valuation needs to be consistent with markets and to be acceptable to the general public.

IFRS 4 (2005) paragraph 34 recommends, but does not require, that a discretionary participation feature be recognized separately from the guaranteed element and be classified as either a liability or a separate component of equity. In our study, we also propose that we split conceptually the value of an insurance policy liability into two parts: protection value and policyholder’s equity value. The policyholder’s equity value represents the right to participate in the distribution of the profits of the insurer, whether the right is explicitly stated or it is given as the right to receive dividends. Note that the bonus loading that was explicitly allowed for in life insurance premium calculation is a part of protection value, since the predefined bonuses are just repayments of loadings for the bonuses included in gross premium income.

The profit margin in excess of the MVM generates a profit at point of sale of the insurance policy (see Hairs et al (2001)), which could be classified as an up-front profit for shareholders. This is not regarded as a component of the policy liability value in our research.

Taking the MVM and policyholders’ participating right into account, we redefine or
define three terms here: core protection value, protection value and the insurance policy liability value.

- The core protection value stems solely from the guaranteed component, excluding MVM.
- The protection value includes the MVM on top of the core protection value.
- Insurance policy liability value is the whole value of the insurance policy. This value includes\textsuperscript{15} the policyholders’ equity value on top of the protection value.

Using these terms, the insurer’s balance sheet is shown in Figure 3-4.

\textbf{Figure 3-4 - Proposed balance sheet of the insurer}

<table>
<thead>
<tr>
<th>Market Value Of Tangible Assets</th>
<th>Put Option Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Core Protection Value</td>
</tr>
<tr>
<td></td>
<td>Protection Value</td>
</tr>
<tr>
<td></td>
<td>MVM</td>
</tr>
<tr>
<td></td>
<td>Policyholders' Equity Value</td>
</tr>
<tr>
<td></td>
<td>Up-Front Profit</td>
</tr>
<tr>
<td></td>
<td>Equityholders' Equity Value</td>
</tr>
</tbody>
</table>

Note that, compared to Figure 3-3, the equity value has decreased by the policyholders’ equity value. When the business is liquidated, the franchise value and the put option value disappear. Assuming that policyholders also lose their participating rights, the liquidation value is the value of the tangible assets with the protection value subtracted as shown on the Figure 3-4.

\textsuperscript{15} This is largely in line with Project Updates (2006) article 39 that states: “The Board has tentatively concluded that policyholder participation rights create a liability when the insurer has an unconditional obligation to transfer economic benefits to policyholders, current or future”.

44
3.7 The scope of our research

One objective of our research is to analyze the equity holders’ value by evaluating the insurance liability under the fair valuation regime. We omit several components that are not essential for our study, as described below, from the insurer’s balance sheet.

- We do not consider any franchise value.
- We do not consider the protection fund. Instead, we require minimum capital for insurers.
- Policyholders have no explicit or implicit participating rights.

Figure 3-5 shows the balance sheet that we will deal with.

![Figure 3-5 - The balance sheet assumed in this study](chart)

The dotted arrow in Figure 3-5 shows the transition of ownership. At the inception of a policy, the MVM will be included in the insurer’s liability. As time passes, the MVM will gradually be realized as profit and come under shareholders’ ownership.

3.8 Summary

In this chapter, we showed that the market value of risk premiums for securities is zero. This leads to difficulty in choosing discount factors that produce present values that are consistent with market prices. However, in the financial economics approach, discount factors are determined taking into account assumed risk premiums so that present values are consistent with market prices. Thus the financial economics approach is expected to be a dominant methodology in fair valuation environment. In addition, we have shown that an insurance liability is generally very complicated but it can be decomposed into several components. We clarified that we will consider the MVM and put option value,
but we do not consider policyholders’ equity value or franchise value in our study.
4 Valuation using deflators

In Chapter 3, we have illustrated that the financial economics approach is expected to be the dominant methodology for valuation of insurance contracts in the fair valuation environment. In this chapter, we will explain deflators\(^ {16}\) and related concepts, in particular a market-price-of-risk process\(^ {17}\), in detail. Deflators are a key concept for modern financial economics and all models based on the modern financial economics approach are built using this concept implicitly or explicitly. When we try to apply existing models to problems other than those that the models were originally designated for, it is very important that we carefully examine assumptions on which these models were built. Therefore, we need to understand deflators and related concepts before we put a financial economics model to use.

Most established models based on the modern financial economics approach were developed in order to price specific derivative products. Each model uses the most appropriate measure for its purpose; therefore the stochastic development of the underlying securities under the model is different from that in the real world. However, this does not mean that the application of these models or the idea behind these models is limited to their original purposes. These models have survived for many years and are still actively used in practical fields. In addition, a significant amount of research has been conducted and there is a large amount of resources available for these models. In other words, these models are already standard in financial modelling work. Therefore these models may be applicable to, or give us some useful insights into, our modelling work in finance or insurance.

\(^{16}\) A deflator is defined in Section 4.3.2.

\(^{17}\) A market-price-of-risk process is defined in Section 4.5.3.
4.1 Introduction

Under the fair valuation regime, valuation models are expected to be built under arbitrage-free assumptions. The reason is that arbitrage-free assumptions ensure that the calculated prices are consistent with market prices of securities; and this consistency is the major requirement under the fair valuation environment. We admit that the assertion that the market is arbitrage-free\textsuperscript{18} is a strong assumption, and in most markets arbitrage trading would not be possible due to various constraints on trading even if there were a chance of a riskless arbitrage profit. However, there is no other methodology at present by which we can secure the consistency. Once we obtain the prices calculated under arbitrage-free assumptions, we would be able to refine these prices taking the constraints in real markets into account.

We will show that prices of financial products are calculated, in a well-functioning and arbitrage-free market, as the expected value of future weighted payoffs whose weights are deflators. We will also show that the market-price-of-risk process alone determines the deflator in a continuous-time setting once the numeraire security\textsuperscript{19} is chosen.

In the course of this discussion, we will see that the pricing method using deflators, which we call “deflator pricing”, is equivalent to the pricing method using the equivalent martingale measure, which is the so-called “risk-neutral pricing”. Both methods are driven by the deflator generator process explicitly or implicitly.

We do not have to assume any utility function in either method. We start from the

\textsuperscript{18} Arbitrage is formally defined in Section 4.2.

\textsuperscript{19} A numeraire security is formally defined in Section 4.3.3.
assumption that market prices of all the securities traded in the market uniquely exist all
the time. The market price of a security exists since the expected returns on the security
have already been discounted by some discount factor that represents market
participants’ collective utility. In other words, utility information is contained in market
prices. This means that financial economics models require fewer assumptions than
traditional econometrics models where utility functions need to be defined in most cases.
Risk-neutral pricing requires even fewer assumptions than deflator pricing. In
risk-neutral pricing an assumption on expected returns is not required; the excess
returns disappear and are implicitly absorbed into the converted measure. This is one
reason why derivative pricing, a typical example of which is the Black-Scholes formula,
is flourishing so well among practitioners and academics.

In our study, we will use deflator pricing, not risk-neutral pricing. The reason why we
use deflator pricing is that we need to look into what will happen in the future in the real
world in order to conduct a return and risk study. Deflator pricing does not change the
measure, and so the projection of prices of financial products shows what will happen in
the real world in the future. On the other hand, the projection in the risk-neutral method
tells us what will happen when the world is risk-neutral. While both methods result in
the same expected present value, by applying a different discount factor in each method,
the risk-neutral method is not applicable when a real-world projection is required.

These two methods are based on the same valuation principle, which underlies modern
financial economic models. This fundamental valuation model is well described by, for
example, Duffie (2001) and Panjer et al. (1998); and underlying ideas in this chapter
were sourced from this literature.

After a description of principal concepts, we will introduce the single-period model. The results under the single-period case will be extended to the multi-period case. Then we will move on to the continuous-time case. After an illustration of a valuation example on a plain call option on equity, we will look into a practical issue of a simulation environment.

4.2 Defining financial markets
Suppose there are $N$ financial products publicly traded in the market, (“securities” hereafter). Here we make two fundamental assumptions:

- We can sell or buy any number, including fractions, of securities at any time without affecting the market price of these securities.
- There always exists only one market price for each security.

We call the former assumption the “friction-free assumption”, and the latter the “one-price principle”.

These assumptions may seem obvious but many authors do not state these assumptions explicitly. However, we should bear in mind that these assumptions are rather artificial and that no such market exists in reality. Indeed, in reality, the tradable amounts are limited, transaction costs are incurred, and trading activities are restricted by technical and regulatory constraints in any market. Nevertheless, the friction-free and one-price assumptions are essential for valuations to be mathematically tractable.
Let $S^{(i)}(t) (t \geq 0, i = 1, 2, \ldots, N)$ denote the price of the $i$-th security at time $t$, and $S(t)(t \geq 0)$ denote the column vector whose $i$-th component is $S^{(i)}(t)$. We assume that there are no dividend payouts and prices are always strictly positive.

The future price of a security involves some uncertainty. This uncertainty is represented by states. A state, which is denoted by $\omega \in \Omega$, is an event, one of which will take place, and $\Omega$ is the whole set of the events. Let $S^{(i)}(t, \omega)$ denote the price of the $i$-th security at time $t$ when the specific event $\omega$ has occurred, and let $S(t, \omega)$ denote the column vector whose $i$-th element is $S^{(i)}(t, \omega)$.

A portfolio is a set of securities. We define the market value of a portfolio to be the sum of the product of the number of units held and the price of the security over the position held by the portfolio. The friction-free assumption allows us to create a portfolio in which each security is composed of any number, including fractions, of units. The constituents of the portfolio may change over time. A trading strategy is a series of portfolios that changes as time progresses. Let the $N$-dimensional row vector $\theta(t)$, which is a random vector, denote the trading strategy at time $t \geq 0$, and let $V^\theta(t)$ denote the market value of the portfolio created by a trading strategy $\theta(t)$ at time $t > 0$. By definition, $V^\theta(t) = \theta(t)S(t)$.

A trading strategy $\theta(t)$ is said to be self-financing if the portfolio created by the strategy involves no interim cash inflows or outflows. In a discrete-time setting, the process $\theta(t)$ is self-financing if:

$$\theta(t_{k-1})S(t_k) = \theta(t_k)S(t_k)$$

(4.1)
is satisfied for \( k = 1, 2, ..., T \), where \( T \) is the maximum time we consider.

In a continuous-time setting, the process \( \theta(t) \) is self-financing if:

\[
d[\theta(t)S(t)] = \theta(t)[dS(t)]
\]  
(4.2)

is satisfied for \( t \geq 0 \).

Arbitrage is defined in the context of the portfolio. An arbitrage is a self-financing trading strategy\(^{20}\), \( \theta(t) \), such that if \( \dot{V}^{0}(0) < 0 \) then \( \dot{V}^{0}(T) \geq 0 \) \(^{21}\) holds, or if \( \dot{V}^{0}(0) \leq 0 \) then \( \dot{V}^{0}(T) > 0 \) \(^{22}\) holds. This means we can make a riskless profit by the strategy \( \theta(t) \). If an arbitrage exists, we say that the model or the market admits an arbitrage.

Suppose there are \( M \)\(^{23}\) states in the market, namely \( \omega_1, \omega_2, ..., \omega_M \). For each security, the payoff can be represented as an \( M \)-dimensional vector whose \( j \)-th element is a specific payoff when the state \( \omega_j \) happens. The market is called\(^{24}\) complete if all the points in \( \mathbb{R}^M \) can be attained by a linear combination of the payoff of the \( N \) securities.

In this study, we assume that markets are arbitrage-free and complete. When the market is arbitrage-free and complete, deflators and equivalent risk-neutral measures are

\(^{20}\) We followed Duffie (2001) for this definition in that an arbitrage refers to a trading strategy. Some literature, for example Panjer et al. (1998), uses the term arbitrage as an act rather than a strategy.

\(^{21}\) This means the left term can never be negative whatever event comes true.

\(^{22}\) This means the left term can never be negative whatever event comes true, and will be strictly positive with some strictly positive probability.

\(^{23}\) In a continuous setting, the number of states is infinite.

\(^{24}\) In a continuous setting, the market is complete when the rank of the volatility matrix is full almost everywhere. See Duffie (2001, p.118).
uniquely defined. When the market is arbitrage-free but not complete, deflators and risk-neutral measures exist but are not unique.

In the market that admits arbitrage, we need to rely on equilibrium pricing where we need to assume some utility function. See, for example, Panjer et al. (1998) for details of equilibrium pricing.

4.3 Single-period model

In this section, we will introduce deflator pricing and discuss its equivalence to risk-neutral pricing in a single-period setting. Most of the underlying concepts are sourced from Panjer et al. (1998) and Duffie (2001). We follow Duffie (2001) for the notation.

Here in the single period model, we assume that there are $M$ states at a fixed time $t$. We define a payoff matrix, which is an $N \times M$ matrix $D(t)$ where the $j$-th column is the vector of prices of $N$ securities when the event $\omega_j$ happens, denoted by $S(t, \omega_j)$. Let $D^{(i,\cdot)}(t)$ denote the $i$-th row of $D(t)$, which is the payoff for the $i$-th security.

4.3.1 The state price vector

The state price vector, following the terminology by Panjer et al. (1998), is an $M$-dimensional strictly positive column vector $\psi(t)$ that satisfies:

$$S(0) = D(t)\psi(t).$$

Let $\psi(t, \omega_i)$ denote the $i$-th element of $\psi(t)$.

---

25 A vector $x$ is said to be positive when all the elements are non-negative and $x$ is not zero. A vector $x$ is said to be strictly positive when all the elements are positive.
Now we consider the simple case where \( N = M = 2 \). Suppose we are at time zero and we know the prices of securities, \( S^{(1)}(0) \) and \( S^{(2)}(0) \). In addition, we know the following payoff matrix that shows what can occur at time \( t \):

\[
D(t) = \begin{pmatrix}
S^{(1)}(t, \omega_1) & S^{(1)}(t, \omega_2) \\
S^{(2)}(t, \omega_1) & S^{(2)}(t, \omega_2)
\end{pmatrix}.
\]

Consider the following equation:

\[
\begin{pmatrix}
S^{(1)}(0) \\
S^{(2)}(0)
\end{pmatrix} = D(t)\boldsymbol{\psi}(t) = \begin{pmatrix}
S^{(1)}(t, \omega_1) & S^{(1)}(t, \omega_2) \\
S^{(2)}(t, \omega_1) & S^{(2)}(t, \omega_2)
\end{pmatrix} \begin{pmatrix}
\psi(t, \omega_1) \\
\psi(t, \omega_2)
\end{pmatrix}.
\]

We will show each element of \( \psi(t) \) is the price of an Arrow-Debreu security\(^{26}\), and therefore strictly positive under the arbitrage-free assumption.

When the market is complete, i.e. when the rank of \( D(t) \) is 2 in this case, \( D(t) \) is invertible. Therefore \( \psi(t) \) is uniquely obtained to be \( D(t)^{-1}S(0) \). Since the market is friction-free, there is a trading strategy \( \theta(0) \) that satisfies \( V^\theta(t) = \theta(0)D(t) = (1 \ 0) \), namely \( \theta(0) = (1 \ 0)D(t)^{-1} \). The portfolio created by this strategy is a synthetic Arrow-Debreu security because the market value of the portfolio at time \( t \) is 1 only when the event \( \omega_1 \) occurs and otherwise the market value of the portfolio at time \( t \) is 0.

The market value of this portfolio at time zero is:

\[
\theta(0)S(0) = \left[(1 \ 0)D(t)^{-1}\right]S(0) = (1 \ 0)\left[D(t)^{-1}S(0)\right] = (1 \ 0)\psi(t) = \psi(t, \omega_1).
\]

This means \( \psi(t, \omega_1) \) is the state price of \( \omega_1 \), which is defined to be the price of the Arrow-Debreu security that pays 1 only when the event \( \omega_1 \) has occurred.

\[\text{---------------------}\]

\(^{26}\) A security that pays 1 in one specific state and pays zero otherwise is called an Arrow-Debreu security.
The assumption that the market is arbitrage-free is needed to ensure that \( \psi(t, \omega) \) is positive. In other words, without this assumption, \( \psi(t, \omega) \) could be negative. We see this by the following simple argument. An Arrow-Debreu security has a positive payoff. In an arbitrage-free market, the market value of a portfolio with positive payoff must be positive. Therefore, \( \psi(t, \omega_1) \) must be positive in order for the market to be arbitrage-free. Similar arguments lead us to the conclusion that \( \psi(t, \omega_2) \) is the state price of \( \omega_2 \), which is positive. Putting these together, \( \psi(t) = \left( \begin{array}{c} \psi(t, \omega_1) \\ \psi(t, \omega_2) \end{array} \right) \) is strictly positive and therefore \( \psi(t) \) is a state price vector.

Now we have shown that the state price vector exists under the arbitrage-free and complete market assumption, and the price of each security is the weighted sum of the state prices, whose weights are the payoffs of the security at time \( t \) when a specific event happens. Note that the state price vector is unique, i.e. it is common for both securities.

Conversely, in a complete market, the market is arbitrage-free if the state price vector exists. This is obvious since any positive payoff of a security results in a positive price when the state price vector exists. By definition, the state price vector is strictly positive. Therefore, the sum of the product of each payoff and each state price is also positive. The sum is, by definition, the price of the security.

In summary, in a complete market, the arbitrage-free condition is equivalent to the
existence of the state price vector. Though we have shown this result when the number of states is two, it is obvious that this equivalence can be expanded to the general case where the number of states is $M$ as long as the market is complete. Duffie (2001) provides a concise mathematical proof for more general cases, including when the market is not complete, using the Separating Hyperplane Theorem. Also, detailed discussions with examples are given by Panjer et al. (1998). The following theorem is known as the fundamental theorem of asset pricing. See, for example, Panjer et al. (1998, p.185).

Theorem: The single-period securities market model is arbitrage-free if and only if there exists a state price vector.

Suppose the state price vector is obtained or given for a set of securities on some states. The state price vector is also the state price vector for any other securities as long as the payoff of these securities is defined on the same states. We will illustrate this using our previous two-dimensional example. Consider a third security whose payoff is defined on the states $\omega_1$ and $\omega_2$. These states are the same as for securities 1 and 2. The payoff of the third security, denoted by $(S^{(3)}(t, \omega_1) \quad S^{(3)}(t, \omega_2))$, should be a linear combination of the payoffs of security 1 and 2, since the two row vectors of the payoff matrix $D(t)$ span $\mathbb{R}^2$. Equivalently, there exists a trading strategy $\theta(0)$ that satisfies $(S^{(3)}(t, \omega_1) \quad S^{(3)}(t, \omega_2)) = \theta(0)D(t)$. The market value of the portfolio created by this trading strategy $\theta(0)$ is $\theta(0)[D(t)\psi(t)] = \theta(0)S(0)$. The price of the third

---

27 This property is sometimes described as “the cash flow of the third security is attainable”. See, for example, Panjer et al. (1998).
security at time zero, denoted by $S^{(3)}(0)$, should also be $\theta(0)S(0)$. If, for example $S^{(3)}(0) > \theta(0)S(0)$, we can obtain a riskless profit simply by selling the third security and buying the portfolio at time zero. Now we have:

$$S^{(3)}(0) = \theta(0)S(0) = \theta(0)[D(t)\psi(t)] = [\theta(0)D(t)]\psi(t) = \left(S^{(3)}(t, \omega_i) S^{(3)}(t, \omega_j)\right)\psi(t),$$

showing $\psi(t)$ is also a state price vector for the third security.

This result expands to the $M$-dimensional case; see, for example, Panjer et al. (1998).

### 4.3.2 Deflator pricing and risk-neutral pricing

In this section, we will introduce the state price deflator and define deflator pricing. In addition, we will show that risk-neutral pricing is equivalent to deflator pricing. Both methods are interchangeable, i.e. we can switch risk-neutral pricing to deflator pricing and vice versa.

We have shown that, given an $M$-dimensional ($M \leq N$) state price vector $\psi(t)$, the price of a security is obtained by:

$$S^{(i)}(0) = D^{(i,\omega)}(t)\psi(t) = \sum_{\omega \in \Omega} S^{(i)}(t, \omega)\psi(t, \omega)$$

(4.4)

for any $i(i = 1, 2, ..., N)$.

Now let $P(t, \omega)$ denote the probability that the event $\omega$ happens at time $t$ and let $P(t, \omega) > 0$ for all $\omega$. We define $\pi(t, \omega)$ to be:

$$\pi(t, \omega) = \frac{\psi(t, \omega)}{P(t, \omega)}.$$  

(4.5)

The expression (4.4) can be written as:

$$S^{(i)}(0) = \sum_{\omega \in \Omega} S^{(i)}(t, \omega)\psi(t, \omega) = \sum_{\omega \in \Omega} S^{(i)}(t, \omega)\pi(t, \omega)P(t, \omega)$$

(4.6)

$$= E[S^{(i)}(t)\pi(t)],$$

(4.7)
where \( \pi(t) \) is a random variable which takes the values \( \pi(t, \omega) \) when the event \( \omega \) happens. Since \( \psi(t, \omega) \) is strictly positive, \( \pi(t, \omega) \) is also strictly positive. We call \( \pi(t) \) a state price deflator, or simply, a deflator. According to Duffie (2001), a state price deflator is also known as a state price density, a pricing kernel, or a marginal-rate-of-substitution process.

The formula (4.7) shows that the prices of financial products are calculated as expected values of the future payoffs weighted by the state price deflator. We call this method of pricing deflator pricing.

Now we will introduce risk-neutral pricing, which is the dominant methodology for pricing derivatives. We will also show that risk-neutral pricing is in fact a transformation of deflator pricing.

Note that the deflator pricing (4.7) requires that we know \( P(t, \omega) \) while (4.4) does not; formula (4.4) simply requires a knowledge of \( S(0) \) and \( D(t) \). In this sense, (4.4) is simpler than (4.7). By introducing a bank account, (4.4) can be a further simplified and can be transformed to a well-known form.

Suppose that security \( N \) is a bank account with \( S^{(N)}(0) = 1 \), and let \( B(t) \) denote the price of the bank account at time \( t \), i.e. \( S^{(N)}(t) = B(t) \). The rate incurred on the bank account is called a short rate. We assume that, in this single-period case, the short rate,
quoted on a continuous compounding basis, is constant\(^{28}\), and let \( r \) denote the short rate. Then \( B(t) = e^{rt} \).

Now, applying (4.4) to \( B(t) \), \( 1 = B(t)\sum_{\omega \in \Omega} \psi(t, \omega) \), then we know \( \sum_{\omega \in \Omega} \psi(t, \omega) = B(t)^{-1} \).

We define \( Q(t, \omega) \) to be:

\[
Q(t, \omega) = B(t)\psi(t, \omega).
\]

Then \( Q(t, \omega) \) defines a new probability over \( \Omega \) since the summation of \( Q(t, \omega) \) over \( \Omega \) is 1 and each \( Q(t, \omega) \) is positive. When the market is complete, \( Q(t, \omega) \) must exist and be unique since the state price vector must exist and be unique. We call this new probability, denoted by \( Q \), a risk-neutral probability measure.

Using \( Q(t, \omega) \), formula (4.4) can be written as:

\[
S^{(i)}(0) = \sum_{\omega \in \Omega} S^{(i)}(t, \omega)\left[B(t)^{-1}Q(t, \omega)\right]
= E^Q\left[B(t)^{-1}S^{(i)}(t)\right]
\]

where \( E^Q[ \ ] \) is the expectation operator over the probability measure \( Q \).

We call this method of pricing risk-neutral pricing. Note that risk-neutral pricing does not require that we know \( P(t, \omega) \). Instead, we need to obtain \( Q(t, \omega) \).

### 4.3.3 Choice of numeraire security

The formula (4.9) shows that \( B(t)^{-1} \) functions as a deflator on the new measure \( Q \). The price of the security whose reciprocal works as a deflator under some measure is

\(^{28}\) Even in the multi-period setting, the short rate over time \( t_i \) to \( t_{i+1} \) is constant once the information at \( t_i \) is given. Therefore a risk-neutral measure can be defined on the conditional basis in the multi-period setting.
called a numeraire, and we call the security a numeraire security. Using these terms, formula (4.9) can be re-stated as: the bank account is a numeraire security and \( B(t) \) is a numeraire under the measure \( Q \).

The state price deflator is the product of the deflator and the Radon-Nikodym derivative of the new measure \( Q \) with respect to the original measure \( P \), as the following simple transformation shows:

\[
S^{(j)}(0) = E^Q \left[ S^{(j)}(t) B(t)^{-1} \right]
\]

\[
= \sum_{\omega \in \Omega} \left[ S^{(j)}(t, \omega) B(t)^{-1} \right] Q(t, \omega)
\]

\[
= \sum_{\omega \in \Omega} S^{(j)}(t, \omega) \left[ B(t)^{-1} \frac{Q(t, \omega)}{P(t, \omega)} \right] P(t, \omega)
\]

\[
= E \left[ \sum_{\omega \in \Omega} S^{(j)}(t) \left[ B(t)^{-1} \xi(t) \right] \right]
\]

where \( \xi(t) \) is the Radon-Nikodym derivative of the measure \( Q \) with respect to the measure \( P \) that takes the value \( \xi(t, \omega) = \frac{Q(t, \omega)}{P(t, \omega)} \) for each \( \omega \).

So far, we have chosen the bank account as the numeraire security. However, as we will show below, we can choose any security as a numeraire security.

Now we choose the \( j \)-th (\( j=1,2,\ldots,N \)) security, and consider a new measure under which this security becomes a numeraire security. We let \( R(t) \) denote the accumulated return

29 For this terminology, see, for example, Duffie (2001) or Hull (1999).

30 This terminology is our original.

31 In general \( \tilde{\xi}(t) \) denotes the density process of the Radon-Nikodym derivative rather than the Radon-Nikodym derivative itself, i.e. \( \tilde{\xi}(t) = E \left[ dQ / dP \big| \mathcal{F}_t \right] \). See Duffie (2001, p.110). In a single period setting, however, these are the same.
of this security, i.e. \( R(t) = S^{(i)}(t) / S^{(i)}(0) \) and \( R(t, \omega) = S^{(i)}(t, \omega) / S^{(i)}(0) \). In addition, we define \( \xi_{/R}(t) \) to be:

\[
\xi_{/R}(t) = \pi(t)R(t),
\]

(4.10)

and,

\[
\xi_{/R}(t, \omega) = \pi(t, \omega)R(t, \omega).
\]

Then expression (4.7) can be rewritten as:

\[
S^{(i)}(0) = E\left[ S^{(i)}(t) \left( R(t)^{-1} \xi_{/R}(t) \right) \right].
\]

(4.11)

Since \( \sum_{\omega \in \Omega} \xi_{/R}(t, \omega)P(t, \omega) = E\left[ S^{(i)}(t)\pi(t) \right] / S^{(i)}(0) = 1 \) and \( \xi_{/R} \) is strictly positive, we can define a new probability measure \( Q_{/R} \) such that \( \xi_{/R}(t, \omega) = Q_{/R}(t, \omega) / P(t, \omega) \).

Here, \( Q_{/R} \) is the measure where \( \xi_{/R}(t) \) is the Radon-Nikodym derivative of \( Q_{/R} \) with respect to \( P \). Then formula (4.11) can be re-expressed using this notation:

\[
S^{(i)}(0) = E\left[ \left( S^{(i)}(t)R(t)^{-1} \right) \xi_{/R}(t) \right] = \sum_{\omega \in \Omega} \left[ S^{(i)}(t, \omega)R(t)^{-1} \right] \frac{Q_{/R}(t, \omega)P(t, \omega)}{P(t, \omega)} = E_{/R} \left[ S^{(i)}(t)R(t)^{-1} \right].
\]

(4.12)

This expression shows that the chosen security \( j \) functions as a numeraire security under the new measure \( Q_{/R} \).

In pricing derivatives, financial economists often start\(^{32}\) by selecting a numeraire security and then apply formula (4.12), as shown in the following example of the forward risk neutral measure. The form of the probability distributions of securities is chosen so that the final expectation calculation is simplified as long as the distributions

---

\(^{32}\) Mostly in the continuous setting that will be discussed in Section 4.5.
are reasonably realistic.

The zero-coupon bond that expires at time $T$ is often used as a numeraire security for pricing simple interest rate derivatives whose payoffs are settled at time $T$. The typical example is a caplet whose payment is made at time $T$. We let $Q^R$ denote this measure. Substituting $R(t) = P(t, T) / P(0, T)$ into (4.12) and letting $t = T$:

$$S^{(i)}(0) = E^{Q^R} \left[ \left( \frac{P(t, T)}{P(0, T)} \right)^{-1} S^{(i)}(t) \right] = P(0, T) E^{Q^R} \left[ S^{(i)}(T) \right],$$

where $P(s, T)$ is the price of a zero coupon bond at time $s$ that expires at $T$, and $E^{Q^R} \left[ \right]$ indicates that the expectation is taken under the measure $Q^R$.

Formula (4.13) shows $E^{Q^R} \left[ S^{(i)}(T) \right] = S^{(i)}(0) / P(0, T)$, i.e. the expected prices are equal to the forward prices under this measure. For this reason, this measure is called a forward risk neutral measure. As shown in (4.13), the advantage of this method is that no $P(s, T)$ terms appear inside the expectation operator in the final form since $P(t, t) = 1$.

The pricing under the forward risk neutral measure is, when interest rates are stochastic, simpler than the risk-neutral pricing (4.9) where the term $B(t)^{-1}$ stays inside the expectation operator.

4.3.4 Deflator generator and summary

Here, we define a new term, a deflator generator, and present a summary of the above discussion.

Under the measure $Q_{/R}$, $R(t)^{-1}$ is a deflator. By the definition of $\xi_{/R}(t)$, we know
that $\pi(t)$ is obtained by multiplying $\xi_{t}^R(t)$ by $R(t)^{-1}$. In other words, $\xi_{t}^R(t)$ generates the state price deflator from $R(t)^{-1}$. Therefore we call $\xi^{(e)}(t)$ a deflator generator associated with the numeraire $R(t)$.

Using this term, we now know that the state price deflator can be obtained by:

1. Choosing an appropriate numeraire security, and,
2. Find the deflator generator, and then,
3. Multiplying the reciprocal of the numeraire with the deflator generator.

Note that the deflator generator is also the Radon-Nikodym derivative of $Q_R$ with respect to $P$. In practical situations, the deflator generator is likely to be set so that the resulting calculation is mathematically tractable, as long as that setting reasonably well resembles the behaviour of the securities.

The discussion on the single period model has been, in fact, just various transformations of formula (4.4) using some new notation and concepts. Since all the transformations are reversible, we see that following statements are equivalent.

In the single period setting in a friction-free and complete market under the one-price principle:

- The market is arbitrage-free.
- There exists a unique state price deflator.
- There exists a unique risk-neutral measure.

In addition, the following is also equivalent to each of above.

- There exists a unique measure in which a selected security $i(i = 1, 2, \ldots, N)$

---

33 This naming is our original.
becomes a numeraire security.

- There exists a unique deflator generator for a selected security $i(i = 1, 2, ..., N)$.

### 4.4 Extension to multi-period model

The single period model can be naturally extended into the multi-period model where there are $L+1$ observation times: $t_0 = 0 < t_1 < t_2 < ... < t_L = T$. Duffie (2001) and Panjer et al. (1998) describe this extension with rigorous mathematical proofs. Here, we try to summarize the results from this literature rather intuitively. We need to redefine some concepts that have been introduced in the single period setting and define new concepts so that the results under the single period setting can be naturally adapted to the multi-period setting. We will see that all the statements made in Section 4.3.4 hold true under the multi-period setting.

In principle, the multi-period model can be regarded as a sequence of connected single period models. In the multi-period setting, prices of securities, deflators, deflator generator and trading strategies are defined to be processes indexed by time. However, on a given state at a given point of time $t_j (j = 0, 1, 2, ..., L-1)$, the one-step movement toward $t_{j+1}$ can be regarded as a single period. Therefore properties we have developed in the single-period setting can be applied on this conditional basis. Therefore we need to redefine some processes so that the conditional properties of these processes are consistent with what we have developed in the single-period setting. In this context, the state price deflator and deflator generators are redefined as density processes.
4.4.1 Information structure

The hardest task would be to implement an information structure where investors learn gradually about the true state of nature at intermediate points of time as time passes. This information structure is expressed as a sequence of subsets of the sample space that becomes finer as time passes. For example, consider a two-period model where a security price moves only one-tick up or down for the first period, and moves in the same way for the second period. Let \( u \) denote the upward movement and \( d \) denote the downward movement. Then all the sample paths of the movement of this security are denoted by \{\{uu\},\{ud\},\{du\},\{dd\}\}. Each path represents a specific event\(^3\). We let \( \omega_1, \omega_2, \omega_3, \omega_4 \) denote sample path \{\{uu\},\{ud\},\{du\},\{dd\}\} respectively. Then the sample space is \( \Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} \). At time zero, we have no information, and so we know that the feasible set of events is \( \{\omega_1, \omega_2, \omega_3, \omega_4\} \). At time 1, for example, we have observed that the price has gone up, then now we know that the feasible set of events in the future has been reduced to \( \{\omega_1, \omega_2\} \). Then at time 2, we finally know that we are in \( \omega_1 \), out of \( \{\omega_1, \omega_2\} \). Here, we have narrowed down sets of feasible events gradually.

Let \( \omega_{12} \) be the set \( \{\omega_1, \omega_2\} \), and let \( \omega_{34} \) be the set \( \{\omega_3, \omega_4\} \). At time 1, we only know that either of \( \omega_{12} \) or \( \omega_{34} \) takes place. In other words, what takes place at time 1 is a set of events such as \( \omega_{12} \) or \( \omega_{34} \) not an underlying event such as \( \omega_1 \). There is a useful mathematical tool to handle sets of events. That is a field whose element is a set of underlying events, not underlying events themselves. We need a sequence of fields indexed by time since the subsets that comprises a field become finer and finer as time passes.

