A study of term structure of interest rates - theory, modelling and econometrics

Author:
Chen, Shuling

Publication Date:
2009

DOI:
https://doi.org/10.26190/unsworks/22853

A Study of Term Structure of Interest Rates — Theory, Modelling and Econometrics

SHULING CHEN

A thesis in Financial Mathematics and Statistics
presented to
The University of New South Wales
for the degree of
Doctor of Philosophy

November 2009

Origination Statement

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed .... (SHULING CHEN)

January 20, 2010

Copyright Statement

Authenticity Statement

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.

Signed .... (SHULING CHEN)

January 20, 2010

Acknowledgements

In 1999, I moved to Sydney and worked with professor William Dunsmuir on a statistical modelling project for the 2000 Sydney Olympic Games. As a collaboration between the University of New South Wales (UNSW) and the Australian Bureau of Meteorology, the project aimed to use statistical techniques to accurately forecast weather. After completion of the project, I decided to stay at UNSW and study for a PhD degree. This dissertation marks the successful completion of the research project. The following comments are an expression of my sincere thanks to a number of people who have offered the unstinting support and generous help necessary for me to successfully complete my PhD research project.

First of all, I would like to thank Professor William Dunsmuir, my co-supervisor, for encouraging me to pursue a PhD degree at UNSW. I appreciate his initial and continuing support and efficient research supervision. He has spent many hours reviewing different drafts of this thesis. Without his perceptive critical comments and incisive ideas, this thesis would never have seen its completion. For all that, I thank him most sincerely.

I would also like to sincerely thank my supervisor A/prof Ben Goldys for his patient assistance in all my endeavours along the way with this project. I am deeply appreciative of his supervision, his many suggestions, ideas, and his supply of the relevant reference literature. His persistence has enabled me to obtain a level of understanding sufficient for this challenging project.

I would like to thank Dr Spiridon Penev for his assistance on study t-copula and meta-elliptical distributions with given marginals.

I wish to thank the School of Mathematics, UNSW, which financed my research and my attendance to various Australian and international conferences.

I would like to express my gratitude to Ms Nino Kordzakhia, of the Commonwealth Bank of Australia (CBA), for her presentation and discussion of the yield curve constructions. Appendix C is based primarily on discussions with Ms Kordzakhia.

I am particularly grateful to Mr William Hood, of the Reserve Bank of Australia (RBA). Through email, he spent a great deal of time answering my requests and questions, especially those concerning the RBA yield curves data sets, their supplements, and the generic yield curves. Section 4.2 and the method of 'Generic Bond Yields of the Reserve Bank of Australia' are based primarily on Mr Hood's advice.

I would like to thank Ms Marilyn Chilvers, my current Executive director at Community Services, the Department of Human Services NSW, for her great support and her valuable time on proof reading and editing the thesis.

I thank the thesis examiners for their interest in reading my thesis and for their valuable time and comments.

Last, but certainly not least, I would like to thank my family and friends for all their love and support through a very busy and demanding period of my life. My special thanks go to my daughter Renee whose love has given me both the spiritual and mental support required for the completion of this task.

Abstract

This thesis is concerned with the modelling of the term structure of interest rates, with a particular focus on empirical aspects of the modelling.

In this thesis, we explore the θ-parameterised (θ being the length of time to maturity) term structure of interest rates, corresponding to the traditional T-parameterised (T being the time of maturity) term structure of interest rates. The constructions of Australian yield curves are illustrated using generic yield curves produced by the Reserve Bank of Australia based on bonds on issue and by constructed yield curves of the Commonwealth Bank of Australia derived from swap rates.

The data used to build the models is Australian Treasury yields from January 1996 to December 2001 for maturities of 1, 2, 3, 5 and 10 years, and the second data used to validate the model is Australian Treasury yields from July 2000 to April 2004 for maturities of all years from 1-10. Both data were supplied by the Reserve Bank of Australia. Initially, univariate Generalised Autoregressive Conditional Heteroskedasticity (GARCH), with models of individual yield increment time series are developed for a set of fixed maturities. Then, a multivariate Matrix-Diagonal GARCH model with multivariate asymmetric t-distribution of the term structure of yield increments is developed. This model captures many important properties of financial data such as volatility mean reversion, volatility persistency, stationarity and heavy tails.

There are two innovations of GARCH modelling in this thesis: (i) the development of the Matrix-Diagonal GARCH model with multivariate asymmetric t-distribution using meta-elliptical distribution in which the degrees of freedom of each series varies with maturity, and the estimation is given; (ii) the development of a GARCH model of term structure of interest rates (TS-GARCH). The TS-GARCH model describes the parameters specifying the GARCH model and the degrees of freedom using simple smooth functions of time to maturity of component series. TS-GARCH allows an empirical de-

scription of complete interest rate yield curve increments therefore allowing the model to be used for interpolation to additional maturity beyond those used to construct the model. Diagnostics of TS-GARCH model are provided using Australian Treasury bond yields.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Background

Over the last two decades, the trading of interest rate derivatives has rapidly increased and a number of new products have been introduced. The prices of interest rate derivatives depend on interest rate levels and interest rate models. Therefore, efficient market practice and financial theory rely on good statistical analysis and appropriate modelling of interest rates.

There are two alternative philosophies for modelling interest rates dynamics. These are: modelling of *spot rates* and modelling of *term structure of interest rates* (also called *yield curves*).

The spot rate is a scalar stochastic process. Spot rate models are relatively simple, and a variety of spot rate models have been proposed. However, most popular spot rate models do not have enough degrees of freedom to fit the observed term structure of interest rates, and empirical analysis of observed spot rate does not provide an adequate description of the dynamics of the term structure of interest rates. Therefore, in recent years the research focus has been on the modelling of term structure of interest rates that describes the movements of the whole yield curves over time (Heath, Jarrow and Morton 1992). The term structure of interest rates refers to the relationships of bonds in different maturity. The interest rates of bonds (called *yields*) plotted against their maturity are called *yield curves*. More generally, the *term structure of interest rates* or *yield curves* is a general term used for bond prices, yields of bonds and forward interest rates of different maturities. Economists and investors believe that the shape of the yield curve reflects the market's future expectation for interest rates and the conditions for monetary policy. (http://www.finpipe.com/termstru.htm, *Financial Pipeline*). Term structure models provide a better way of capturing the dynamics of the term structure of interest rates or yield curves. Since it is a function of the maturity level and the time of evolution, the term structure of interest rates can be viewed as a function-valued stochastic process.

1.2 Objective

This thesis is devoted to the modelling of the term structure of interest rates with a particular focus on empirical aspects of the modelling. Bond yields data is used in the models.

It is well known that financial time series contain trends and are strongly autocorrelated. Statistical models of financial and economic time series are often based on the *returns* or *increments* (one-lag difference). The discrete-time model of the increments also has the advantage of allowing the mean and variance of increments to depend directly on the level of the financial time series in a way consistent with the continuous-time model based on diffusion processes (See Section 3.6). Modelling the increments of financial time series needs to account for certain empirical facts of financial time series. These include heavy tailed distributions, volatility clustering or persistence, volatility mean reversion, asymmetric impact on volatility by innovations, and possible influence of exogenous variables on volatility (Engle and Patton, 2001). In this thesis, yield increment series are fitted with Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models, using Australian Treasury yield data 1996-2001 from the Reserve Bank of Australia. And the Australian Treasury yields from July 2000 to April 2004 for maturities of all years from 1-10 are used to validate the models.

1.3 History and Motivation

The term structure of interest rates can be modelled by dynamic stochastic processes or an empirical statistical approach. Firstly, it is necessary to clarify the source of data, i.e. the derivation of term structure of interest rates. When dealing with interest rates derivatives, the underlying variable, i.e. interest rates, cannot be traded in the financial market. As a matter of fact, it is only possible to trade interest rate instruments such as bonds, options and swaps. Thus interest rates are derived from the market prices of bonds, options or swaps. Chapter 4 describes one of the term structures of interest rates available in Australia, that is, the generic yield curves produced by the Reserve Bank of Australia (RBA) based on the bonds on issue. Chapter 4 analyses the statistical properties of the yield curves, finding appropriate distributions and possible exogenous innovation variables which can explain the variations of the yield increments. Another constructed yield curves is by the Commonwealth Bank of Australia (CBA) derived from swap rates, which is introduce in Appendix C.

The traditional parameters of term structure of interest rates are the time of maturity T and the time of evolution t, which we refer to as the T-parameterisation. Based on this parameterisation, the terminology of interest rates and interest rate derivatives have been defined, and corresponding theory and models have been developed. They are reviewed in Chapter 2, based on Musiela and Rutkowski (1997) and Hull (2003). In the world of financial markets, media financial reports and existing financial databases, most of the term structure of interest rates data are stored and displayed using two parameters: the length of time to maturity θ and the time of evolution t. We refer to this as the θ-parameterisation. Both yield data sets obtained from the Reserve Bank of Australia and the Commonwealth Bank of Australia were stored in this θ-parameterisation. The θ-parameterisation is mathematically convenient since all yield curves $R(t, \cdot)$ evolve over the same domain $[\theta_{min}, \theta_{max}]$ as time t varies. Brace, Gatarek and Musiela (1997) proposed this θ-parameterisation. They transformed the T-parameterised term structure of interest rates models to the θ-parameterised term structure of interest rates models by a transformation of $T = t + \theta$, and studied the interest rate derivatives pricing by the θ-parameterised term structure of interest rates. We have further developed the θ-parameterisation of term structure of interest rates. Chapter 3 introduces the terminology of interest rates and interest rate derivatives under the θ-parameterisation, and develops corresponding pricing theory, including the no-arbitrage condition, martingale, and term structure of interest rate modelling. This chapter systematically investigates the θ-parameterised term structure of interest rates and their derivatives, which provides a convenient language for statistical analysis of yield curves and modelling in the following chapters.

Reviewing the yield curves theory and diffusion process of yields, Chapter 2 and Chapter 3 for the T-parameterised term structure of interest rates and the θ-parameterised term structure of interest rates respectively, it has been known that volatility modelling is a key aspect in yield curve's modelling (Heath, Jarrow and Morton 1992). Section 3.6 addresses that the mean and variance of interest rate increments in the discrete-time model depend directly on the level of the interest rates in a way consistent with the continuous-time model based on diffusion processes. The GARCH model and its extensions have been widely applied in volatility models over the last twenty years (Engle 1982). Referring to Engle and Patton (2001), Chapter 5 presents a univariate GARCH model of the yield increments, incorporating the indicator variables representing the RBA decisions of lowering and raising the target cash rate. Using Australian Treasury bond yields over the period 1996-2001 provided by the RBA, GARCH models were derived for individual yield increment time series with maturity levels of 1, 2, 3, 5, and 10 years. The empirical results show that the GARCH models capture the most important phenomena of financial yield series. It is observed that the estimated model parameters vary functionally with times to maturity of the yields, especially the degrees of freedom of t-distribution of yield increments linear increase along the times to maturity. This functional dependence incorporates to a new model proposed, a GARCH model of term structure of interest rates. We refer to this model as the *GARCH model of Term Structure of Interest Rates* (TS-GARCH). The estimation of TS-GARCH involves the multivariate GARCH modelling, using the whole data sets of term structure of interest rates (yield increments).

Chapter 6 develops a multivariate GARCH model for term structure of yield increments. The multivariate GARCH model in available literatures and statistical softwares are in normal or multivariate t-distribution. Univariate GARCH models for each individual yield increments in a certain maturity showed, in Chapter 5, that the t-distributions are more appropriate than normal assumption and degrees of freedom of t-distributions of yield increments in different maturity were different. The multivariate asymmetric t-distribution using meta-elliptical distributions concepts (Fang, Fang and Kotz 2002) is defined, that extends and modifies the multivariate asymmetric t-distribution presented in Fang, Fang and Kotz (2002). The multivariate asymmetric t-distribution presented in Fang, Fang and Kotz (2002) was in zero mean and a dispersion matrix specified as a correlation matrix, and the original random vector and the constructed random vector have the same dispersions. For our purpose of modelling the volatility of yield increments, we define the multivariate asymmetric t-distribution with general mean and covariance matrix, and moreover, the original random vector and the constructed random vector have the same variances. With the general multivariate asymmetric t-distribution, a Matrix-Diagonal GARCH(1,1) model (Matrix-Diagonal GARCH-AMt model) is developed that allows the different distributions of marginals. The estimation is given based on Australian Treasury bond yields data provided by the RBA. The diagnostics of the Matrix-Diagonal GARCH-AMt model are presented and the key aspects of the model are discussed such as the volatility mean reversion, volatility persistent, asymmetric impact of lowering and raising target rate, and residuals stationarity. Likelihood ratio tests are used to compare the nested models, Matrix-diagonal GARCH(1,1) with multivariate asymmetric t-distribution (different t-df of marginals) vs. Matrix-diagonal GARCH(1,1) with simple multivariate t-distribution (same t-df of marginals). The Akaike Information Criterion (AIC) is used to measure the goodness of fit of the models, univariate GARCH(1,1) models (The correlation of yield curves is ignored, and less parameters) vs. Matrix-Diagonal GARCH(1,1) with multivariate asymmetric t-distribution (The correlation of yield curves is taken into account, and more parameters).

Chapter 7 builds up the TS-GARCH model, corresponding to the Matrix-Diagonal GARCH model with multivariate asymmetric t-distribution developed in Chapter 6. The estimation of TS-GARCH model is given. The goodness of fit of the model is discussed. The out-of-sample assessment of the model is presented using the TS-GARCH model. Based on the concept of TS-GARCH model proposed in Chapter 5 and the estimators of conditional covariance GARCH model from multivariate Matrix-Diagonal GARCH-AMt modelling of Chapter 6, Chapter 7 extends the TS-GARCH model proposed from Chapter 5 to conditional covariance processes. Estimations of the TS-GARCH model base on multivariate Matrix-Diagonal GARCH-AMt modelling is given. The out-of-sample assessment of TS-GARCH is examined based on diagnostics of the model for both interpolation and forecasting in any possible middle-to-long-term maturity.

1.4 Thesis Outline

The thesis is organised as follows:

Chapter 2 reviews the fundamentals of derivatives pricing theory, the definitions of interest rates and derivatives under the traditional notation, the fundamental spot rate models and term structure of interest rate models.

Chapter 3 presents a θ-parameterisation of the term structure interest rates proposed by Brace and Musiela (1994) and Brace, Gatarek and Musiela (1997), which we will develop further. Starting from the definition and terminology of interest rate and interest rate derivatives under the θ-parameterisation, no-arbitrage condition and martingales are developed upon the θ-parameterisation term structure of interest rates.

Chapter 4 studies the construction of the yield curves and the statistics of the yield curves.

Chapter 5 summarises the properties of yield increments and presents an appropriate GARCH(1,1)-t model for individual yield increment series. A model for the dependence of GARCH parameters on times to maturity is proposed. It is called GARCH model of term structures (TS-GARCH).

Chapter 6 defines the multivariate asymmetric t-distribution and develops a Matrix-Diagonal GARCH model with multivariate asymmetric t-distribution, which allows different marginals. The estimation is given.

Chapter 7 extends the TS-GARCH model proposed in Chapter 5, and the estimation of TS-GARCH is given.

Chapter 8 summarises the thesis. Possible topics of further research are pointed out.

Chapter 2

Fundamentals of Derivatives Pricing and Interest Rates Models

This chapter commences with a summary of the theory of derivatives pricing. This is followed by a review of various concepts of interest rates and derivative securities pricing. It ends with a review of models for spot interest rates and term structure of interest rates as identified in the literature.

Section 1 is a summary of the fundamentals of pricing, including the no-arbitrage and martingale theory that more precisely characterises an arbitrage-free economy from a mathematical perspective. Section 2 reviews the definitions of interest rate and interest rate derivatives including swap and swap pricing. Swap rate and swaps valuation specialise the theory developed in section 1, which will be used for the construction of the CBA yield curves in section 4.3. Section 3 considers a number of spot rate models. Most spot rate models are described by a stochastic differential equation with simple time invariant parameters. These do not have enough degrees of freedom to fit the observed term structure of interest rates. The Hull and White (1990) spot rate model is given by a stochastic differential equation with time changing parameter functions giving increased degrees of freedom to fit the term structure of interest rates. Section 4 reviews the approach of Heath, Jarrow and Morton (HJM)(1992) of modelling forward interest rates, which aims to match the initial yield curve with observed market data. HJM showed that the no-arbitrage condition leads necessarily to a model of forward rates driven by the volatility process only. Concepts such as the risk-neutral martingale measure and the T-forward martingale measure are discussed with the arbitrage-free condition.

Interest rates characterise the bond market. Issued bonds are of two different types: (i) *coupon* or *coupon-bearing* instruments paying interest periodically with a principal amount paid at maturity, and (ii) *zero coupon* or *discount bonds* that are discount securities bearing no coupon and paying only the principal at maturity. Treasury issued bonds with maturity of one year or less are zero-coupon bonds or discount bonds, and Treasury bonds with maturity longer than one year are coupon-bearing bonds. In the United States, zero coupon bonds are called *bills*, and coupon Treasury securities are called *notes* if their maturity at issuance is from 2 to 10 years or *bonds* if maturity at issuance is longer than 10 years. From a pricing point view, coupon-bearing bonds are equivalent to a portfolio of zero coupon bonds. Without loss of generality, we assume that the face value of the bond is 1 dollar and consider zero-coupon bonds only.

We also simply assume that the source of randomness in the markets is generated by a single Brownian motion. The arbitrage-free condition is assumed. Much of the material of this chapter is a review based primarily on the books of Musiela and Rutkowski (1997), Hull (2003) and Moraleda (1997). Some extensions with proof are presented.

2.1 Theory of Derivative Pricing

An *economy* (or a *market*) consists of (d+1) assets with prices $S_0(t), \dots S_d(t)$ at time t, considered on a time interval $[0, T^*]$, where T^* is a time horizon that is fixed once and for all. All assets are continuous predictable processes defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where Ω denotes the sample space of all possible states, ω, of the world, $\mathcal{F}$ is a σ-algebra of measurable events and P is a probability measure. We will assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \le T^*}$ is a completed filtration of a standard Wiener process (W_t), that is $W_1 \sim N(0, 1)$. Intuitively, the filtration $\mathbb{F}$ can be thought of as the flow of information in the economy that is, $\mathcal{F}_t$ includes all information on the economy available at time t. We will also assume that all processes S_i, $i = 0, \dots, d$ are square integrable semi-martingales.