\(^3\) Note that, intuitively, each event \( \omega \) represents a route in the multi-period and continuous-time setting whereas each event \( \omega \) rather represents a point in the single-period setting.
passes. This is because we acquire more and more information of the movement of security prices as time passes.

Formally, this information structure is formulated as a filtration. A filtration is a time-series sequence of a $\sigma$–field that satisfies $\mathcal{F}_t \subset \mathcal{F}_s (0 \leq t < s)$ where $\mathcal{F}_t$ denotes the $\sigma$–field at time $t$. On the above example, the information is modelled as:

$$
\mathcal{F}_0 = \sigma\{\{\omega_1, \omega_2, \omega_3, \omega_4\}\}^{36},
$$

$$
\mathcal{F}_1 = \sigma\{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}, \text{ and,}
$$

$$
\mathcal{F}_2 = \sigma\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\} = \sigma\{\Omega\}.
$$

We assume that the relevant stochastic processes are adapted$^{37}$ to the above information structure, which means, intuitively, at time $t$, values of the processes at time $t$ can be deduced from the information available at time $t$.

The underlying probabilities are assigned to each element on $\Omega$. When $\mathcal{F}_j (j = 0, 1, 2, \ldots, L - 1)$ is given, we are already in a smaller space than $\Omega$; therefore probabilities need to be redefined on the condition that we are already in this small subspace. If in the single period between $t_j$ and $t_{j+1}$ the market is arbitrage-free when

$^{35}$ A $\sigma$–field $\mathcal{F}$ on $\Omega$ is a collection of subsets of $\Omega$ that satisfies: $\phi \in \mathcal{F}$, $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, and $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A_1 \cup A_2 \cup \ldots \in \mathcal{F}$.

$^{36}$ $\sigma\{A, B, C, \ldots\}$ is a $\sigma$–field generated by $\{A, B, C, \ldots\}$, which is the intersection of all $\sigma$–fields that contains $\{A, B, C, \ldots\}$.

$^{37}$ See, for example, Panjer et al. (1998) for details of the term “adapt”.

66
is given, there should exist a random variable $\tilde{\pi}(t_{j+1})$ that satisfies:

$$S(t_j) = E \left[ S(t_{j+1}) \tilde{\pi}(t_{j+1}) \mid \mathcal{F}_{t_j} \right].$$

Using this recursively for $t_{j-1}, \ldots, 0$, we have:

$$S(0) = E \left[ S(t_{j+1}) \tilde{\pi}(t_1) \tilde{\pi}(t_2) \cdots \tilde{\pi}(t_{j+1}) \right].$$

This suggests that $S(t_j)\pi(t_j) (j = 0, 1, 2, \ldots, L)$ where $\pi(t_j) = \tilde{\pi}(t_1)\tilde{\pi}(t_2)\cdots\tilde{\pi}(t_j)$ is a martingale\(^{38}\). In the next section, we define the state price deflator taking this property into account.

### 4.4.2 Deflators and deflator generators

In the multi-period setting, we define the state price deflator $\pi(t)$ to be the process that makes the process:

$$S^{(i)}(t)\pi(t) \quad (t = t_j, j = 0, 1, 2, 3, \ldots, L)$$

a martingale for $i = 1, 2, 3, \ldots, N$, and $\pi(0) = 1$.

If this state price deflator exists, there also exists a state price deflator in each single period component that satisfies (4.7) on the conditional basis.

Since $S^{(i)}(t)\pi(t)$ is a martingale, for any $t$ and $s$ ($t = t_k, s = t_m, 0 \leq k < m \leq L$):

$$S^{(i)}(t) = E \left[ S^{(i)}(s) \frac{\pi(s)}{\pi(t)} \mid \mathcal{F}_t \right]$$

holds for any security $i$.

This formula shows that $\pi(s)/\pi(t)$ is the state price deflator in the period between $t$

---

\(^{38}\) For a martingale, the conditional expectation at time $t$ of any future values of the martingale is the value of the martingale at $t$. 

67
and \( s \) when \( \mathcal{F}_t \) is given.

Now we choose the \( j \)-th security as a numeraire security. We defined the deflator generator using formula (4.10) when \( t \) is fixed. In the multi-period setting we interpret \( t \) as a variable, and we redefine the deflator generator using (4.10) again, i.e.,

\[
\xi_{j,R}(t) = \pi(t) R(t),
\]

where \( \pi(t) \) is the process defined above, and \( R(t) = S^{(j)}(t) / S^{(j)}(0) \) \( (t = t_j, j = 0,1,2,3,\ldots,L) \). Now \( \pi(t) = R(t)^{-1} \xi_{j,R}(t) \) and \( S^{(j)}(t) \pi(t) \) is a martingale, so that:

\[
S^{(j)}(t) R(t)^{-1} \xi_{j,R}(t) = E \left[ S^{(j)}(s) R(s)^{-1} \xi_{j,R}(s) \mid \mathcal{F}_t \right]
\]

for any \( t \) and \( s \) \( (t = t_k, s = t_m, 0 \leq k < m \leq L) \).

Substituting \( R(t) \) by \( S^{(j)}(t) / S^{(j)}(0) \), we have:

\[
S^{(j)}(t) = E \left[ S^{(j)}(s) \left( \frac{S^{(j)}(s)}{S^{(j)}(t)} \right)^{-1} \left( \frac{\xi_{j,R}(s)}{\xi_{j,R}(t)} \right) \mid \mathcal{F}_t \right]. \tag{4.14}
\]

Formula (4.14) suggests that \( \xi_{j,R}(s) / \xi_{j,R}(t) \) is the deflator generator for the period between \( t \) and \( s \) when \( \mathcal{F}_t \) is given. In fact:

\[
E \left[ \frac{\xi^{(R)}(s)}{\xi^{(R)}(t)} \mid \mathcal{F}_t \right] = E \left[ \frac{\pi(s) R(s)}{\pi(t) R(t)} \mid \mathcal{F}_t \right] = R(t)^{-1} E \left[ \frac{\pi(s)}{\pi(t)} \right] R(s) \mid \mathcal{F}_t = 1.
\]

The final transformation holds since \( \pi(s) / \pi(t) \) is the state price deflator in the period between \( t \) and \( s \) when \( \mathcal{F}_t \) is given.

\( \xi_{j,R}(t) \) \( (0 \leq t \leq T) \) is a strictly positive martingale with \( \xi_{j,R}(0) = 1 \). If \( \xi_{j,R}(T) \) has a
finite variance\(^{39}\), we can define a new measure \( Q_R \) that is equivalent to \( P \). Under the measure \( Q_R \), \( \xi_R(t) \) is the density process for the Radon-Nikodym derivative with respect to the measure \( P \). Therefore, we can rewrite\(^{40}\) (4.14) as:

\[
S^{(j)}(t) = E^{Q_x}\left[ \frac{S^{(j)}(s)}{S^{(j)}(t)} \right]^{-1} | \mathcal{F}_t .
\]

### 4.4.3 Trading gain and arbitrage

Now we have defined the state price deflator, a deflator under a new measure, and a deflator generating process. Before we move onto arbitrage in a multi-period setting, we introduce the concept of a trading gain.

The change of market value of a self-financing portfolio is caused solely by the change of prices of component securities. Using (4.1),

\[
V^0(t_k) - V^0(t_{k-1}) = \theta(t_k)S(t_k) - \theta(t_{k-1})S(t_{k-1})
\]

\[
= \theta(t_{k-1})[S(t_k) - S(t_{k-1})]
\]

\[
= \theta(t_{k-1})\Delta S(t_k)
\]

(4.15)

where \( \Delta S(t_k) = S(t_k) - S(t_{k-1}) \).

We call the value \( \theta(t_{k-1})\Delta S(t_k) \) a trading gain between times \( t_{k-1} \) and \( t_k \). Expression

\[^{39}\text{Duffie (2001) emphasizes this condition, in the continuous setting, in Theorem E in Chapter 6. Duffie (2001) states the product of two random variables of finite variance is of finite expectation. Therefore this condition guarantees the existence of the expectation value of a process with a finite variance under the new measure.}\]

\[^{40}\text{When } \xi(t)(0 \leq t \leq T) \text{ is a density process of the measure } Q, \ dQ / dP = \xi(T) \text{ and } E^Q[z(s)|\mathcal{F}_t] = E^P[z(s)\xi(s)|\mathcal{F}_t] / \xi(t) \text{ (0 \leq t < s \leq T) for any } \mathcal{F}_t \text{-measurable random variable } z(s).\]
(4.15) can be written as:

\[ V^\theta(t_k) = V^\theta(t_{k-1}) + \theta(t_{k-1}) \Delta S(t_k) . \]

Using this recursively, we get:

\[ V^\theta(t_k) = V^\theta(0) + \sum_{i=1}^{k} \theta(t_{i-1}) \Delta S(t_i) . \]  

This formula indicates that the market value of a self-financing portfolio is obtained by adding the total trading gain to the initial market value of the portfolio. This property is useful in the continuous-time setting illustrated in Section 4.5.

Recall that arbitrage refers to a self-financing strategy that may result in riskless profits. Since there is no leakage or injection of funds for a self-financing portfolio, it is obvious that a self-financing portfolio cannot produce any riskless profits at any time if there is no arbitrage chance in any of the single-period components of the extended period.

We have redefined some processes so that the properties in the single-period model are preserved on a conditional basis in the multi-period setting. In principle, as far as arbitrage is concerned, what is true in a single period setting is also true in the multi-period setting. In other words, the results stated in Section 4.3.4 still hold in a multi-period setting. See, for example, Panjer et al. (1998) for formal proofs for general cases.

### 4.5 Extension to continuous-time setting

Here we extend the results from the discrete-time setting to a continuous-time setting. The most important result is that a deflator generator is determined solely by market-price-of-risk processes. We refer to Duffie (2001) Chapter 6, on which most of
this material is based, for technical details and formal proofs.

4.5.1 Model specifications

The continuous-time model could be regarded as the limiting case of the discrete-time model. In fact, the processes for security prices, trading strategies, the state price deflator and deflator generators can be defined in the same way as in the multi-period setting, except that time $t$ takes any value between 0 and $T$ in the continuous-time setting. In the continuous-time setting, integration replaces summation over a finite number of elements, and the infinitesimal interval $dt$ replaces $\Delta t = t_j - t_{j-1}$. Therefore, the results in the discrete-time case are expected to hold even in the continuous-time case. We should, however, note that expected values or variances may not converge to finite values in the continuous-time setting. This is the major problem we face when we extend discrete-time and finite-state models to continuous-time models. In order to avoid the divergence of moments, we will constrain ourselves to deal with well-behaved processes. Under this constraint, the results stated in 4.3.4 almost\footnote{Some technical conditions need to be applied} hold even in the continuous-time setting.

Suppose there are $N$ securities in the market. Each security price is modelled using the $d$-dimensional standard Wiener process. For $t \in [0,T]$, let $W(t) = (W^{(1)}(t), \ldots, W^{(d)}(t))^T$ denote the column vector of $d$ standard independent Wiener processes and let $\mathcal{F}_t$ denote a $\sigma$-field that is an element of the standard filtration of $W(s)(0 \leq s \leq T)$. The price of security $j$ ($j = 1, 2, \ldots, N$), denoted by $S^{(j)}(t)$, is assumed to follow the Ito process that satisfies:
\[ dS^{(j)}(t) = \mu^{(j)}(t)dt + \sigma^{(j\cdot)}(t)dW(t), \]

where \( \mu^{(j)}(t) \) is an adapted scalar process in \( \mathcal{X}^{i} \), and \( \sigma^{(j\cdot)}(t) \) is a \( d \)-dimensional row vector of which each element is an adapted process in \( \mathcal{X}^{2} \).

\( \mathcal{X}^{i} (i=1,2) \) is defined to be the space of adapted processes, represented by \( z \), that satisfy:

\[
\mathcal{X}^{i} \equiv \left\{ z : \int_{0}^{T} |z(t)|^2 dt < \infty \text{ almost surely for any } T > 0 \right\}.
\]

Duffie (2001) imposed additional restrictions that each security price \( S^{(j)}(t) \) is in the space \( H^{2} \). This means:

\[
E\left[ \left( \int_{0}^{T} \mu^{(j)}(t)dt \right)^2 \right] < \infty, \text{ and}
\]

\[
E\left[ \int_{0}^{T} \sigma^{(j\cdot)}(t) \cdot \sigma^{(j\cdot)}(t)dt \right] < \infty, \text{ for } T > 0.
\]

Using the column vector \( \mu(t) \) whose \( j \)-th element is \( \mu^{(j)}(t) \) and the \( N \times d \) matrix \( \sigma(t) \) whose \( j \)-th row is \( \sigma^{(j\cdot)}(t) \), the prices of the \( N \) securities are collectively expressed as a column vector as follows:

\[ dS(t) = \mu(t)dt + \sigma(t)dW(t). \]

Let \( \pi(t) \) denote, assuming it exists, the state price deflator for the \( N \) securities. We define a new process denoted by \( S^{(j)}(t) \) that is the price process multiplied by the state price deflator: i.e. \( S^{(j)}(t) = S^{(j)}(t)\pi(t) \). Similarly, let \( S^{(j)}(t) \) denote the price process deflated\(^{42} \) by the numeraire security \( i \), i.e. \( S^{(j)}_{i}(t) = S^{(j)}(t) / S^{(i)}(t) \).

\(^{42}\) The verb “deflate” simply means “divide”. Being deflated does not always mean the deflated price process is a martingale.
We also define the vector of these as: $S_{i/}(t) = S(t)/S^{(i)}(t)$, and $S_{x}(t) = S(t)\pi(t)$.

Given the trading strategy $\theta(t)$ and $S(t)$ $(0 \leq t \leq T)$, the gain process is defined by $\int_{0}^{t} \theta(s)dS(s), (0 \leq t \leq T)$, which is the continuous version of the total trading gain that appeared in (4.16). In order for this gain process to be well-defined, we assume that $\theta(t)$ is in $\mathcal{K}(S)$. An adapted process $\theta(t)$ is said to be in $\mathcal{K}(S)$ when $\theta(t)\mu(t)$ is in $\mathcal{K}$, and each element of the row vector $\theta(t)\sigma(t)$ is in $\mathcal{K}^2$.

Duffie (2001, p.95) states\(^{43}\) that the gain process has a finite variance when $\theta(t)$ is an Ito process in $\mathcal{H}^2(S)$. $\theta(t)$ is said to be in $\mathcal{H}^2(S)$, if, for $T > 0$,

$\theta(t)$ is in $\mathcal{K}(S)$, and,

$$E\left[\left(\int_{0}^{T} \theta(t)\mu(t)dt\right)^2\right] < \infty,$$

and,

$$E\left[\int_{0}^{T} \left(\theta(t)\sigma^{(j)}(t)\right)^2 dt\right] < \infty \text{ (for all } j = 1, 2, ..., d),$$

where $\sigma^{(j)}(t)$ is the $j$-th column vector of the matrix $\sigma(t)$.

Note that $\mathcal{H}^2(S)$ is different from $\mathcal{H}^2(S_x)$ and $\mathcal{H}^2(S_{i/})$ in general.

If $\theta(t)$ is self-financing, integrating formula (4.2), we have:

$$\theta(t)S(t) = \theta(0)S(0) + \int_{0}^{t} \theta(s)dS(s).$$

(4.17)

This shows that the market value of the self-financing portfolio is obtained by adding the total trading gain to the initial market value of the portfolio, as we have shown in the discrete setting. Note that the expected value of the stochastic integral that appears in

\(^{43}\) Relating to $\mathcal{H}^2(S)$, Duffie (2001) also defines another space $\Theta(S)$. When $\theta(t) \in \Theta(S)$, there is some constant $k$ with $\theta(t)S(t) \geq k$ for all $t \geq 0$ almost surely.
(4.17) does not always exist. The expected value, however, exists\textsuperscript{44} if $\theta(t)$ is in $\mathcal{H}^2(S)$.

If $\theta(t)$ is self-financing for $S(t)$, $\theta(t)$ is also self-financing for $S_\pi(t)$ and $S_{\dot{\iota}}(t)$, i.e. self-financing status is preserved\textsuperscript{45} after any deflating operation. Therefore, for a self-financing strategy $\theta(t)$:

$$\theta(t)S_\pi(t) = \theta(0)S_\pi(0) + \int_0^t \theta(s)dS_\pi(s),$$  \hspace{1cm} (4.18)

also holds.

The expected value of the stochastic integral that appears in (4.18) exists if $\theta(t)$ is in $\mathcal{H}^2(S_{\pi})$.

4.5.2 Some sufficient conditions for no arbitrage

Consider a strategy $\theta(t)$ that is self-financing for $S(t)$. Since $\theta(t)$ is also self-financing for $S_\pi(t)$ and $S_{\dot{\iota}}(t)$, and both $\pi(t)$ and $S^{\dot{\iota}}(t)$ are strictly positive, we know\textsuperscript{46} that $\theta(t)$ is an arbitrage for $S(t)$ if and only if it is an arbitrage for $S_\pi(t)$ and $S_{\dot{\iota}}(t)$.

Now we can show that if the state price deflator exists, there is no arbitrage in $\mathcal{H}^2(S_{\pi})$.

Suppose a state price deflator $\pi(t)$ exists. If a self-financing strategy $\theta(t)$ is in $\mathcal{H}^2(S_{\pi})$:

\textsuperscript{44} See Duffie (2001) Proposition 5B.

\textsuperscript{45} This property is called “Numeraire Invariance Theorem”. See Duffie (2001).

\textsuperscript{46} This is because the sign is the issue in an arbitrage, and multiplying or dividing by a strictly positive value does not change the sign.
\[
E \left[ \int_0^T \theta(s) dS_s(s) \right] = 0,
\]
holds since \( S_\pi(t) \) is a martingale and the stochastic integral is well-defined.

Now taking expectation on both sides of (4.18), we have:
\[
E \left[ \theta(T)S_\pi(T) \right] = \theta(0)S_\pi(0). \tag{4.19}
\]
Looking at (4.19), we know that \( \theta(0)S_\pi(0) \) is positive when \( \theta(T)S_\pi(T) \) is positive.

Therefore there is no arbitrage in \( H^2(S_\pi) \) for \( S_\pi(t) \) and therefore for \( S(t) \) also.

Similarly, if \( S_i(t) \) admits\(^{47}\) the equivalent martingale measure under which \( S_i(t) \) becomes a martingale, for a self-financing strategy \( \theta(t) \) in \( H^2(S_i) \):
\[
E^{Q_i} \left[ \theta(T)S_i(T) \right] = \theta(0)S_i(0),
\]
holds. Then there is no arbitrage in \( H^2(S_i) \).

In summary, the existence of the state price deflator or the equivalent martingale measure is a sufficient condition for no arbitrage. In the next section, we will consider the necessary condition for no arbitrage. In due course, we will find the concrete shape of deflator generators in this continuous-time case.

### 4.5.3 The market-price-of-risk

Suppose the deflated process \( S_i(t), (0 \leq t \leq T) \) has the form:
\[
dS_i(t) = \mu_i(t)dt + \sigma_i(t)dW(t). \tag{4.20}
\]

\(^{47}\) The existence of the equivalent martingale measure implicitly assumes that the Radon-Nikodym derivative has a finite variance. Therefore the expectation exists under the new measure. See Duffie (2001) Theorem E in Chapter 6.
The market-price-of-risk process, \( \eta_i(t) \), for the process \( S_i(t) \) is an \( \mathbb{R}^d \) -valued column vector process that satisfies\(^{49}\):

\[
\mu_i(t, \omega) = \sigma_i(t, \omega) \eta_i(t, \omega), \quad \text{for } t \in [0, T] \text{ and } \omega \in \Omega,
\]

where \( \bullet(t, \omega), (\bullet \text{ is one of } \mu, \sigma, \eta) \) is the value of \( \bullet(t) \) when the event \( \omega \) happens at time \( t \).

If there is no arbitrage in \( \mathcal{H}^2(S_i) \) \( (i = 1, 2, \ldots, N) \), there is a market-price-of-risk process for \( S_i(t) \), i.e. the existence of the market-price-of-risk is necessary for no arbitrage. We can discern\(^{50}\) this by the following simplified argument.

Suppose the market-price-of-risk process does not exist. Then, using the fundamental linear algebra theorem, there should exist\(^{51}\) a trading strategy \( \theta(t, \omega) \) that satisfies:

\[
\theta(t, \omega) \sigma_i(t, \omega) = 0 \quad \text{and} \quad \theta(t, \omega) \mu_i(t, \omega) \neq 0 \quad \text{for some } (t, \omega).
\]

By applying \( -\theta(t, \omega) \) if \( \theta(t, \omega) \mu_i(t, \omega) < 0 \), we can see that the resulting market value of the portfolio generated by this strategy has zero volatility and positive drift at

---

\(^{48}\) The subscript suggests that it depends on the choice of the numeraire security.

\(^{49}\) The solution is not always unique unless the rank of \( \sigma \) is \( d \).

\(^{50}\) See, for example, Duffie (2001) for a rigorous proof.

\(^{51}\) If this does not hold, \( \theta(t, \omega) \mu_i(t, \omega) = 0 \) holds for all \( \theta(t, \omega) \) belonging to \( \mathcal{N}\left(\sigma_i(t, \omega)^T\right) \), which is the nullspace of \( \mathcal{H}\left(\sigma_i(t, \omega)\right) \), which is the space spanned by the column vectors of

\( \sigma_i(t, \omega) \). That means \( \mu_i(t, \omega) \) is orthogonal to \( \mathcal{N}\left(\sigma_i(t, \omega)^T\right) \), therefore \( \mu_i(t, \omega) \) is in \( \mathcal{H}\left(\sigma_i(t, \omega)\right) \), i.e. \( \mu_i(t, \omega) = \sigma_i(t, \omega) \eta_i(t, \omega) \) should hold for some \( \eta_i(t, \omega) \), resulting in the contradiction.
these points. This means that we can make a riskless profit at these points.

Conversely, we show that when the market-price-of-risk process is given, there is no arbitrage. As we have shown in the multi-period setting, if the market admits an equivalent martingale measure, there is no arbitrage. We will show that if the given market-price-of-risk process satisfies some technical requirement, the process generates a density process for the Radon-Nikodym derivative of a new measure and \( S_{h}(t) \) becomes a martingale under the new measure.

Suppose the market-price-of-risk process \( \eta_{h}(t) \), for the process \( S_{h}(t) \) is given. Then the expression (4.20) can be written as:

\[
\begin{align*}
    dS_{h}(t) &= \mu_{h}(t)dt + \sigma_{h}(t)dW(t) = \sigma_{h}(t)\eta_{h}(t)dt + \sigma_{h}(t)dW(t) \\
    &= \sigma_{h}(t)[\eta_{h}(t)dt + dW(t)].
\end{align*}
\]

The Girsanov Theorem states that \( W^{Q_{h}}(t) \) which satisfies:

\[
W^{Q_{h}}(0) = 0,
\]

and,

\[
dW^{Q_{h}}(t) = \eta_{h}(t)dt + dW(t),
\]

becomes a standard Wiener process under a new measure \( Q_{h} \) if \( \eta_{h}(t) \) satisfies Novikov’s condition:

\[
E\left[ \exp\left( \frac{1}{2} \int_{0}^{T} \eta_{h}(s) \cdot \eta_{h}(s)ds \right) \right] < \infty.
\]

The Girsanov Theorem also tells us the explicit form of the density process for the Radon-Nikodym derivative of the measure \( Q_{h} \) with respect to \( P \), denoted by \( \xi_{h}(t) \)

---

\[52\] This also depends on the choice of the numeraire security.
That is the stochastic exponential of $\eta_i(t)$:

$$\xi_i(t) = \exp\left( -\int_0^t \eta_i(s)^T d\mathbf{W}(s) - \frac{1}{2} \int_0^t \eta_i(s) \cdot \eta_i(s) ds \right),$$

or, equivalently,

$$d\xi_i(t) = -\xi_i(t) \eta_i(t)^T d\mathbf{W}(t) \text{ with } \xi_i(0) = 1.$$  

Duffie (2001) shows that if $\eta_i(t)$ is $L^2$-reducible, $S_i(t)$ admits an equivalent martingale measure $Q_i$ with $dS_i(t) = \sigma(t) d\mathbf{W}^{Q_i}(t)$. We say that $\eta_i(t)$ is $L^2$-reducible when $\xi_i(t)$ has a finite variance and $\eta_i(t)$ satisfies Novikov’s condition. As we have shown in Section 4.5.2, when $S_i(t)$ admits an equivalent martingale measure, there is no arbitrage in $\mathcal{H}^2(S_i)$.

Note that the existence of an $L^2$-reducible market-price-of-risk process $\eta_i(t)$ for $S_i(t)$ is a sufficient condition for no arbitrage in $\mathcal{H}^2(S_i)$, and hence is not exactly equivalent to no arbitrage in $\mathcal{H}^2(S_i)$. If there is no arbitrage in $\mathcal{H}^2(S_i)$, a market-price-of-risk process should exist, but that market-price-of-risk process does not have to be, strictly speaking, $L^2$-reducible.

### 4.5.4 Forms of deflator generator

In this section, we show the form of deflator generators and also the form of deflators for $S(t)$ when an $L^2$-reducible market-price-of-risk process is given.

The process $\xi_i(t)$, which is defined by (4.22) makes $S_i(t)$ a martingale under the
measure $Q_i$. The process $\xi_i(t)$ also makes $S_i(t)\xi_i(t)$ a martingale under the original measure $P$ since $E^{Q_i}[S_i(t)] = E[S_i(t)\xi_i(t)]$ by the definition of the Radon-Nikodym derivative.

Recall that $S_i(t) = S(t)/S^{(i)}(t)$. Since $S(t)[S^{(i)}(t)^{-1}\xi_i(t)]$, which is $S_i(t)\xi_i(t)$, is a martingale under the measure $P$, the state price deflator $\pi(t)$ is:

$$\pi(t) = [S^{(i)}(t)^{-1}\xi_i(t)].$$

(4.23)

Here, $\xi_i(t)$ is the deflator generator for the numeraire security $i$. Substituting (4.22) into (4.23), we have:

$$\pi(t) = [S^{(i)}(t)^{-1}\exp\left(-\int_0^t \eta_i(s)\eta_i^\top(s)ds - \frac{1}{2}\int_0^t \eta_i(s)\eta_i^\top(s)ds\right)].$$

(4.24)

Now we have the form of the state price deflator when an $L^2$-reducible market-price-of-risk process is given. The market-price-of-risk process $\xi_i(t)$ that appears in formula (4.24) depends on the choice of the numeraire security $i$.

Conventionally, however, when the term “market-price-of-risk” is used without reference to any numeraire security; the bank account is implicitly assumed to be the numeraire security in defining the market-price-of-risk process. Therefore, now we will transform formula (4.24) given that we know the market-price-of-risk process when the bank account is used as a numeraire security.

Suppose the $N$-th security among the $N$ securities is the bank account, i.e. $S^{(N)}(t) = \exp\left(\int_0^t r(s)ds\right)$. We assume that there exists an $L^2$-reducible

---

53 For example, Hull (2000, p508) uses this term in this context.
market-price-of-risk process, \( \eta_N(t) \), for \( S_N(t) \). Recall that \( S_N(t) \) denotes the security price vector deflated by \( S^{(N)}(t) \). From now on, we use the notation \( \eta(t) \) for \( \eta_N(t) \) for simplicity. Substituting this notation into formula (4.24), we have:

\[
\pi(t) = \exp \left( \int_0^t \left[ -r(s) - \frac{1}{2} \eta(s) \cdot \eta(s) \right] ds - \int_0^t \eta(s) \cdot dW(s) \right).
\]

Similarly, using this notation, the price of the particular security \( j \) deflated by the bank account security is described as:

\[
dS^{(j)}_N(t) = \sigma^{(j)}_N(t) [ \eta(t) dt + dW(t) ], \tag{4.25}
\]

using formula (4.21).

Suppose we want security \( i \) to be the numeraire security. Formula (4.24) gives the form of the deflator once we know the market-price-of-risk process associated with the numeraire security \( i \), i.e. \( \eta_i(t) \). Since we already know \( \eta(t) \), here we consider expressing \( \eta_i(t) \) using \( \eta(t) \). For this purpose, at first we will transform formula (4.25) that shows the absolute change of price so that the new form shows the relative change of price.

Dividing\(^5\) both sides of formula (4.25) by \( S^{(i)}_N(t) \), then omitting the superscript for simplicity, we rewrite the above expressions as:

\[
\frac{dS_N(t)}{S_N(t)} = \sigma_S(t) [ \eta(t) dt + dW(t) ],
\]

where \( \sigma_S(t) \) is defined by \( \frac{\sigma^{(i)}_N(t)}{S^{(i)}_N(t)} \).

Applying Ito’s lemma to \( S(t) = S_N(t)S^{(N)}(t) \), we have:

\(^5\) This is allowed since the prices are strictly positive.
\[
\frac{dS(t)}{S(t)} = \left[ r(t) + \sigma_S(t)\eta(t) \right] dt + \sigma_S(t)dW(t).
\] (4.26)

Note that the volatility of \( S_{N}(t) \) is the same as the volatility of \( S(t) \).

Now, we choose security \( i \) as the numeraire security, and we let \( R(t) \) denote \( S^{(i)}(t)/S^{(i)}(0) \). Following the expression shown in (4.26), we assume that \( R(t) \) follows:

\[
\frac{dR(t)}{R(t)} = \left[ r(t) + \sigma_R(t)\eta(t) \right] dt + \sigma_R(t)dW(t).
\]

Using Ito’s lemma, the process \( S_R(t) \), which is defined to be \( S(t)/R(t) \), follows:

\[
\frac{dS_R(t)}{S_R(t)} = \left( \sigma_S(t) - \sigma_R(t) \right) \left[ \eta(t) - \sigma_R(t)^T \right] dt + dW(t).
\]

The above result shows, by definition of the market-price-of-risk, that \( \eta(t) - \sigma_R(t)^T \) is the market-price-of-risk when the numeraire security is security \( i \), i.e.:

\[
\eta_i(t) = \eta(t) - \sigma_R(t)^T.
\] (4.27)

Applying this result to formula (4.22), the form of the deflator generator is obtained as:

\[
\xi_{/R}(t) = \exp \left\{ -\int_{0}^{t} \left[ \eta(s)^T - \sigma_R(s) \right] dW(s) - \frac{1}{2} \int_{0}^{t} \left\| \eta(s)^T - \sigma_R(s) \right\|^2 ds \right\},
\] (4.28)

where \( \|a\| = \sqrt{a \cdot a} \) for a vector \( a \).

Formula (4.27) is equivalent to \( \eta(t) = \eta_i(t) + \sigma_R(t)^T \), and substituting this into (4.26), we have

\[
\frac{dS(t)}{S(t)} = \left[ r(t) + \sigma_S(t)\sigma_R(t)^T \right] dt + \sigma_S(t)\left[ \eta_i(t)dt + dW(t) \right].
\]

Consider the new measure \( Q_{/R} \) such that \( dW^{Q_{/R}}(t) = \eta_i(t)dt + dW(t) \). Under this measure \( \xi_{/R}(t) \) in (4.28) is the density process for the Radon-Nikodym derivative \( dQ_{/R}/dP \). Then, we have:

81
\[
\frac{dS(t)}{S(t)} = \left[ r(t) + \sigma_s(t)\sigma_R(t)^T \right] dt + dW^{Q_R}(t),
\]  
(4.29)

and \( S_t(t) \) is a martingale under this new measure \( Q_R \).

Formula (4.29) is obtained by setting \( \eta(t) \) to \( \sigma_R(t)^T \) and \( dW(t) \) to \( dW^{Q_R}(t) \) in formula (4.26). This result shows that if we want to obtain a form of Ito process of a security under the measure where security \( i \) is the numeraire security, we simply set the market-price-of-risk to the volatility of the relative price change of the numeraire security \( i \), i.e. we just set \( \eta(t) \) to \( \sigma_R(t)^T \). This property also tells us that if \( \eta(t) \) happens to be \( \sigma_R(t)^T \) in the real world, then security \( i \) is a numeraire security in the real world.

We have assumed \( W(t) = \left( W^{(1)}(t), W^{(2)}(t), \ldots, W^{(d)}(t) \right)^T \) where the dimension of \( W(t) \) is \( d \). Suppose the numeraire security \( i \) is affected only by first \( k \) \((1 \leq k \leq d) \) elements of \( W(t) \). Then the \( d \)-dimensional vector \( \sigma_R(t) \) has the form: \( (\sigma^{(1)}, \sigma^{(2)}, \ldots, \sigma^{(k)}, 0, \ldots, 0) \).

In this case, \( \xi_R(t) \) in (4.28) can be separated into two parts:

\[
\xi_R(t) = \exp \left\{ -\int_0^t \left[ \eta(s)^T(l-k+1)dW(s)^{(l-k)} - \frac{1}{2} \int_0^t \| \eta(s)^T(l-k+1) \|^2 ds \right] 
\cdot \exp \left\{ -\int_0^t \left[ \eta(s)^T(l-k+1)dW(s)^{(l-k)} - \frac{1}{2} \int_0^t \| \eta(s)^T(l-k+1) \|^2 ds \right] \right\},
\]

where we let \( a^{(l-m)} \) \((l \leq m) \) denote the vector with the dimension of \( l-m+1 \) whose components are the \( l, l+1, \ldots, m \)-th elements of the vector \( a \).