A *trading strategy* is any locally bounded predictable vector process $h(t) = (h^0(t), \cdots, h^k(t))$, where $h^i(t)$ is the number of units of asset i (eg. currencies, bonds, stock indices) held in the portfolio at time t. The *value process* corresponding to the trading strategy h is defined by

$$V_t(h) = \sum_{i=0}^{k} h^i(t) S_i(t).$$

Among all trading strategies we are interested in those that are *self-financing*, which requires that

$$V_t(h) = V_0(h) + \sum_{i=0}^{k} \int_0^t h^i(s) dS_i(s), \quad t \leq T,$$

where $T \leq T^*$. Therefore, the changes in the portfolio value in a self-financing trading strategy are only due to changes in the values of assets and not because of net buying or selling.

So far all prices are given in terms of the local currency. Relative prices are given in terms of a *numeraire asset*, $S_0(t)$, which is a (non-dividend paying) traded asset with strictly positive values over the entire trading horizon $t \in [0, T]$. A *relative price* is defined as

$$Z_i(t) = S_i(t)/S_0(t), \quad i = 0, 1, \ldots, d.$$

The value of the first relative asset is 1 at all times, i.e. $Z_0(t) = 1$. The ratio $1/S_0(t)$ is called a *discount factor*. The *relative value process* corresponding to a portfolio h is defined as

$$V_t^Z(h) = V_t(h)/S_0(t).$$

An *arbitrage opportunity* is a self-financing trading strategy h such that $V_0(h) = 0$, $V_T(h) \geq 0$ and $P(V_T(h) > 0) > 0$. An arbitrage opportunity requires no investment at time $t = 0$, its value at time T is non-negative with probability one, and strictly positive with positive probability. It is reasonable to assume that there are no arbitrage opportunities in the economy. Economies without arbitrage opportunities are called "no-arbitrage", "free of arbitrage" or "arbitrage-free". The no-arbitrage assumption is a key feature in option pricing theory. A no-arbitrage economy can be characterised in an alternative way using the concept of an equivalent martingale measure. An *equivalent martingale measure* Q is a probability measure on $(\Omega, \mathcal{F})$, such that

1. P and Q are equivalent, that is $P(A) = 0$ if and only if $Q(A) = 0$, for every $A \in \mathcal{F}$.

2. The relative price processes Z_i are Q-martingales, that is

   $$E^Q(Z_i(s)|\mathcal{F}_t) = Z_i(t)$$

   for all $i = 0, 1, \cdots, k$ and all $0 \leq t \leq s \leq T$, where E^Q stands for expectation with respect to the measure Q.

A derivative security or a claim Y is *attainable* if there exists a self-financing strategy h such that $V_T(h) = Y$. Such a strategy h is called a *replicating* or *hedging* strategy for a given derivative security Y.

An economy (or market) is said to be *complete* if each derivative security or claim is attainable. The following result, fundamental to the theory of derivatives pricing, shows the importance of the concepts of completeness and equivalent martingale measure. See Harrison and Pliska (1981, 1983) and Musiela and Rutkowski's (1997) Corollary 3.12 and Proposition 3.1.5.

Proposition 2.1 *(I) An economy is arbitrage-free if and only if there exists an equivalent martingale measure Q.*
(II) An arbitrage-free economy is complete if and only if there exists a unique equivalent martingale measure Q.

A *derivative security* or a *contingent claim* is an $\mathcal{F}_T$-measurable random variable Y such that $E^Q(Y) < \infty$. Options, futures, and forwards are examples of derivative securities.

2.2 Interest Rates and Interest Rate Derivatives

In this section we introduce the concepts of interest rates and interest rate derivatives including swaps. We will also discuss pricing of swaps.

A bond is a contract between the issuer (borrower) and the investor (lender) with

- a *face value*, which is the amount that will be paid at the end of loan,
- a *maturity date*, which is the date on which the loan will be repaid to the holder, and
- a *coupon rate*, which is the annual rate (percentage of the bond's face value) of the bond's interest payment.

Because the face value, maturity date, and coupon rate of a bond are normally fixed for the term of the loan, bonds are also known as *fixed-income securities*.

2.2.1 Fixed-Income Securities and Their Derivatives

A *zero coupon bond* with maturity date T, is a contract which guarantees to pay holder \$1 at time T. The price of such an instrument at time t is denoted by $B(t,T)$, with $0 \leq t \leq T \leq T^*$ and $B(T,T) = 1$. We assume that, for any fixed maturity T, the coupon price $B(\cdot, T)$ is a strictly positive and adapted process on a filtered probability space $(\Omega, I\!F, P)$.

The market value of a bond assumed to be *risk free* varies through the loan period. Investors who purchase bonds desire to obtain a certain *yield* (or rate of return) from the investment. The desired yield is affected by the market value of the bond and financial climate. *Yield to maturity* represents the percentage rate of return paid if the security is held until its maturity date. This calculation is based on the coupon rate, the length of time to maturity and the market price of the security.

The *yield-to-maturity* $R(t,T)$ of a zero-coupon bond maturing at time T (called also a *zero-coupon yield*) is an adapted process defined on $(\Omega, I\!F, P)$ by the formula

$$R(t,T) = -\frac{1}{T-t} ln(B(t,T)), \quad \forall t \in [0,T], \tag{2.1}$$

or equivalently,

$$B(t,T) = e^{-R(t,T)(T-t)}, \quad \forall t \in [0,T], \tag{2.2}$$

which implies that the yield-to-maturity $R(t,T)$ corresponds to the fixed rate of return at time t over the period $[t,T]$.

Denoted by $F(t,S,T)$, the *forward interest rate* for a period $[S,T]$, prevailing at time t, $t \leq S \leq T$, satisfies,

$$e^{F(t,S,T)(T-S)} = \frac{B(t,S)}{B(t,T)},$$

which implies

$$F(t,S,T) = -\frac{1}{T-S} \log \frac{B(t,T)}{B(t,S)}.$$

$F(t,S,T)$ corresponds to the rate contracted at time t, on a risk-free loan over the future period time [S,T].

The *instantaneous forward interest rate*, denoted by $f(t,T)$, is the forward interest rate prevailing at the date t, $t \leq T$, for instantaneous risk-free borrowing or lending at date T. It can be interpreted as the interest rate over the infinitesimal time interval $[T, T+dT]$ as seen from time t. The instantaneous forward interest rate is a mathematical idealisation rather than a quantity observable in practice. Given a family of instantaneous forward interest rates $f(t,T)$, the zero coupon is

$$B(t,T) = \exp\left(-\int_t^T f(t,u)du\right), \quad \forall t \in [0,T], \tag{2.3}$$

provided that the integral on the right side of (2.3) exists for almost all ω.

On the other hand, if the family of zero coupon bonds $B(t,T)$ is sufficiently smooth with respect to maturity T, the implied instantaneous forward interest rate $f(t,T)$ is given by

$$f(t,T) = -\frac{\partial \log B(t,T)}{\partial T}. \tag{2.4}$$

This formula may be taken as a definition of instantaneous forward rate $f(t,T)$. It is easy to see that $f(t,T)$ is a limit of forward interest rates $F(t,S,T)$ as $S \uparrow T$. The *instantaneous forward interest rate* will be called the *forward rate* later, because typical models of interest rates are formulated in terms of instantaneous forward rates.

Leaving the technical assumptions aside, from either one of $B(t,T)$, $R(t,T)$ or $f(t,T)$, we can recover the others. The *term structure of interest rates*, known also as the *yield curve*, is the function that relates the yield $R(0,T)$ and the time to maturity T.

The *instantaneous spot interest rate* or *short rate*, or simply *spot rate*, $r(t)$ is defined as

$$r(t) = f(t,t)$$

if it exists. Spot rate is the risk-free borrowing or lending rate prevailing at time t over the infinitesimal time interval $[t, t+dt]$.

Suppose the dynamics for spot rate and forward rate are given by the following stochastic differential equations:

$$df(t,T) = \mu_f(t,T)dt + \sigma_f(t,T)dW_t,$$

$$dr(t) = \mu_r(t)dt + \sigma_r(t)dW_t,$$

where (W_t) is a standard Brownian motion. We will assume that all three functions $f(0,\cdot)$, μ_f and σ_f are jointly continuous together with their t derivatives. Then the definition $r(t) = f(t,t)$ yields the following relations:

$$\begin{cases} \mu_r(t) = \frac{df(0,t)}{dt} + \mu_f(t,t) + \int_0^t \frac{\partial \mu_f}{\partial t}(s,t)ds + \int_0^t \frac{\partial \sigma_f}{\partial t}(s,t)dW_s \\ \sigma_r(t) = \sigma_f(t,t) \end{cases}, \tag{2.5}$$

which means that the dynamics of the spot rate can be derived from the dynamics of the forward rate.

Although an explicit expression for μ_f and σ_f cannot be derived from μ_r and σ_r, the following relation holds. Let $B(t,T)$ be the time t price of a zero-coupon with maturity at time T. Then the spot rate and zero coupon bonds are related by

$$B(t,T) = E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]$$

E^Q denotes expectation with respect to an appropriate equivalent martingale measure. Taking into account that

$$f(t,T) = -\frac{\partial \log B(t,T)}{\partial T},$$

we have found that the evolution of forward interest rate $f(t,T)$ is completely determined by the spot interest rate process $\{r(t) : t \leq T\}$. It follows that the modelling of the spot rate and of the forward rate (modelling of the term structure of interest rates modelling) are, given some technical assumptions, equivalent.

2.3 Spot Interest Rate Models

In this section we review some popular spot rate models. We concentrate on the spot rate processes r_t that are solutions of diffusion equations with time independent coefficients:

$$dr_t = \mu(r_t)dt + \sigma(r_t)dW_t, \tag{2.6}$$

where the functions μ and σ are, respectively, the *drift* and the *diffusion* of the process. W_t is a standard Brownian motion under the real measure P.

The choice of drift and diffusion in (2.6) is often arbitrary. Despite an array of models, relatively little is known about how these models compare in terms of their ability to capture the actual behaviour of the spot rate and the accuracy of pricing the interest rate derivatives. Table 2.1 below presents a list of spot interest rate models that are popular and often referred to in the literature.

In Table 2.1, Model 1 (Merton 1973), is a stochastic process with a simple Brownian motion. Model 2, used by Vasicek (1977), is an Ornstein-Uhlenbeck process. Volatility is constant in both models 1 and 2. Model 3 is a square root process proposed by Cox, Ingersoll and Ross (CIR) (1985). Model 4 is used by Dothan (1978). Model 5 is a

Table 2.1: List of Spot Rate Models

Model	Reference	Drift function $\mu(r)$	Diffusion function $\sigma(r)$
1	Merton (1973)	α	σ
2	Vasicek (1977)	$\alpha + \beta r$	σ
3	CIR SR (1985)	$\alpha + \beta r$	$\sigma r^{1/2}$
4	Dothan (1978)		σr
5	GBM(Black and Scholes 1973)	βr	σr
6	Brennan and Schwartz(1980)	$\alpha + \beta r$	σr
7	CIR VR (1980)		$\sigma r^{3/2}$
8	CEV (Cox and Ross 1976)	βr	σr^{γ}
9	Chan (1992)	$\alpha + \beta r$	$\sigma r^{\gamma}(\gamma \geq 1)$
10	Ait-Sahalia (1996b)	$\alpha_0 + \alpha_1 r + \alpha_2 r^2 + \alpha_3/r$	$\beta_0 + \beta_1 r + \beta_2 r^{\gamma}$
11	Hull and White (1990)	$\theta(t) - \beta r$	σ
12	Hull and White (1990)	$\theta(t) - \beta r$	$\sigma\sqrt{r}$
13	Ait-Sahalia (1996a)	$\alpha + \beta r$	$\sigma(t)$

Geometric Brownian Motion (GBM) of Black and Scholes (1973). Model 6 is used by Brennan and Schwartz (1980) and also by Courtadon (1982).

The first six models in the table above assume a linear mean reverting drift of the form $\mu(r_t) = \alpha + \beta r_t$, while some specifically assume $\alpha = 0$ or $\beta = 0$. It has been recognised in the finance literature that one of the most important problems in financial modelling is the specification of the diffusion function $\sigma(\cdot)$ correctly. In the first six models the diffusion functions are assumed either to be constant, or proportional to the square root of the spot rate or to the spot rate itself. In model 7, introduced by Cox, Ingersoll and Ross (CIR) (1980), the diffusion depends on spot rate with a power of three-halves. Model 8 is the Constant Elasticity of Variance (CEV) process introduced by Cox and Ross (1976).

In order to estimate and compare a variety of continuous-time models of the spot interest rate using Generalised method of moments, Chan et al. (1992) proposed model 9 that provides a common framework in which models 1-8 in Table 2.1 can be nested. This is given by

$$\mu(r_t) = \alpha + \beta r, \quad \sigma(r_t) = \sigma r_t^{\gamma}. \tag{2.7}$$

Using annualised one-month US Treasury bill yield data from 1964 to 1989, he found that the most successful models in capturing the dynamics of short-term interest rates are those that allow the volatility of interest rate changes to be highly sensitive to the level of the spot rate. This means that the models with $\gamma \geq 1$ capture the dynamics of the spot rate better than those with $\gamma < 1$. Chan's unconstrained estimate of γ is 1.5.

Ait-Sahalia (1996b) tests parametric models by comparing their parametric density to the same density estimated non-parametrically, and a Generalised parametric specification of the spot rate model (model 10) was given which nested the single-factor diffusion models 1-9 in Table 2.1. The general model is given by

$$\mu(r, \theta) = \alpha_0 + \alpha_1 r + \alpha_2 r^2 + \alpha_3/r, \quad \sigma(r, \theta) = \beta_0 + \beta_1 r + \beta_2 r^{\gamma}, \tag{2.8}$$

where

$\beta_0 \geq 0$ (and $\beta_2 > 0$ if $\beta_0 = 0$ and $0 < \gamma < 1$, or $\beta_1 > 0$ if $\beta_0 = 0$ and $\gamma > 1$),

$\beta_2 > 0$ if either $\gamma > 1$ or $\beta_1 = 0$, and $\beta_1 > 0$ if either $0 < \gamma < 1$ or $\beta_2 = 0$,

$\alpha_2 \leq 0$ and $\alpha_1 < 0$ if $\alpha_2 = 0$,

$\alpha_3 > 0$ and $2\gamma \geq \beta_0 \geq 0$, or $\alpha_3 = 0$, $\alpha_0 > 0$,$\beta_0 = 0$,$\gamma > 1$ and $2\alpha_0 \geq \beta_1 > 0$.

These natural restrictions are necessary for σ^2 to be positive in the neighbourhood of zero boundary, to ensure that the drift is mean-reverting at high interest rate values, and to guarantee that σ is not zero.

Using the 7-day Eurodollar deposit rate 1973-1995, Ait-Sahalia empirically tested several models in 2.8. He found the strong non-linearity of the drift and higher volatility when away from mean, which rejected the existing liner drift models and constant elasticity of variance (CEV) model.

Models 1-10 in Table 2.1 have time-independent drift and diffusion and therefore do not have enough degrees of freedom to fit the complete term structure of interest rates (including the observed initial yield curve). Models 11-13 belong to the class denoted as exogenous term structure models. They can be accurately fitted to the initial yield curve as observed in the market through the time-dependent parameter function in their drift. It is desirable for an interest rate model that it can match the initial term structure of interest rate or initial yield curve. Another important property for an interest rate model is that it can be used to price derivative securities. Hull and White (1990)'s spot rate model is specified by a stochastic differential equation with time changing parameter functions, which allows more freedom to fit the term structure of interest rates. The Hull and White model, given by a stochastic differential equation, is

$$dr_t = [\theta(t) - \beta(t)r_t]dt + \sigma(t)r_t^{\rho}dW_t, \tag{2.9}$$

where $\theta(t)$, $\beta(t)$ and $\sigma(t)$ are deterministic functions, and W_t is a Brownian Motion. This spot rate model exogenously incorporates both the initial structure of interest rates and the derivative security pricing. Also, the Hull and White (1990) model is Generalised to the model 1-10 in Table 2.1, and it has an affine term structure when $\rho = 0$ or $\rho = 0.5$ that corresponds to model 11 (the extended Vasicek model) or model 12 (the extended CIR model) respectively. By an affine term structure model we mean a diffusion model with the drift and square of the diffusion term described by linear functions.

Ait-Sahalia (1996a) proposed a nonparametric estimation procedure for a continuous - time stochastic model to price interest rate derivatives, based on model 13 from Table 2.1. Because pricing of derivative securities depends crucially on the form of the instantaneous volatility of the underlying process, he did not assume a model for $\sigma(t)$ but rather estimated it non-parametrically. The drift and volatility functions are forced to match the stationary density π of the process, which is assumed to exist. The procedure for estimation he used is : Assuming $\mu(r_t) = \alpha + \beta r_t$, firstly, estimate α, β

by ordinary least square(OLS); Secondly, calculate the non-parametric Kernel estimator of the density $\pi(r)$ and, finally, using well known formulae for the stationary density of one-dimensional diffusion, put $\sigma^2(r) = \frac{2}{\pi(r)} \int_0^r \mu(u, \theta)\pi(u)du$. Sahalia developed this procedure taking the discrete character of the data into account while maintaining the attractiveness of the continuous time model, and estimated non-parametrically the diffusion function. The spot rate data used in his article is seven-day Eurodollar deposit rate, bid-ask midpoint, from Bank of America (1/6/73 — 25/2/95). Discount bonds and options on discount bonds were computed by Sahalia's nonparametric model, compared with CIR and Vasicek model.

2.4 Modelling of the Term Structure of Interest Rates

Most spot rate models, like models 1-10 in Table 2.1, are described by a stochastic differential equation with simple, time invariant parameters. These spot rate models do not have enough degrees of freedom to obtain realistic behaviour for the term structure of interest rates. Some denoted as exogenous spot rate models, like the Hull and White model, can fit the initial term structure as observed in the market through the complicated time-dependent parameter functions. However, estimation of general time-dependent parameter functions is difficult.