If \( \eta(s)^{(l-k)} \) is independent of \( \eta(s)^{(l-k+1)} \), the deflator is the product of the two independent parts:

\[
\pi(t) = R(t)^{-1} \xi_R(t)
\]
\[
= \left[ R(t)^{-1} \exp \left( -\int_0^t \left[ \eta(s)^T \left[ 1_{[1-k]} - \sigma_R(s)^{[1-k]} \right] dW(s)^{[1-k]} - \frac{1}{2} \int_0^t \| \eta(s)^T \left[ 1_{[1-k]} - \sigma_R(s)^{[1-k]} \|^2 \right] ds \right) \right] \\
\cdot \exp \left( -\int_0^t \eta(s)^T \left[ 1_{[(k+1)-d]} \right] dW(s)^{[(k+1)-d]} - \frac{1}{2} \int_0^t \| \eta(s)^T \left[ 1_{[(k+1)-d]} \right] \|^2 ds \right). \tag{4.30}
\]

In particular, if we choose the measure \( Q/R \) such that:

\[
dW_{Q,R}^{[1-k]}(t)^{[1-k]} = \left[ \eta(s)^T \left[ 1_{[1-k]} - \sigma_R(s)^{[1-k]} \right] dW(t)^{[1-k]} \right], \text{ and,}
\]

\[
dW_{Q,R}^{[(k+1)-d]}(t)^{[(k+1)-d]} = dW(t)^{[(k+1)-d]},
\]

or, equivalently, if we assume that \( \eta(s)^T \left[ 1_{[1-k]} = \sigma_R(s)^{[1-k]} \right] \) in the real world, (4.30) is simplified as:

\[
\pi(t) = R(t)^{-1} \exp \left( -\int_0^t \eta(s)^T \left[ 1_{[(k+1)-d]} \right] dW(s)^{[(k+1)-d]} - \frac{1}{2} \int_0^t \| \eta(s)^T \left[ 1_{[(k+1)-d]} \|^2 ds \right). \tag{4.31}
\]

### 4.6 An example: pricing a plain call option

In this section we will derive a pricing formula for a plain call option in deflator pricing and risk-neutral pricing as an application of the general pricing principles we have shown in this chapter. During the derivation process, we will discuss the equivalence between these two pricing methods. The resulting pricing formula is the well-known Black-Scholes formula.

Jarvis et al. (2001) suggested that a pricing formula for plain options, which is in fact the Black-Scholes formula, can be derived using the real-world measure without converting the measure to a risk-neutral measure although they did not show the derivation process in detail. Here, we actually derive a pricing formula for a simple call option both under the real-world measure and under the risk-neutral measure, and then we show both calculations are equivalent.
We assume that the price of a with-dividend equity follows the following process as in the Black-Scholes setting:

\[
\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t), \tag{4.32}
\]

where the volatility \(\sigma\) and the drift \(\mu\) are constant and \(W(t)\) is the one-dimensional standard Wiener process.

In addition, we assume that the short rate, which is denoted by \(r\) is constant.

We choose the bank account as the numeraire security. When we assume no arbitrage, the market-price-of-risk \(\eta\) exists. By the assumptions, \(\eta\) is constant and \(\mu = r + \sigma \eta\).

Substituting this into (4.32), we get:

\[
S(t) = S(0) \cdot \exp \left( (r + \sigma \eta - \frac{1}{2} \sigma^2) t + \sigma W(t) \right). \tag{4.33}
\]

Clearly \(\eta\) is \(L^2\)-reducible. Therefore the deflator generator \(\xi(t)\) exists and

\[
\xi(t) = \exp \left( -\frac{1}{2} \eta^2 t - \eta W(t) \right) \text{ using (4.22).}
\]

The state price deflator is:

\[
\pi(t) = B(t)^{-1} \xi(t) = \exp \left( -rt - \frac{1}{2} \eta^2 t - \eta W(t) \right).
\]

In fact \(S(t)\pi(t)\) is a martingale since a simple calculation leads to:

\[
S(t)\pi(t) = S(0) \exp \left( -\frac{1}{2} (\sigma - \eta)^2 t + (\sigma - \eta) W(t) \right),
\]

which is a martingale.

Now, let \(C(T)\) denote the price of a European call option at time zero with strike price \(K\) that expires at time \(T\).

First, we calculate \(C(T)\) using deflator pricing.

\[
C(T) = E \left[ \pi(T) \max(S(T) - K, 0) \right] \tag{4.34}
\]

84
By setting \( W(T) = \sqrt{T}z \) where \( z \) follows the standard normal distribution,

\[
C(T) = E \left[ \exp \left( -rT - \frac{1}{2} \eta^2 T - \eta \sqrt{T} z \right) \cdot \max \left( S(0) \cdot \exp \left( (r + \sigma \eta - \frac{1}{2} \sigma^2)T + \sigma \sqrt{T} z \right) - K, 0 \right) \right]
\]

\[
= \int_{-\infty}^{\infty} \exp \left( -rT - \frac{1}{2} \eta^2 T - \eta \sqrt{T} z \right) \cdot \max \left( S(0) \cdot \exp \left( (r + \sigma \eta - \frac{1}{2} \sigma^2)T + \sigma \sqrt{T} z \right) - K, 0 \right) \cdot f_z(z) dz,
\]

where \( f_z(z) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} z^2 \right) \).

By straightforward integration, we get:

\[
C(T) = S(0) \Phi[d_1] - Ke^{-\eta} \Phi[d_2],
\]

where \( d_2 = \frac{\log(\frac{S(0)}{K}) + rT - \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}} \), \( d_1 = d_2 + \sigma \sqrt{T} \), and \( \Phi[\ ] \) is the cumulative standard normal distribution function.

This is the Black-Scholes formula. Note that \( \eta \) has disappeared during the interim calculation.

Next, we derive the same formula using risk-neutral pricing, and show that risk-neutral pricing is in fact a transformation of deflator pricing.

Going back to (4.34), we have:

\[
C(T) = E \left[ B(T)^{-1} \xi_z(T) \max(S(T) - K, 0) \right]
\]

\[
= \int_{-\infty}^{\infty} B(T)^{-1} \xi_z(z) \max(S(T) - K, 0) f_z(z) dz,
\]

where \( \xi_z(z) = \exp \left( -\frac{1}{2} \eta^2 T - \eta \sqrt{T} z \right) \).
Now by a simple calculation, we know:

\[ \xi(z)f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(z + \eta \sqrt{T})^2\right). \]

Therefore, formula (4.35) becomes:

\[ C(T) = \int_{-\infty}^{\infty} B(T)^{-1} \max(S(T) - K, 0) f_\xi(\tilde{z}) d\tilde{z}, \]

(4.36)

where \( \tilde{z} = z + \eta \sqrt{T} \), i.e. \( z = \tilde{z} - \eta \sqrt{T} \).

This \( \tilde{z} \) defines a new measure \( Q \). Under this measure, (4.33) becomes:

\[
S(T) = S(0) \cdot \exp \left( (r + \sigma \eta - \frac{1}{2} \sigma^2) T + \sigma \sqrt{T} \tilde{z} \right) \\
= S(0) \cdot \exp \left( (r + \sigma \eta - \frac{1}{2} \sigma^2) T + \sigma \sqrt{T} \left( \tilde{z} - \eta \sqrt{T} \right) \right) \\
= S(0) \cdot \exp \left( (r - \frac{1}{2} \sigma^2) T + \sigma \sqrt{T} \tilde{z} \right).
\]

Now \( z \) was completely replaced by \( \tilde{z} \) in (4.36) as:

\[
C(T) = \int_{-\infty}^{\infty} \exp(-rT) \cdot \max \left( S(0) \cdot \exp \left( (r - \frac{1}{2} \sigma^2) T + \sigma \sqrt{T} \tilde{z} \right) - K, 0 \right) \cdot f_\xi(\tilde{z}) d\tilde{z} \\
= E^Q \left[ B(T)^{-1} \max(S(T) - K, 0) \right].
\]

The measure \( Q \) indicates that the integration is calculated with respect to \( \tilde{z} \). Computing this leads us to the Black-Scholes formula again.

As seen, the risk-neutral pricing and deflator pricing are equivalent; effectively the same integrations are calculated in both methods. Note that the calculation under risk-neutral pricing is much simpler than that under deflator pricing. The market-price-of-risk disappeared before calculating the expected value in risk-neutral pricing, while it disappeared during the calculation of the expected value in deflator pricing. In risk-neutral pricing, however, the probability distribution represented by \( \tilde{z} \) is not the
true distribution in the real world. The true distribution in the real world is represented by \( z \), and hence we need to use \( z \) when we need to handle, for some reason, the real-world distribution.

### 4.7 Practical issues in simulation

In this section, we will quantify errors in calculating the expected value of a deflator generator in a simulation environment.

Under the assumption of constant volatility \( \sigma \) and constant market-price-of-risk \( \eta \);

\[
\xi(t) = \exp\left(-\frac{1}{2}\eta^2 t - \eta W(t)\right)
\]

is a martingale for \( t > 0 \), where \( W(t) \) is the standard one-dimensional Wiener process.

Now, for a fixed \( t \), we try to obtain \( E[\xi(t)] \) by \( N \) simulations with \( N \) random samples

by the formula:

\[
\overline{\xi}(t) = \frac{\sum_{n=1}^{N} \exp\left(-\frac{1}{2}\eta^2 t - \eta Z(n)\sqrt{t}\right)}{N},
\]

where \( Z(n) \) is the \( n \)-th independent standard normal random sample.

Then, \( E[\overline{\xi}(t)] = 1 \) and the standard deviation of \( \overline{\xi}(t) \), denoted by \( \sigma_{\overline{\xi}(t)} \), is:

\[
\sigma_{\overline{\xi}(t)} = \sqrt{\frac{e^{\eta^2 t} - 1}{N}}.
\]

Since \( \overline{\xi}(t) \) is asymptotically normal, a 95% confidence interval is approximately:

\[
1 \pm 1.96 \sigma_{\overline{\xi}(t)} = 1 \pm 1.96 \sqrt{\frac{e^{\eta^2 t} - 1}{N}}.
\]
The following table illustrates $1.96\sigma_{\bar{z}(t)}$ when $t=50$. We have chosen $t=50$ since we need to deal with long-term cash flows in evaluating long-term financial products such as annuity contracts. We have not put much importance on the cash flows that occur in years over 50 since the present value of these cash flows is, in practice, very small.

Table 4-1 - 95% confidence range of the simulated average

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>1,000</th>
<th>2,000</th>
<th>3,000</th>
<th>4,000</th>
<th>5,000</th>
<th>10,000</th>
<th>100,000</th>
<th>1,000,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>5.0%</td>
<td>3.5%</td>
<td>2.9%</td>
<td>2.5%</td>
<td>2.2%</td>
<td>1.6%</td>
<td>0.5%</td>
<td>0.2%</td>
</tr>
<tr>
<td>15%</td>
<td>8.9%</td>
<td>6.3%</td>
<td>5.2%</td>
<td>4.5%</td>
<td>4.0%</td>
<td>2.8%</td>
<td>0.9%</td>
<td>0.3%</td>
</tr>
<tr>
<td>20%</td>
<td>15.7%</td>
<td>11.1%</td>
<td>9.0%</td>
<td>7.8%</td>
<td>7.0%</td>
<td>5.0%</td>
<td>1.6%</td>
<td>0.5%</td>
</tr>
<tr>
<td>25%</td>
<td>28.9%</td>
<td>20.4%</td>
<td>16.7%</td>
<td>14.5%</td>
<td>12.9%</td>
<td>9.1%</td>
<td>2.9%</td>
<td>0.9%</td>
</tr>
<tr>
<td>30%</td>
<td>58.5%</td>
<td>41.4%</td>
<td>33.8%</td>
<td>29.2%</td>
<td>26.2%</td>
<td>18.5%</td>
<td>5.8%</td>
<td>1.8%</td>
</tr>
<tr>
<td>40%</td>
<td>33.3%</td>
<td>239.2%</td>
<td>195.3%</td>
<td>169.2%</td>
<td>151.3%</td>
<td>107.0%</td>
<td>33.8%</td>
<td>10.7%</td>
</tr>
<tr>
<td>50%</td>
<td>321.7%</td>
<td>2270.3%</td>
<td>1853.7%</td>
<td>1605.3%</td>
<td>1435.9%</td>
<td>1015.3%</td>
<td>321.1%</td>
<td>101.5%</td>
</tr>
<tr>
<td>70%</td>
<td>1295279.4%</td>
<td>915900.9%</td>
<td>747829.9%</td>
<td>647639.7%</td>
<td>579266.6%</td>
<td>409603.3%</td>
<td>129527.9%</td>
<td>40960.3%</td>
</tr>
</tbody>
</table>

Table 4-1 suggests that the simulation approach usually allows for at least several per cent pricing errors with a 95% confidence level. The errors increase dramatically as the market-price-of-risk increases for the same number of simulations. Several per cent errors would be acceptable, depending on the purpose of the valuation. We remark that several per cent error over 50 years is translated into a 10 basis point, i.e. 0.1%, error per year, which is not significant as an error per year.

There is always a limit on the number of simulations, namely $N$, due to the restriction of CPU power or time constraints. When $N=5,000$, The upper limit of $\eta$, practically, would be 20%. When N is 100,000 or beyond, $\eta$ that is slightly more than 30% could be handled. However, $\eta$ beyond 40% is not, in practice, tractable even with 1,000,000
4.8 Summary
In this section, we have illustrated that, under the arbitrage-free assumption, expected present values that are consistent with the market are calculated using a deflator. In the continuous-time setting, if the market-price-of-risk process is given for a market, the market is arbitrage-free and the deflator generator is obtained as the stochastic exponential of the market-price-of-risk process. However, high market-price-of-risks tend to cause large calculation errors in long-term simulation.

Deflators are one of the most essential concepts in modern financial economics, and some important techniques stems from this concept. Using the deflator generator as a Radon-Nikodym derivative, a different measure can be defined and we can change a measure. Under a measure where a specific security is a numeraire security, the market-price-of-risk is the volatility of the price of the numeraire security. In risk-neutral pricing, the market-price-of-risk is set to zero.
5 Interest rate term structure models

In this chapter, we overview some established interest rate term structure models. Modelling the interest rate term structure is critical for valuation and risk management of insurance contracts under the fair valuation regime. This is primarily because the current market yield curve strongly influences the current value of an insurance contract, and any change in the yield curve afterwards keeps affecting the value of the insurance contract until the contract expires. The spot rates are derived from the yield curve observed in the market; and under the fair valuation regime, the spot rates are used to discount cash flows that arise from an insurance contract with or without some spread over the spot rates. When the market yield curve changes, all the spot rates for all terms are likely to change. For example, suppose current interest rates are 5% for all terms; and this 5%-flat yield curve has just changed to a slightly sloped yield curve where the short rate is 4.75% and 10-year interest rate is 5.25%. Then all spot rates for all terms are affected, and the present value of the cash flows also changes.

Another reason why modelling interest rate term structure is important is that interest rates are important variables that represent macro economic conditions; and these affect most economic activities. It is well known that long-term equity returns are also highly influenced by interest rates. In other words, interest rates are a key driver of a macro economy, and their term structure contains essential information to model the future macro economics.

Both practitioners and academics have not put great importance on interest rate term structure models in insurance studies, primarily because only a long-term average
interest rate, not current fluctuating interest rates, is an issue under cost accounting. Therefore, a comprehensive understanding of an interest rate term structure and models is still under way, even though some models are applied to some specific products such as options on an investment product. Most interest rate term structure models are built on common conceptual frameworks that have been developed historically. We clarify these conceptual frameworks by reviewing some established interest rate term structure models. In particular, we explain the positive interest approach in detail that we will use for our simulation.

5.1 Classification of interest rate term structure models
There are plenty of term structure models used by academics and practitioners. These models are in general not consistent because these models were developed for different purposes, under different assumptions and different technical platforms. This inconsistency often bewilders us. For example, one model might assume that interest rates follow a lognormal distribution while another model assumes that they follow a normal distribution; and either model may be regarded as correct and be well used in practice. Therefore, getting a bird’s eye view of the universe of interest rate term structure models and understanding the general concepts behind these models is important when we attempt to choose a model for our specific purpose. Here, we start our review of these models by classifying them into three categories: economic hypotheses, pricing models, and descriptive models.

The economic hypotheses try to explain why the term structure exists from the macro economic point of view, i.e. why the short-term and long-term interest rates are different.
Pricing models are mainly used for evaluating financial products, in particular, derivatives. The objective of descriptive models is a reasonably accurate description of the yield curve, using a spanning set of elementary functions.

Wilkie’s model (1986, 1995) is widely used in actuarial circles. This model is a return model, rather than a term structure model, that deals with stochastic returns on ordinary shares with inflation being the driving factor. We do not use this model because this model is not designed to produce a current interest rate term structure that is consistent with the market or to produce stochastic projections of interest rate term structures for the future.

5.1.1 Economic hypotheses
There are three well-known economic hypotheses. They are called expectations, liquidity preference, and market segmentation theory. According to Cairns (2004), the expectations hypothesis states that the current yield curve reflects market expectations without incorporating risk premiums; liquidity preference theory insists that the longer-dated bonds include some risk premiums; and the market segmentation hypothesis states that bond markets are segmented by investors who have specific concerns for specific expiry dates. These three hypotheses are rather conceptual frameworks to which specific mathematical models are tied.

5.1.2 Pricing models
Following Chaplin (1997), pricing models are further divided into two classes; equilibrium models and evolutionary models.

The equilibrium models explain the levels of yield over time in relation to some, usually
one or two, exogenous factors. The exogenous factors are usually well-known economic variables such as money-market rates. An evolutionary model does not allow arbitrage between the model output prices and the market prices. The parameters of the evolutionary models are set to fit the given initial term structure.

Pricing models can also be classified into three categories purely from the technical point of view. A pricing model is based either on a classical approach, on a Heath, Jarrow and Morton (1992) ("HJM", hereafter) approach, or on a positive interest approach.

The models under the classical approach deal with the development of an instantaneous spot rate, while models under the HJM approach deal with forward rates. The models under the positive interest approach model the price process of zero coupon bonds in a way that any forward rate can never be negative.

Most models under the classical approach are equilibrium models, but there are some evolutionary models under this approach. All models under the HJM and positive interest approach are evolutionary models.

5.1.3 A descriptive model

The typical descriptive model is Cairns’ (1998) model. The formula of this model is:

$$f(t, t+s) = b_0 + \sum_{i=1}^{4} b_i e^{-c_i s}$$

(5.1)

where $f(t, t+s)$ is the continuously compounding instantaneous forward rate observed at $t$ that expires at the time $t+s$ for $s>0$, and \{\(b_i\)\}_{i=0}^{4}$ and \{\(c_i\)\}_{i=1}^{4} are constants.
5.2 Consistency between pricing models and economic hypotheses

In this section, we raise one interesting question: whether established pricing models are consistent with the expectation hypothesis or liquidity preference theory. We do not consider the market segmentation theory since we know that nowadays many market players compare yields of some bonds to other bonds that are on different expiry dates to find whether there are over-priced or under-priced bonds in the market. If the over-priced bonds are sold and under-priced bonds are bought, then the yield curve tends to reach a smoothed shape. For example, an investor whose investment horizon term is long, such as a pension fund, is likely to buy long-dated bonds and hold them until expiration. In other words, long-dated bonds are supported by this kind of investor. However, it is unlikely that the long-dated bonds are priced in the market reflecting solely such specific interests by these investors. The specific demand due to the specific interests would be smoothed out by the transactions made by the market players who look at the entire yield curve.

All of the established pricing models are based on the financial economics approach that we have discussed in Chapter 4 under the assumption that the market is complete and arbitrage-free. Cox, Ingersoll and Ross (1981) showed that under such a mathematical framework, pricing models can be consistent with liquidity preference theory, but can only be consistent with the local expectation theory, which is one of three versions of expectation theory. Pricing models cannot be consistent with the other two versions of expectation theory, which are return-to-maturity expectation theory and unbiased expectation theory.
Model developers who focus only on pricing have no interest in the market-price-of-risk in the real world and make no assumption on the market-price-of-risk in the real world simply because it is unnecessary; i.e. they do not reject or accept economic hypotheses. In other words, if we use pricing models solely for the purpose of calculating expected present value, we do not need to know the market-price-of-risk process in the real world. This is because, as we have shown in Chapter 4, whatever market-price-of-risk, or risk premium, process may be assumed, the resulting expected present value is the same. The deflator is and should be modified for the choice of the market-price-of-risk process so that the prices of tradable securities fit to the market. Therefore, each pricing model uses an arbitrary measure, i.e. assumes an arbitrary market-price-of-risk process, which is most convenient for the purpose of implementing that model.

An established pricing model uses some measure only for mathematical tractability, without considering what the risk premium in the real world is. However, this does not mean that we cannot make any assumption on the risk premium in the real world. In fact, any assumption on the risk premium in the real world can be made as long as some technical conditions are satisfied without causing any contradiction with the model.

Pricing models can be consistent with liquidity preference theory if we assume the market-price-of-risk process in the real world is such that the resulting risk premiums for longer-dated bonds are higher than for shorter-dated bonds. The simplest example is when all bond prices for any expiry dates are driven by only one Wiener process and yield volatility is the same for all these bonds. In this case, since longer-maturity bonds have higher volatilities than shorter-maturity bonds, a positive market-price-of-risk
process results in positive risk premiums that increase as the time-to-expiry increases.

If we assume that the market-price-of-risk in the real world is zero, the market is consistent with the local (or risk-neutral) expectation theory where all bonds generate the same instantaneous expected rate of return.

However, in general, a market-price-of-risk process that is consistent with the return-to-maturity expectation theory does not exist. This hypothesis states that the total return generated from the long-term bonds will be equal to that of a series of short-term bonds continually rolled over. Cox, Ingersoll and Ross (1981) showed that this hypothesis can hold for a specific long-term bond by setting the market-price-of-risk to the volatility of the long-term bond. However, this result also states that the hypothesis fails to hold for all the long-term bonds simultaneously since volatilities for different bonds are in general different. For example, if we set the market-price-of-risk to the volatility of the 10-year bond, the return-to-maturity expectation theory holds for a 10-year term. However, the theory does not hold, for example, for a 15-year term. For the same reason, a market-price-of-risk process that is consistent with the unbiased (or pure) expectation theory does not exist. This hypothesis states that current forward rates are unbiased predictors of future spot rates.

5.3 Philosophies behind pricing models
In this section we discuss the philosophies behind pricing models rather than mathematical properties of them.
5.3.1 Equilibrium models

The conspicuous feature of equilibrium models is that parameters for these models are found through historical analysis in most cases, and the resulting interest rate term structure is unlikely to be consistent with the current term structure observed in the market. As Fitton and McNatt (2002) state, one philosophy underlying the equilibrium models is that the interest rate term structure must be estimated, rather than taken as given, employing a statistical approach allowing for some statistical errors. Since the exogenous factors are usually observable, the parameter estimation is usually easy and intuitive. Sometimes even the views of the user are reflected in the parameters. Using these parameters, we can measure how the snapshot term structure in the marketplace deviates from the term structure that would appear in equilibrium. Therefore, the equilibrium models are particularly useful in making investment decisions.

The equilibrium models assume that the market will eventually converge to its equilibrium form; thus any current deviation from equilibrium is temporary and will disappear. This philosophy indirectly tells us that the term structure produced by a model is more correct than the term structure currently observed in the market. Although such a relative view on the market is essential in making investment decisions, the equilibrium models are not directly applicable for fair valuation or derivative pricing where the initial term structure is required to be matched to the market. The equilibrium models, however, are worth studying since they incorporate the most important features of the behaviour of interest rates.

5.3.2 Evolutionary models

Evolutionary models are the natural development of equilibrium models. With the rapid
development of interest rate derivatives, attempts have been made to fit the model term structure to the market term structure by adding more parameters to some equilibrium models. A typical example is the Hull and White (1990) model. In this model, the rate to which the money-market rates converge is time-dependent. The time-dependent convergence rates are chosen so that the resulting model term structure fits to the term structure that is observed at the pricing date. Since evolutionary models are constructed with reference to the term structure that is observed at only a single point in time, these models tend to fail to describe the actual long-term behaviour of the term structure, as Fitton and McNatt (2002) argue. This deficiency has not been an issue since evolutionary models had primarily been developed for pricing derivatives, without any active attempt to explain the yield levels by some economic variables. In fact, Hull and White (1990) admit that the same parameter set does not describe the term structure at two different times. Similarly, Fitton and McNatt (2002) argue that the snapshot term structure is not always in equilibrium and contains some pricing errors; therefore we should rely on historically estimated parameters not those implied in the current snapshot term structure for projection purposes.

5.3.3 Choosing between equilibrium and evolutionary models

The equilibrium models are suitable for long-term projection and the evolutionary models can fit the current market term structure. At present there is no model that completely satisfies both requirements. Hull and White (1990) emphasize that it is important to distinguish between the goal of developing a model for describing term structure movement and the goal of developing a model for practical derivative pricing. Fitton and McNatt (2002) concluded that the evolutionary models are appropriate only
for calculating expected present values where market prices are reliable, while the
equilibrium models are most appropriate for projection work such as stress testing or
reserve adequacy testing.

However, it is not desirable to use different models for valuation and projection
purposes. Using different models is likely to produce inconsistent or even contradictory
results. Therefore, recent efforts to develop an evolutionary model that is also suitable
for long-term projection have been made, for example, by Cairns (2004b). In our study,
we conduct the valuation, risk calculation and business appraisal based on the modelling
framework proposed by Cairns (2004b). We will explain this model in detail in Section
5.7.

5.4 Models under the classical approach
We introduce established models under the classical approach in this section. This class
of models involves modelling instantaneous spot rates, or short rates, in the one-factor
case. Two-factor models use another random variable in addition to short rates, for
example, the yield on consols. The primary advantage of this approach is that close
proxies for the underlying variables exist in financial markets and hence data is readily
available in public for statistical analysis. Therefore, a great deal of research on these
models has been conducted for decades, in particular for the past two decades by both
academics and practitioners. Now an enormous amount of material on these models is
available. Table 5-1 lists some commonly-used models under this category.
Table 5.1 - Commonly-used classical approach models

<table>
<thead>
<tr>
<th>Number of Factors</th>
<th>Equilibrium Models</th>
<th>Evolutionary Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-factor</td>
<td>• Vasicek (1977)</td>
<td>• Hull and White (1990)</td>
</tr>
<tr>
<td></td>
<td>• Cox-Ingersoll-Ross (1985)</td>
<td>(Extended Vasicek)</td>
</tr>
<tr>
<td></td>
<td>• Brennan-Schwartz (1979)</td>
<td>• Black-Karasinski (1991)</td>
</tr>
<tr>
<td></td>
<td>• Black-Karasinski (1991)</td>
<td>(general case)</td>
</tr>
<tr>
<td>Two-factor</td>
<td>• Brennan-Schwartz (1979)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Longstaff-Schwartz (1992)</td>
<td></td>
</tr>
</tbody>
</table>

5.4.1 One-factor models

Now we will introduce one-factor models in detail because these models are very important; the conceptual framework behind these models is still at the heart of almost all interest rate term structure models.

The four one-factor equilibrium models have the same form:

\[ dY(t) = \alpha [\mu - Y(t)] dt + Y(t)^\beta \sigma dW(t), \quad (5.2) \]

where \( Y(t) \) is \( r(t) \) or \( \ln(r(t)) \), \( \alpha \) and \( \sigma \) are some positive constants, and \( \beta \) is one of 0, 0.5, or 1. The four models are described by assigning the values in Table 5-2 to \( Y(t) \) and \( \beta \) in formula (5.2).

Table 5.2 - One-factor equilibrium models

<table>
<thead>
<tr>
<th>Model</th>
<th>( Y(t) )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vasicek (1977) (“Vasicek”, hereafter)</td>
<td>( r(t) )</td>
<td>0</td>
</tr>
<tr>
<td>Cox-Ingersoll-Ross (1985) (“CIR”, hereafter)</td>
<td>( r(t) )</td>
<td>0.5</td>
</tr>
<tr>
<td>Brennan-Schwartz (1979) (“BS”, hereafter)</td>
<td>( r(t) )</td>
<td>1</td>
</tr>
<tr>
<td>Black-Karasinski (1991) (“BK”, hereafter)</td>
<td>( \ln(r(t)) )</td>
<td>0</td>
</tr>
</tbody>
</table>

Note that the BK model can be rewritten using Ito’s lemma as:
\[dr(t) = \alpha \left( \mu + \frac{\sigma^2}{2\alpha} r(t) - r(t) \ln(r(t)) \right) dt + r(t) \sigma dW(t).\]

This expression tells us that it is expected that \( r(t) \) under the BK model behaves similarly to that under the BS model since the BK and BS models have the same volatility of \( r(t) \), which is \( r(t)\sigma \).

All of the above four models have the mean-reverting feature where \( r(t) \) or \( \ln(r(t)) \) is pulled back to the convergence rate \( \mu \) in the long run. In particular, under the Vasicek model, \( r(t) \) follows the Ornstein-Uhlenbeck process. It is well known (see, for example, Rolksi et al. (1998, p562)) that, for given \( r(0) \):

\[r(t) = \mu + (r(0) - \mu)e^{-at} + \sigma \int_0^t e^{-a(t-s)}dW(s),\]

and hence \( r(t) \) given \( r(0) \) is normally distributed with the mean \( \mu + (r(0) - \mu)e^{-at} \) and the variance \( \sigma^2 \frac{1-e^{-2at}}{2\alpha} \).

Figure 5-1 shows sample paths of the four models on the same random numbers. The smooth lines show the mean and plus and minus one standard deviation away from the mean based on the Vasicek model.
We set $r(0)$ to 1%, the convergence rate to 4%, and $\alpha$ to 0.1386 ($= \ln(2)/5$) so that the current short rate converges to the convergence rate by half in 5 years. Other used parameters are shown in Table 5-3.

**Table 5-3 - Parameters used for Figure 5-1**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vasicek</td>
<td>4%</td>
<td>1%</td>
</tr>
<tr>
<td>CIR</td>
<td>4%</td>
<td>5% ($= \frac{1%}{4^{0.5}}$)</td>
</tr>
<tr>
<td>BS</td>
<td>4%</td>
<td>20% ($= \frac{1%}{4%}$)</td>
</tr>
<tr>
<td>BK</td>
<td>ln(4%)</td>
<td>20% ($= \frac{1%}{4%}$)</td>
</tr>
</tbody>
</table>

The parameter $\sigma$ was chosen so that the volatility is 1% when $r(t)$ is 4%.
Figure 5-1 illustrates the fundamental characteristics of these models. When short rates are close to the convergence rate, 4% in this case, all four models move almost in parallel. When short rates are higher than the convergence rate, the BK model produces the most volatile movements of interest rates, but the four models still exhibit similar movement patterns. A clear distinction appears when the short rates are low, in particular when the short rates are close to zero. In this environment, the Vasicek model produces much more volatile short rates than other models do, and only the Vasicek model sometimes produces negative short rates.

In spite of the deficiency of producing negative short rates, the Vasicek model is still one of the most important interest rate models. The advantage of the Vasicek model lies in its simplicity; there are analytical formulae for the prices of zero coupon bonds and for the price of European options on these bonds. This mathematical tractability stems from the fact that the total interest from time 0 to time t is easily calculated as:

\[
\int_{0}^{t} r(s)ds = \mu t + (r(0) - \mu)k(\alpha,t) + \int_{0}^{t} \sigma k(\alpha,t-s)dW(s),
\]

where \( k(\alpha,t) = \frac{1-e^{-\alpha t}}{\alpha} \),

and this total interest follows a normal distribution.

In the BS model, the resulting instantaneous spot rates are always positive, but there is no analytical formula for bond or option prices under the BS model. This drawback was overcome by the CIR model. Under this model, instantaneous spot rates follow the non-central Chi-squared distribution, and there are analytical formulae for bonds and options. See, for example, Cairns (2004). In the recent global low interest rate
environment, the BK model has been gaining more credit. Cairns (2004) quotes Chan et al. (1992) as stating that the volatility form $r(t)\sigma$ fitted historical data significantly better than $r(t)^{0.5} \sigma$ or $\sigma$. In addition, the BK model is mathematically tractable since \( \ln(r(t)) \) follows the Ornstein-Uhlenbeck process and hence $r(t)$ is lognormally distributed.

5.4.2 Two-factor models

Cairns (2004) gives a concise description of two-factor models and a discussion related to them. The Longstaff and Schwartz (1992) model is in principle a two-factor version of the CIR model. Under this model, the instantaneous spot rate and instantaneous variance of it are defined to be linear combination of two independent state variables, each of which follows its own CIR process. Using this model, we can focus on the level of volatilities rather than the shapes of yield curves. Under the Brennan and Schwartz (1992) consol model, the instantaneous spot rate and the consol yield follow interlinked Ito processes. We need to be careful in the parameterization of this model since the yield on consols could become infinite in finite time under this model.

5.4.3 Evolutionary models

Evolutionary models under classical models are natural extensions of equilibrium models. The Hull and White (1990) model is an extension of the Vasicek model, and it is formulated as:

$$dr(t) = \alpha [\mu(t) - r(t)] dt + \sigma dW(t),$$

where $\mu(t)$ is a deterministic function of time.

The general BK model has a similar form to the Hull and White (1990) model, but for
The form of this model is:

\[
d \ln(r(t)) = \phi(t) \left[ \ln(\mu(t)) - \ln(r(t)) \right] dt + \sigma(t) dW(t),
\]

where \( \phi(t), \mu(t), \) and \( \sigma(t) \) are deterministic functions of time.

The functions \( \phi(t), \mu(t), \) and \( \sigma(t) \) are chosen so that the resulting initial interest rate term structure produced by the model fits the term structure that is observed in the market. Black and Karasinski (1991) proposed that the deterministic functions be calibrated using the yield curve, the yield volatility curve of zero coupon bonds, and the cap curve.

### 5.5 Heath, Jarrow and Morton approach

In this section, we explain the fundamental aspects of the HJM approach. We introduce this approach because this approach is one of the dominant methodologies for modelling interest rate term structure, and this approach is different from the classical approach. Even though we do not use this approach directly, understanding the basics of this approach helps us to grasp the concepts that underlie modern term structure modelling under the financial economics approach.

Forward rates and zero coupon bond prices are almost equivalent concept since they are interrelated by the following simple formula:

\[
f(t, T_1, T_2) = \frac{\ln[P(t, T_1)] - \ln[P(t, T_2)]}{T_2 - T_1}, \quad (T_1 < T_2),
\]

where \( f(t, T_1, T_2) \) is the continuous compounding forward rate at time \( t \) for the period from \( T_1 \) to \( T_2 \).