In recent years, the research focus has been on the modelling of the term structure of interest rates that is much more appropriate to capture the yield curves movements. Moraleda and Pelsser (1997) presented an empirical comparison of term structure models by forward interest rates versus spot interest rates. A significant contribution to term structure modelling has been made by Heath, Jarrow and Morton (HJM). In this section, we review the Heath, Jarrow and Morton (HJM) (1992) model, which allows the modelling of the entire term structure at once. The HJM approach for modelling forward interest rates matches perfectly the initial yield curve as observed in the markets. With this approach, HJM showed that correct volatility modelling is the key point of the interest rate modelling. Concepts such as the risk-neutral martingale measure and the T-forward martingale measure are reviewed with the arbitrage-free condition.

2.4.1 HJM Methodology and Risk-Neutral Measure P^*

As defined in Section 2.2, the forward rate $f(t,T)$ is a strictly positive adapted process on a filtered probability space $(\Omega, \mathbb{F}, P)$. Suppose the term structure modelling is based on an exogenous specification of dynamics of instantaneous forward rates $f(t,T)$. We assume that, for any fixed maturity time $T \leq T^* < \infty$, the dynamics of the forward rate $f(t,T)$ are described by

$$df(t,T) = \mu_f(t,T)dt + \sigma_f(t,T)dW(t), \tag{2.10}$$

where $\mu_f(\cdot,T)$ and $\sigma_f(\cdot,T)$ are real valued $(\mathcal{F}_t)$ adapted stochastic processes with $t \in [0,T]$ and $\sigma_f > 0$,

$$\sup_{t,T\leq T^*} (|\mu_f(t,T) + |\sigma_f^2(t,T)|^2) < \infty.$$

Consequently, using formula (2.3) and the Ito formula, it can be shown that the process of zero-coupon bond prices $B(t,T)$ is described by

$$dB(t,T) = \mu_B(t,T)B(t,T)dt + \sigma_B(t,T)B(t,T)dW(t). \tag{2.11}$$

More precisely, we have the following lemma.

Lemma 2.1 *The relation of the dynamics are*
(I) if $f(t,T)$ satisfies (2.10), and the functions $T \to \mu_f(t,T)$ and $T \to \sigma_f(t,T)$ are locally integrable for every t, then the dynamics of discount bonds $B(t,T)$ are given by (2.11) with

$$\begin{cases} \mu_B(t,T) = r(t) - \int_t^T \mu_f(t,s)ds + \frac{1}{2}(\int_t^T \sigma_f(t,s)ds)^2, \\ \sigma_B(t,T) = -\int_t^T \sigma_f(t,s)ds; \end{cases} \tag{2.12}$$

(II) conversely, by (2.4), if $B(t,T)$ satisfies (2.11), and the functions $T \to \mu_B(t,T)$ and $T \to \sigma_B(t,T)$ are continuously differentiable for every t, then the dynamics of forward-rates $f(t,T)$ are given by (2.10) with

$$\begin{cases} \mu_f(t,T) = \sigma_B(t,T)\frac{\partial\mu_B(t,T)}{\partial T} - \frac{\partial\mu_B(t,T)}{\partial T}, \\ \sigma_f(t,T) = -\frac{\partial\sigma_B(t,T)}{\partial T}. \end{cases} \tag{2.13}$$

The proof of this Lemma can be found in Musiela and Rutkowski (1997).

The fundamental theorems for an arbitrage-free economy are stated in, for example, Musiela and Rutkowski (1997) and are briefly reviewed here.

Theorem 2.1 *The market is arbitrage-free (or, no-arbitrage) if and only if the relative bond price* $Z(t,T) = \frac{B(t,T)}{B(t)}$*, for any fixed* T*,* $T > 0$*, is a martingale under the unique martingale measure* P^* *equivalent to the actual measure* P*, where* $B(t) = e^{\int_0^t r(u)du}$ *is the risk-free asset (being a numeraire) and* $r(t)$ *is a spot rate.*

The measure P^* is called the *risk-neutral measure.* The Radon-Nikodym derivative of measure P^* with respect to P is given in the next theorem.

Theorem 2.2 *The relative bond price* $Z(t,T) = \frac{B(t,T)}{B(t)}$ *is a martingale under the risk-neutral measure* P^**equivalent to* P*, if and only if there exists an adapted process* $\varphi(t)$*,* $0 \le t \le T^*$*, which is independent of* $T \le T^*$*, satisfying*

$$\mu_B(t,T) + \varphi(t)\sigma_B(t,T) - r(t) = 0, \tag{2.14}$$

or equivalently satisfying

$$\mu_f(t,T) = \sigma_f(t,T)\left(-\varphi(t) + \int_t^T \sigma_f(t,s)ds\right). \tag{2.15}$$

The unique risk-neutral measure P^**, equivalent to the actual measure* P*,* $P^* \sim P$ *on* $\mathcal{F}_T$*, is given by*

$$P^*(A) = \int_A Z(T)dP, \quad A \in \mathcal{F}_T$$

where

$$Z(t) = \exp\left\{\int_0^t \varphi(u)dW(u) - \frac{1}{2}\int_0^t \varphi^2(u)du\right\},$$
$$\mathbb{E}_P(Z(T)) = 1.$$

And the process

$$W_t^* = W_t - \int_0^t \varphi(u)du$$

is a Brownian motion under the risk-neutral measure P^**.*

Proof. The relative bond price is $Z(t,T) = \frac{B(t,T)}{B(t)}$ and the numeraire is $B(t) = e^{\int_0^t r(u)du}$.

Then

$$dB(t) = r(t)B(t)dt,$$

$$\begin{aligned} dZ(t,T) &= \frac{1}{B(t)}dB(t,T) - \frac{B(t,T)}{B^2(t)}r(t)B(t)dt \\ &= \frac{B(t,T)}{B(t)}\left(\frac{dB(t,T)}{B(t,T)} - r(t)dt\right) \\ &\overset{(2.11)}{=} Z(t,T)[(\mu_B(t,T) - r(t))dt - \sigma_B(t,T)dW_t, \end{aligned}$$

Thus $Z(t,T) = \frac{B(t,T)}{B(t)}$ is a martingale under the equivalent martingale measure P^*, if and only if there exists the function $\varphi(t)$ satisfying

$$\varphi(t) = -\frac{\mu_B(t,T) - r(t)}{\sigma_B(t,T)},$$

where $\varphi(t)$ is independent of T. By the relation of bond price and forward rates (lemma 2.1), we have

$$-\int_t^T \mu_f(t,s)ds + \frac{1}{2}(\int_t^T \sigma_f(t,s)ds)^2 - \varphi(t)\int_t^T \sigma_f(t,s)ds = 0,$$

and differentiating with respect to T yields,

$$\mu_f(t,T) = \sigma_f(t,T)\left(-\varphi(t) + \int_t^T \sigma_f(t,s)ds\right).$$

Thus, we get the equation of (2.15), which means, in an arbitrage-free market, the drift coefficient of forward rates is uniquely determined by the volatility coefficient and a stochastic process interpreted as the *market risk premium* $\varphi(t)$. ■

The next theorem was proved initially by Heath Jarrow and Morton (1992), see also Musiela and Rutkowski (1997).

Theorem 2.3 (HJM) *Under the risk-neutral measure P^*, the stochastic differential equation for $f(t,T)$, obtained for every maturity T, is*

$$df(t,T) = \sigma_f(t,T)\sigma_f^*(t,T)dt + \sigma_f(t,T)dW_t^*, \tag{2.16}$$

and stochastic differential equation for $B(t,T)$, obtained for every maturity T, is

$$dB(t,T) = r(t)B(t,T)dt - \sigma_f^*(t,T)B(t,T)dW_t^*, \tag{2.17}$$

where W_t^* is a Brownian motion under the risk-neutral measure P^*, and

$$\sigma_f^*(t,T) = \int_t^T \sigma_f(t,u)du.$$

The approach described above was proposed in Heath, Jarrow and Morton (1992); it shows that the volatility process of the forward rates plays a crucial role in the dynamics of the term structures. The derivative securities can be calculated by the observable forward rates at the initial time and the volatility process $\sigma_f(t,T)$.

2.4.2 T-forward Martingale Measure and No-Arbitrage

In the previous section, we have introduced the actual measure P and the equivalent risk-neutral martingale measure P^*, $P^* \sim P$, under which all relative bond price processes

$$\frac{B(t,T)}{B(t)}, \quad t \in [0,T],$$

are strictly positive martingales. By Theorem 2.1 and Theorem 2.2, it can be concluded that the market does not admit arbitrage if and only if there exists an adapted process $\varphi(t)$ satisfying (2.14) or (2.15).

Now we focus on the equivalent martingale measure P^* and introduce T-Forward Measures.

Definition 2.1 *An equivalent probability measure $P^T \sim P^*$ on $\mathcal{F}_T$ is called a T-forward measure if*

$$\frac{dP^T}{dP^*} = \frac{1}{B(0,T)B(T)}.$$

This definition is based the observation that, for fixed $T > 0$,

$$\frac{1}{B(0,T)B(T)} > 0, \qquad I\!E_{P^*}[\frac{1}{B(0,T)B(T)}] = 1,$$

and for $t \leq T$, we have

$$I\!E_{P^*}[\frac{dP^T}{dP^*}|\mathcal{F}_t] = \frac{B(t,T)}{B(0,T)B(t)}.$$

Lemma 2.2 *For any $S > 0$,*

$$\frac{B(t,S)}{B(t,T)}, \quad t \in [0, S \wedge T]$$

is a P^T-martingale.