Letting \( T_2 \to T_1 \), we have:
\[ f(t, T) = -\frac{\partial}{\partial T} \ln[P(t, T)] = \frac{1}{P(t, T)} \left( -\frac{\partial P(t, T)}{\partial T} \right) \quad (t < T), \]

where \( f(t, T) \) is the instantaneous forward rate at time \( t \) for the term \( T \).

Bonds are actively traded and the prices of these are available in public, and an enormous amount of derivative products based on forward rates, such as caps, swaptions, and forward rates agreements, are traded in the market. Forward rates can be obtained directly from the market without any intermediate mathematical modelling or estimations as opposed to the future short rates. Market practitioners are likely to prefer models that deal with forward rates directly. From that background, Heath, Jarrow and Morton (1992) developed a modelling framework that uses the forward rates as the fundamental building blocks.

Models under the HJM approach are evolutionary, i.e. the initial model curve is fitted to the initial market observed yield curve. The HJM approach, however, is fundamentally different from the evolutionary models under the classical approach such as the HW model. Unlike the classical approach where yield curves, including the initial yield curve, are the output, the HJM approach starts from the given initial yield curve. Therefore, there is no problem in choosing parameters that conform to the initial yield curve. This is the notable feature and the first advantage of the HJM approach. The second advantage is its flexibility in specification of volatilities and in the number of factors, namely the dimension of the driving Wiener processes. The volatility parameters can be set flexibly and, more importantly, separately for each expiry date; and the HJM approach can be extended to as many factors as required without increasing mathematical difficulty. Under the classical approach, such a flexible setting
or extension to multi-factors leads to extreme mathematical intractability.

However, there are several drawbacks in the HJM approach. Firstly, in general no analytical solution exists for prices of zero coupon bonds, or of plain options on these bonds. Secondly, the HJM approach is not, in general, non-Markov. Therefore all the development history of yield curves over time, rather than the history of only the state variables, needs to be recorded in a simulation. Thirdly, short rates may diverge or become negative in finite time, depending on the volatility function. See, for example, Chaplin (1997).

The HJM approach deals with a series of an uncountable number of instantaneous forward rates. On the contrary, the Brace, Gatarek and Museila (1997) (“BGM”, hereafter) model deals with a series of a finite number of effective forward rates that are observed in the market. The primary objective of the BGM model is to price derivatives consistently with the Cap market. Therefore, little attention is paid on the long-term behaviour of the interest rate term structure. Nevertheless, this model is popular amongst practitioners because this model is understandable, parameters are easily obtained, and it is automatically consistent with Cap prices.

We briefly introduce Caps here. Caps are liquid derivative instruments and thus we have a market quotation, i.e. volatilities, for short maturity, for example, one year, to long maturity, for example 20 years. In the Cap market volatilities are quoted instead of prices, and market participants calculate prices using the standardized market model. A Cap is a series of caplets whose reset dates are $T_0, T_1, \ldots, T_n$ and payment dates are
For a caplet whose reset date is \( T_k \) \((k = 0, 1, 2, \ldots, n)\), the difference between a predetermined cap rate and the London Interbank Offer Rate (“LIBOR”, hereafter) at time \( T_k \) is paid at time \( T_{k+1} \) only when the difference is positive. The market model assumes that the future LIBOR at time \( T_k \) follows a lognormal distribution under the forward risk-neutral measure under which the numeraire security is the zero coupon bond that expires at \( T_{k+1} \). It should be noted that the measure used for valuing the caplet whose reset date is \( T_k \) is different from the measure used for the caplet whose reset date is \( T_l \) for \( k \neq l \).

Following the market model, the BGM model also assumes that forward rates are lognormally distributed. The BGM model, however, uses a measure called the rolling forward risk-neutral measure, which is different from the measure that is used in the market model. Under the rolling forward risk-neutral measure, the world is always forward risk-neutral with respect to a bond maturing at the next caplet reset date. Given caplet volatilities, the drifts and volatilities of all forward rates under the rolling risk-neutral measure are obtained solely by the caplet volatilities. That means the development of all the forward rates can be described solely by caplet volatilities. See, for example, Hull (1999) for the formulae and mathematical derivation.

### 5.6 Positive interest

In this study, we use a multi-factor model (“CMF model”, hereafter) proposed by Cairns (2004b). The CMF model is based on the Flesaker and Hughston (1996) (“FH”, hereafter) approach, which is a special case of the general positive interest approach developed by Rogers (1997) and Rutkowski (1997) (“RR”, hereafter). The RR approach
is also called a positive interest approach since it ensures that interest rates remain positive. Ensuring positive interest rates is particularly important when interest rates on Japanese Yen whose short-term rates are close to zero are studied. We will use the positive interest approach in our simulation later; thus we will introduce this approach in detail in this section. Most of the material in this section is sourced from Cairns (2004).

5.6.1 Rogers and Rutkowski framework

The RR framework uses general properties of supermartingales. Suppose the process \( \{M(t)\} (t \geq 0) \) is a strictly positive supermartingale. Now we define the process \( G(t,T) \) to be \( E[M(T) | M(t)]/M(t) \) for \( T \geq t \). Then \( G(t,T) \) satisfies \( G(t,T) \leq 1 \), and \( G(T,T) = 1 \). We also see \( G(t,T_1) \geq G(t,T_2) \) for \( t \leq T_1 < T_2 \). Therefore, the process \( G(t,T) \) is a reasonable candidate for \( P(t,T) \), which is the price of a zero coupon bond at time \( t \) that expires at time \( T \). In particular, \( G(t,T_1) \geq G(t,T_2) \) for \( t \leq T_1 < T_2 \) guarantees that all the instantaneous forward rates observed at time \( t \), which are calculated by

\[
- \ln \left[ \frac{G(t,T_2)}{G(t,T_1)} \right] \left( T_2 - T_1 \right),
\]

are non-negative and hence the short rates are non-negative all the time. Bearing these properties of supermartingales in mind, we will introduce the RR framework based on the summary provided by Cairns (2004).

On the probability triple \( (\Omega, \mathcal{F}, \mathbb{P}) \), let \( A(t) \) be a strictly positive diffusion process adapted to \( \{\mathcal{F}_t\} (t \geq 0) \) that satisfies:

\[
dA(t) = A(t) \left[ \mu_A(t) dt + \sigma_A(t)^T d\tilde{W}(t) \right],
\]

(5.3)
where \( \hat{W}(t) \) is a \( d \)-dimensional column-vector standard Wiener process, \( \mu_d(t) \) is a non-positive scalar and \( \sigma_d(t) \) is a \( d \)-dimensional column-vector previsible process.

Then we assume that \( P(t,T) \) is obtained as:

\[
P(t,T) = \frac{E^{\hat{P}}[A(T)|\mathcal{F}_t]}{A(t)} \quad \text{for any } T \geq t.
\]

Now we show that the term-structure given by this formula is arbitrage-free. Since \( A(t) \) is strictly positive, for a fixed \( T \), \( D(t,T) = E^{\hat{P}}[A(T)|\mathcal{F}_t] \) is a strictly positive martingale under \( \hat{P} \). By the martingale representation theorem, we can write:

\[
dD(t,T) = D(t,T)\sigma_D(t,T)^T d\hat{W}(t),
\]

where \( \sigma_D(t) \) is some \( d \)-dimensional column-vector process.

Using Ito’s lemma:

\[
dP(t,T) = d\left( \frac{D(t,T)}{A(t)} \right)
\]

\[
= P(t,T) \left[ \left[ -\mu_d(t) - \sigma_d(t) \cdot (\sigma_D(t,T) - \sigma_A(t)) \right] dt + (\sigma_D(t,T) - \sigma_A(t))^T \, d\hat{W}(t) \right]
\]

\[
= P(t,T) \left[ \left[ r(t) + \hat{\eta}(t) \cdot S(t,T) \right] dt + S(t,T)^T \, d\hat{W}(t) \right]
\]

where

\[
r(t) = -\mu_d(t),
\]

\[
\hat{\eta}(t) = -\sigma_d(t),
\]

and \( S(t,T) = \sigma_D(t,T) - \sigma_A(t) \).

We note that \( \hat{\eta}(t) \) can be interpreted as the market-price-of-risk for the term structure since \( \hat{\eta}(t) \) is a function solely of \( t \), independent of \( T \), i.e. \( \hat{\eta}(t) \) is common for all \( T \).
Recall that we have shown in Section 4.5 that the existence of an \( L^2 \)-reducible market-price-of-risk process is sufficient for the market to be arbitrage-free. Therefore, as long as \( \tilde{\eta}(t) \) is \( L^2 \)-reducible, the term-structure is arbitrage-free.

We interpret \( r(t) \) as the instantaneous spot rate. If \( A(t) \) is a supermartingale \( \mu_\lambda(t) \) is non-positive, thus \( r(t) \) is non-negative.

Under the measure \( \hat{P} \), the deflator is already known and expected present values are easily calculated without converting the measure. That is why this measure is called a pricing measure. To see this, under the measure \( \hat{P} \),

\[
P(t,T) = \frac{E^\hat{P}[A(T)|\mathcal{F}_t]}{A(t)} = E^\hat{P}[P(T,T) \left\{ \frac{A(T)}{A(t)} \right\} |\mathcal{F}_t] \quad \text{for all} \quad T \geq t.
\]

This formula shows that \( A(T)/A(t) \) is the deflator given \( \mathcal{F}_t \) for the price process \( P(t,T) \) under \( \hat{P} \), and \( A(t) \) is the density process of the deflator. Therefore, the price of a financial asset, which is denoted by \( V(t) \), is obtained by:

\[
V(t) = E^\hat{P}[V(T) \left\{ \frac{A(T)}{A(t)} \right\} |\mathcal{F}_t],
\]

where \( V(T) \) is the payoff of the financial asset at time \( T \).

This formula holds as long as the payoff of the financial asset can be replicated by zero coupon bonds. See, for example, Cairns (2004b) for the formal mathematical proofs.

Under the measure \( \hat{P} \), we know the market-price-of-risk for this term structure \( \tilde{\eta}(t) \) is \( -\sigma_\lambda(t) \). Now let \( \eta(t) \) denote the market-price-of-risk in the real world. We need to estimate \( \eta(t) \) since our objective is not only to price a financial asset but also to obtain a realistic projection of the future development of the value of the financial asset.
Following Cairns (2004b), we assume that \( \eta(t) \) is expressed using an \( L^2 \)-reducible process \( \theta(t) \) as \( \eta(t) = \hat{\eta}(t) + \theta(t) \). Then:

\[
d\hat{W}(t) = dW(t) + \theta(t)dt
\] (5.9)

should hold where \( W(t) \) is a standard Wiener process under the real-world measure \( P \).

In summary, we have:

\[
dP(t,T) = P(t,T)\left[\left[r(t) + [\hat{\eta}(t) + \theta(t)] \cdot S(t,T)\right] dt + S(t,T)^\gamma dW(t)\right]. \tag{5.10}
\]

The process \( \theta(t) \) should be chosen so that the risk premium \( \hat{\eta}(t) + \theta(t) \cdot S(t,T) \) is as realistic as possible.

The risk neutral measure \( Q \) is often used for pricing financial products. By choosing the new measure \( \hat{W}(t) \) that satisfies \( d\hat{W}(t) = \hat{\eta}(t)^\gamma dt + d\hat{W}(t) \) in formula (5.5), the pricing measure \( \hat{P} \) is converted to the risk-neutral measure \( Q \). This conversion is equivalent to setting \( \theta(t) \) to \( -\hat{\eta}(t) \) in formula (5.10). However, as Cairns (2004b) suggests, pricing is much easier under \( \hat{P} \) than under the measure \( Q \) in the RR framework where the form of the deflator is already known; therefore there would be no need to convert the measure from \( \hat{P} \) to the risk-neutral measure \( Q \) in the RR framework.

### 5.6.2 Flesaker and Hughston approach

The Flesaker and Hughston (1996) (“FH”) approach is a special case of the general positive interest approach. The FH approach can be interpreted as the modelling of the

---

55 Strictly, Cairns (2004b) just assumed that \( -\theta(t) \) satisfies Novikov’s condition.
slope of the price curve of zero coupon bonds, even though FH do not explicitly characterize the model as a “slope model”. The development of the initial term structure of the slope of the price curve of zero coupon bonds is driven by a family of martingales. Let \( \bar{\phi}(t,T) \) denote the slope of the zero coupon bond yield curve at time \( t \) for the bond that expires at \( T \), i.e. \( \bar{\phi}(t,T) \equiv -\frac{\partial}{\partial T} P(t,T) \). Since \( P(t,\infty) = 0 \) holds for all \( t \), \( P(t,T) = \int_T^\infty \bar{\phi}(t,s) ds \) holds for all \( T \). In other words, the zero coupon bond price is the integrated slope from its expiry date to infinity. Note the similarity between the FH approach and the HJM approach. The HJM approach models the instantaneous forward rates \( f(t,T) \), and forward rates are related to the slope of the price curve of zero coupon bonds by the following simple relationship:

\[
(0) \quad f(t,T) = -\frac{1}{P(t,T)} \ln(P(t,T)) = \frac{1}{P(t,T)} \bar{\phi}(t,T).
\]

Suppose the zero coupon bond price curve at time 0, \( \{P(0,T)\}_{T>0} \), is given. We define the function of the slope of the curve at time 0 as \( \phi(T) \equiv \bar{\phi}(0,T) \), and we assume \( \phi(T) \) is strictly positive for all \( T>0 \).

Now we consider the evolution of the slope. The evolution of the slope with time is assumed to be driven by a strictly positive family of martingales \( \{M(t,T)\} \) with:

\[
\hat{d}M(t,T) = M(t,T)\sigma_M(t,T)^\top d\hat{W}(t),
\]

where \( \hat{W}(t) \) is a \( d \)-dimensional column-vector standard Wiener process, and \( \sigma_M(t) \) is a \( d \)-dimensional column-vector previsible process, and \( M(0,T) = 1 \).

As a crude and instinctive idea, we would be tempted to model the slope at \( t \) for the zero coupon bond that expires at \( T \) as \( \bar{\phi}(t,T) = \phi(T)M(t,T) \). Although conceptually this is
in line with the FH framework, \( \phi(T)M(t,T) \) on its own does not satisfy the condition that \( P(t,t) \), which is \( \int_t^\infty \phi(s)M(t,s)ds \), should be one. Therefore, we need to normalize \( M(t,T) \) by dividing it by \( \int_t^\infty \phi(s)M(t,s)ds \). That normalization leads us to model the slope at \( t \) for the zero coupon bond that expires at \( T \) as:

\[
\overline{\phi}(t,T) = \phi(T) \frac{M(t,T)}{\int_t^\infty \phi(s)M(t,s)ds}.
\] (5.12)

Then, the zero coupon bond price is obtained by integrating the slope after \( T \), i.e.:

\[
P(t,T) = \int_t^\infty \overline{\phi}(t,u)du = \int_t^\infty \phi(u)\frac{M(t,u)}{\int_t^\infty \phi(s)M(t,s)ds}du = \int_t^\infty \phi(u)M(t,u)du \int_t^\infty \phi(s)M(t,s)ds.
\] (5.13)

Cairns (2004) starts the introduction of the FH model from formula (5.13). Then he moves on to the introduction of the formulae for forward and spot rates as follows.

Substituting (5.12) and (5.13) into (5.11), the instantaneous forward rates are given by:

\[
f(t,T) = \phi(T)\frac{M(t,T)}{\int_t^\infty \phi(u)M(t,u)du}.
\] (5.14)

Therefore, the instantaneous spot rate is:

\[
r(t) = f(t,t) = \frac{\phi(t)M(t,t)}{\int_t^\infty \phi(u)M(t,u)du}.
\]

In order to see that the FH formulation is a special case of the RR approach, let:

\[
A(t) = \int_t^\infty \phi(s)M(t,s)ds.
\]

Firstly the numerator of (5.13) is \( E[A(T)\mid \mathcal{F}_t] \) since:

\[
E[A(T)\mid \mathcal{F}_t] = E[\int_t^\infty \phi(s)M(T,s)ds\mid \mathcal{F}_t]
= \int_t^\infty \phi(s)E[M(T,s)\mid \mathcal{F}_t]ds.
\]
\[ = \int_0^\infty \phi(s)M(t,s)ds. \]

Secondly, \( A(t) \) is a supermartingale. For \( t < T \):
\[
A(t) = \int_t^\infty \phi(s)M(t,s)ds = \int_T^T \phi(s)M(t,s)ds + \int_T^\infty \phi(s)M(t,s)ds
\]
\[
> \int_T^\infty \phi(s)M(t,s)ds = E[A(T)|\mathcal{F}_t],
\]
since \( \phi(s)M(t,s) \) is strictly positive.

The short rates, and volatility of zero coupon bond prices can be described by \( \phi(t) \) and \( M(t,T) \). Cairns (2004b) shows that \( \sigma_\phi(t,T), \sigma_\Delta(t) \) and hence \( S(t,T) \) in (5.3) and (5.4) can be expressed by \( \phi(t) \) and the volatility of \( M(t) \) as follows. We have added some intermediate steps to the original work by Cairns (2004b). Let:
\[
A(t,T) \equiv \int_T^\infty \phi(s)M(t,s)ds, \text{ and } dM(t,T) = M(t,T)\sigma(t,T)^T d\tilde{W}(t) \quad (t \leq T).
\]

We defined \( A(t) \) in the RR approach. Now \( A(t,T) \) corresponds to \( A(t) \) and \( A(t,T) \) corresponds to \( E^\beta[A(T)|\mathcal{F}_t] \). For \( t < T \),
\[
dA(t,T) = \int_T^\infty \phi(s)dM(t,s)ds = \int_T^\infty \phi(s)M(t,s)\sigma(t,s)^T d\tilde{W}(t)ds
\]
\[
= A(t,T)V(t,T)^T d\tilde{W}(t) \quad (t < T),
\]
where
\[
V(t,T) = \frac{\int_T^\infty \phi(s)M(t,s)\sigma(t,s)ds}{A(t,T)} = \frac{\int_T^\infty \phi(s)M(t,s)\sigma(t,s)ds}{\int_T^T \phi(s)M(t,s)ds}. \quad \text{(5.15)}
\]

This is the volatility of \( A(t,T) \), which corresponds to \( \sigma_\phi(t,T) \) in (5.4), and \( V(t,t) \) corresponds to \( \sigma_\Delta(t) \) in (5.3). The volatility of \( P(t,T) \) is, using (5.8):
\[
S(t,T) = V(t,T) - V(t,t). \quad \text{(5.16)}
\]

We emphasize again that these volatilities are independent of the choice of measure, or
equivalently, the choice of the market-price-of-risk. By taking the logarithm of and differentiating both sides of the formula \( f(t, T) = \phi(T)M(t, T)/A(t, T) \), we obtain the volatility of the instantaneous forward rate \( f(t, T) \) to be:

\[
\sigma_M(t, T) = V(t, T). \tag{5.17}
\]

Using (5.7), the market-price-of-risk under the measure \( \hat{P} \), is:

\[
\hat{\eta}(t) = -V(t, t).
\]

The drift of \( A(t, t), \mu_A(t) \), can be calculated reversely using (5.6) since we have already obtained the formula of the instantaneous spot rate. We, however, try to obtain \( \mu_A(t) \) by a more straightforward method. Since

\[
d\left( \int_t^\infty g(t, s)ds \right) = -g(t, s)dt + \int_t^\infty dg(t, s)ds
\]

in general,

\[
dA(t, t) = d\left( \int_t^\infty \phi(s)M(t, s)ds \right) = -\phi(t)M(t, t)dt + A(t, t)V(t, t)\hat{\eta}d\hat{W}(t).
\]

Recall formula (5.3):

\[
dA(t) = A(t)\left[ \mu_A(t)dt + \sigma_A(t)\hat{\eta}d\hat{W}(t) \right].
\]

Comparing the above two formulae, the drift of \( A(t, T) \) is expressed as:

\[
\mu_A(t) = -\frac{\phi(t)M(t, t)}{A(t, t)} = -\frac{\phi(t)M(t, t)}{\int_t^\infty \phi(s)M(t, s)ds}.
\tag{5.18}
\]

5.7 Cairns’ Integrated Gaussian model

Cairns (2004b) proposes a model based on the FM framework whose driving factors are Ornstein-Uhlenbeck processes. The model aims to meet the demands from banks as well as insurers for good models to price long-term derivatives and to calculate the fair-value of liabilities. Cairns (2004b) points out that such models have become very
important mainly because the value of some interest-rate guarantees has significantly increased in the historically low interest rate environment. Even though there are some models that aimed to describe the realistic dynamics in the long run, see, e.g. Wilkie (1995), Yakoubov, Teeger, and Duval (1999), these were not designed for short-term risk management or derivative pricing. Cairns (2004b) is proposing a model (“CMF model”) that can be used for short-term risk management or derivative pricing without sacrificing realistic dynamics in the long run.

5.7.1 Basic characteristic of the model

Based on the observation of historical interest rates in the U.K. for over 100 years and of those in U.S. for approximately 50 years, Cairns (2004b, p.418-419) drew up the nine desirable characteristics required for a term-structure model. Among them, his point 8 describes how the CMF model stands out from most of the other existing models. His point 8 states:

“The model should give rise to sustained periods of both high and low interest rates of all maturities in a manner consistent with what we have observed in the past.”

He also points out that most, but not all, of the existing continuous-time arbitrage-free models do not allow for long cycles of interest rates due to too strong mean-reversion. Cairns (2004b, p.419) states:

“what is important in any modelling exercise is that the model gives rise to a wide and realistic range of future scenarios.”
In fact, the CMF model is shown to be capable of producing a variety of shapes of yield curves and volatility term structures for some sustained periods. In particular, the model can reproduce the flattened S-shape yield curve experienced in Japan in 2002, in addition to the classical rising, falling, humped and dipped curves. The flattened S-shape yield curve is a yield curve that starts at and remains close to zero for some time and which gradually picks up afterwards. Recall that the Black and Karasinski (1991) model, which is fundamentally a lognormal model, is usable even in a very low interest rate environment. When interest rates are very low, the interest rates are expected to follow a lognormal distribution since there should be a lower boundary of zero for the interest rates. In fact, Cairns (2004b) showed, mathematically, in the one-factor case, that the CMF model can act like the Black and Karasinski (1991) model in a low interest rate environment and act like the Vasicek (1977) model in a high interest rate environment.

The CMF model has, we believe, two other major advantages. The first advantage is that the model is Markov where only the values of state variables are required to describe the evolution of the term structure through time. This feature suggests that the CMF model has overcome the drawback of the HJM framework where the entire curve at all times needs to be recorded. The second advantage is the necessity of only one-dimensional integration. While numerical integration is required, it involves only one-dimensional integration whatever number of factors may be used. This can be easily implemented using, for example, Simpson’s rule.
5.7.2 Model description

The material of this section is sourced from Cairns (2004b). Note that we added intermediary calculation and explanation and we changed the order of the presentation. We start by describing the two fundamental components of the model. The first component is the correlated \( n \)-dimensional column-vector Wiener process \( \dot{Y}(t) \) that satisfies:

\[
\dot{Y}(t) = C \dot{W}(t),
\]

where \( C \) is a \( n \)-square matrix for which \( CC^T \) is the instantaneous correlation matrix of \( d\dot{Y}(t) \).

Let \( \dot{Y}_i(t) \) denote the \( i \)-th element of \( \dot{Y}(t) \), and \( \rho_{ij} \) denote the \( ij \)-th element of the correlation matrix, and \( c_{ij} \) denote the \( ij \)-th element of \( C \). Then the \( ij \)-element of \( CC^T \), denoted by \( (CC^T)_{ij} \), is \( \rho_{ij} \).

The second component is the \( n \)-dimensional column-vector volatility \( \sigma(t,T) \) whose \( i \)-th element, denoted by \( \sigma_i(t,T) \), decreases exponentially with \( T \) with the decay parameter \( \alpha_i \) for a given \( t \), i.e.:

\[
\sigma_i(t,T) = \sigma_i \exp[-\alpha_i(T-t)] \quad (t \leq T),
\]

where \( \sigma_i \) is some positive constant.

Next, we specify \( M(t) \) for the FH formulation. In Cairns (2004b) model, \( M(t) \) is defined using the above two fundamental components. \( M(t) \) satisfies:

\[
dM(t,T) = M(t,T) \sigma(t,T)^T d\dot{Y}(t) \quad \text{with} \quad M(0,T) = 1.
\] (5.19)
Using Ito’s lemma, we obtain:

\[ M(t,T) = \exp \left[ \sum_{i=1}^{n} \sigma_i e^{-\alpha_i(T-t)} \hat{X}_i(t) - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j e^{-(\alpha_i + \alpha_j)(T-t)} (1 - e^{-(\alpha_i + \alpha_j)T}) \right], \]

where \( \hat{X}_i(t) = \int_0^t e^{-\alpha_i(t-s)} d\hat{Y}_i(s) \) with \( \hat{X}(0) = 0 \).

By definition, we see \( \hat{X}_i(t) \) satisfies:

\[ \hat{X}(0) = 0 \text{ and } d\hat{X}_i(t) = -\alpha_i \hat{X}_i(t)dt + d\hat{Y}_i(s). \]

For the FM formulation to be complete, we need to specify the deterministic function \( \phi(s) \). Cairns (2004b) has chosen to use a predetermined form of \( \phi(s) \) so that the term \( \phi(s)M(t,s) \) is significantly simplified. This approach is different from what FH originally suggested. FH suggested to use the slope of zero coupon prices observed at time 0 as \( \phi(s) \), not a predetermined function. Cairns (2004b) admits that the resulting initial theoretical bond prices may differ from those observed. By choosing this predetermined \( \phi(s) \), however, \( P(t,T) \) becomes Markov and time-homogeneous, and there will be no need for recalibration of \( \phi(s) \). These merits seem to overwhelm the deficiency of the initial mismatch as long as the mismatch is not significant. In fact, the mismatch between the market prices of bonds and the initial theoretical bond prices would be minimal if parameters are carefully chosen.

Preparing to define \( \phi(s) \), we at first define the following deterministic function:

\[ H(s,\hat{x}) \equiv \exp \left[ -\beta s + \sum_{i=1}^{n} \sigma_i \hat{x} e^{-\alpha_i s} - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j e^{-(\alpha_i + \alpha_j)s} \right], \]
for some constant parameters $\beta$ and $\hat{x} = (\hat{x}_1, ..., \hat{x}_n)^T$.

Then, we define $\phi(s)$ as $\phi(s) = \phi(s, \hat{x})$ for some constant $\phi$.

Now we calculate:

$$\phi(s)M(t, s) = \phi \exp \left[ -\beta s + \sum_{i=1}^{n} \sigma_i e^{-(\alpha_i + \alpha_j)X_i(t)} - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j e^{-(\alpha_i + \alpha_j)X_i(t)} \right]$$

$$= \phi e^{-\beta t} H(s - t, X(t)),$$

where $X(t) = (X_1(t), ..., X_n(t))^T$ and $X_i(t) = \hat{x}_i e^{-\alpha_i t} + \hat{X}_i(t)$.

Now $X_i(t)$ is an Ornstein-Uhlenbeck process under $\hat{P}$ with $X_i(0) = \hat{x}_i$ and satisfies:

$$dX_i(t) = -\alpha_i X_i(t)dt + d\hat{Y}_i(t)$$

$$= -\alpha_i X_i(t)dt + \sum_{j=1}^{n} c_{ij} d\hat{W}_j(t), \quad (5.20)$$

by the definition of $\hat{Y}_i(t)$.

Now, applying this specific form of $\phi(s)M(t, s)$ to the definition of $A(t, T)$ under the FH formulation, we have:

$$A(t, T) = \int_{T}^{T} \phi(s)M(t, s)ds = \phi e^{-\beta t} \int_{T}^{T} H(s - t, X(t))ds$$

$$= \phi e^{-\beta t} \int_{T}^{T} H(s, X(t))ds. \quad (5.21)$$

It follows that:

$$P(t, T) = \frac{A(t, T)}{A(t, t)} = \frac{\int_{T}^{T} H(s, X(t))ds}{\int_{0}^{T} H(s, X(t))ds}, \quad (5.22)$$

and the deflator for the cash flow at $T$ given $\mathcal{F}_t$ is:
\[
A(T,T) = \frac{\phi e^{-rT}}{A(t,t)} = \frac{\phi e^{-rT}}{\int_0^\infty H(s,X(t))ds} = e^{-r(T-t)} \int_0^\infty H(s,X(t))ds.
\] (5.23)

These formulae tell us that \( P(t,T) \) is Markov and time-homogeneous, and the constant \( \phi \) is irrelevant. From now on, we set \( \phi = 1 \). Using (5.14), the instantaneous forward rate at time \( t \) for the settlement date \( T \) is:

\[
f(t,T) = \frac{\phi(T)M(t,T)}{\int_T^\infty \phi(s)M(t,s)ds} = \frac{H(T-t,X(t))}{\int_T^\infty H(s,X(t))ds}.
\]

Next, we obtain the form of the volatility of \( A(t,T) \), i.e. \( V(t,T) \) in formula (5.15).

Since \( d\tilde{Y}(t) = Cd\tilde{W}(t) \) by definition, the term \( \sigma(t,T)^T d\tilde{Y}(t) \) can be written as:

\[
\left( C^T \sigma(t,T) \right)^T d\tilde{W}(t).
\] (5.24)

Therefore, a column vector \( V(t,T) \) is obtained as:

\[
V(t,T) = \frac{\int_T^\infty \phi(s)M(t,s)C^T \sigma(t,s)ds}{A(t,T)} = \frac{e^{-rT} \int_T^\infty H(s-t,X(t))C^T \sigma(t,s)ds}{e^{-rT} \int_T^\infty H(s-t,X(t))ds}
\]

\[
= \frac{\int_T^\infty H(s,X(t))\bar{\sigma}(s)ds}{\int_T^\infty H(s,X(t))ds}
\] (5.25)

where \( \bar{\sigma}(s) = C^T \sigma(0,s) \).

The \( j \)-th element of \( \bar{\sigma}(s) \), denoted by \( \bar{\sigma}_j(s) \), is \( \sum_{i=1}^n \sigma_i e^{-\alpha_i s} c_{ij} \).

As shown in formula (5.17), the volatility of the instantaneous forward rate \( f(t,T) \) is \( \sigma_M(t,T) - V(t,T) \). Applying (5.24) to (5.19), the volatility is:

\[
C^T \sigma(0,T-t) - V(t,T).
\] (5.26)
5.7.3 Choice of the market-price-of-risk

As shown in formula (5.7), \( \hat{n}(t) \), the market-price-of-risk under the measure \( \hat{P} \) is

\( -V(t,t) \). Now we need \( \eta(t) \), which is the market-price-of-risk in the real world. Recall
that we assumed that \( \eta(t) \) is expressed as \( \eta(t) = \hat{n}(t) + \theta(t) \), and therefore

\( \eta(t) = -V(t,t) + \theta(t) \).

Under the measure \( P \), the process \( X_i(t) \) behaves differently from that under the
measure \( \hat{P} \). Under this measure \( P \), \( X_i(t) \) satisfies;

\[
dX_i(t) = \left[ -\alpha_i X_i(t) + \sum_{j=1}^{n} c_{ij} \theta_j(t) \right] dt + \sum_{j=1}^{n} c_{ij} dW_j(t),
\]

by substituting (5.9) into (5.20).

This formula indicates that \( X_i(t) \) is in general not normal any more under the measure
\( P \) unless \( \theta(t) \) has certain specific forms. Cairns (2004b) proposes to use a constant
vector \( \theta \) whose \( j \)-th element is \( \theta_j \) for \( \theta(t) \). When \( \theta(t) \) is a constant \( \theta \), then
\( X_i(t) \) satisfies:

\[
dX_i(t) = \alpha_i \left[ \mu_i - X_i(t) \right] dt + \sum_{j=1}^{n} c_{ij} dW_j(t), \quad (5.27)
\]

where

\[
\mu_i = \alpha_i^{-1} \sum_{j=1}^{n} c_{ij} \theta_j.
\]

In the matrix form, \( \mu = (\mu_1, ..., \mu_n)^T = \text{diag}(\alpha_i)^{-1} C \theta \) where \( \text{diag}(\alpha_i) \) is the diagonal
matrix with elements \( \{\alpha_i\}_{i=1,2,...,n} \).

Cairns (2004b) argues that choosing constant \( \theta \) is pragmatic since \( X_i(t) \) are still
normally distributed under $P$. He also proposes to specify $\mu_i$ first and to use the formula:

$$\theta = C^{-1} \text{diag}(\alpha_i) \mu,$$  \hspace{1cm} (5.28)

to derive the $\theta_j$.

Referring to formula (5.10) the risk premium on $P(t,T)$ is expressed as:

$$[ -V(t,t) + \theta(t) ] \cdot S(t,T) = [ -V(t,t) + C^{-1} \text{diag}(\alpha_i) \mu ] \cdot [ V(t,T) - V(t,t) ]$$

using formula (5.16).

Once parameters are set, $V(t,t)$ and $V(t,T)$ depend only on the value of $X(t)$ as shown in formula (5.25). Cairns (2004b) investigated the shape of the risk premiums on the 30-year zero-coupon bond in a two-dimensional setting where $X_1(t)$ and $X_2(t)$ are state variables. He has found that $\mu = (-2,6)^T$ gives a reasonable outcome when $\alpha = (0.6,0.06)^T, \sigma = (0.6,0.4)^T, \rho_{12} = -0.5$ and $\beta = 0.04$.

5.8 Summary

Building an interest rate term structure model that can both describe the realistic dynamics in the long run and fit the market is very difficult. The CMF model is a model that can achieve this difficult task. This model can produce a variety of shapes of yield curves and realistic behaviour of short rates in the long run while minimizing the mismatch between the market bond prices and theoretical bond prices. This model seems to contain as much merit of both equilibrium and evolutionary models as practically possible. This model suits our study since the fair valuation requires consistency with the market and we need a long-term realistic projection of the interest
rate term structure.
6 Parameter estimation, equity modelling and market scenario generation

In this chapter, we fix parameters of the models that will be used in our simulation in Chapter 7. We fit the parameters of the CMF model to the term structure of the Japanese Yen (“JPY”, hereafter) interest rates. The parameters are, in principle, chosen to fit the market on a specific date. In addition, our model for the equity portfolio value is explained. We estimate parameters of this equity model for the Japanese equity market.

6.1 Estimating parameters of the CMF model

We need to decide what interest rates we will use for our simulation before we start estimating parameters. There are two major fixed-income markets in Japan: the interest rate swap market and the government bond market. We use the interest rate swap rates, rather than the government bond yields, as the discount curve for this study. An interest rate swap is a derivative on the LIBOR. The floating rate payer pays the LIBOR and receives a predetermined fixed rate half-yearly. We also use the interest rate Cap, which is also a derivative on the LIBOR, to fit the volatilities. The reasons why we use the interest rate swaps are as follows.

Firstly, as we introduced in Chapter 2, there is an argument that implicitly supports the use of the swap rate curve rather than the government bond yield curve amongst insurance industries. This argument states that the investment products provided by insurers are usually priced with a spread equivalent to corporate bonds with A ratings added over government bond yields. The spread does not reflect credit standings of insurers because insurers’ credit standings are usually high; most insurers are rated AA or higher. A likely explanation of why this spread exists is that the spread represents the liquidity of the products. Products provided by insurers are usually much less liquid.
than government bonds. Thus investors require some compensation for this illiquidity.

Secondly, the swap rate curve is much smoother than the government bond yield curve, free from the strong demand for specific bonds with high coupons or bonds that are deliverable for futures settlement. Therefore the swap curve is much more analytically tractable than the government bond yield curve. The empirical study conducted by Kikugawa and Singleton (1994) shows that the Japanese government bond prices are affected by factors such as the benchmark status, the nominal coupon and how far the price is from par, and demands for bonds that are deliverable for futures settlement. Accordingly the Japanese government bond yield curve is bumpy and hence applying mathematical models to it is more difficult.

Thirdly, there are liquid derivative markets on the LIBOR and interest rate swaps, i.e. Caps, floors, and swaptions. These derivatives provide essential information on how the market forecasts the future development of the yield curve. In particular, we can fix the volatilities of our model by matching the price of these derivatives produced by the model to the market price of these derivatives.

For estimating parameters, we have used the daily Euro Yen LIBOR for terms of 1 year and less, and interest rate swap rates for terms of 2 years to 30 years. The sample period is from 8\textsuperscript{th} March 1991 to 8\textsuperscript{th} March 2006, and the data source is DataStream.

**6.1.1 Historical behaviour of JPY yield curve**

In this section, we discuss some characteristics of the historical behaviour of the JPY
yield curve. Since the “collapse of the bubble” in 1989, which means the overall rapid asset price decline and resulting contraction of the economy, Japan has been experiencing very low interest rates over the entire maturity spectrum as shown in Figure 6-1.

![Figure 6-1 - Historical JPY interest rates](image)

In Figure 6-1, “2Yr - 10Yr spread” shows the difference between the 2-year interest rate and 10-year interest rate. This quantity is often called a slope, which represents how steep a yield curve is. “Concavity” shows the difference from the linearly interpolated point between the 2-year rate and the 10-year rate to the 5-year swap rate. We can see several features of the historical development of the JPY yield curve through observation of the above figure. Firstly, the interest rates have kept falling over the past 15 years, and they have just started increasing recently. In particular, the overnight rates
were stuck at almost zero since 1999 when the Bank of Japan introduced its zero rate policy. Secondly, the yield curve has been upward-sloping, measured by the spread between the 2-year rate and the 10-year rate, since 5th September 1991. The spread was more than 1% for the most of the period, and its fluctuation does not seem to have a high correlation with the movement of the swap rates. Thirdly, the yield curve has been concave for most of the sample period. During the mid 1990s the concavity is relatively high, and the interest rates are also high then. The concavity seems to fluctuate within a fixed band, and seems not to have high correlation with the movements of slopes and interest rates. These observations support the arguments, for example stated by Hull (1999, p.358), that the parallel shift, slope and concavity are the three factors that dominate the yield curve movement.

Ito (2005) conducted an empirical study of the Japanese interest rate term structure, using the daily 1 to 12-month LIBOR and Japanese 2, 3, 4, 5, 7, and 10-year interest rate swap rates for the period from 1st March 1993 to 31st March 1998. He calculated spot rates from the original interest rate swap rates, which are par rates, and studied the term structure of spot rates. Using the likelihood ratio test method proposed by Johansen (1988), he determined the number of common trends among the LIBOR and the interest rate swap rates. He concluded that a set of rates whose expiry is not more than 2 years fluctuates with one common trend, i.e. one independent factor, and a set of rates whose expiry is not more than 7 years fluctuates with two common trends. Moreover, a whole set of rates fluctuates with three common trends. Ito (2005) argues that the result could be explained by the segmentation of markets into three groups. The first group is the “money dealing” group, where cash managers deal with products
whose expiry date is from overnight to 2 years. The second group is the “ALM operation” group, where Japanese mega-banks actively transact interest rate swaps whose expiry dates are from 3 to 7 years. The third group is the “capital market” group, where bond dealers are actively involved in capital-market-oriented transactions. Therefore the market is driven by three groups each of which represents a market participant whose perspective of the market is different from the others.

6.1.2 Empirical structure of forward rate volatilities
In this section, we show that, through empirical analysis, a JPY forward rate is largely lognormally distributed with volatilities that are locally constant, but these volatilities tend to decrease as the interest rate level increases. We investigated the effective forward rate for the period from 1 year to 1.5 years. Let $fwd(t)$ denote the effective forward rate for the 6-month period from 1 year to 1.5 years at time $t$, i.e. $fwd(t) = (P(t,1)/P(t,1.5) - 1) \times 2$. The reason why we have chosen the effective forward rate, and not the continuously compounding rate, is to have consistency with the cap market where the payoff is made for the difference between the strike rate and the effective short rate.

Let $\Delta fwd(t)$ denote $fwd(t + \Delta t) - fwd(t)$. If the forward rate is lognormally distributed with constant volatility, the following relationship holds:

$$\text{Std} \left( \frac{\Delta fwd(t)}{fwd(t)} \right) = \sigma_1 \sqrt{\Delta t},$$

i.e.,

$$\text{Std} \left( \frac{\Delta fwd(t)}{\sqrt{\Delta t}} \right) = \sigma_1 \cdot fwd(t),$$

(6.1)
where \( \text{Std}(x) \) denotes the standard deviation of a random variable \( x \) and \( \sigma_1 \) is a constant.

If the forward rate is normally distributed with constant volatility, the following relationship holds:

\[
\text{Std}(\Delta fwd(t)) = \sigma_2 \sqrt{\Delta t},
\]

i.e.,

\[
\text{Std}\left(\frac{\Delta fwd(t)}{\sqrt{\Delta t}}\right) = \sigma_2,
\]

where \( \sigma_2 \) is a constant.

Note that \( \sigma_2 \) is the volatility of the absolute change of the forward rate while \( \sigma_1 \) is the volatility of the change relative to the forward rate. Following the market convention, we call the former a “basis point volatility” and the latter a “percentage volatility”.

Formulae (6.1) and (6.2) tell us that the standard deviation of \( \frac{\Delta fwd(t)}{\sqrt{\Delta t}} \), the annualized change of the forward rate, is proportional to the forward rate when its distribution is lognormal; and is constant regardless of the forward rate level when its distribution is normal. Putting together formulae (6.1) and (6.2), we have:

\[
\text{Std}\left(\frac{\Delta fwd(t)}{\sqrt{\Delta t}}\right) = \sigma_2 + \sigma_1 \cdot fwd(t),
\]

where \( \sigma_2 \) is zero when the forward rate follows a lognormal distribution, or \( \sigma_1 \) is zero where the forward rate follows a normal distribution.

This expression suggests that, when the standard deviation of the annualized change of the forward rate is plotted against the forward rate, the points constitute a line from the origin whose slope is the percentage volatility when the distribution is lognormal, while the points constitute a horizontal line whose y-axis is the basis point volatility when the
distribution is normal.

We will investigate historical data to see which distribution the forward rate follows. We analyze two kinds of volatility; one is historical and the other is the volatility implied in cap prices. The historical basis point volatility for a specific date is calculated as the standard deviation of annualized absolute change of the forward rate over the immediate past 30 days. The historical implied volatility for the forward rate is calculated from the market cap flat volatility\(^{56}\). Historical Cap volatilities are available from DataStream for 1, 2, 3, 4, 5, 7 and 10 year expiry terms. We obtain spot volatilities for every half-year expiry from 0.5 to 10 years from these flat volatilities as follows, and hence we obtain the spot volatility for the caplet that resets in one year and is settled in 1.5 years.

A price of a Cap is calculated as the total price of all the caplets that comprise the Cap once the flat volatility for the Cap is given. The price of each caplet is easily calculated from the flat volatility using a standard formula. See, for example, Hull (1999, p.540). Let \(P_{cap}(T)\) denote the price of a Cap whose term is \(T\) years. When \(T=2\), for example, this Cap is composed of four caplets that are settled in 0.5, 1, 1.5 and 2 years. The price of this Cap is obtained by adding prices of those caplets that are calculated using the standard formula from the given flat volatility. Note that we do not know the market price of each independent caplet at this stage while we know the price of Caps. The flat volatility is the quotation for Cap prices, not caplet prices. In other words, when

\(^{56}\) A Cap is composed of a series of caplets that expire half yearly in a row. The same flat volatility is applied to all the caplets comprising any particular cap, while the spot volatility is applied to each caplet separately.
the flat volatility for a Cap is 20%, this does not mean that the volatility of all the caplets that comprise the Cap is 20%. This tells us that the price of the Cap is obtained by applying the standard formula for all the caplets as though the volatility of all the caplets were 20%. For example, suppose the flat volatility for a 2-year Cap is 15% and that for a 3-year Cap it is 20%. When the 2-year Cap is priced, we use 15% volatility for all the caplets, while we use 20% volatility for all the caplets when the 3-year Cap is priced. For example, the caplet that is settled is 1 year is included in both Caps. The price of this caplet when the 2-year Cap is priced is different from that when the 3-year cap is priced. Therefore, we need to estimate the market price of each caplet, or the spot volatility of this particular caplet, from the Cap prices we have obtained.

We estimate the spot volatilities for caplets whose expiries are 1, 2, 3, 4, 5, 7 and 10 years by the following procedure.

1. We guess the spot volatilities for caplets whose expiries are 1, 2, 3, 4, 5, 7 and 10 years. Initial guesses are given by the flat volatilities.

2. We obtain spot volatilities for every half year expiry from 0.5 to 10 years using cubic spline interpolation.

3. Using these spot volatilities, we calculate all the caplet prices, then the prices of the caps whose expiries are 1, 2, 3, 4, 5, 7 and 10 years are calculated. Let $P_{cap}(T)$ denote the price of the cap whose expiry is $T$ years that was calculated using the obtained spot volatilities.

4. We minimize the error:

$$\sum_{T=1,2,3,4,5,7,10} \left( P_{cap}(T) - \overline{P_{cap}(T)} \right)^2,$$

by changing the estimation of the spot volatilities.
This process is illustrated in Figure 6-2.

Figure 6-2 - Spot volatility determination process

These spot volatilities are percentage volatilities. We then get the basis point volatilities by multiplying the forward rate by the spot volatilities.

Historical and implied basis point volatilities of the forward rate are plotted in Figure 6-3 and Figure 6-4 against the corresponding forward rates on the x-axis.
Figure 6-3 - Historical 30-day volatility and forward rates

Figure 6-4 - Implied volatilities and forward rates
Observing these figures, roughly speaking, we see that the plotted points constitute a line from the origin when the forward rates are low; typically the forward rates are less than 2%. However, the slope declines gradually to the level where the line is horizontal as the forward rates increase. In fact, implied volatilities constitute a nearly horizontal line when the forward rates are more than 2%. This observation suggests that the forward rate follows a lognormal distribution when the rate is low, and then gradually approaches a normal distribution as the rate increases.

In order to assess this quantitatively, we conducted a regression analysis. The dependent variable is the historical 30-day or implied Cap annualized basis point volatility. The explanatory variable is the forward rate. We have chosen the sub-periods where the relationship between the forward rates and volatilities seems, by observation, stable. We expect that a high coefficient of determination will be obtained within each sub-period.

\[ \text{The variance of the dependent variable is the sum of the variance of the projections that are on the regression line and the variance of the errors from the projections. The coefficient of determination shows the ratio of the variance of the projections to the variance of the dependent variable. See, for example, Davidson and MacKinnon (2004).} \]
Table 6.1 - One-year forward rates and volatility

<table>
<thead>
<tr>
<th>Period</th>
<th>Range (mostly)</th>
<th>Average</th>
<th>Historical</th>
<th>Implied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\beta$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>8/March/91 ~ 20/July/92</td>
<td>More than 4%</td>
<td>5.86%</td>
<td>0.16</td>
<td>0.95</td>
</tr>
<tr>
<td>21/July/92 ~ 12/May/95</td>
<td>1.5% ~ 4%</td>
<td>3.27%</td>
<td>0.42</td>
<td>0.87</td>
</tr>
<tr>
<td>15/May/95 ~ 27/Feb/01</td>
<td>0.25% ~ 1.5%</td>
<td>0.91%</td>
<td>0.97</td>
<td>0.89</td>
</tr>
<tr>
<td>28/Feb/01 ~ 18/Aug/03</td>
<td>Less than 0.25%</td>
<td>0.19%</td>
<td>0.87</td>
<td>0.78</td>
</tr>
<tr>
<td>19/Aug/03 ~ 8/March/06</td>
<td>0.25% ~ 1.5%</td>
<td>0.31%</td>
<td>0.80</td>
<td>0.76</td>
</tr>
<tr>
<td>ALL</td>
<td>1.59%</td>
<td>0.33</td>
<td>0.60</td>
<td>0.63</td>
</tr>
</tbody>
</table>

We let $R^2$ denote a coefficient of determination, and $\beta$ denote the slope of the regression. The entries with a high coefficient of determination are highlighted. We observe that $\beta$, which can also be interpreted as percentage volatility as indicated in formula (6.1), decreases as the forward rate increases. The $\beta$ for historical volatility decreases to 0.42 and then 0.16 from 0.97 as the forward rate range increases from 0.25%~1.5% to 1.5%~4% and more than 4%. A lower $\beta$ value shows the distribution is closer to normal.

In each sub-period, correlations are high. The lowest $R^2$ of 0.80 is scored for historical volatility for the period when forward rates are 0.25%~1.5% starting from 19th August.
2003. Even for that period, the implied volatility produces high correlation with \( R^2 \) being 0.95. The highest \( R^2 \) of 0.99 is scored for implied volatility for the period when forward rates are less than 0.25%. The correlation of cap implied volatility with forward rates tends to be higher than that of the historical volatility. The likely explanation of this is that the implied volatility does not suffer from a bid-offer spread or other type of data deficiency, and it is determined theoretically using models anticipating future fluctuations. In the sub-period when the forward rate is more than 4%, even \( R^2 \) on historical volatility is 0.95. These high correlations show that the volatilities do not change dramatically in each sub-period. In other words, the percentage volatility is largely determined by the level of forward rate, and percentage volatility decreases as forward rate increases. Accordingly, the shape of the distribution of forward rate approaches a normal distribution as the level of forward rate increases.

### 6.1.3 Fitting the CMF model parameters to the market

We will now show the CMF model parameters that were fitted to the JPY interest rate market and explain the fitting process. The parameters have been determined as follows, fitting to the market swap rates and market cap prices:

\[
\beta = 4\%, \quad \alpha = \begin{bmatrix} 0.69 \\ 0.14 \\ 0.035 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 42.9\% \\ 84.4\% \\ 23.1\% \end{bmatrix}, \quad \rho = \begin{bmatrix} 1 & -0.4 & 0.2 \\ -0.4 & 1 & -0.5 \\ 0.2 & -0.5 & 1 \end{bmatrix}, \quad X(0) = \begin{bmatrix} -4.58 \\ 0.37 \\ -4.47 \end{bmatrix}.
\]

Figure 6-4 and Figure 6-5 show the market and model yield curves and the one-year instantaneous forward rate percentage volatilities as a function of the forward rates. The pairs of percentage volatility and forward rate were obtained by simulation. The model yield curve is well fitted to the market yield curve. The model also exhibits the feature
that the percentage volatility decreases as the forward rate increases. The extent of the
decrease is, however, much more moderate than what we have observed in the historical
data. After numerous trial calculations, we concluded that such a rapid volatility
decrease with the increase of the forward rate cannot be implemented under any
parameter set consistent with the market. In fact, it was not feasible to find parameters
that produced the market price of Caps that have short expiry terms, for example 2 years.
Therefore we included only a 10-year Cap price in the fitting process. This is acceptable
since we are more interested in the long-term behaviour of the interest rate term
structure rather than the short-term.
Figure 6·5 - Market and model yield curve

Figure 6·6 - Model forward rate percentage volatilities
We outline how we have fixed the parameters.

Number of Factors: We followed the study of Ito (2005). The empirical study by Ito (2005) suggested that the JPY yield curve fluctuates with three independent factors when the maximum expiry term of the interest rates is 10 years. This result requires us to use not less than three factors since we will deal with a yield curve with an expiry term of more than 30 years. It is, however, practically very difficult to deal with 4 or more factors. Therefore, we use three factors in our study.

\( \beta \): We set this parameter to 4% by observing the historical forward rate for the period from 30 years to 30.5 years. This forward rate was obtained through fitting the market curve to Cairns’ (1998) descriptive model using formula (5.1) with the parameters: 
\[ \{c_i\}_{i=1}^{5} = \{0.2, 0.4, 0.8, 1.6, 3.2\} \]. We added the last parameter \( c_5 = 3.2 \) to the first four parameters that Cairns (1998) found to be appropriate, in order to meet the objective of describing the long end of the curve. The forward rate is \( \beta \) when the expiry date is infinite (Cairns (2004)). In Figure 6-7, we see that the 30-year forward rate is more than 3% for the most of the time even after the zero rate policy took effect in 1999.
The study by Ito (2005) suggests that the short-end (not more than two years to expiry), middle-end (three years to seven years) and the long-end (ten years) of the Japanese spot rate curve are influenced by different factors. Therefore we assume, under the CMF model framework, there are three factors whose half lives are 1, 5 and 20 years. Recall that a half life is defined to be the $T$ that satisfies $f(T) = 1/2$ for the function $f(t) = e^{-\alpha t}$ where $\alpha$ is a positive constant. When $T$ is given, $\alpha = \ln(2)/T$. Accordingly, $\alpha$ is calculated as: $\alpha_1 = \ln(2)/1 = 0.69$, $\alpha_2 = \ln(2)/5 = 0.14$, and $\alpha_3 = \ln(2)/20 = 0.03$. Under this setting, factor 1 has little effect on the middle and long-end of the curve, and factor 2 has little effect on the long-end of the curve.

\( \rho \): Formula (5.26) gives a vector volatility process of instantaneous forward rates for
any expiry. The first, second and third element of this vector are the volatility for the second and third independent Wiener processes that drive the model. Therefore we have three volatility curves for each independent Wiener process. Correlations are chosen so that the three independent Wiener processes represent the three independent factors that govern parallel shift, slope and concavity of yield curve as well as possible. In other words, correlations are chosen so that horizontal, monotone and humped volatility curves are produced. Our fitting of $\rho$ and $\sigma$ was conducted in an integrated process based on trial and error. Figure 6-8 shows the resulting three volatility curves.

The shape of the curves suggests that the first Wiener process makes the yield curve more humped, the second more sloped and the third more parallel-shifted.
Given $\beta, \alpha$, and $\rho$, $\sigma$, and $X(0)$ were fitted to the current, as of 8th March 2006, market swap rates and Cap prices. These parameters were chosen so that the following quantity is minimized:

$$\sum_{T=2,10,20} \left[ Rswap(T) - \overline{Rswap}(T) \right]^2 + \left[ Pcap^{(CMF)}(10) - \overline{Pcap}(10) \right]^2,$$

where $Rswap(T)$ is the $T$-year swap rate generated by the model, $\overline{Rswap}(T)$ is the $T$-year swap rate from the market, and $Pcap^{(CMF)}(10)$ is the 10-year Cap price generated by the model. Recall that $\overline{Pcap}(10)$ is the 10-year Cap price calculated from the market data. The market data provides only flat volatilities for Caps. We calculated the market Cap price from the given flat volatility data for the 10-year Cap. We have chosen these expiry terms for the swap rates because 2-year, 10-year and 20-year swap rates are benchmarks of interest rates for short, long and super-long term sectors; and therefore these swaps are more actively traded than interest rate swaps whose terms are other than those. For the same reason we have chosen the 10-year term for Caps, taking into account the fact that Caps whose terms are more than 10 years are not actively traded.

The model Cap prices $Pcap^{(CMF)}(10)$ are calculated using the simulation method under the pricing measure using 1,000 sample paths. The outline of the process of this pricing of the Cap is as follows. Note that we reset the seed to the same number every time when a new simulation starts. This results in using the same set of random samples for each pricing.

1. We temporarily fix $\sigma$ and $X(0)$.

---

58 The data source is DataStream
2. We repeat the following procedure 1,000 times and take the average of Cap prices calculated for each sample path.

3. Suppose we already know \( X(t) \). Recall that the CMF model is a Markov model. Then, we can calculate \( P(t,T) \) for all \( T \geq t \) using formula (5.22) and the value of the deflator at time \( t \) using formula (5.23). Then we know the price of the Cap that is settled at time \( t \) discounted by the deflator.

4. We generate three independent normal random numbers. Using these sample numbers, we calculate three increments of each element of \( X(t) \) using the discretised version of formula (5.20), i.e.:

\[
\Delta X_i(t) = -\alpha_i X_i(t) \Delta t + \sum_{j=1}^{3} c_{ij} \sqrt{\Delta t} Z_j(t) \quad (i, j = 1, 2, 3),
\]

where \( Z_j(t)(j = 1, 2, 3) \) are random samples generated at time \( t \).

Then we know \( X(t + \Delta t) \). Repeating this, we extend a sample path until \( t = 10 \).

We have chosen the number 1,000 as a practical choice. Numerical minimization requires a large amount of calculation to find derivatives for each variable. Even though the simulation was implemented as a binary dynamic-link library module, after a number of trial calculations, we have found that 1,000 is almost the maximum number that Excel can handle on our PC equipped with a Pentium 4 processor.

We used the same set of random samples to generate 1,000 sample paths. This means if two 10-year Caps had been priced twice using the identical parameters, the resulting prices would have been identical. Fixing random samples enabled the minimization process to converge. If different sets of random samples are used for pricing the Cap
each time, the numerical minimization never converges, since derivatives for each parameter cannot be numerically calculated precisely enough to move onto the next estimation.

The overall fitting result is shown in Table 6-2 and Table 6-3 with the fitted items are highlighted.

**Table 6-2 - Yield curve fitting result**

<table>
<thead>
<tr>
<th>Term</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Rate</td>
<td>0.1325%</td>
<td>0.2575%</td>
<td>0.5850%</td>
<td>1.2375%</td>
<td>1.7900%</td>
<td>2.0800%</td>
<td>2.2700%</td>
<td>2.4375%</td>
</tr>
<tr>
<td>Model Rate</td>
<td>0.2650%</td>
<td>0.3652%</td>
<td>0.5964%</td>
<td>1.2327%</td>
<td>1.8060%</td>
<td>2.0864%</td>
<td>2.2594%</td>
<td>2.4792%</td>
</tr>
<tr>
<td>Difference</td>
<td>0.1325%</td>
<td>0.1077%</td>
<td>0.0114%</td>
<td>-0.0048%</td>
<td>0.0160%</td>
<td>0.0064%</td>
<td>-0.0106%</td>
<td>0.0417%</td>
</tr>
</tbody>
</table>

**Table 6-3 - Cap price fitting result**

<table>
<thead>
<tr>
<th>Cap term</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Cap Price</td>
<td>0.073%</td>
<td>0.504%</td>
<td>1.070%</td>
<td>1.750%</td>
<td>2.437%</td>
<td>3.921%</td>
<td>6.218%</td>
</tr>
<tr>
<td>Model Cap Price</td>
<td>0.089%</td>
<td>0.364%</td>
<td>0.851%</td>
<td>1.543%</td>
<td>2.341%</td>
<td>3.820%</td>
<td>6.209%</td>
</tr>
<tr>
<td>Difference</td>
<td>0.016%</td>
<td>0.139%</td>
<td>0.219%</td>
<td>0.207%</td>
<td>0.096%</td>
<td>0.101%</td>
<td>0.009%</td>
</tr>
</tbody>
</table>

The model curve is well fitted to the market curve except for 0.5 and 1 year terms. For terms from 2 to 20 years, the differences are around 0.01%. Even for the 30-year term, the difference is 0.04%. We see differences more than 0.10% for 0.5 to 1 year terms. However, we do not put much importance on differences for these very short terms. Since the interest rates for the short terms are very close to zero and the slope of the
yield curve for these terms is very steep, it may be difficult to find a model to produce such a curve. In addition, our objective is long-term projection rather than short-term. Therefore inconsistency between the market and model at very short terms is not an important issue.

For the same reason, we do not put much importance on differences on 1, 2, 3 and 4 year Caps. We have a good fitting result for 5,7 and 10 year Caps. The differences are 0.10%, 0.10% and 0.01% respectively. Considering a Cap is an option, that is leveraged in nature, these figures are reasonable. Differences for 3 and 4 year caps are around 0.21%. This is still not an extremely large figure, but is not acceptable for pricing derivatives. However, we do not reject the parameters we have obtained for the reason that our objective is long-term projection.

These fitting results confirm our view that it is almost impossible to specify a model and parameters that perfectly satisfy both requirements for long-term simulation and precise derivative pricing, in particular, under an extremely steep JPY yield curve with a zero overnight rate on the chosen date.

The Cap prices shown in Table 6-3 (“Fitted Cap prices”, hereafter) were obtained using the fixed set of random samples. With the parameters fixed, a different set of random numbers produces Cap prices different from Fitted cap prices. This dispersion is not expected to be negligible since the number of sample paths we used is 1,000, which is not large. We have examined the dispersion by sampling 101 Cap prices using different sets of random samples with the parameters we have obtained. Table 6-4 shows the
average of the sample Cap prices, estimates of the standard deviation of Cap price and the T values that are calculated to be:

\[
\frac{\text{Average cap price} - \text{Fitted cap price}}{\text{Standard deviation}}.
\]

### Table 6.4 - Cap price dispersion

<table>
<thead>
<tr>
<th>Cap term</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitted cap price</td>
<td>0.073%</td>
<td>0.504%</td>
<td>1.070%</td>
<td>1.750%</td>
<td>2.437%</td>
<td>3.921%</td>
<td>6.218%</td>
</tr>
<tr>
<td>Average cap price</td>
<td>0.088%</td>
<td>0.362%</td>
<td>0.838%</td>
<td>1.458%</td>
<td>2.139%</td>
<td>3.632%</td>
<td>6.126%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.003%</td>
<td>0.018%</td>
<td>0.049%</td>
<td>0.101%</td>
<td>0.159%</td>
<td>0.283%</td>
<td>0.473%</td>
</tr>
<tr>
<td>T value</td>
<td>-0.278</td>
<td>-0.155</td>
<td>-0.261</td>
<td>-0.834</td>
<td>-1.273</td>
<td>-0.663</td>
<td>-0.174</td>
</tr>
</tbody>
</table>

The critical value of the student-\(t\) distribution with 100 degrees of freedom for a 5% two-sided probability is 1.984. Therefore, we cannot reject the hypothesis that the Fitted cap prices are true cap prices with 5% probability.

### 6.1.4 The market-price-of-risks

The market-price-of-risks cannot be obtained by fitting them to the market snapshot because the market snapshot does not contain any information on the market-price-of-risks for the future. The market-price-of-risks, or risk premiums, for the future are essentially estimates. We have determined the market-price-of-risks based on the common recognition prevailing among academics and practitioners. As a practical choice, we assume that the market-price-of-risks are constant.

As explained in Section 5.7.3, fixing the market-price-of-risk is equivalent to fixing
\( \mu = (\mu_1, \mu_2, \mu_3)^T \) under the assumption of a constant market-price-of-risk. Here, \( \mu \) is the convergence level of \( X(t) \) as \( t \) goes infinity, as Formula (5.27) shows. Therefore, the exercise of fixing the appropriate market-price-of-risk is equivalent to finding the ultimate shape of a yield curve that is achieved when \( X(t) \) converges to \( \mu \). Note that \( \mu \) does not affect the fixing of other parameters, i.e. the parameters we have already fixed are independent of the choice of \( \mu \). We decided to use \( \mu = (0, 0, 0)^T \), i.e. we assume that the pricing measure \( \hat{P} \) is exactly the real-world measure \( P \). Figure 6-9 shows the shape of the converged volatility, yield and risk premium curves, i.e. volatility, yield and risk premium curves when \( \lim_{t \to \infty} X(t) = \mu \).

**Figure 6-9 - Converged shape of the volatility, yield and risk premium curve**
The first reason why we have chosen $\mu = (0, 0, 0)^T$ is that the shape of the converged model curves is reasonable. The curves exhibit the following reasonable features:

- the spread between 2-year and 10-year par rates is 0.84%,
- the concavity is 0.10%, and
- the risk premiums are 0.91%, 1.29%, and 1.44% for 10, 20 and 30-year zero coupon bonds.

The second reason is that, assuming $\mu = (0, 0, 0)^T$, the risk premiums implied in the model curves at time zero are also reasonable. The risk premiums depend on $X(t)$, therefore, the risk premiums at time zero when $X(0) = (-4.58, 0.37, -4.47)^T$ are different from those when $X(t) = \mu = (0, 0, 0)^T$. Risk premiums at time zero for 10, 20 and 30-year zero coupon bonds are calculated to be 0.36%, 0.58%, and 0.69%. Under the very low interest rate environment at time zero, these risk premiums seem reasonable.

The third reason is about practicality. The smaller the market-price-of-risk is, the more accurate our calculation is, as we showed in Section 4.7. In other words, there is always a limit on the size of the market-price-of-risk that can be practically dealt with in a simulation environment. We could come up with a slightly improved shape of the converged model curve assuming $\mu$ that is other than $(0, 0, 0)^T$. Such a $\mu$, however, is likely to result in a market-price-of-risk that is beyond the limit. For example, $\mu = (-1.0, 0.5, 2.0)^T$ seems to improve the shape of the converged curves in that:

- the spread between 2-year and 10-year par rates is 0.94%,
the concavity is 0.17%, and
risk premiums are 1.13%, 1.42%, and 1.60% for 10, 20 and 30-year zero coupon bonds.
This choice of \( \mu \), however, results in market-price-of-risks that are too high to be practically handled. The market-price-of-risks are calculated to be \( \theta = (-69\%, -23\%, -4\%)^T \) using Formula (5.28). This is far beyond the practical limitation of 40% stated in Section 4.7.
The small improvement in the ideal shape for the ultimate curve would not justify the practical calculation problems.

6.2 Some sample paths
Using the parameters obtained above, we have generated 5,000 sample paths of future development of the yield curve by simulation. These generated paths seem reasonable and realistic. As examples, we will show some sample paths in this section.

For each path, we calculated and stored the zero coupon bond prices half yearly up to 100 years from now. We tracked the 0.5, 2, 5, 10, and 20-year continuous compounding spot rates of the 5,000 sample paths. Scenario 836 produced the maximum interest rate of 55.4%, after 50 years, of all the 5,000 sample paths. The entire development of this scenario is shown on Figure 6-10.
Even though extremely high rates were simulated, we can see it happened only once in 100 years, and although that high rate period continued for 10 years from 45 years to 55 years, the rates stayed rather low at all other times. The result that the maximum interest rate was 55.4% among 5,000 simulations does not mean that this rate is the maximum this model would ever produce. That result, however, suggests to us that the model produces extremely high interest rates with very low probability, and the rates tend to end up as a temporary phenomena.

Under scenario 4,532, interest rates stayed below 3.0% over the whole period as shown in Figure 6-11.
This result shows that the model is capable of producing a scenario where a low-interest environment continues for several decades. Most of scenarios that the model has produced were, however, moderate. The average maximum interest rate over the 5,000 simulations was 14.7%. We show as an example scenario 917 whose maximum interest rate was 14.7% in Figure 6-12.
Recall that the converging rate, $\beta$, was set to be 4%. This means that interest rates are expected to become 4% in the long run regardless of the current interest rate environment. Therefore, the curve is expected to be normal, i.e. upward sloping, when rates are less than 4%, and is expected to be inverted, i.e. downward sloping, when rates are more than 4%. By observation, we confirmed that the generated 5,000 sample paths clearly exhibited this feature.

6.3 The stochastic process of the value of the equity portfolio
We assumed that the stochastic process for the value of the with-dividend equity portfolio held in the business unit process follows a simple geometric Wiener process. In our simulation, some portion of earned premiums is assumed to be invested in the equity portfolio. We assume that the equity portfolio is well diversified and that the
performance of the portfolio is close to that of equity indices. For the discrete time interval \( \Delta t \), we assume the value of the equity portfolio follows the following process for given \( \Delta t > 0 \):

\[
S(t + \Delta t) = S(t) \cdot \frac{1}{P(t, t + \Delta t)} \cdot \exp\left[ \eta \sigma \Delta t \cdot \exp\left[ -\frac{1}{2} \sigma^2 \Delta t + \sigma \Delta W_s(t) \right] \right] \quad (t = 0, \Delta t, 2\Delta t, 3\Delta t, ...) \tag{6.3}
\]

where \( S(t) \) is the value of the equity portfolio, \( \sigma \) is the volatility which is constant, \( \eta \) is the market-price-of-risk which is constant on the equity risk, \( \Delta W_s(t) \) is the standard Wiener process for the time period from \( t \) to \( t + \Delta t \) that is independent of the Wiener processes associated with interest rate movements.