Proof. Let $s \leq t \leq S \wedge T$. Bayes' rule gives

$$\begin{aligned}
I\!E_{P^T}[\frac{B(t,S)}{B(t,T)}|\mathcal{F}_s] &= \frac{I\!E_{P^*}[\frac{1}{B(0,T)B(T)}\frac{B(t,S)}{B(t,T)}|\mathcal{F}_s]}{I\!E_{P^*}[\frac{1}{B(0,T)B(T)}]|\mathcal{F}_s} \\
&= \frac{I\!E_{P^*}[I\!E_{P^*}(\frac{1}{B(0,T)B(t)}\frac{B(t,S)}{B(t,T)}|\mathcal{F}_t)|\mathcal{F}_s]}{\frac{B(t,T)}{B(0,T)B(t)}} \\
&= \frac{I\!E_{P^*}[\frac{B(t,T)}{B(0,T)B(t)}\frac{B(t,S)}{B(t,T)}|\mathcal{F}_s]}{\frac{B(t,T)}{B(0,T)B(t)}} \\
&= \frac{\frac{B(s,S)}{B(s)}}{\frac{B(s,T)}{B(s)}} \\
&= \frac{B(s,S)}{B(s,T)}.
\end{aligned}$$

■

Lemma 2.3 *For any $S > 0$,*

$$\frac{B(t,S)}{B(t,T)}, \quad t \in [0, S \wedge T)],$$

is a P^T-martingale, i.e. the market is no-arbitrage, if and only if, there exists a process $\psi(t)$ satisfying

$$\mu_B(t,S) - \mu_B(t,T) - \sigma_B(t,T)[\sigma_B(t,S) - \sigma_B(t,T)] = -\psi(t)[\sigma_B(t,S) - \sigma_B(t,T)]. \quad (2.18)$$

If such a process ψ exists then

$$\frac{dP^T}{dP} = Z(T),$$

where

$$Z(t) = \exp\left\{\int_0^t \psi(u)dW_u - \frac{1}{2}\int_0^t \psi^2(u)du\right\}, \quad t \in [0,T],$$

and

$$W_t^T = W_t - \int_0^t \psi(s)ds$$

is a Brownian motion under the forward measure P^T. By (2.12) the equivalence of bond and forward rates, (2.18) can be expressed as

$$\psi(t) = \frac{\int_t^T \mu_f(t,s)ds - \int_t^S \mu_f(t,s)ds}{\int_t^T \sigma_f(t,s)ds - \int_t^S \sigma_f(t,s)ds} - \frac{1}{2}(\int_t^T \sigma_f(t,s)ds - \int_t^S \sigma_f(t,s)ds). \quad (2.19)$$

Proof. Let

$$F_B(t,S,T) = \frac{B(t,S)}{B(t,T)} = \frac{B_1}{B_2} = g(B_1,B_2).$$

By the Ito Lemma we obtain

$$\begin{aligned}
dF_B(t,S,T) &= dg(B_1,B_2) \\
&= \frac{\partial g(B_1,B_2)}{\partial B_1}dB_1 + \frac{\partial g(B_1,B_2)}{\partial B_2}dB_2 \\
&\quad + \frac{\partial^2 g(B_1,B_2)}{2\cdot\partial B_2^2}(dB_2)^2 + \frac{\partial^2 g(B_1,B_2)}{\partial B_1 \partial B_2}dB_1 dB_2 \\
&= \frac{1}{B_2}B_1[\mu_B(t,S)dt + \sigma_B(t,S)dW_t] - \frac{B_1}{B_2^2}B_2[\mu_B(t,T)dt + \sigma_B(t,T)dW_t] \\
&\quad + \frac{2B_1}{2B_2^3}B_2^2\sigma_B^2(t,T)dt - \frac{1}{B_2^2}B_1 B_2 \sigma(t,S)\sigma(t,T)dt \\
&= F_B\{[\mu_B(t,S) - \mu_B(t,T) - \sigma_B(t,T)(\sigma_B B(t,S) - \sigma_B(t,T))]dt \\
&\quad + [\sigma_B(t,S) - \sigma_B(t,T)]dW_t\},
\end{aligned}$$

and $F_B(t,S,T) = \frac{B(t,S)}{B(t,T)} = \frac{B_1}{B_2}$ is a martingale, if and only if, there is a T-independent process $\psi(t)$ such that

$$\mu_B(t,S) - \mu_B(t,T) - \sigma_B(t,T)[\sigma_B(t,S) - \sigma_B(t,T)] = -\psi(t)[\sigma_B(t,S) - \sigma_B(t,T)],$$

and then

$$dF_B(t,S,T) = F_B(t,S,T)[\sigma_B(t,S) - \sigma_B(t,T)]dW_t^T,$$

where

$$W_t^T = W_t - \int_0^t \psi(u)du$$

is a Brownian motion under the measure P^T. By the Girsanov theorem

$$\frac{dP^T}{dP} = \exp\left\{\int_0^t \psi(u)dW_u - \frac{1}{2}\int_0^t \psi(u)du\right\}.$$

■

Chapter 3

θ-parameterisation of Term Structure of Interest Rates

The definitions and the modelling of the term structure of interest rates with parameters of time t and maturity time T are reviewed in Chapter 2. In financial markets, media, financial reports and financial databases, the term structure of interest rates data are often stored and displayed using two other parameters: the time t and the length of time to maturity θ. Both yield data sets obtained from the Reserve Bank of Australia and from the Commonwealth Bank of Australia are stored under (t, θ) format. We will refer to the usual parameterisation of term structure of interest rates with time t and maturity time T as a *T-parameterisation*, and to the parameterisation with time t and length of time to maturity θ as a *θ-parameterisation*. θ-parameterisation is mathematically convenient since all yield curves $R(t, \cdot)$, evolve over the same "space" domain $[\theta_{min}, \theta_{max}]$ as time t varies. θ-parameterisation was proposed in Musiela (1993). We develop the θ-parameterisation approach to the modelling of term structure of interest rates. This chapter introduces the terminology of interest rates and interest rate derivatives under the θ-parameterisation, and develops the corresponding pricing theories, including the no-arbitrage condition and modelling of term structure of interest rates.

The theory and modelling of the term structure of interest rates in the T-parameterisation in this chapter follows the book by Musiela and Rutkowski (1997). We investigate the relationship of the dynamics of term structures under the θ-parameterisation vs. the T-parameterisation, and then systematically investigate the θ-parameterised term structure of interest rates and their derivatives in order to provide a convenient language for statistical θ-parameterisation vs. the (t, T)-parameterisation.

As an application we will consider the volatility of the Australian Treasury bond yields in θ-parameterisation. Using some ideas of Cont (1998) we will demonstrate that the volatility has a maximum around $\theta = 1$. This fact was also reported in other studies (Bouchaud et al. 1997, Cont 2001) using data from European markets.

3.1 Background

A standard way to report the term structure of interest rates is to provide forward rates $f(t, T)$ that depend on two variables, the prevailing time t and the maturity time T. The term structure models and an appropriate non-arbitrage theory have been developed using this parameterisation (See Chapter 2). For any fixed maturity $T \leq T^*$, the stochastic process of the forward rate $f(t, T)$ is described in HJM by a stochastic differential equation

$$df(t, T) = \mu(t, T)dt + \sigma(t, T)dW(t), \tag{3.1}$$

where the functions μ and σ satisfy technical conditions given in Section 2.4.1. HJM showed that, in an arbitrage-free market, the drift coefficient $\mu(t, T)$ of (3.1) is uniquely determined by the volatility process $\sigma(t, T)$ and a stochastic process interpreted as the *market risk premium*. Under the equivalent martingale risk-neutral measure P^*, P^* is independent of $T \leq T^*$, the process of forward rates is

$$df(t, T) = \sigma(t, T)\sigma^*(t, T)dt + \sigma(t, T)dW_t^*, \tag{3.2}$$

where

$$\sigma^*(t, T) = \int_t^T \sigma(t, \tau)d\tau$$

and W_T^* is Brownian motion under the martingale measure P^*. Hence the evolution of the forward curve $f(t, \cdot)$ is completely determined by the initial curve and the volatility structure

$$f(t, T) = f(0, T) + \int_0^t \sigma(u, T)\sigma^*(u, T)dt + \int_0^t \sigma(u, T)dW_u^*. \tag{3.3}$$