The process:

\[
D(t) = A(t) \cdot \exp\left[ -\frac{1}{2} \eta^2 t - \eta W_s(t) \right] \quad (t = 0, \Delta t, 2\Delta t, 3\Delta t, ...) \tag{6.4}
\]

is the state price deflator for \( S(t) \); i.e. \( \{ S(t)D(t) \} \) is a martingale.

This can be shown as follows. Suppose \( \mathcal{F}_t \), which is the standard filtration of \((W(t), W_s(t))\) at time \( t \), is given. Then:

\[
E\left[ S(t + \Delta t) \frac{D(t + \Delta t)}{D(t)} \mid \mathcal{F}_t \right] = S(t) \cdot \frac{A(t + \Delta t)}{A(t)} \cdot E\left[ \exp\left[ \eta \sigma \Delta t - \frac{1}{2} \sigma^2 \Delta t + \sigma \Delta W_s(t) \right] \cdot \exp\left[ -\frac{1}{2} \eta^2 \Delta t - \eta \Delta W_s(t) \right] \right]
\]

by independence, and, and,

\[
E\left[ \frac{A(t + \Delta t)}{A(t)} \right] = E\left[ \frac{A(t + \Delta t)}{A(t)} \times 1 \right] = P(t, t + \Delta t).
\]

Also,

\[
E\left[ \exp\left[ \eta \sigma \Delta t - \frac{1}{2} \sigma^2 \Delta t + \sigma \Delta W_s(t) \right] \cdot \exp\left[ -\frac{1}{2} \eta^2 \Delta t - \eta \Delta W_s(t) \right] \right]
\]
\[ E \left[ \exp \left( -\frac{1}{2} (\sigma - \eta)^2 \Delta t + (\sigma - \eta) \Delta W_t(t) \right) \right] = 1. \]

We use constant volatility for the change of the value of the equity portfolio because it is sufficient to meet our objective, i.e. to compare risk-return profiles between various business strategies. We start from a simple model, and then we may switch to more sophisticated models if necessary. More sophisticated models such as those with stochastic volatility will be needed only when we conclude that investing in equities is a sensible strategy. As we will show in Section 6.4, historical equity volatility is not a stable quantity. We will estimate a long-term average based on the observation of historical behaviour of equity volatility.

For the same reason, we assume that the Wiener process that drives equities is independent of the Wiener process that drives bonds. Formula (6.3) states that the average return is dominated by the current interest rate, i.e. both of the return on bonds and the return on equities are high in high inflation periods; and both of them are low in low inflation periods. What are independent are the random fluctuations from the mean return. This assumption seems consistent with the notion perceived by market participants. In addition, the assumption of independence is largely in line with historical observations. Figure 6-13 shows the historical rolling 250-day correlation between the daily change ratios, using logarithms, of the Nikkei 225\textsuperscript{59} index and the daily change ratios of interest rates from 3rd of May 1991 to 8th of March 2006. We investigated overnight call rates and 10-year swap rates.

\textsuperscript{59} Nikkei 225 is the Dow-method average of 225 Japanese representative stocks, published by Japan Economic Journal Inc.
The correlation between the overnight rates and equities fluctuates within the band between -0.1 to 0.1 most of the time. This observation is consistent with our assumption. On the other hand, we observe a small positive correlation between long-term rates and equity returns. The correlation between the 10-year swap rates and equities fluctuates within the band between 0 to 0.3 most of the time. However, the observed correlation, which is around 0.15, is not significantly large enough to force us to incorporate it into the model for our study.

### 6.4 Volatility estimation

We will rely on historical analysis to determine the equity volatility, $\sigma$. The equity volatility could be obtained as the volatility implied by the market equity option prices using the Black-Scholes model. However, we cannot collect appropriate equity
volatility data from the current public derivatives market. Unlike the interest rate derivatives market, the expiry terms of publicly-traded options are very short. For example, the furthest expiration of listed Nikkei 225 options published in the Nikkei newspaper is three months. The Nikkei newspaper publishes the average implied volatilities and historical volatilities. The Nikkei newspaper as of 31st March 2006 shows, for example, that the average of implied volatilities of listed Nikkei 225 options and historical Nikkei 225 volatility are 18.6% and 18.4% respectively. Even though these figures give us some information on volatility of equities, these are not useful for our study on their own. We need data for long-term expiry terms, based on long-term historical data, while the implied volatilities published are for short-term options and the historical volatilities published are based on the short observation period. Therefore we rely on historical analysis for the volatility parameter. We set our equity volatility, $\sigma$, to 20% by our historical analysis which is explained below.

Figure 6-14\textsuperscript{61} shows the historical 60-month annualized volatility of the Nikkei 225 index since the start of 1961.

\textsuperscript{61} Data source is DataStream.
We observe that the volatility fluctuates mostly within the range from 10% to 25%. In the 1990s, the volatility moves within a relatively narrow range from 20% to 25%, in line with the global high volatility phenomena of that time. Assuming that the world will never go back to the low-equity-volatility era, we could set our volatility assumption to some figure between 20% and 25%.

We need, however, to consider the characteristic of our assumed equity portfolio and the appropriateness of the sample period over which the volatility is calculated. The volatility of our assumed equity portfolio could be, however, smaller than that of the Nikkei 225. The constituents of a typical equity portfolio held by institutional investors
is closer to that of a capital-weighted index, such as TOPIX\textsuperscript{62}. Institutional investors need to secure liquidity in their portfolio. Therefore they tend to invest more into equities that have large outstanding numbers of shares. On the other hand, the Nikkei 225 is a simple average, which is more likely to be influenced by some constituent equities that have high prices in absolute terms. In fact, the TOPIX 500\textsuperscript{63} index shows an almost consistently lower volatility than the Nikkei 225. Figure 6-15 shows the historical 250-day volatility of both indices since the start of 1993.

\textbf{Figure 6-15 - Historical daily volatility of Nikkei 225 and TOPIX 500 index}

The volatility of the TOPIX 500 is almost consistently lower than that of the Nikkei 225 by approximately 3\%. The average volatility of the TOPIX 500 over the observation

\textsuperscript{62} TOPIX is the capital-weighted equity index published by Tokyo Stock Exchange.

\textsuperscript{63} TOPIX 500 is one of the sub-index of TOPIX index group.
period is 20.6%.

Considering that the volatilities had been relatively low before the 1990s, we set the volatility to 20%, as the long-term average, which is slightly lower than the average historical daily volatility of the TOPIX 500.

6.5 Setting the equity market-price-of-risk and scenario generation

The market-price-of-risk on the equity risk, or the risk premium given in exchange for taking equity risk, is not observable in a market snapshot. Here we assume that the risk premium is 4%. Then the market-price-of-risk is calculated to be 20% by dividing the 4% by the volatility of 20%. An equity risk premium of 4% is in line with the empirical study by Campbell et al. (1997, p.307) that states that the mean excess log return of stocks over commercial paper is 4.2% for the period from 1889 to 1994.

Now we have determined two parameters of our equity model for the value of our equity portfolio. Using Formula (6.3), we can attach a development of the value of the equity portfolio to each yield curve development scenario by generating another independent random sample. In addition, formula (6.4) gives us the value of the deflator on each sample path.

6.6 Summary

In this chapter we have determined parameters of the CMF model for the JPY yield curve, and parameters of our model for the value of our equity portfolio. In principle, the parameters were chosen so that the resulting yield curve and Cap prices fit the current market data. However, we also used deductive methods based on observations of historical behaviour of yield curves and equity indices.
Using these parameters, now we have 5,000 scenarios of development of yield curves and the value of the equity portfolio, and 5,000 associated scenarios of the value of the deflator.
7 Simulation of a business

In this chapter, we will investigate the risk-return profile of a particular business unit that carries a hypothetical annuity portfolio. The premium income is assumed to be invested in bonds and in a portfolio of equities. We use the market scenarios obtained in Chapter 6.

7.1 The annuity portfolio

We consider a hypothetical independent business unit within a life insurance company. The business unit is assumed to have issued a large number, for example 100,000, of identical annuities that are mutually independent. All of the analysis in this chapter is done for the whole business unit (i.e. for the entire annuity portfolio) rather than for one specific annuity contract; therefore the results of the analysis should always be interpreted on the portfolio basis. However, for readability, we will show all the figures per one contract from now on.

We have chosen the annuity product for our study because annuity products have become increasingly popular in Japan after these products started to be sold at commercial banks’ counters. The life insurance association of Japan (2006) showed that the number of policies in-force has exceeded the number of term assurances, as shown in Table 7-1. Another reason why we have chosen the annuity product is that, given a mortality table, the exposure of long term fixed assurances to market risks is in essence same as that of the annuity product. Whether a contract is an assurance or an annuity, as long as the payment amount is fixed, the contract constitutes a long term liability whose payment amounts are fixed. Japanese insurers are still carrying on their own books enormous amounts of long term assurances whose sums assured are fixed, as a result of persistent marketing for the past several decades. All of these contracts still in force are
subject to fair valuation once it begins.

Table 7-1 - Number of policies in-force of major Japanese insurance products

<table>
<thead>
<tr>
<th>(in millions)</th>
<th>Year 2004</th>
<th>Year 2005</th>
<th>Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term assurance</td>
<td>12.86</td>
<td>12.80</td>
<td>-0.06</td>
</tr>
<tr>
<td>Whole-life with-term assurance</td>
<td>19.93</td>
<td>17.84</td>
<td>-2.09</td>
</tr>
<tr>
<td>Whole-life assurance</td>
<td>12.93</td>
<td>13.45</td>
<td>+0.52</td>
</tr>
<tr>
<td>Annuity (fixed)</td>
<td>12.67</td>
<td>13.05</td>
<td>+0.38</td>
</tr>
<tr>
<td>Annuity (variable)</td>
<td>0.56</td>
<td>0.96</td>
<td>+0.4</td>
</tr>
</tbody>
</table>

The surrender and lapse ratio for annuities that were in force in Japan in the fiscal year 2005 was 3.2% and this ratio has been decreasing continually from 5.8% in 2003 and 4.9% in 2004. We therefore ignore surrenders and lapses to simplify our analysis even though we recognize that surrenders and lapses are sometimes essential for analyzing annuities precisely. The inclusion of surrenders and lapses in our study is in fact almost infeasible due to a lack of detailed data. However, their inclusion would contribute little to our analysis of the efficiency of capital usage and to finding the best investment strategy. In our view, 3.2% is insignificant enough for us to ignore surrenders and lapses.

We assume that the business unit issued the annuities with the following terms:

- Issued to a male aged exactly 35,
- Premiums are paid annually in advance for 30 years,
- JPY\(^{64}\) 1,000,000 of annuity is paid annually at the end of the year, the first payment

---

\(^{64}\) From now on, we omit the currency identifier “JPY” for readability. All currency is in JPY hereafter.

As of June 2007, 1 AUD buys JPY 100 and 1 USD buys JPY 120 approximately.
being due in 30 years.

We use the 2004 Japanese Male Mortality Table, which is shown in the Appendix I, and we have extended the maximum age of the Table to 105 by extrapolation assuming Gompertz’s Law. We also use the JPY interest rate term structure produced by the CMF model using the parameters obtained in Chapter 5.

The net premium \( \Pi \) is calculated as:

\[
\Pi = 1,000,000 \cdot \sum_{t=30}^{70} \frac{P(0,t) \cdot l_{35+t}}{l_{35}} = 244,816.
\]

We assume 700,000 of initial one-off expenses, and 100,000 of ongoing expenses that are levied as long as the contract is in force. To cover these costs, we need 44,908 of loading on top of the net premium, which is calculated by the formula:

\[
700,000 + 100,000 \cdot \sum_{t=0}^{69} \frac{P(0,t) \cdot l_{35+t}}{l_{35}}.
\]

Recall that a market value margin (MVM) is a series of hypothetical liabilities added to each expected benefit payment. A MVM for an expected benefit payment disappears immediately when the payment is made, and the MVM is recognized as revenue. The MVM is a kind of cushion against an adverse mortality experience. There is no standard method to calculate MVMs at present. Here, we assume that a MVM is calculated to be 10% of the expected benefit payments. The premium loading for a MVM, which we simply call a MVM loading hereafter, can be calculated as though the MVM were additional annuity payments. Therefore the MVM loading for our annuity is calculated.
to be 10% of $\Pi$, which is 24,482.

In addition to the MVM, we charge the up-front profit. The profit arising from MVMs is not recognized until the associated benefit payment is made since a MVM will be treated as a pseudo long-term liability until it is released. In other words, without the up-front profit, no profit would be recognized while considerable capital is required at the inception of the business; therefore, there would be little incentive to enter this business. We also charge 10% of the expected annuity payment as the up-front profit. The additional loading on the premium for this up-front profit is calculated to be 10% of $\Pi$, which is 24,482.

Including these charges, the gross premium $G$ is calculated as 338,686 as shown in Table 7-2.

<table>
<thead>
<tr>
<th></th>
<th>Margin on the premium</th>
<th>Expected Present Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net premium</td>
<td>244,816</td>
<td>5,279,511</td>
</tr>
<tr>
<td>Expense loading</td>
<td>+44,908</td>
<td>968,448</td>
</tr>
<tr>
<td>MVM loading</td>
<td>+24,482</td>
<td>527,951</td>
</tr>
<tr>
<td>Up-front profit loading</td>
<td>+24,482</td>
<td>527,951</td>
</tr>
<tr>
<td>Gross premium</td>
<td>338,686</td>
<td>7,303,861</td>
</tr>
</tbody>
</table>

The MVM loading and the up-front profit loading generate profits for the business unit, while other premium components offset the benefit payments and expenses. The generated profits are expressed as positive EPVs. On the above figures, the total EPV is
1,055,902 when the EPV of the benefit payments and expenses is included. Out of that, 527,951 is recognized as a profit immediately, and the remaining 527,951 will be gradually discharged and recognized as profits as benefits are actually paid. Mathematically:

\[ \sum_{t=0}^{70} P(0,t) \cdot EACF(t) = 1,055,902 , \]

and also,

\[ \sum_{t=0}^{70} P(0,t) \cdot MVM(t) = -527,951 \]

where \( EACF(t) \) is the expected actual cash flows to be paid at time \( t \) and \( MVM(t) \) is a MVM levied on the benefit payment at time \( t \).

The asset-liability mismatch arises between the MVM and the MVM loading. This mismatch arises from the difference between the timing when a MVM is released and the timing when MVM loadings give rise to cash. A MVM will not be released until payment dates while we earn cash every time premiums are paid long before the payment dates. Therefore, the cash flows arising from the MVM loadings are subject to investment risk. In other words, a MVM needs to be treated as a fixed-amount liability in the risk calculation.

Figure 7-1 shows the expected cash flows (“Contractual Cash Flows”, hereafter), including the MVM.
7.2 Profit emergence under cost accounting and fair value accounting

In this section, we will compare the emergence of profit arising from the annuity portfolio under cost accounting and fair value accounting. For this purpose, we assume that premiums are invested in cash and there are no dividend payouts. In both accounting methods, the cash accumulates from time $t-1$ to time $t$, growing by $1/P(t-1,t)$. Under cost accounting, we assume that the liability values are obtained by discounting the expected cash flows by the forward rates implied in the initial term structure, i.e., at time $t$, the cash flow at time $T$ is discounted by $P(0,T)/P(0,t)$ whatever value $P(t,T)$ takes. On the other hand, under fair value accounting, we use $P(t,T)$ to discount the cash flow.

We have chosen scenario 5 as a scenario for the development of the interest rate term structure. The reason we have chosen this scenario is that the scenario exhibits reasonably but not excessively volatile development of the term structure. Figure 7-2
shows how the interest rates develop as time passes.

**Figure 7-2 · Interest rate term structure development (Scenario 5)**

The liability values under cost accounting and fair value accounting are different. The latter is subject to the fluctuation of the yield curve as well as the interest earned on the cash. The former, however, is influenced only by the interest earned on the cash. Figure 7-3 shows the liability values under both accounting methods.

**Figure 7-3 · Development of liability values under both accounting methods**
The profit per year for each accounting method is calculated based on the liability value shown in the above figure. Under cost accounting, we assume that the 527,951 of up-front profit is deferred and spread over the lifetime of the contract. In accounting terminology, the up-front profit at inception is reserved and is amortized evenly each year until the end of the contract term, to be realized as revenue. Table 7-3, Table 7-4, and Figure 7-4 show the yearly profits under this interest rate scenario under both methods.
Table 7-3 - Yearly profits for both accounting methods

<table>
<thead>
<tr>
<th>Year</th>
<th>Cost accounting yearly</th>
<th>Fair value accounting yearly</th>
<th>Year</th>
<th>Cost accounting yearly</th>
<th>Fair value accounting yearly</th>
<th>Year</th>
<th>Cost accounting yearly</th>
<th>Fair value accounting yearly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-Front</td>
<td>0</td>
<td>527,949</td>
<td>1</td>
<td>9,473</td>
<td>240,258</td>
<td>26</td>
<td>-19,487</td>
<td>694,685</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>11,881</td>
<td>92,229</td>
<td>27</td>
<td>224,869</td>
<td>-748,133</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>14,621</td>
<td>136,121</td>
<td>28</td>
<td>-7,363</td>
<td>965,833</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>31,704</td>
<td>103,176</td>
<td>29</td>
<td>204,515</td>
<td>-606,885</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>28,389</td>
<td>-378,701</td>
<td>30</td>
<td>88,865</td>
<td>-650,359</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>14,392</td>
<td>-105,603</td>
<td>31</td>
<td>-16,495</td>
<td>-1,054,936</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>7,901</td>
<td>379,503</td>
<td>32</td>
<td>-124,218</td>
<td>-337,649</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>36,941</td>
<td>-124,355</td>
<td>33</td>
<td>-166,649</td>
<td>549,925</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>11,470</td>
<td>-560,147</td>
<td>34</td>
<td>-93,978</td>
<td>346,734</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>-31,206</td>
<td>-34,550</td>
<td>35</td>
<td>-28,149</td>
<td>-653,884</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>-44,422</td>
<td>899,553</td>
<td>36</td>
<td>-132,533</td>
<td>85,051</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>-6,678</td>
<td>84,193</td>
<td>37</td>
<td>-127,845</td>
<td>249,339</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>-31,028</td>
<td>-533,745</td>
<td>38</td>
<td>-93,479</td>
<td>179,215</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>-39,836</td>
<td>-588,238</td>
<td>39</td>
<td>-52,812</td>
<td>869,112</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>-80,724</td>
<td>-511,804</td>
<td>40</td>
<td>123,830</td>
<td>374,988</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>-103,466</td>
<td>729,972</td>
<td>41</td>
<td>273,944</td>
<td>-809,216</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td>-85,598</td>
<td>-389,114</td>
<td>42</td>
<td>14,732</td>
<td>45,493</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>-79,323</td>
<td>29,430</td>
<td>43</td>
<td>16,295</td>
<td>188,902</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td>-39,073</td>
<td>-410,359</td>
<td>44</td>
<td>92,231</td>
<td>-348,355</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>-70,487</td>
<td>-1,330,676</td>
<td>45</td>
<td>-13,298</td>
<td>423,794</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>21</td>
<td>-158,051</td>
<td>-1,154,628</td>
<td>46</td>
<td>86,357</td>
<td>114,333</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td>-182,232</td>
<td>442,089</td>
<td>47</td>
<td>83,186</td>
<td>164,473</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>23</td>
<td>-172,208</td>
<td>1,238,376</td>
<td>48</td>
<td>63,479</td>
<td>11,187</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td>-27,422</td>
<td>-601,751</td>
<td>49</td>
<td>33,403</td>
<td>134,966</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>-86,905</td>
<td>655,278</td>
<td>50</td>
<td>51,188</td>
<td>116,514</td>
</tr>
</tbody>
</table>
Table 7-4 - Basic statistics of the yearly profits

<table>
<thead>
<tr>
<th></th>
<th>Cost accounting yearly profit</th>
<th>Fair value accounting yearly profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total:</td>
<td>-268,325</td>
<td>-268,325</td>
</tr>
<tr>
<td>Average:</td>
<td>-3,779</td>
<td>-3,779</td>
</tr>
<tr>
<td>Standard deviation:</td>
<td>81,843</td>
<td>456,542</td>
</tr>
</tbody>
</table>

Figure 7-4 - Yearly profits under both accounting methods

As can be seen from Table 7-4, even though the total profits are same for both methods, the profits under fair value accounting are much more volatile than under cost accounting. The standard deviation of the yearly profits suggests that yearly profit under fair value accounting is, on average, more than 5 times as volatile as that under cost accounting. We have chosen a reasonably realistic interest rate development scenario. It means that such a degree of fluctuation of yearly profits is quite possible under fair value accounting.
7.3 Assumptions on the behaviour of the business unit

From now on, we assume that the business unit acts like an independent firm. In this section, we define the structure of the business unit and how the capital requirement is calculated. The business starts by having the initial capital injected at inception. The dividends will be paid or additional capital injections will be made when necessary at the end of each period over the whole lifetime of the business. The business unit will be dissolved when all the annuity contracts expire. Following the closed book principle, the business is not allowed to enter into any new contracts until it is dissolved.

The adequacy of the capital of the business unit is monitored by a minimum capital and a target capital. We assume that the minimum capital is set to be equal to the regulatory capital. We assume some regulatory regime (discussed in the next section) on capital requirements for our hypothetical business, and that the business unit complies with this regulatory capital requirement. Here we define the actual capital to be the net asset value held by the business unit, i.e. the difference between the market value of assets and the fair value of liabilities held by the business on the spot then. The actual capital shows what the real capital position of the business is now while the minimum and the target capital show the capital which the business unit needs to keep. When the actual capital is less than the minimum capital, the business unit needs to replenish its capital at least to the minimum capital level immediately, or the business needs to be dissolved immediately. The target capital is the capital amount which the owner of the business unit deems an adequate level. We apply a predetermined formula to determine the target capital amount, and we discuss the regulatory and target capital in detail in the next section.
At times 0,1,2,3,…, the cash balance of the business unit changes due to benefit and expense payments and additional equity injections or dividend payments. The resulting cash balance is invested into short or long-term bonds, or equities. We now set out rules for the investments as follows:

- The fund backing the regulatory capital is allowed only to be invested in safe assets which are cash or short bonds whose term to expiry is at most one year.
- We call the excess cash beyond the regulatory capital “free cash”. Free cash will be invested in equities or bonds. We fix the proportion between the amounts invested into bonds or equities; for example, 80% for bonds and 20% for equities. In this case, the ratio between the market value of the equity portfolio and the bond portfolio at the time of allocation should be 80% to 20%. Note that this ratio changes according to the market. However, the ratio is readjusted to 80% to 20% at the next allocation date.
- The interest rate risk can be hedged using bonds. The bond positions arising from this hedging activity are not included in the predetermined asset allocation proportion between bonds and equities.

All of the assets of the business unit, including cash that is obtained as a capital injection, are rebalanced at the end of every period so that the above rules are strictly adhered to at the very beginning of the next period. We assume that we can execute any amount of transactions including fractions without any costs.

7.4 The regulatory and target capital
Under the fair valuation regime, the regulatory capital is expected to be risk-based, and will be determined using some risk measures. At present, Value at Risk (VaR) is the
most commonly used risk measure, in particular among non-insurance financial institutions. VaR is a quantile-based risk measure, which is defined to be the amount more than which we will not lose with some (small) probability in some fixed time horizon.

Despite its popularity, VaR has recently been criticized as a defective risk measure by, for example, Dowd and Blake (2006). It is well known that, in particular for insurance risks, the events that incur extreme losses happen with much higher probabilities than the normal distribution would suggest. In other words, loss distributions in the real world tend to have fat tails\(^{65}\). However, VaR ignores the magnitude of all the tail events beyond the VaR level; this is the primary reason why VaR is criticized.

Dowd and Blake (2006) recommend other coherent\(^{66}\) quantile-based risk measures such as Conditional Tail Expectation (CTE). CTE is the average of the losses that exceed some predetermined percentile, and therefore it is a coherent measure. CTE has already been chosen as an appropriate risk measure by insurers and their regulators in some countries such as the United States; and it is expected to gain in popularity in the next-generation risk regimes globally.

\(^{65}\) The mean excess function \(e(u)=E[X-u|X>u]\) is often used to assess how fat the tail of the distribution of a random variable \(X\) is. If the tail is lighter than that of any exponentially bounded distribution (one example is the normal distribution), \(e(u)\) is decreasing. If the tail is heavier than that of any exponential (examples are the lognormal or Pareto distribution), \(e(u)\) is increasing. See, for example, Rolski et al. (1998).

\(^{66}\) A risk measure \(m(X)\) is called a coherent risk measure if this measure satisfies the following four conditions for any two risky positions \(X\) and \(Y\) with values given by \(V(X)\) and \(V(Y)\):

\[V(Y) \geq V(X) \Rightarrow m(Y) \leq m(X), m(X + Y) \leq m(X) + m(Y), m(hX) = hm(X) \text{ for } h > 0, \text{ and } m(X + n) = m(X) + n \text{ for a certain amount } n.\]

See, for example, Dowd and Blake (2006).
Although we admit that CTE is a better risk measure than VaR, we still use VaR as the risk measure in our study for the following practical reasons.

- We are not using a model that produces fat tails since the primary objective of our study is a comparison between business strategies, not the exact replication of the behaviour of asset prices in the real world.

- CTE is calculated predominantly by a simulation method. However, we cannot use a simulation method for calculating risk measures in this study. We need to calculate the risk measure at each year end until all contracts expire, i.e. we need to calculate risk measures 70 times per one market scenario. For 5,000 market scenarios, we need to calculate approximately 350,000 risk measures in total. If we were to try to calculate each risk measure by simulation using 5,000 sample paths, we would need 1,750,000,000 sample paths in total. This is not feasible.

- Therefore we need to rely on some approximate methods in calculating each risk measure. The approximation we make is that we assume that we already know the shape of the distribution of the change of the portfolio value over one year. Under this assumption, once we know the parameters of the distribution, risk measures are obtained as a function of these parameters and there is equivalence between the CTE calculation and VaR calculation. For example, the 95% CTE of a random variable, which follows the normal distribution with mean 0 and standard deviation 1, is $0.418$. Since the 95% percentile of this distribution is 1.645, the risk measure is calculated to be $2.063 (0.418 + 1.645)$. Since the probability that this random

\[ e(u) = \sigma \Phi \left( \frac{u - \mu}{\sigma} \right) \left( 1 - \Phi \left( \frac{u - \mu}{\sigma} \right) \right) - (u - \mu) \]  

where \( \Phi(\cdot) \) is a cumulative probability function of a standard normal distribution.

\[ \text{For a random variable that follows the } N(\mu, \sigma) \text{ distribution,} \]

176
variable exceeds 2.063 is 2%, 2.063 is also the 98% VaR. In other words, 95% CTE is same as the 98% VaR in this case.

For these practical reasons, we use VaR for the regulatory capital calculation in this study.

The time horizon and the required level of probability need to be specified in order for VaR requirements to be defined. We will discuss these in Section 7.5. Given that the probability is set to the level which the regulator is comfortable with, we assume that the regulatory capital requirement for the business unit is VaR of all assets and liabilities held by the business unit.

We next define the target capital using a concept called the Solvency Margin Ratio ("SMR", hereafter). The regulatory capital is a threshold below which the actual capital should never go. Therefore, the target capital should be set so that some cushion beyond the regulatory capital is maintained. For this reason, we define the target capital to be the regulatory capital multiplied by some constant which is greater than 1. We call this constant our target SMR, which is different from the legal SMR which the Japanese Financial Services Agency ("JFSA", hereafter) stipulates in the Japanese Insurance Act. We use the term SMR because the SMR represents, in principle, the ratio of a net asset value to a regulatory capital requirement; and the SMR is a familiar concept in the Japanese life insurance industry.

The legal SMR is the key figure by which the JFSA measures how solvent insurers are, i.e. how adequately insurers hold funds for expected payments. The JFSA orders corrective actions by any insurer whose legal SMR goes under 200%. The legal SMR is
calculated according to the following formula:

\[ \text{Legal SMR} = \frac{\text{Total Solvency Margin Amount}}{\text{Risk Amount}} \times 2. \]

The Total Solvency Margin Amount is in essence the actual capital of the insurer and the Risk Amount is in essence the risk-based regulatory capital requirement. However, since insurers’ current financial reports are based on cost accounting, current reported book capital cannot be used for this purpose, and the risk-based capital is not reported. Therefore, in order to measure the current solvency of insurers, both the actual capital amount that reflects current asset prices and a risk-based capital requirement need to be calculated. The Total Solvency Margin Amount is calculated by adding some unrecognized capital items, for example unrealized capital gains multiplied by some coefficients, to the book capital. The Risk Amount is calculated for several risk factors such as the insurance risk, the interest rate risk, market risk and operational risk by a prescriptive method. The JFSA does not explicitly state that the Risk Amount is a regulatory capital requirement. However, the fact that the JFSA orders corrective actions when the legal SMR goes under 200% indicates that the Risk Amount is a de facto minimum regulatory capital requirement.

In summary, in essence:

\[ \text{Legal SMR} = \frac{\text{Actual capital}}{\text{Risk-based regulatory capital requirement}} \times 2. \]

According to the statistics\(^{68}\) from the JFSA as of November 2006, all life insurance companies operating in Japan now keep a legal SMR more than 400%; and the simple average of SMR of the life insurance companies is 1,607%. Table 7-5 shows the

\(^{68}\) http://www.fsa.go.jp/singi/solvency/siryou/20061120/01-04.pdf
distribution of the legal SMR of life insurance companies operating in Japan.

Table 7-5 · Legal SMRs

<table>
<thead>
<tr>
<th>Legal SMR range</th>
<th>Number of companies</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>From</td>
<td>To</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>400</td>
<td>0</td>
</tr>
<tr>
<td>400</td>
<td>600</td>
<td>3</td>
</tr>
<tr>
<td>600</td>
<td>800</td>
<td>3</td>
</tr>
<tr>
<td>800</td>
<td>1000</td>
<td>4</td>
</tr>
<tr>
<td>1000</td>
<td>1200</td>
<td>10</td>
</tr>
<tr>
<td>1200</td>
<td>1400</td>
<td>2</td>
</tr>
<tr>
<td>1400</td>
<td>1600</td>
<td>6</td>
</tr>
<tr>
<td>1600</td>
<td>1800</td>
<td>2</td>
</tr>
<tr>
<td>1800</td>
<td>2000</td>
<td>2</td>
</tr>
<tr>
<td>2000</td>
<td>3000</td>
<td>4</td>
</tr>
<tr>
<td>3000</td>
<td>5000</td>
<td>1</td>
</tr>
<tr>
<td>5000</td>
<td>1</td>
<td>2.6%</td>
</tr>
</tbody>
</table>

The way of calculating the legal SMR is being updated as the technology for quantifying risks develops. It is still unclear what definition of SMR will be applied under the fair valuation regime.

In our study, we ignore the coefficient 2 that is used in the definition of the legal SMR, and then we use the term SMR, as intuitively as possible, as a synonym of a ratio of an actual capital to regulatory capital requirement. Then our SMR is defined to be:
Our SMR = \frac{\text{Actual capital of the business unit}}{\text{The regulatory capital requirement}}.

We assume that the business unit sets a target SMR level. The target capital is the capital amount where the target SMR is satisfied. Suppose the target SMR is set to be 200% (which is largely equivalent to a 400% legal SMR). When the minimum capital requirement is, for example, 10 million, the target capital is 20 million. The target capital changes to 14 million if the minimum capital requirement changes to 7 million. From now on, we use the term SMR for the target SMR.

In our simulation, we consider conservative, standard and aggressive strategies. In the conservative strategy, the SMR is 300%. The SMR is 200% on the standard strategy and 110% on the aggressive strategy.

The choice of target SMR is a trade-off between the security and efficiency of the business unit. The higher SMR enhances the credibility of the business unit, but reduces the return on capital on the business unit. On the aggressive strategy, the capital is more effectively used than other strategies, but the capital providers are expected to suffer from frequent additional capital injection requests. On the other hand, on the conservative strategy, the additional capital requests are expected to be much less frequent than those on the aggressive strategy. However, the return on capital would be much more diluted than that on the aggressive strategy. Therefore, we compare the results produced on the three strategies to find which strategy is most desirable.

7.5 VaR calculation method

In order to calculate the VaR, we need to fix the time horizon and the required level of probability. We use one year as the time horizon and 99.86% as the probability. We have chosen one year because a major review of the balance sheet structure commonly takes
place yearly at the beginning of the fiscal year; and because it usually takes several months to complete raising capital. The 99.86% probability was chosen so that the default probability of the business unit is equivalent to the default probability implied by a corporate bond rated AAA. According to Hull (1999), the average cumulative default rate of a AAA-rated bond is 1.4% over ten years. The default probability per year is calculated to be \(1 - (1 - 0.014)^{0.1} \approx 0.0014\); therefore we use 99.86% as the probability for our VaR calculation; and we call the VaR a one-year 99.86% VaR.

If the distribution of the change of the value of all the assets and liabilities including cash over one year is normally distributed, the one-year 99.86% VaR is obtained as:

\[
3.0 \times \sigma_p,
\]

where \(\sigma_p\) is the standard deviation of the change of the value of all the assets and liabilities over one year. This is because the probability point that gives 99.86% in the normal distribution is 2.99, and the expected change is conventionally assumed to be zero in VaR calculations. In fact, as the expected asset returns are largely offset by expected liability growth, the expected return for the net position is small. In summary, what we need to know in order to calculate VaR is the standard deviation of the change of the value of all the assets and liabilities under the assumption that this change follows the normal distribution,

At time \(t\), our model gives instantaneous volatilities for all assets and liabilities and instantaneous covariances between them. For small \(\Delta t\), for example 10 days, the probability distribution of the change of the portfolio value can be approximated by a

\(^{69}\text{See, for example, Hull (1999) Chapter 14.}\)
normal distribution with the standard deviation obtained using these volatilities and covariances. Let \( \Delta P(t, T_i) = P(t + \Delta t, T_i) - P(t, T_i) \). We have modelled
\[
dP(t, T) / P(t, T) = \text{(some drift)} dt + S(t, T) d\hat{W}(t)
\]
. Therefore the covariance between \( \Delta P(t, T_i) / P(t, T) \) and \( \Delta P(t, T_j) / P(t, T) \) is \( S(t, T_i) \cdot S(t, T_j) \Delta t \) for small \( \Delta t > 0 \) for any \( i, j = 0, 1, 2, \ldots \). This gives the covariance matrix whose \( ij \)-th element shows the covariance between the proportional\(^70\) change of value of each bond. We add one row and column that shows the covariances between the proportional change of the value of our equity portfolio and that of the other bond. Then we get the covariance matrix between all bonds and equities. We let \( \Sigma \cdot \Delta t \) denote this covariance matrix when the state variables \( X(0), X(1) \) and \( X(2) \) are 0, which is the assumed convergence value of these variables. We show some selected entries of \( \Sigma \) and corresponding correlation matrix in Table 7-6 and Table 7-7. The tables show reasonable features that we observe in real markets. The absolute yield volatilities are decreasing gradually as the term increases; the absolute yield volatilities for the 1-year, 10-year and 20-year terms are calculated\(^71\) to be 1.0\%, 0.85\% and 0.56\% respectively. In addition, the correlation between any two bonds decreases as the difference between the terms of these bonds increases.

\[^70\]The “proportional” means the change is measured relative to its price. On the other hand, the “absolute” means the change amount itself. If the change is +1 and the price was 90, the absolute change is +1 and the proportional change is 11.1%.

\[^71\]The absolute yield volatility for a term is approximated by dividing the proportional volatility of the bond for the term by the duration of the bond. See, for example, Hull (1999) Page 536.
Table 7-6 · Covariance matrix of selected bonds and equity portfolio

<table>
<thead>
<tr>
<th>Years to maturity of zero coupon bonds</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>70</th>
<th>100</th>
<th>Equity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.011%</td>
<td>0.022%</td>
<td>0.050%</td>
<td>0.082%</td>
<td>0.105%</td>
<td>0.106%</td>
<td>0.100%</td>
<td>0.097%</td>
<td>0%</td>
</tr>
<tr>
<td>2</td>
<td>0.022%</td>
<td>0.043%</td>
<td>0.103%</td>
<td>0.170%</td>
<td>0.218%</td>
<td>0.220%</td>
<td>0.209%</td>
<td>0.202%</td>
<td>0%</td>
</tr>
<tr>
<td>5</td>
<td>0.050%</td>
<td>0.103%</td>
<td>0.253%</td>
<td>0.425%</td>
<td>0.549%</td>
<td>0.556%</td>
<td>0.530%</td>
<td>0.514%</td>
<td>0%</td>
</tr>
<tr>
<td>10</td>
<td>0.082%</td>
<td>0.170%</td>
<td>0.425%</td>
<td>0.721%</td>
<td>0.940%</td>
<td>0.966%</td>
<td>0.929%</td>
<td>0.906%</td>
<td>0%</td>
</tr>
<tr>
<td>20</td>
<td>0.105%</td>
<td>0.218%</td>
<td>0.549%</td>
<td>0.940%</td>
<td>1.247%</td>
<td>1.319%</td>
<td>1.298%</td>
<td>1.278%</td>
<td>0%</td>
</tr>
<tr>
<td>40</td>
<td>0.106%</td>
<td>0.220%</td>
<td>0.556%</td>
<td>0.966%</td>
<td>1.319%</td>
<td>1.461%</td>
<td>1.490%</td>
<td>1.488%</td>
<td>0%</td>
</tr>
<tr>
<td>70</td>
<td>0.100%</td>
<td>0.209%</td>
<td>0.530%</td>
<td>0.929%</td>
<td>1.298%</td>
<td>1.490%</td>
<td>1.558%</td>
<td>1.571%</td>
<td>0%</td>
</tr>
<tr>
<td>100</td>
<td>0.097%</td>
<td>0.202%</td>
<td>0.514%</td>
<td>0.906%</td>
<td>1.278%</td>
<td>1.488%</td>
<td>1.571%</td>
<td>1.590%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Equity: 0% 0% 0% 0% 0% 0% 0% 0% 4.0%

Table 7-7 · Correlation matrix of selected bonds and equity portfolio

<table>
<thead>
<tr>
<th>Years to maturity of zero coupon bonds</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>70</th>
<th>100</th>
<th>Equity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.994</td>
<td>0.959</td>
<td>0.929</td>
<td>0.900</td>
<td>0.839</td>
<td>0.772</td>
<td>0.739</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>0.994</td>
<td>1.000</td>
<td>0.985</td>
<td>0.963</td>
<td>0.938</td>
<td>0.875</td>
<td>0.804</td>
<td>0.770</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>0.959</td>
<td>0.985</td>
<td>1.000</td>
<td>0.995</td>
<td>0.976</td>
<td>0.915</td>
<td>0.844</td>
<td>0.810</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>0.929</td>
<td>0.963</td>
<td>0.995</td>
<td>1.000</td>
<td>0.991</td>
<td>0.941</td>
<td>0.877</td>
<td>0.846</td>
<td>0.0</td>
</tr>
<tr>
<td>20</td>
<td>0.900</td>
<td>0.938</td>
<td>0.976</td>
<td>0.991</td>
<td>1.000</td>
<td>0.977</td>
<td>0.931</td>
<td>0.907</td>
<td>0.0</td>
</tr>
<tr>
<td>40</td>
<td>0.839</td>
<td>0.875</td>
<td>0.915</td>
<td>0.941</td>
<td>0.977</td>
<td>1.000</td>
<td>0.988</td>
<td>0.976</td>
<td>0.0</td>
</tr>
<tr>
<td>70</td>
<td>0.772</td>
<td>0.804</td>
<td>0.844</td>
<td>0.877</td>
<td>0.931</td>
<td>0.988</td>
<td>1.000</td>
<td>0.998</td>
<td>0.0</td>
</tr>
<tr>
<td>100</td>
<td>0.739</td>
<td>0.770</td>
<td>0.810</td>
<td>0.846</td>
<td>0.907</td>
<td>0.976</td>
<td>0.998</td>
<td>1.000</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Equity: 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0

The covariance between \( \Delta P(t,T_j)/P(t,T_j) \) and \( \Delta P(t,T_j)/P(t,T_j) \), strictly, depends on \( X(0), X(1) \) and \( X(2) \). However, we have found by trial calculations that the change of volatilities with changes of \( X(0), X(1) \) and \( X(2) \) is insignificant. In addition, the constant variance assumption saves significant calculation load and time. Therefore we
use a constant matrix $\Sigma$ for our simulation.

Using this covariance matrix, the standard deviation of the absolute change of the value of the portfolio is calculated as $\sqrt{V^T \Sigma V} \sqrt{\Delta t}$ where $V$ is the vector of the current value of the asset and liability components. We let $\sigma_{\Delta t}$ denote this standard deviation.

We have obtained the standard deviation of the absolute change of the value of all assets and liabilities over a short period $\Delta t > 0$ . However, what we need is the standard deviation of this value over one year. Here, solely for the purpose of calculating one-year 99.86% VaR, we assume that the probability distribution of the absolute change of the value of the portfolio over one year follows a normal distribution with standard deviation:

$$\sigma_{\Delta t} \sqrt{\frac{1}{\Delta t}}.$$  \hspace{1cm} (7.2)

This approximation is reasonable for the following reasons:

1. The bond and equity prices are modelled as Ito processes with very smooth drifts and volatilities. This means that extreme skew is not possible in a short time, i.e. the probability distribution over one year is expected to be close to normal.

2. The factor $\sqrt{1/\Delta t}$ multiplied to $\sigma_{\Delta t}$ implies that the subsequent changes of the value of the portfolio are independent of the past history of the value and of the current value. This is true for our model for equities, but is not strictly true for bonds since interest rates exhibit a mean-reversion feature. It takes, however, several years for mean-reversion to cause perceivable effects on the bond prices. For example, if the instantaneous short rate $r(t)$ follows the Vasicek (1977) model $dr(t) = a(b - r(t)) + \sigma dW(t)$, the standard deviation of
\( r(T) \) given \( r(0) \) is \( \sigma \sqrt{\frac{1 - e^{-2aT}}{2a}} \), which is approximately \( \sigma \sqrt{T - aT^2} \), which is \( \sigma \sqrt{T} \sqrt{1 - aT} \), when \( aT \) is small. Even when \( T = 1 \), this is approximately \( \sigma \sqrt{T} \) unless \( a \) is significantly large.

3. Volatilities of bond prices decrease as time passes due to a shortening of time to expiry. The effect of this decrease is small enough to be ignored for a one year time horizon. For example, suppose the current yield of a zero coupon bond is 5\%, and the volatility of the absolute change of the yield of this bond is 1\%. Using the formula \( \Delta P(t,T) / P(t,T) \approx -(T - t) \Delta y \) where \( \Delta y \) is the change of the yield of this bond, the volatility of the proportional change of the price of this bond is 30\% when \( T - t \) is 30 years. Suppose one year has passed. Then the term to expiry \( T - t \) becomes 29 years and volatility of the bond price becomes 29\%, which is only 1\% smaller than that one year ago.

4. Our approximation results in a slightly larger standard deviation than what the exact calculation, which considers the mean-reversion and bond volatility decrease, would produce. Therefore, our approximation leads to higher VaR than what an exact calculation would produce, and hence stricter capital requirements.

In summary, the VaR is calculated as follows.

Applying (7.2), the standard derivation of the change of the value of the portfolio over one year is given by \( \sqrt{V^T \Sigma V} \), and the one-year 99.86\% VaR is given as:

\[
3 \sqrt{V^T \Sigma V},
\] applying (7.1). For example, when we have only two zero coupon bonds that expire at
times $T_i$ and $T_j$, the standard deviation of the change of this simple portfolio over one year is obtained as

$$\sqrt{[m_1 P(t,T_i)] [S(t,T_i) \cdot S(t,T_j)] [m_2 P(t,T_j)]}$$

when we hold $m_1$ units of bonds that expires at time $T_i$ and $m_2$ units of bonds that expire at time $T_j$. The one-year 99.86% VaR is given by

$$3\sqrt{[m_1 P(t,T_i)] [S(t,T_i) \cdot S(t,T_j)] [m_2 P(t,T_j)]}.$$ 

The accuracy of this approximation will be examined in Section 7.8.3.

Table 7-8 shows the standard deviation of the present value of Contractual Cash Flows over one year using $\Sigma$, VaR that is calculated by applying (7.3), and the target capital when the SMR is 200%.

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation</td>
<td>220,311</td>
</tr>
<tr>
<td>VaR ( = Minimum capital)</td>
<td>660,934</td>
</tr>
<tr>
<td>Target capital</td>
<td>1,321,869</td>
</tr>
</tbody>
</table>

Note that investment risks are not considered in the calculation above. In the actual simulation, the VaR and capital requirements change all the time according to any changes in cash flows and asset constituents.

### 7.6 Investment, risk-hedging and dividend strategy and interim capital injection

We now set strategies that the business unit will adopt on investment, risk hedging, dividends and interim capital injection. The comparison from investment strategies, in particular, is essentially important for us to find the most efficient way of using the
capital.

Investment strategy:
Each year we rebalance the portfolio, i.e. we liquidate all the assets and reallocate the excess cash into assets after the annuity, expense and dividend payments are made. Recall that the statutory fund, which backs the regulatory capital, is required to be held in cash or short-term bonds. The cash that remains after the statutory fund is reserved from the excess cash is called the free cash. We consider three investment strategies in our simulation: aggressive, standard and passive strategy that are determined by the equity investment ratio (“EIR”, hereafter). EIR shows what percentage of the free cash is invested into equities. The EIR of the aggressive strategy is 50%, i.e. we invest 50% of the free cash into equities. The EIR of the standard strategy is 25%, and the EIR of the passive strategy is 0%. The remaining cash after the equity investment is allocated to a zero coupon bond that expires in 20 years. If the 20 year date exceeds the contract’s expiration date, we choose the zero coupon bond that expires at the contract’s expiration date. We have chosen 20 years as the maximum bond investment term because 20 years is the maximum expiry date for Japanese government bonds that are practically available in the market.

Risk hedging:
We consider two cases: one with risk-hedging and the one without it. In the risk-hedging case, we minimize the interest rate risk by buying or selling a zero coupon bond (“hedge bond” hereafter) that expires in 20 years on a one-year forward settlement basis. This transaction is executed independently of the investment transactions. The hedge transactions are forward-settled so that the hedge transactions only change the
risk profile of the business unit without affecting the cash position of it. As is the case with investment, we choose the zero coupon bond that expires at the contract’s expiration date if the 20 year date exceeds the contract’s expiration date.

We can easily calculate how much we should hold of the hedge bond at a specific year end on a specific scenario. Suppose the asset allocation process has finished and we have fixed how much we will invest in a bond and the equity portfolio. Adding the expected cash flows arising from the annuity portfolio, we know a column vector whose \(i\)-th element is the expected present value of the cash flow that arises at time \(T_i\). We let \(V\) denote this column vector. The total variance of the change of the value of the asset and liabilities is obtained as \(V^T \Sigma V\). Suppose \(T_j\) corresponds to the expiry date of the hedge bond. Consider a column vector denoted by \(V(x)\) whose elements are the same as those of \(V\) apart from the \(j\)-th element, which is a variable \(x\). The solution \(x^*\) which minimizes the variance \(V(x)^T \Sigma V(x)\) can easily be found. See Appendix II for details. The cash amount to be invested into the hedge bond is obtained as \(x^*-V_j\), where \(V_j\) is the \(j\)-th element of \(V\).

**Dividend Strategy:**

We simulate full, aggressive and standard dividend strategies. We define the excess capital to be the portion of the actual capital that exceeds the target capital. On the full dividend policy, at the end of each year, we pay out 100% of the excess capital to the capital contributors. On the standard dividend policy, we pay out 75% of the excess capital, and we pay out 50% of the excess capital on the standard dividend policy. The dividend payment takes place after annuity and expense payments are made. Therefore, some portion of the planned dividend may not be paid due to a shortage of cash.
In summary, we consider 54 combinations of above strategies as shown on Table 7-9.

<table>
<thead>
<tr>
<th>Number of cases</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMR</td>
<td>3 110%, 200%, or 300%</td>
</tr>
<tr>
<td>Dividend payout ratio</td>
<td>3 50%, 75%, or 100%</td>
</tr>
<tr>
<td>Equity investment ratio</td>
<td>3 0%, 25%, or 50%</td>
</tr>
<tr>
<td>Hedging Y/N</td>
<td>2 Yes or No</td>
</tr>
</tbody>
</table>

**Interim capital injection:**

In addition to the combinations above, we consider two cases when we encounter a capital shortage. The actual capital may fall short of the minimum capital requirement. This happens when the VaR increases as well as when the asset or liability prices suffer from an adverse movement. In one case (“injection option” hereafter), we raise the capital up to the target capital level in the same way as we did at the inception of the business. In another case (“liquidation option” hereafter), we liquidate all the assets and liabilities and terminate the business. We assume that we can transfer assets and liabilities at our evaluation prices.

Note that a capital injection may also happen if the business unit does not have enough cash to make annuity and expense payments, even when the actual capital is above the minimum capital.

**7.7 The process to determine capital inflow or outflow amounts**

The dividend payout or additional capital injection changes the amount of free cash
which can be invested in a bond or an equity portfolio. The VaR and the regulatory capital requirement change accordingly and the change of the regulatory capital produces some additional capital requirement or surplus. This process continues until the total capital inflow or outflow amount converges to some level. At the end of each year, we find the converged capital inflow or outflow amount using the following process.
At the end of each year, we start this process by calculating VaR based on the latest position of assets and liabilities assuming that this position remains intact in the next year. Once we have the VaR assumption, we can calculate the regulatory capital and target capital amount for the next year. We can also determine the divided or capital injection amount since we know all the capital amounts, i.e. the regulatory, target and actual capital, at this stage. Then, we liquidate all the assets into cash and pay dividends or receive a capital injection. The resulting total cash is allocated to each asset category in accordance with the predetermined rules. Based on the new position of the assets, we recalculate the VaR and compare this revised VaR with the VaR we have calculated at the start of this cycle. If the difference between these VaRs is significant, we estimate the new VaR by linear approximation and then we go back the beginning to repeat this process again until VaR values match.
7.8 Simulation

We investigate the risk-return profile of the business unit in this section. We have already generated 5,000 sample market scenarios over next 70 years in Chapter 6. Using these 5,000 market scenarios, we will simulate the development of the balance sheets, dividends and capital injections for each strategy. We will call a series of cash flows arising from the initial capital injection, dividends and additional capital injections a capital flow. We will analyze the capital flows in detail.

7.8.1 Precision of the simulation

The price of a zero coupon bond at time $t$ that expires at time $T$ is given by formula (5.22). Then we can calculate $P(0,T)$ for all $T>0$. At the same time, the zero coupon bond price $P(0,T)$ can also be calculated by simulation using the deflator $D(t)$ that is given by formula (6.4). In order to check the precision of the simulation, we compare prices of zero coupon bonds obtained using simulation with those obtained using formula (5.22).

Consider a zero coupon bond that expires at time $T$. This bond always pays 1 at time $T$. Therefore, the expected value at time 0 of the price of this bond discounted by the deflator $D(T)/D(0)$ should be $P(0,T)$ for $T>0$, i.e. $P(0,T) = E[D(T)/D(0)\times1] = E\left[D(T)/D(0)\right]$. Now let $D(t,j)$ denote the simulated deflator for time $t$ on the $j$-th simulation where $t_0=0, t_1=1, t_2=2,\ldots$, and $j=1,2,\ldots,5000$. Letting $\bar{D}(t_i)$ denote the simple average of $D(t_i,j)$ over $j$, we compare $\bar{D}(t_i)/D(0)$ to $P(0,t_i)$ for each $t_i$. Figure 7-6 shows the pricing errors, which are $\bar{D}(t_i)/D(0) - P(0,t_i)$. 

192
There are 70 bonds under investigation. Errors on most of the bond prices are within 1%, while errors above 1% are observed for 14 bonds. Errors on all bonds are within 2%. We believe that this order of error is not large enough to invalidate results based on the simulation. This is primarily because we are more interested in a comparison between strategies rather than the absolute figures. In our study, we will compare various strategies using the same 5,000 market scenarios. Therefore, when comparing results under strategy A and those under strategy B, the pricing errors under strategy B are expected to be close to the pricing errors under strategy A. That implies that the comparison is still valid even though the calculated values contain some errors, as long as our comparisons are subject to similar errors. The second reason why we use 5,000 sample paths is that we think 5,000 is a practical figure. It is not feasible to reduce the errors dramatically by increasing the number of simulations, since our simulation requires significant calculations including a large number of numerical integrations. In addition, note that the errors are actually smaller than it appears in terms of figures per
annum. For example, an error of 2% on a bond that expires in 50 years involves 4 basis point (=0.04%) error per annum.

The difference between the present value of Expected Actual Cash Flows (EACF) calculated by simulation using the 5,000 market scenarios and that using formula (5.22) is small. We have shown in section 7.1 that

\[ \sum_{i=0}^{70} P(0, t_i) \cdot EACF(t_i) = 1,055,902 \]

where \( EACF(t_i) \) is the actual cash flow expected at \( t_i \). By simulation, the expected value is calculated to be

\[ \sum_{i=0}^{70} \bar{D}(t_i)/D(0) \cdot EACF(t_i) = 953,226 \].

The pricing error is -102,675, which is about 10% of the present value of EACF. This figure appears large, but we need to take into account the duration of the business and the fact that this figure is the net figure of assets and liabilities. This business continues for 70 years. Therefore the 10% difference is translated into a difference less than 0.03% per annum. This per-annum figure is so small that we could ignore the error. In addition, 1,055,902, which is the present value of EACF, is a net figure of 7,303,861, which is the present value of expected premium income, and 6,247,959, which is the present value of expected benefit and expense payments. The pricing error of 102,675 is approximately 1.5% of the present value of premium income or the present value of expected benefit or expense payments.

### 7.8.2 Simulation example

For each strategy and for each market scenario, at the end of each year, we conduct the following calculations.

- We evaluate all the assets and liabilities based on the position at the start of the current term.
- Then we calculate the current actual capital amount.
We calculate the dividend or capital injection amount and the VaR at the beginning of the next term, using the method we have explained in section 7.7. The regulatory and target capital amount are also obtained.

During the above process of finding the dividend amount, allocation of cash to each asset including the hedge bond is completed.

We will illustrate ideas for the strategy where the SMR is 200%, the dividend payout ratio is 50%, the share investment ratio is 25%, and there is hedging. For this illustration purpose, we use the market scenario number 1. For simplicity, we show all monetary figures in units of one thousand, adding the sign “K” after the figure.

At time 0, the business unit received a capital injection of 873K. The injected cash is, together with the first premium, used to fund the payment for the initial expenses. At the very beginning of the first year, i.e. at time 0+, the VaR is calculated to be 502K and that amount of cash is invested in a one-year bond as a backing asset for the regulatory capital. No cash remains for investment at this time. The hedging position is made to minimize the variance of the portfolio. Note that the VaR is determined so that the VaR based on the resulting portfolio at time 0+ is 502K. That means that the VaR, after all the cash payments including the capital injection at time 0 are made, is 502K. Evaluating all the assets and liabilities using the market prices at time 0, the actual capital is calculated to be 1,401K. The capital amount of 1,401K is the sum of the injected capital of 873K and the up-front profit of 528K.

At time 1, our actual capital has become 1,298K due to market movement. After receiving the premium paid at time 1 and paying expenses, dividends and VaR are
calculated to be 200K and 449K respectively. The target capital is 449K×2=898K, therefore the dividend amount was calculated to be (1,298K-898K)×50%=200K, since our dividend payout ratio is 50%. All assets and hedge positions are liquidated into cash, and then the cash is reallocated to assets. After 449K is allocated to the short bond as the backing asset for the regulatory capital, 75% of the remaining cash is invested in a 20-year zero coupon bond and 25% of the remaining cash is invested in equities. At time 1+, the actual capital is 1,098K, the regulatory capital is 449K, the market value of bonds is 49K and the market value of the equities is 16K.

Table 7-10 shows these figures for times 0+,1+,2+,3+,4+.

<table>
<thead>
<tr>
<th>Time</th>
<th>Capital flow</th>
<th>Actual capital</th>
<th>VaR</th>
<th>Target capital</th>
<th>MV of bonds</th>
<th>MV of Equities</th>
</tr>
</thead>
<tbody>
<tr>
<td>0+</td>
<td>873</td>
<td>1401</td>
<td>502</td>
<td>1004</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1+</td>
<td>-200</td>
<td>1098</td>
<td>449</td>
<td>899</td>
<td>49</td>
<td>16</td>
</tr>
<tr>
<td>2+</td>
<td>-128</td>
<td>1023</td>
<td>447</td>
<td>895</td>
<td>219</td>
<td>73</td>
</tr>
<tr>
<td>3+</td>
<td>-95</td>
<td>936</td>
<td>420</td>
<td>841</td>
<td>369</td>
<td>123</td>
</tr>
<tr>
<td>4+</td>
<td>-81</td>
<td>995</td>
<td>457</td>
<td>913</td>
<td>676</td>
<td>225</td>
</tr>
</tbody>
</table>

Note that positive figures for capital flows mean that the business has received some capital from the capital provider, i.e. positive figures mean that the capital provider has invested some money in the business. Similarly, negative figures mean the capital provider has received some dividends.
The first graph of Figure 7-7 shows the above figures together with figures for the terms extended to the whole lifetime of the business. For comparison purposes, two graphs of the capital flows under strategies whose EIRs are zero are also shown in Figure 7-7. Those figures were also obtained using market scenario 1.
Figure 7-7 - Illustration of capital flows under three strategies

**SMR 200%, EIR 25%:**

![Graph showing capital flows, actual capital, VaR, and target capital for SMR 200%, EIR 25%.](image)

**SMR 200%, EIR 0%:**

![Graph showing capital flows, actual capital, VaR, and target capital for SMR 200%, EIR 0%.](image)

**SMR 110%, EIR 0%:**

![Graph showing capital flows, actual capital, VaR, and target capital for SMR 110%, EIR 0%.](image)
When 25% is invested into equities, we observe that VaR keeps increasing for the first three decades and decreasing thereafter. The VaR is also volatile. Accordingly, frequent capital injections are required for the first three decades, while considerable dividends are paid out after three decades in accordance with the release of capital. On the other hand, when there is no investment into equities, VaR decreases for the first two decades to almost zero, then increases again but not significantly. This observation highlights the change in risk profile of the business brought by equity investments; holding equities makes the portfolio far riskier, which causes much larger capital requirements.

While investment in equities leads to higher capital requirements and also higher dividend payments, the issue is whether the profile of the strategy that involves equity investments is better or worse. We will use the average figures over all the market scenarios to look into this issue.

For each strategy, we simulate capital flows using the method we have illustrated. Then we obtain the average capital flow at each year end over the 5,000 simulations to find how the strategy choice affects the capital flows as a whole.

7.8.3 The accuracy of the VaR approximation

In Section 7.5, we have assumed that the distribution of the change in the value of all the assets and liabilities including cash over one year is normally distributed. In this section, we will assess how accurate this assumption is.

Consider the strategy with a 50% dividend, 200% SMR and a 25% equity investment. Let $NAV(4+, i)$ denote the net asset value that includes all the assets and liabilities at time 4 (in years) immediately after the capital inflow or outflow occurs for the $i$-th
market scenario \((i=1,2,\ldots,5000)\), and let \(STD(4+, i)\) denote the approximate standard deviation then, i.e. one-third of the VaR then, for the \(i\)-th market scenario \((i=1,2,\ldots,5000)\). Likewise, let \(NAV(5-, i)\) denote the net asset value at time 5 (in years) immediately before the capital inflow or outflow occurs for the \(i\)-th market scenario. We have chosen year 5 because it takes several years for the funding for the initial expense payment to be fully repaid, and normal investment operations start only after this funding is fully repaid.

Now we define a quantity \(Q(i)\) to be:

\[
Q(i) = \frac{NAV(5-, i) - NAV(4+, i)}{STD(4+, i)}.
\]

If the change of the value, \(NAV(5-, i) - NAV(4+, i)\), follows a normal distribution with the mean zero and the standard deviation \(STD(4+, i)\), \(Q(i)\) \((i=1,2,\ldots,5000)\) are samples from a standard normal distribution.

In order to check the accuracy of our assumption about the distribution, we look into the Q-Q plot of \(Q(i)\) \((i=1,2,\ldots,5000)\) against a standard normal distribution.
In our Q-Q plot, the $x$-axis value of a point shows the percentile of a random variable that follows a standard normal distribution for a particular probability, while the $y$-axis value of the point shows the percentile of our sample data for that probability.

We focus on the left side of the figure since we are trying to measure risks. Even though our data shows a slightly fatter tail than the normal distribution in the range beyond three standard deviations, largely points are aligned on the straight line. This means our assumption is reasonable under our model setting.

We also check some standard statistics of these samples.
Table 7.11 - Statistics of standardized value change

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.14</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.993</td>
</tr>
<tr>
<td>Minimum</td>
<td>-4.38</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.24</td>
</tr>
</tbody>
</table>

The mean is 0.14. This tells us that the mean annual growth rate of the value is positive and small, which is only 14% of its standard deviation. For the purpose of measuring risk figures, our assumption of zero growth is reasonable because this assumption results in slightly bigger, therefore more conservative, risk figures and the error is small anyway.

7.8.4 Internal Rate of Returns

We let $\overline{C}(t_i)$ denote the average capital flow at time $t_i$ ($t_0 = 0, t_1 = 1, t_2 = 2, ..., t_{70} = 70$). In order to compare each strategy, at first we look at the internal rate of return ("IRR" hereafter) on the injection option. The IRR is the rate that satisfies $\sum_{i=0}^{70} (1 + IRR)^{t_i} \overline{C}(t_i) = 0$. We remark that a more appropriate analysis would be to compute the IRR for each scenario, then study the distribution of the IRRs over 5,000 scenarios. We had tried this method at first, but we found that some cash flows are so extreme that we could not calculate an IRR for such cash flows. This told us that this method is not technically feasible.

Table 7-12 shows IRRs for all the strategies on the injection option, i.e. we assumed that no interim liquidation of the business unit is allowed. The shaded cells indicate that we will look into that case in detail.
We make several observations about the table above.

- The IRR is increased by the hedging activity without exception. There are greater increases when the SMR is lower and the EIR is lower, as illustrated by the strategies where the dividend payout ratio is 50%. When SMRs are 300%, 200% and 110%, and the EIR is 50%, the hedging transaction increased IRR by 0.2%, 0.4% and 2.6% respectively. When the EIR is 0%, however, the corresponding increases in IRR are much larger, which are 1.5%, 2.5% and 4.5% respectively.

- The higher ratio of equity investment does not always leads to higher IRR. In particular, when we adopt the hedging strategy, the higher EIR results in a lower
IRR without exception, as illustrated by the strategies where the dividend payout ratio is 50% with hedging. When the SMR is 110%, the IRR decreases to 9.9% and 9.2% from 10.0% as the EIR increases to 25% and 50% respectively from 0%. Similarly, when the SMR is 200%, the IRR decreases to 6.3% and 5.8% from 7.1% as the EIR increases to 25% and 50% respectively from 0%.

- The higher SMR requirement leads to lower IRRs without exception. The degree of decrease in IRR when IRR is relatively large, for example 10%, tends to be larger than the decrease when IRR is small, as illustrated by the strategies where the dividend payout ratio is 50%, and the EIR is 0%. Without hedging, the IRR decreased by 0.9% to 4.6% from 5.5% as the SMR increased to 200% from 110%. In contrast, with hedging, the IRR decreased by 2.9% to 7.1% from 10.0% as the SMR increased to 200% from 110%.

- The higher dividend payment causes high IRR without exception. The extent of increase in IRR is, however, small. Only under the strategies where SMR is 110% with hedging, IRR increased around 1% when the dividend payment ratio increased from 50% to 75%. However the increase in IRR is around 0.2% for other strategies. Since the effect of the dividend policy choice is small, we will look into only the strategies with a 50% dividend payout ratio from now on. This reduces the total number of strategies under comparison to 18 from 54.

These results confirm that the primary revenue source of the business unit is the profit margin, and that the contribution from risk premium is relatively small. By taking market risks, the additional revenue that stems from risk premiums is obtained. In particular, the investment into equities produces considerable additional revenues as we will show in section 7.8.5. The risk premiums, however, cannot overwhelm the
disadvantage of higher capital requirements from the viewpoint of capital efficiency. Taking market risks dilutes the revenue sourced from the profit margins, then results in a lower return on capital.

More investment into equities tends to lead, but not always, to increases in IRR when interest rate risk is not hedged. This is because the total standard deviation of the portfolio composed of bonds and equities is not the weighted sum of the standard deviation of each asset when these assets are not perfectly correlated. In other words, the risks of each asset are offset when they are mixed. However, the offset effect is not big enough to exceed the disadvantage of the higher capital requirement.

7.8.5 Expected Present Values

In this section, we examine the expected present value (“EPV”, hereafter) of the capital flows. Given the Contractual Cash Flows, the future capital flows depends only on the status of the interest rate term structure and equities. Therefore, under the no-arbitrage assumption, the EPV of the capital flows can be calculated as the average of the capital flows discounted by the deflator $D(t)$, which is defined in formula (6.2), over all simulations, i.e.:

$$\text{EPV} = \frac{1}{5,000} \sum_{j=1}^{5,000} \sum_{i=0}^{70} \frac{D(t_i, j)}{D(0)} C(t_i, j),$$

where $C(t_i, j)$ is the simulated capital flow at time $t_i$ on the $j$-th simulation, and $D(t_i, j)$ is the simulated deflator for the time $t_i$ on the $j$-th simulation.

We have determined the premiums so that the EPV of the Contractual Cash Flows is 1,056K. The business unit starts from zero capital and will be liquidated when all the annuities expire, i.e. the business unit is just a conduit between the annuitants and the
capital provider. This means that the Contractual Cash Flows are eventually passed onto the capital outflows in some way as time passes. Therefore the EPV of the capital flows should be -1,056K, which means that the EPV of all the dividends taking into account capital injections is 1,056K from the capital provider’s point of view. Note that this EPV is independent of the choice of investment strategy. The EPV of the capital flows is -1,056K whatever assets the business unit holds.

While the EPV using the deflator \( D(t) \) is consistent with market prices, that EPV does not provide any key information for equity investment. We have already discussed this in Section 3.2. The deflator \( D(t) \) removes all the risk premiums during the discounting process. However, the risk premiums are what equity investors are planning to earn, and thus the risk premiums are what equity investors are most interested in as future revenue. In order to show how much revenue the equity risk premium is expected to produce, we introduce an interest-deducted value (“IDV”, hereafter). We calculate IDV using \( A(t) \), that is the deflator solely for bonds, for discounting cash flows instead of \( D(t) \), i.e.:

\[
\text{IDV} = \frac{1}{5,000} \sum_{j=1}^{5,000} \sum_{i=0}^{n} \frac{A(t_i, j)}{A(0)} C(t_i, j),
\]

where \( A(t_i, j) \) is the simulated deflator solely for bonds for time \( t_i \) on the \( j \)-th simulation.