By the relation between discount bond price and forward interest rates, see Section.4.1, the stochastic process of the discount bond price $B(t, T)$ is

$$dB(t, T) = r(t)B(t, T)dt - \sigma^*(t, T)B(t, T)dW_t^*. \tag{3.4}$$

The solution of this equation is therefore determined by the initial condition $B(0,T)$, the volatility structure, and the spot rate. The term structure of interest rates or *yield curve* model is the model describing the *evolution of the yield curve.* The term *yield curve* is ambiguous, since it could refer to a zero-coupon curve $B(t,\cdot)$, a zero-coupon yield curve $R(t,\cdot)$, or a forward-rate curve $f(t,\cdot)$. The domain of any yield curve $X(t,S)$ is $S : S \in [t,T]$ that shrinks with time t, which is sometimes mathematically inconvenient. It is also well known that yields are often reported on the market in terms of time to maturity $\theta = T - t$. It was proposed in Musiela (1993) to study the yield curves in terms of parameters (t,θ) instead of (t,T). He defined a re-parameterised forward rate $r(t,\theta) = f(t,t+\theta)$, and bond prices $P(t,\theta) = B(t,t+\theta)$. This parameterisation has the advantage that the term structure process $r(t,\cdot)$ or $P(t,\cdot)$ evolves on the same "space" domain $[\theta_{min},\theta_{max}]$. Later, this approach was successfully applied in deriving the famous BGM model of LIBOR rates in Brace, Gatarek and Musiela (1997). It was shown in BGM that

$$dr(t,\theta) = \frac{\partial}{\partial\theta}[r(t,\theta) + \frac{1}{2}(\sigma^*(t,\theta))^2]dt + \sigma(t,\theta)dW_t^*, \tag{3.5}$$

and

$$dP(t,\theta) = [r(t,0) - r(t,\theta)]P(t,\theta)dt - \sigma^*(t,\theta)P(t,\theta)dW_t^*. \tag{3.6}$$

Due to the fact that the new variable $\theta = T - t$, is a function of the time t, the drift and diffusion coefficients of the models for $r(t,\theta)$ and $P(t,\theta)$ are different from those for $f(t,T)$ and $B(t,T)$. In order to study the dynamics of the term structure of interest rates in θ-parameterisation, the fixed-income securities and derivatives will be defined using θ-parameterisation. Under the θ-parameterisation, the corresponding term structure theory and models will be developed. One advantage of θ-parameterisation, as mentioned above, is that the yield curve movements are defined on the same domain. On a more practical level, the θ-parameterised term structure allows statistical analysis and modelling to be consistent with the data set formatted because, in the real world, most term structure data are stored or formatted in terms of time of recording t and the time to maturity θ. Obviously, we assume here that trading is continuous in time t and time to maturity θ.

3.2 Fixed-Income Securities and Derivatives under θ- parameterisation

This section is devoted to defining various fixed-income securities and derivatives using θ-parameterisation, which depends on the pricing time t and the time to maturity θ. The definitions correspond to these defined in (t, T) format in Section 2.2.1

We define the interest rates on the filtered probability space $(\Omega, \mathbb{F}, P)$ described in Chapter 2.

Definition 3.1 *The price of a zero coupon at time t with maturity in θ years ahead is denoted by $P(t, \theta)$, $t \in [0, T^{**}]$, $\theta \in [0, \theta^*]$ and $P(t, 0) = 1$. T^{**} is a fixed time horizon and θ^* is a fixed number of years, while $T^{**} + \theta^* = T^*$. For any fixed length to maturity θ, $0 \leq \theta < \theta^*$, $P(\cdot, \theta)$ follows a strictly positive and adapted process on a filtered probability space $(\Omega, \mathbb{F}, P)$.*

Definition 3.2 *The forward interest rate in θ years contracted at time t for the period $[t+\theta, t+\theta+\tau]$, over the duration of τ years, is denoted by $F(t, \theta, \tau)$ and defined to satisfy*

$$e^{F(t,\theta,\tau)\tau} = \frac{P(t, \theta)}{P(t, \theta + \tau)}, \tag{3.7}$$

or equivalently

$$F(t, \theta, \tau) = -\frac{1}{\tau} \log \frac{P(t, \theta + \tau)}{P(t, \theta)}. \tag{3.8}$$

The forward interest rate $F(t, \theta, \tau)$ corresponds to the rate of return of a contract agreed at time t for an investment of 1 dollar in θ years over duration of $[t + \theta, t + \theta + \tau]$. To understand this point, we can establish the following portfolio at time t that sells 1 discount bond with maturity in θ years and buys $P(t, \theta)/P(t, \theta + \tau)$ bonds with maturity in $\theta + \tau$ years. Thus the net investment is zero at time t because $1 \cdot P(t, \theta)$ is received and $P(t, \theta)P(t, \theta + \tau)/P(t, \theta + \tau)$ is paid. This strategy implies the contract issued at time t that invests 1 dollar in θ years and yield $\frac{P(t,\theta)}{P(t,\theta+\tau)}$ dollars in $\theta + \tau$ years. So the continuously compounded rate of return of this strategy is (3.7).

Definition 3.3 *The yield-to-maturity $R(t, \theta)$ on a zero-coupon bond with maturity in θ years is defined by the formula*

$$P(t, \theta)e^{R(t,\theta)\theta} = 1, \quad t, \theta \geq 0, \tag{3.9}$$

which implies

$$P(t, \theta) = e^{-R(t,\theta)\theta}, \tag{3.10}$$

$$R(t, \theta) = -\frac{1}{\theta} \log P(t, \theta). \tag{3.11}$$

The yield-to-maturity $R(t, T)$ corresponds to the rate of return of an investment at time t over the period θ years. By (3.8), (3.11) and $P(t, 0) = 1$, we also have $R(t, \theta) = F(t, 0, \theta)$.

Definition 3.4 *The instantaneous forward interest rate, $r(t, \theta)$, prevailing at time t, for the interest rate in θ years, $t \geq 0$, $\theta \geq 0$, is defined as*

$$r(t, \theta) = -\frac{\partial \log P(t, \theta)}{\partial \theta}, \tag{3.12}$$

or

$$r(t, \theta) = -\frac{1}{P(t, \theta)} \frac{\partial P(t, \theta)}{\partial \theta}, \tag{3.13}$$

provided the family of zero coupon bonds $P(t, \theta)$ is sufficiently smooth with respect to maturity θ.

We assume that $r(t, \theta)$ is locally bounded,

$$\sup_{t,\theta \geq 0} |r(t, \theta)| < \infty, \quad I\!\!P - a.s.$$

If $\partial P(t, \theta)/\partial \theta$ exists and is continuous, then there exists the limit of forward interest rate $F(t, \theta, \tau)$ as $\tau \downarrow 0$, and by equation (3.8), we have

$$r(t, \theta) = \lim_{\tau \to 0} F(t, \theta, \tau),$$

which means that the instantaneous forward interest rate corresponds to the forward interest rate prevailing at time t, for the interest rate in θ years.

The relationship between discount bonds and instantaneous forward rates follows from solving the differential equation (3.12) as

$$P(t, \theta) = \exp\{-\int_0^\theta r(t, \tau) d\tau\}. \tag{3.14}$$

We simply denote *instantaneous forward interest rate* as *forward rate*, because almost all interest rate models in the literature consider instantaneous rates.

Under some technical assumptions, on either $P(t,\theta)$, $R(t,\theta)$ or $r(t,\theta)$, we can recover the others. Each of them characterises interest rates as a function of time and time to maturity. This function, in any of its possible forms, is known as the *term structure of interest rates* or *yield curve.*

Definition 3.5 *The spot interest rate, or short rate,* $r(t)$*, is defined as*

$$r(t) = r(t,0). \tag{3.15}$$

3.3 Modelling of the θ-parameterised Term Structure of Interest Rates

In analogy to the HJM modelling of the T-parameterised term structure of interest rates, the θ-parameterised term structures modelling is supposed to be based on an exogenous specification of dynamics of instantaneous forward rates $r(t,\theta)$. We assume that for any fixed $\theta \in (0,\theta^*)$, the dynamics of the forward rate $r(t,\theta)$ is described by the equation

$$dr(t,\theta) = \mu_r(t,\theta)dt + \sigma_r(t,\theta)dW(t), \tag{3.16}$$

where (W_t) is Brownian motion with respect to the measure P, $\mu_r(\cdot,\theta)$ and $\sigma_r(\cdot,\theta)$ are real valued $(\mathcal{F}_t)$ adapted stochastic process, $t,\theta \geq 0$, $\sigma_r(t,\theta) > 0$, and

$$\int_0^t |\mu_r(u,\theta)|du + \int_0^t \sigma_r^2(u,\theta)du < \infty, \quad P - a.s. \tag{3.17}$$

$\mu(t,\theta)$ is called *drift function* and $\sigma(t,\theta)$ is called *volatility* function or diffusion function.

Another possible starting point for building a term structure model is to assume that for any θ the process of bond prices $P(t,\theta)$ is described by

$$dP(t,\theta) = \mu_P(t,\theta)P(t,\theta)dt + \sigma_P(t,\theta)P(t,\theta)dW(t). \tag{3.18}$$

The following lemma establishes, under some technical conditions, the equivalence between the dynamics (3.16) of forward rates $r(t,\theta)$ and the dynamics (3.18) of zero coupon prices $P(t,\theta)$.