Discounting cash flows by \( A(t) \) removes all interest and risk premiums on bonds, but does not remove the risk premiums on equities. Actually, using (6.3),
\[
E \left[ \frac{A(t + \Delta t)}{A(t)} \cdot S(t + \Delta t) \right] \\
= E \left[ \frac{A(t + \Delta t)}{A(t)} \cdot S(t) \cdot \frac{1}{P(t, t + \Delta t)} \cdot \exp[\eta \sigma \Delta t] \cdot \exp \left[ -\frac{1}{2} \sigma^2 \Delta t + \sigma \Delta W_s(t) \right] \right] \\
= S(t) \exp[\eta \sigma \Delta t] \cdot \frac{1}{P(t, t + \Delta t)} \cdot E \left[ \frac{A(t + \Delta t)}{A(t)} \cdot 1 \right] \cdot \exp \left[ -\frac{1}{2} \sigma^2 \Delta t + \sigma \Delta W_s(t) \right] \\
= S(t) \exp[\eta \sigma \Delta t] \cdot \frac{1}{P(t, t + \Delta t)} \cdot P(t, t + \Delta t) \\
= S(t) \exp[\eta \sigma \Delta t].
\]

This result shows that the equity risk premium \( S(t) \left[ \exp[\eta \sigma \Delta t] - 1 \right] \) still remains after discounting by \( A(t) \).

Now we show EPVs and IDVs for some selected strategies in Table 7-13.

**Table 7-13 - EPV and IDV**

<table>
<thead>
<tr>
<th>SMR</th>
<th>Investment Ratio</th>
<th>Hedging</th>
<th>IRR</th>
<th>EPV (in 1,000)</th>
<th>IDV (in 1,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110%</td>
<td>0%</td>
<td>No</td>
<td>5.5</td>
<td>-1,114</td>
<td>-1,050</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>10.0</td>
<td>-1,062</td>
<td>-1,050</td>
</tr>
<tr>
<td>200%</td>
<td>0%</td>
<td>Yes</td>
<td>7.1</td>
<td>-1,068</td>
<td>-1,049</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>Yes</td>
<td>6.3</td>
<td>-1,071</td>
<td>-2,881</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>Yes</td>
<td>5.8</td>
<td>-1,073</td>
<td>-5,792</td>
</tr>
<tr>
<td>300%</td>
<td>0%</td>
<td>Yes</td>
<td>5.8</td>
<td>-1,078</td>
<td>-1,050</td>
</tr>
</tbody>
</table>

All figures are negative. However, we use positive figures by changing the sign for readability. That means we look at the figures from the viewpoint of the capital provider. When there is no investment in equities, both EPV and IDV are approximately equal to the correct value, which is 1,056K. IDV shows that approximately 1,800K and 4,700K
of additional revenues are expected to arise from the equity risk premiums when 25% and 50% of free cash are invested into equities respectively. Though these figures are significant compared to the EPV that is 1,056K, these absolute additional revenues have not overwhelmed the disadvantage of higher capital requirements as IRR indicates.

In general, the present value concepts behind EPV or IDV cannot perfectly describe the profile of the business primarily because present values do not take into account the duration of the underlying business.

7.8.6 Payback Period

Here we will measure the efficiency of the capital usage by looking at how quickly the injected capital is paid back to the capital provider. We define the payback period as the time when the accumulated average interest-deducted capital flow becomes negative. Table 7-14 shows how many years it will take for the initial capital to be paid back. Formally, we first find:

$$\min(t_i): \sum \bar{AC}(t_i) < 0,$$

where

$$\bar{AC}(t_i) = \frac{1}{5,000} \sum_{j=1}^{5,000} \frac{A(t_i, j)}{A(0)} C(t_i, j),$$

and then Payback Period is obtained by linear interpolation.
Table 7-14 - Payback Period and Total Capital Injection

<table>
<thead>
<tr>
<th>SMR</th>
<th>Equity Investment</th>
<th>Hedging</th>
<th>IRR</th>
<th>Payback Period (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110%</td>
<td>0%</td>
<td>No</td>
<td>5.5</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>10.0</td>
<td>4.9</td>
</tr>
<tr>
<td>200%</td>
<td>0%</td>
<td>Yes</td>
<td>7.1</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>Yes</td>
<td>6.3</td>
<td>32.6</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>Yes</td>
<td>5.8</td>
<td>35.0</td>
</tr>
<tr>
<td>300%</td>
<td>0%</td>
<td>Yes</td>
<td>5.8</td>
<td>15.0</td>
</tr>
</tbody>
</table>

The Payback Period for the strategy where the SMR is 110% with no equity investment and with hedging (“Base Strategy” hereafter) is significantly better than other strategies. Figure 7-9 shows the average capital flows on the base strategy, and Figure 7-10 shows the average capital flows on the strategy where the SMR is 110% with no equity investment but without hedging. The payback period when the interest rate risk is hedged is 4.9 years. This is approximately one third of that when the interest rate risk is not hedged, which is 14.8 years.
Figure 7-9 - Average capital flows for the strategy with SMR 110%, no equity investment and with hedging

Figure 7-10 - Average capital flows for the strategy with SMR 110%, no equity investment and without hedging
Comparing these capital flows, we see that VaR and hence capital requirements when interest rate risk is hedged are much smaller than those when it is not hedged. On both strategies VaR largely decreases for the first two decades. When the interest rate risk is hedged, the VaR decreases steadily for the first decade since inception. Thus the capital that was initially injected was paid back at a fast pace without being disturbed much by adverse movements of the interest rate term structure. When the interest rate risk is not hedged, VaR decreases little for the first decade. This led to the much longer payback period.

At time 22 years, the interest rate risk on the future annuity payments is almost completely offset by the interest rate risk on the combined position of future premium income flows and bonds. After this timepoint the interest rate risk on the future annuity payments gradually dominates the whole interest rate risk as both the amounts and the duration of future premium incomes decrease. Accordingly VaR starts to increase and additional capital injections are required for several consecutive years, whether the interest rate risk is hedged or not. When the interest rate risk is hedged, however, the required capital amount is much smaller than when it is not hedged. This smaller capital requirement has resulted in the higher IRR.

Consider the strategy where the SMR is 200% with no equity investment and with hedging. Figure 7-11 shows average capital flows under this strategy. This strategy is almost the same as the Base Strategy but the difference is that SMR is 200% in this strategy. The payback period on this strategy almost doubled to 11.6 years. This result seems reasonable since repayment of the injected capital is intuitively expected to take twice as long when the capital requirement almost doubled from 110% SMR to 200%
Next, we consider the strategy where the SMR is 110% with EIR 25% and with hedging. This strategy is different from the Base Strategy in that 25% of free cash is invested into equities. The equity investment significantly increases VaR, and the resulting capital requirement prevents the injected capital from being released at early stages of the business. The resulting payback period for this strategy is 32.6 years. Figure 7-12 shows that VaR keeps increasing as the premium income is invested in equities; therefore there is no room for dividends to be paid for the first three decades except for the first several years. Note that the y-axis scale in Figure 7-12 has been changed from the former two strategies. The first several dividends are in fact just the repayments of the funding for the cash payments for the initial expenses. The business unit needed the cash to pay the expenses that occur at time zero; therefore this cash was raised as a part of the initial capital. This injection is in fact a kind of loan beyond the target capital requirement.
from the capital provider. Thus it was repaid as soon as the business earns premium income.

**Figure 7-12 - Capital flows for the strategy with SMR 200%, 25% equity investment and with hedging**

7.8.7 Additional capital injections

In this section we analyze additional capital injections, which capital providers must provide when actual capital is short of target capital. The capital provider of the business unit is an investor who is more interested in long-term return rather than short-term profit taking. Therefore, the capital provider would not be much concerned about the short-term change in the value of the capital. However, the capital provider would be very anxious to avoid additional capital injections. This is not only because additional capital injection itself is a further financial burden, but also because the occurrence of the capital injection means the minimum capital requirement has been temporarily breached. Here, we examine the following values for each strategy to quantify the extent to which the capital provider is exposed to additional capital injections.
- Time to the first additional capital injection (“First Injection”, hereafter)
- Number of additional capital injections (“Injection Number”, hereafter)
- Total interest-deducted additional capital injection amount (“Injection Amounts”, hereafter)

Table 7-15 shows the values for selected strategies. On all of these strategies interest rate risk is hedged.

<table>
<thead>
<tr>
<th>Value</th>
<th>Statistics</th>
<th>SMR 110%, No equities</th>
<th>SMR 200%, No equities</th>
<th>SMR 200%, 25% equities</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Injection (in years)</td>
<td>Average</td>
<td>7.9</td>
<td>18.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Injection Number</td>
<td>Average</td>
<td>7.7</td>
<td>7.2</td>
<td>23.2</td>
</tr>
<tr>
<td>Injection Amount (in 1,000)</td>
<td>Average</td>
<td>411</td>
<td>312</td>
<td>2,611</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>174</td>
<td>172</td>
<td>760</td>
</tr>
<tr>
<td></td>
<td>95% Percentile</td>
<td>712</td>
<td>641</td>
<td>3,943</td>
</tr>
</tbody>
</table>

It is clearly shown that investment in equities has significantly increased the frequency and amount of additional capital injections. The figures for the first strategy and the second strategy, from left, are similar except that the First Injection for the second strategy is 18.9 years while that on the first strategy is 7.9 years. This is reasonable since the only difference between the first and second strategy is the SMR. The SMR on the second strategy is 200% while the SMR on the first strategy is 110%. That means

---

72 We use the same approach as that used for calculating IDV.
that the second strategy has more free cash than the first strategy. The greater interest earned on greater free cash created a thicker cushion against adverse market movements, so that the second strategy withstood the adverse market movements in the early stages of the business. Note that the portfolios on the first and second strategy are exposed to a small interest rate risk. Even though the interest rate risk is significantly reduced by hedging, a small interest rate risk still remains since we have used a simple hedging technique using one 20-year bond.

The first and the second strategies are comfortable strategies. Additional capital injections are required 7 to 8 times on average on the first and second strategy over the whole lifetime of the business, and the total interest-deducted injection amount is approximately 400K on average. Considering the duration of the business is 70 years and the annual gross premium is 339K, 8 capital injections of total amount 400K are not large figures. In addition, the fact that the standard deviation of the total capital injection amount is approximately 170K shows that additional capital injection amounts do not fluctuate much. We observe that the distribution of the total capital injection amount has a fatter tail than a normal distribution has. The 95% percentile is 1.9 times the standard deviation, on the second strategy, away from the average, which would be 1.64 times the standard deviation in the case of a normal distribution. However, the fat tail feature is too small to prevent us from concluding that the first and second strategies are comfortable strategies with respect to exposure to additional capital injections.

On the contrary, the figures on the third strategy show that an investment into equities significantly increases the exposure to additional capital injections. On the third strategy where 25% of free cash is invested in equities, the additional injections will be expected
in 4 years from the inception of the business on average. Over the lifetime of the business, 23 additional injections will be required on average. The average additional injection amount is 2,611K, which is approximately 8 times as much as that on the second strategy; and the amount is approximately 7.7 times as much as the annual gross premium. The standard deviation of the total additional injection amount is 760K, which is 2.3 times the annual gross premium. Even though the 95th percentile is 1.75 times the standard deviation away from the average, suggesting that the distribution of the additional injection amount could be normal, clearly investment into equities leaves the capital provider far more exposed to additional capital injections.

7.8.8 Liquidation Option

In this section we will look into the effect of the liquidation option. Now we investigate the three strategies that we have analyzed in section 7.8.7. but with a liquidation option this time. On the liquidation option, no additional capital is provided, and the business unit is dissolved after liquidating all the assets and liabilities if the minimum capital requirement is breached. Table 7-16 shows IRR, EPV and IDV of the capital flows on the selected strategies. The column “(IRR)” shows, for comparison purpose, the IRR on the original strategies where SMR and EIR are the same but additional capital is required to be injected, with interim liquidation not being allowed.

<table>
<thead>
<tr>
<th>SMR</th>
<th>EIR</th>
<th>(IRR)</th>
<th>IRR</th>
<th>EPV (in 1,000)</th>
<th>IDV (in 1,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110%</td>
<td>0%</td>
<td>10.0</td>
<td>16.6</td>
<td>-523</td>
<td>-521</td>
</tr>
<tr>
<td>200%</td>
<td>0%</td>
<td>7.1</td>
<td>7.1</td>
<td>-527</td>
<td>-531</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>6.3</td>
<td>8.3</td>
<td>-561</td>
<td>-1,295</td>
</tr>
</tbody>
</table>

Table 7-16 - EPV and IDV on liquidation option
The first observation is that early liquidation produces a higher, or almost equal in the case of the second strategy, IRR. This can be explained by the existence of the up-front profit. At the inception of the business, we have already recognized the up-front profit whose amount is 528K. The remaining profits are gradually recognized as MVMs are released after the actual payments of annuities start in 30 years. In the IRR calculation, the up-front profit is spread over a shorter period as the liquidation happens earlier; therefore the IRR tends to increase as the lifetime of the business is shortened. For example, if liquidation happens in 5 years, the up-front profit will be spread over 5 years, i.e. the up-front profit allocated to each year is approximately 106K. If the liquidation is not allowed, the up-front profit is allocated over the lifetime of the business, which is 70 years. In this case, the up-front profit allocated to each year is approximately 7.5K.

The second observation is that the calculated EPVs are approximately the same as the up-front profit of 528K. Compared to the EPVs calculated before on strategies where interim liquidation is not allowed, EPVs are decreased by 528K. This is because, as we have already stated, when the business is liquidated before annuity payments start, the future profits included in the MVM are discarded while the up-front profit has already been recognized at the start of the business. The lost EPV is the PV of MVM, which is 528K. Note that the IRR is still higher even though the profits including MVM are lost when liquidation is allowed, compared with when liquidation is not allowed. As we have stated, this is due to the shortened lifetime of the business.

Compared to the cases where interim liquidation is not allowed, the EPV decreased to 528K from 1,056K. This result shows that the shareholder’s call option value is negative.
in our setting. As we have discussed in Chapter 3, allowing interim liquidation should create a call option on assets of the firm to the owner of the business. Therefore, one would expect that the EPV would be higher than 1,056K, or at least higher than 528K taking into account the lost profits included in MVM, if liquidation is allowed. In fact, however, EPV is approximately equal to 528K. This can be explained by our specific business model that involves the minimum capital requirement. The minimum capital prevents the actual capital from plunging to a negative level. Therefore, the expected value of policyholders’ benefits is secured even in the liquidation situation. The benefits would be eroded if there were no minimum capital requirement, in which case some shareholders’ call option value would be created. As a result, the call option would never be in-the-money. In addition, when the business is liquidated before annuity payments are made, the business will lose the profit arising from unreleased MVM associated with the dismissed annuity payments. Therefore, the MVM system creates the incentive for the shareholders to carry on the business until the annuities are paid and profits included in the MVM associated with the annuity payments are recognized.

The third observation is that the IDV decreased when 25% of free cash is invested into equities. The IDV decreased from 2,881K to 1,295K. This result is simply because the future equity risk premiums are lost by interim liquidations, as IDV includes equity risk premiums while EPV excludes equity risk premiums. If the objective of this business is to obtain equity risk premiums, the early liquidation causes negative effects on the objective.

7.9 Longevity risk
In this section, we consider the effects of longevity on the annuity business we have
investigated. So far we have assumed that the cash flow projection does not change and cash flows are realized as initially expected over time. In reality, however, mortality rates change; the mortality experience may differ from what was initially expected, and the mortality expectation in 2007 will be different from that in 2006. Wilkie et al. (2003) compare various base tables, PMA68Base, PMA80Base, PMA92B, RMC92Base, PA(90)Males for males aged 65 to 119, and found that mortality rates in the UK had decreased significantly over a period of about 30 years. For example, the mortality rate for a male aged 65 on PMA92Base is approximately half of that on PMA68Base. The improvement effect gradually decreases as age increases; the mortality rate for a male aged 90 on PMA92Base is, for example, approximately 80% of that on PMA68Base. Wilkie et al. (2003) were interested in the effect of mortality improvements on Guaranteed Annuity Options. They concluded that improvements in mortality had a more significant effect in increasing the value of Guaranteed Annuity Options than a reduction in interest rates had.

Now we try to measure the effect of a change in mortality rates on our hypothetical annuity business.

Let \( q_x(t) \) denote the mortality rate for a male aged \( x(=0,1,2,...) \) at time \( t=0,1,2,... \). The mortality table we have used so far gives \( q_x(0) \) for all \( x \). The mortality table at time \( t \) will be tabulated using \( q_x(t) \) for all \( x \). Here, we assume:

\[
q_x(t) = q_x(0) \cdot f(x,t), \tag{7.4}
\]

where \( f(x,t) \) is a deterministic function of \( x \) and \( t \), and we call this a mortality improvement ratio.

We anticipate that mortality rates around retirement ages will improve more than at
younger ages due to improved living conditions and medical developments. On the other hand, it is also impossible for the improvement to continue forever since humans are not immortal. Despite those obvious prospects, precise forecasting of mortality improvement in the distant future is very difficult and in fact is not practical. Therefore, we use a simple formula for $f(x,t)$ here. At first, we assume that an ultimate improvement ratio, $\beta$, for a particular age $x_0$ exists. The mortality improvement ratio gradually reaches this level as time passes. Let $\beta(t)$ denote the mortality improvement ratio at time $t$. Then, we model:

$$\beta(t) = \beta + (1 - \beta) e^{-\alpha t}.$$ 

We also assume that the improvement ratio gradually decreases as the age departs from $x_0$. Then we define the function $f(x,t)$ as:

$$f(x,t) = \begin{cases} 
\beta(t) + \left[1 - \beta(t)\right] \left(1 - e^{-\alpha_1|x-x_0|}\right) & \text{when } x \geq x_0 \\
\beta(t) + \left[1 - \beta(t)\right] \left(1 - e^{-\alpha_2|x-x_0|}\right) & \text{when } x < x_0.
\end{cases}$$

At first we try the parameters such that: $x_0 = 65, \beta = 0.5, \alpha_1 = \ln(2)/10$ so that $f(65,10) = 3/4$ and hence the mortality rate of a 65 year old is cut by 25% in 10 years; $\alpha_2 = \ln(2)/20$ so that $f(85,10) = 7/8$, i.e. the mortality rate of the person aged 85 improves in 10 years by half the amount of a person aged 65; and $\alpha_3 = \ln(5)/5$ so that $f(60,10) = 7/8$, i.e. the mortality of a person aged 60 improves in 10 years by half the amount of a person aged 65.

Figure 7-13 shows how mortality rates improve according to this formula. Note that $q_{104}(t)$ is always 1, therefore $f(104,t) = 1$ for any $t$. 

220
Let $p_x(t)$ denote that the survival probability for a male aged $x$ at time $t$. We have already specified $q_x(t)$ for each $t = 0, 1, 2, \ldots$ by (7.4). Using this, $p_x(t)$ can be calculated as $1 - q_x(t)$.

Now let $N_x(0, 0)$ ($x = 0, 1, 2, \ldots$) denote a large number of people, for example 1,000,000, who are aged $x$ at time 0. We consider the development of this population as time passes. We also define $N_x(t, s)(t = 0, 1, 2, \ldots, s = 0, 1, 2, \ldots)$ as follows depending on whether $s \leq t$ or $t < s$.

- When $s \leq t$, $N_x(t, s)$ represents the number of people who have survived up to time $s$ given the information available at time $t$.
- When $t < s$, $N_x(t, s)$ represents the number of people who are expected to survive up to time $s$ using the mortality table available at time $t$. 
Clearly, \( N_x(0,s) = N_x(0,0) \prod_{k=0}^{s-1} p_{x+k}(0) \) (\( s = 1, 2, 3, \ldots \)).

For \( t = 1, 2, 3, \ldots \), we have:
\[
N_x(t,s) = N_x(0,0) \prod_{k=0}^{t-1} p_{x+k}(\min(k,t)) \quad (s = 1, 2, 3, \ldots),
\]
assuming that people die exactly as expected one year before, i.e. the mortality experience between time \( t - 1 \) and time \( t \) for the people aged \( x \) at \( t - 1 \) is \( q_x(t - 1) \).

To understand this formulae, suppose we are at time 1 now. Since we have assumed people die as expected, \( N_x(0,0) \cdot p_x(0) \) people have survived to time 1, i.e. \( N_x(1,1) = N_x(0,0) \cdot p_x(0) \). The people are all aged \( x + 1 \) now. At this stage, survival probabilities for these survivors are given by the mortality table available at time 1 as \( p_{x+1}(1), p_{x+2}(1), \ldots \). Therefore, for example, the expected number of people who will survive up to the time 2, denoted by \( N_x(1,2) \), is \( N_x(0,0) \cdot p_x(0) \cdot p_{x+1}(1) \). Similarly, we know \( N_x(1,3) = N_x(0,0) \cdot p_x(0) \cdot p_{x+1}(1) \cdot p_{x+2}(1) \). Thus, generally:
\[
N_x(1,s) = N(0,0) \prod_{k=0}^{s-1} p_{x+k}(\min(k,1)) \quad (s = 1, 2, 3, 4, \ldots),
\]
i.e. the formula holds when \( t = 1 \). At time 2, \( N_x(0,0) \cdot p_x(0) \cdot p_{x+1}(1) \) people have survived. As at time 1, survival probabilities for these survivors are given by the mortality table available at time 2 as \( p_{x+2}(2), p_{x+3}(2), \ldots \). Therefore \( N_x(2,s) = \prod_{k=0}^{s-1} p_{x+k}(\min(k,2)) \), i.e. the formula holds when \( t = 2 \). Repeating this process, we know the formula holds for all \( t \).

Now we define \( p_x^0(t) \) to be \( N_x(t,s)/N_x(0,0) \). Then \( p_x^0(t) \) gives the probability that a life who is aged \( x \) at time 0 survives for \( s \) years from time 0 given the information available at time \( t \). Note that this survival probability is for a person who is initially
aged \( x \), not for the person who is aged \( x \) at time \( t \), and that the probability is not conditional on surviving to aged \( x + t \). Accordingly, the term for which the survival probability is calculated is measured from the time 0, not from the time \( t \).

This definition may seem strange since it allows a person to partially die, and it contains both the past mortality experience and the expected mortality in the future. However, this definition is meaningful and convenient when we apply the probability to a large population, as we use it in the following study.

Let \( G(l)(l = 0,1,2,3,...) \) denote the net cash flow at time \( l \) per contract, which can be premium income or an annuity payment including expenses per contract. Under the assumption that the mortality experience is as expected, at time \( t \), \( p^0_x(t) \cdot G(t) \) is the actual cash flow that has just occurred, and \( \cdot \cdot \cdot p^0_x(t) \cdot G(t + k) \) is the cash flow expected at time \( t + k \) \((k = 1,2,3,...)\).

Now we investigate the effects of mortality improvements on our hypothetical business. The mortality improvement affects only the value of annuities, and does not affect the value of assets. In addition, here we assume that the economic variables such as interest rates do not affect the cash flows arising from the annuities. In other words, the analysis on the mortality improvement is conducted independently from the analysis on market risks arising from the asset-liability mismatch. In order to investigate the effects solely arising from the mortality improvements and to exclude any effects arising from the asset-liability mismatch, we make the following assumptions this time. Note that we have assumed deterministic development of mortality rates; therefore no stochastic simulation is involved in this analysis.
• Interest rates develop as initially expected, i.e. $P(t, T) = P(0, T) / P(0, t)$ for all $t < T$.

• Premium income is invested only into cash.

• The business unit takes no market risks. We do not consider any capital support or regulatory capital requirement.

• The profits, which may be negative, are paid out or replenished immediately.

In this setting, when there is no improvement in mortality, we have 527,951 of up-front profit and a sequence of yearly profits due to the release of the MVM after 30 years of no profit periods. Figure 7-14 shows the yearly profit emergence and the accumulated profit. Note that these figures are not discounted. Discounting profits by the initial yield curve, we get 1,055,902 of profit.

**Figure 7-14 - Profit emergence when there is no mortality improvement**

![Profit emergence graph](image)

When we assume that mortality improves according to the formula and parameters we
have specified, the business suffers loss from year 1 for 30 years until the release of MVM starts. Figure 7-15 shows the yearly profit emergence and the accumulated profit in this case.

**Figure 7-15 - Profit emergence when there is mortality improvement**

The accumulated profit is 1,349,429 and the total discounted profit is 349,726. Both of these are positive but the discounted figure shows that up-front profit, which is 527,951, is partly taken away. The accumulated profit plunges below zero after 13 years, but moves back above zero after 36 years as the MVMs are released.

One may well expect that the decrease in mortality rates should cause much higher annuity payments and therefore a much heavier financial burden on the insurer. This counter-intuitive result is due to the instalment of the premium. When mortality improvement occurs in the early stage of the contract, we collect more premiums from more survivors even though expected annuity payments increase. In other words, only the mortality improvements at the later stage of the contract term would produce
significant effects on the final performance of the business. In the case of the study by Wilkie et al. (2003), the mortality improvement has caused much more adverse effects than our case since they assumed a single premium for the guaranteed annuity option contracts.

Table 7-17 shows how the total discounted profit changes for various $\alpha_1$ and $\beta$. When the half life is 10, for example, it means that $\alpha_1 = \ln(2)/10$ so that the mortality improves by half in 10 years.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Half life</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>0.5</td>
<td>349,726</td>
</tr>
<tr>
<td>0.25</td>
<td>-77,815</td>
</tr>
<tr>
<td>0.1</td>
<td>-364,613</td>
</tr>
</tbody>
</table>

For the same $\beta$, a larger half life results in a higher profit without exception. This is because the overall magnitude of improvement in mortality diminishes as the half life increases, even though the improvement occurs in the later stages of the contract’s term. The profit becomes negative only when mortality improves significantly and quickly. Among the cases shown in Table 7-17, the profit became negative only when mortality improves by 37.5%, that is $(100\% - 25\%)/2$, in 10 years.

We have charged a loading on premiums for 10% of the expected annuity payments as the up-front profit and the same amount for the MVM. In conclusion, under our mortality improvement model, the up-front profit and MVM provide a reasonable cushion against the longevity risk. It should, however, be emphasized that there is no
way to hedge out the longevity risk; therefore some capital should also be set aside for the longevity risk as well as the capital for the market risk. The regulatory capital for the expected loss exceeding the MVM incurred due to the mortality improvement is likely to be required under the fair valuation regime. However, the method to calculate the capital requirement related to the mortality risk has not yet been proposed to our knowledge at present.

7.10 Conclusion
The solvency margin ratios of insurance companies operating in Japan shown on Table 7-5 indicate that insurance companies are generally healthy with regard to capital. If insurance companies just seek absolute returns ignoring the efficient usage of capital, the investment into equities may be a desirable strategy that is expected to produce significantly high revenues as the IDV values in Table 7-13 suggest.

However, in the case of our simple setting of the business under the fair valuation regime, taking greater market risks by increasing the equity investment ratio simply reduces the relative return on the injected capital. This is due to the strict capital requirement for taking market risks. Note that we have assumed the price of equities is log-normally distributed. It is well known that the distribution of the price of equities in the real world is more fat-tailed than the normal distribution. If we had used a probability distribution with a fatter tail as our model for the price of equities, the capital requirements would have been even larger and hence the return on capital would have been even lower on strategies where equities are held. By contrast, the reduction of the market risk through asset-liability management, even by our simple method using only one bond, significantly improves the return on the capital.
We have studied simplified business strategies and we have made some important assumptions, in particular, about the market-price-of-risk for equities. A higher market-price-of-risk, for example 30%, would increase the IRR even though the EPVs are not affected. In addition, in the real world, fund managers could be able to achieve investment returns higher than what we have assumed using more flexible and better equity investment techniques.

Whatever techniques are used, however, investment into risky assets is very likely to lead to frequent and significant capital injections to support the business. In a practical situation, the general public as well as regulators may regard such frequent injections as questions about the safety of the business.

We could conclude, in general, that efforts should be made to minimize the risk, instead of taking risks searching for a higher return, under the risk-based capital regime in the fair value principle. This incentive is what the introduction of the fair value accounting is aiming at.
References


International Actuarial Association (2000a, May 31). *Comments on basic issues of the IASC insurance issues paper*. Downloaded from the homepage of the International Actuarial Association.


Appendix I

Mortality Table (Male)

<table>
<thead>
<tr>
<th>x</th>
<th>l_x</th>
<th>x</th>
<th>l_x</th>
<th>x</th>
<th>l_x</th>
<th>x</th>
<th>l_x</th>
<th>x</th>
<th>l_x</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100000</td>
<td>1</td>
<td>99692</td>
<td>26</td>
<td>98914</td>
<td>51</td>
<td>95050</td>
<td>76</td>
<td>65905</td>
</tr>
<tr>
<td>2</td>
<td>99654</td>
<td>27</td>
<td>98847</td>
<td>52</td>
<td>94652</td>
<td>77</td>
<td>63293</td>
<td>102</td>
<td>594</td>
</tr>
<tr>
<td>3</td>
<td>99626</td>
<td>28</td>
<td>98778</td>
<td>53</td>
<td>94215</td>
<td>78</td>
<td>60510</td>
<td>103</td>
<td>377</td>
</tr>
<tr>
<td>4</td>
<td>99605</td>
<td>29</td>
<td>98707</td>
<td>54</td>
<td>93735</td>
<td>79</td>
<td>57565</td>
<td>104</td>
<td>231</td>
</tr>
<tr>
<td>5</td>
<td>99588</td>
<td>30</td>
<td>98634</td>
<td>55</td>
<td>93215</td>
<td>80</td>
<td>54455</td>
<td>105</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>99572</td>
<td>31</td>
<td>98559</td>
<td>56</td>
<td>92653</td>
<td>81</td>
<td>51199</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>99558</td>
<td>32</td>
<td>98481</td>
<td>57</td>
<td>92046</td>
<td>82</td>
<td>47807</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>99537</td>
<td>34</td>
<td>98308</td>
<td>59</td>
<td>90685</td>
<td>84</td>
<td>40651</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>99528</td>
<td>35</td>
<td>98215</td>
<td>60</td>
<td>89928</td>
<td>85</td>
<td>36939</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>99518</td>
<td>36</td>
<td>98117</td>
<td>61</td>
<td>89118</td>
<td>86</td>
<td>33201</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>99509</td>
<td>37</td>
<td>98014</td>
<td>62</td>
<td>88258</td>
<td>87</td>
<td>29489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>99498</td>
<td>38</td>
<td>97903</td>
<td>63</td>
<td>87343</td>
<td>88</td>
<td>25880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>99485</td>
<td>39</td>
<td>97782</td>
<td>64</td>
<td>86366</td>
<td>89</td>
<td>22350</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>99467</td>
<td>40</td>
<td>97649</td>
<td>65</td>
<td>85320</td>
<td>90</td>
<td>19025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>99443</td>
<td>41</td>
<td>97502</td>
<td>66</td>
<td>84192</td>
<td>91</td>
<td>15936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>99410</td>
<td>42</td>
<td>97342</td>
<td>67</td>
<td>82962</td>
<td>92</td>
<td>13124</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>99368</td>
<td>43</td>
<td>97168</td>
<td>68</td>
<td>81613</td>
<td>93</td>
<td>10618</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>99320</td>
<td>44</td>
<td>96977</td>
<td>69</td>
<td>80134</td>
<td>94</td>
<td>8431</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>99268</td>
<td>45</td>
<td>96767</td>
<td>70</td>
<td>78519</td>
<td>95</td>
<td>6564</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>99214</td>
<td>46</td>
<td>96539</td>
<td>71</td>
<td>76766</td>
<td>96</td>
<td>5006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>99158</td>
<td>47</td>
<td>96292</td>
<td>72</td>
<td>74873</td>
<td>97</td>
<td>3735</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>99100</td>
<td>48</td>
<td>96025</td>
<td>73</td>
<td>72838</td>
<td>98</td>
<td>2724</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>99039</td>
<td>49</td>
<td>95733</td>
<td>74</td>
<td>70667</td>
<td>99</td>
<td>1939</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>98977</td>
<td>50</td>
<td>95410</td>
<td>75</td>
<td>68358</td>
<td>100</td>
<td>1345</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- For the year 2004.
- Published by the Ministry of Health, Labour and Welfare in Japan.
- Mortality after age 100 was obtained by extrapolation assuming Gompertz’s Law.
Appendix II

Minimizing a simple quadratic form

Let $A$ be an $n$-dimensional symmetric matrix with its $ij$-th element denoted by $a_{ij}$, with $a_{ii} > 0$ for all $i$. Let $w(x)$ be a $n$-dimensional weight vector whose elements are constant apart from the $i$-th element which is variable. That is, $x \in \mathbb{R}^1$, i.e. $w(t) = (w_1, w_2, \ldots, w_{i-1}, x, w_{i+1}, \ldots, w_n)^T$. Then the function $f(x) = w(x)^T A w(x)$ is minimized at:

$$x^* = -\frac{w_0 \cdot a_i}{a_{ii}},$$

where $w_0 = (w_1, w_2, \ldots, w_{i-1}, 0, w_{i+1}, \ldots, w_n)^T$, and $a_i$ is the $i$-th column of $A$.

This is easily proved as follows.

Let $1_i$ be the $n$-dimensional vector whose elements are zero apart from the $i$-th element, which is 1. Then, $w(x) = w_0 + x1_i$. Therefore,

$$f(x) = w(x)^T A w(x) = [w_0 + x1_i]^T A [w_0 + x1_i]$$

$$= x^2 1_i^T A 1_i + x [1_i^T A w_0 + w_0^T A 1_i] + w_0^T A w_0$$

$$= a_{ii} x^2 + 2 (w_0 \cdot a_i) x + w_0^T A w_0$$

since $A$ is symmetric

$$= a_{ii} \left[ x + \frac{w_0 \cdot a_i}{a_{ii}} \right]^2 + w_0^T A w_0 - \left( \frac{w_0 \cdot a_i}{a_{ii}} \right)^2.$$

As $a_{ii}$ is positive, $f(x)$ takes its minimum when $x = -\frac{w_0 \cdot a_i}{a_{ii}}$. 

235
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Sato, Manabu

Title:
Fair valuation of insurance liabilities - a case study

Date:
2007-12

Citation:

Publication Status:
Unpublished

Persistent Link:
http://hdl.handle.net/11343/39364

File Description:
Fair valuation of insurance liabilities - a case study

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.