Lemma 3.1 *The relations of the dynamics of forward rate* $r(t,\theta)$ *and zero coupon bond price* $P(t,\theta)$ *are as follows:*

(I) If $r(t,\theta)$ satisfies (3.16) and (3.17) holds then the discount bond dynamics are given by (3.18) with

$$\begin{cases} \mu_P(t,\theta) = -\int_0^\theta \mu_r(t,\tau)d\tau + \frac{1}{2}\left(\int_0^\theta \sigma_r(t,\tau)d\tau\right)^2, \\ \sigma_P(t,\theta) = -\int_0^\theta \sigma_r(t,\tau)d\tau. \end{cases} \tag{3.19}$$

Putting

$$\begin{cases} \mu^*(t,\theta) = \int_0^\theta \mu_r(t,\tau)d\tau, \\ \sigma^*(t,\theta) = \int_0^\theta \sigma_r(t,\tau)d\tau, \end{cases} \tag{3.20}$$

we obtain equivalent formulae for μ_P and σ_P:

$$\begin{cases} \mu_P(t,\theta) = -\mu^*(t,\theta) + \frac{1}{2}(\sigma^*(t,\theta))^2, \\ \sigma_P(t,\theta) = -\sigma^*(t,\theta). \end{cases} \tag{3.21}$$

(II) Let $P(t,\theta)$ satisfy (3.18), and assume $\mu_P(t,\theta)$ and $\sigma_P(t,\theta)$ are continuously differentiable in θ and

$$\int_0^t \left(\left|\frac{\partial \mu_P}{\partial\theta}(s,\theta)\right| + \left|\frac{\partial \sigma_P}{\partial\theta}(s,\theta)\right|^2 \right) ds < \infty, \quad P-a.s.$$

Then the forward rate dynamics are given by (3.16) with

$$\begin{cases} \mu_r(t,\theta) = -\frac{\partial\mu_P(t,\theta)}{\partial\theta} + \sigma_P(t,\theta)\frac{\partial\sigma_P(t,\theta)}{\partial\theta}, \\ \sigma_r(t,\theta) = -\frac{\partial\sigma_P(t,\theta)}{\partial\theta}. \end{cases} \tag{3.22}$$

Proof.

(I) Since

$$P(t,\theta) = \exp\{-\int_0^\theta r(t,\tau)d\tau\},$$

by the Itô formula in Appendix B, we have

$$dP(t,\theta) = -e^{-\int_0^\theta r(t,\tau)d\tau} \cdot d\left(\int_0^\theta r(t,\tau)d\tau\right) + \frac{1}{2}e^{-\int_0^\theta r(t,\tau)d\tau} d\left\langle \int_0^\theta r(\cdot,\tau)d\tau \right\rangle_t$$

Thus by (3.16)

$$\begin{aligned} \frac{dP(t,\theta)}{P(t,\theta)} &= -\left(\int_0^\theta \mu_r(t,\tau)d\tau\right)dt - \left(\int_0^\theta \sigma_r(t,\tau)d\tau\right)dW_t + \frac{1}{2}\left(\int_0^\theta \sigma_r(t,\tau)d\tau\right)^2 dt \\ &= \left(-\int_0^\theta \mu_r(t,\tau)d\tau + \frac{1}{2}(\int_0^\theta \sigma_r^2(t,\tau)d\tau)^2\right)dt - \left(\int_0^\theta \sigma_r(t,\tau)d\tau\right)dW_t. \end{aligned}$$

So

$$\begin{aligned}\mu_P(t,\theta) &= -\int_0^\theta \mu_r(t,\tau)d\tau + \frac{1}{2}(\int_0^\theta \sigma_r^2(t,\tau)d\tau)^2 \\ &= -\alpha^*(t,\theta) + \frac{1}{2}(\sigma^*(t,\theta))^2.\end{aligned}$$

$$\begin{aligned}\sigma_P(t,\theta) &= -\int_0^\theta \sigma_r(t,\tau)d\tau \\ &= \sigma^*(t,\theta).\end{aligned}$$

(II)

$$r(t,\theta) = -\frac{\partial \log P(t,\theta)}{\partial\theta},$$

$$\begin{aligned}dr(t,\theta) &= -\frac{\partial}{\partial\theta}[d\log P(t,\theta)] \\ &\overset{It\hat{o}}{=} -\frac{\partial}{\partial\theta}\left[\frac{1}{P(t,\theta)}dP(t,\theta) - \frac{1}{2}\frac{1}{P^2(t,\theta)}d\left\langle P(\cdot,\theta)\right\rangle_t\right] \\ &= -\frac{\partial}{\partial\theta}\left[\mu_p dt + \sigma_P dW - \frac{1}{2}\sigma_P^2 dt\right] \\ &= \left[-\frac{\partial\mu_p}{\partial\theta} + \sigma_P\frac{\partial\sigma_P}{\partial\theta}\right]dt - \frac{\partial\sigma_P}{\partial\theta}dW_t.\end{aligned}$$

So the forward rate $r(t,\theta)$ dynamics are given by

$$dr(t,\theta) = \mu_r(t,\theta)dt + \sigma_r(t,\theta)dW(t),$$

with

$$\begin{cases}\mu_r(t,\theta) = -\frac{\partial\mu_P(t,\theta)}{\partial\theta} + \sigma_P(t,\theta)\frac{\partial\sigma_p}{\partial\theta}, \\ \sigma_r(t,\theta) = -\frac{\partial\sigma_P(t,\theta)}{\partial\theta}.\end{cases}$$

■

This lemma shows that, under the θ-parameterisation, the drift and volatility of zero coupon prices $P(t,\theta)$ can be recovered from the drift and volatility of forward rate prices $r(t,\theta)$. It is a simpler result than the relationship (2.12) between the dynamics of T-parameterised forward rates $f(t,T)$ and those of zero coupon prices $B(t,T)$ which involves the spot rate process $r(t)$.

Since short rate $r(t) = r(t,0)$, the dynamics of short rates can be described by

$$dr(t) = \mu_r(t)dt + \sigma_r(t)dW(t),$$

where

$$\begin{cases} \mu_r(t) = \mu_r(t,0), \\ \sigma_r(t) = \sigma_r(t,0), \end{cases} \tag{3.23}$$

which shows that, under the θ-parameterisation, the drift and volatility of spot rates $r(t)$ can be recovered from the drift and volatility of forward rates $r(t,\theta)$. It is a much simpler result than the relationship (2.5) between the dynamics of spot rates $f(t)$ and those of T-parameterised forward rates $f(t,T)$ that involves the $\frac{df(0,t)}{dt}$, $\int_0^t \frac{\partial \mu_f}{\partial t}(s,t)ds$ and $\int_0^t \frac{\partial \sigma_f}{\partial t}(s,t)dW_s$.

3.4 Relationships of the Dynamics of the Term Structures Models: θ-parameterised term structures and T- parameterised term structures

Martingale measures and no-arbitrage conditions based on T-parameterised term structures were reviewed at Section 2.4. Accordingly, we are going to investigate the martingale measures and no-arbitrage conditions with θ-parameterised term structures in the next section. Before doing so, we will study the relationships between the dynamics of θ-parameterised term structure and T-parameterised term structure models in this section.

Let us recall that

$B(t,T)$ is the zero coupon prices priced at time t and with maturity T,

$f(t,T)$ is the forward rate prevailing at time t for interest rate borrowing at time T,

$P(t,\theta) = B(t,t+\theta)$ is the zero coupon price at time t and with maturity at $t+\theta$, and

$r(t,\theta) = f(t,t+\theta)$ is the forward rate prevailing at time t for interest rate borrowing in θ years after time t.

The dynamics of T-parameterised forward rates $f(t,T)$ and zero coupon prices $B(t,T)$ are described by

$$df(t,T) = \mu_f(t,T)dt + \sigma_f(t,T)dW_t,$$

$$dB(t,T) = \mu_B(t,T)B(t,T)dt + \sigma_B(t,T)B(t,T)dW_t.$$

With $\theta = T - t$, we let

$$\mu_f(t,T) = \mu_f(t, t+\theta) \stackrel{def}{=} \tilde{\mu}_f(t,\theta), \tag{3.24}$$

$$\sigma_f(t,T) = \sigma_f(t, t+\theta) \stackrel{def}{=} \tilde{\sigma}_f(t,\theta), \tag{3.25}$$

$$\mu_B(t,T) = \mu_B(t, t+\theta) \stackrel{def}{=} \tilde{\mu}_B(t,\theta), \tag{3.26}$$

$$\sigma_B(t,T) = \sigma_B(t, t+\theta) \stackrel{def}{=} \tilde{\sigma}_B(t,\theta). \tag{3.27}$$

Thus, the dynamics of $f(t,T)$ and $B(t,T)$ can be expressed as

$$df(t,T) = \tilde{\mu}_f(t,\theta)dt + \tilde{\sigma}_f(t,\theta)dW_t,$$

$$dB(t,T) = \tilde{\mu}_B(t,\theta)B(t,\theta)dt + \tilde{\sigma}_B(t,\theta)B(t,T)dW_t.$$

Correspondingly, the dynamics of θ-parameterised forward rates $r(t,\theta)$ and zero coupon prices $p(t,\theta)$ are described by

$$dr(t,\theta) = \mu_r(t,\theta)dt + \sigma_r(t,\theta)dW_t,$$

$$dP(t,\theta) = \mu_P(t,\theta)P(t,\theta)dt + \sigma_P(t,\theta)P(t,\theta)dW_t.$$

And we have

Lemma 3.2 *The relations of the dynamics of θ-parameterised term structures and T-parameterised term structures are*

$$\begin{cases} \mu_r(t,\theta) = \tilde{\mu}_f(t,\theta) + \frac{\partial r(t,\theta)}{\partial \theta}, \\ \sigma_r(t,\theta) = \tilde{\sigma}_f(t,\theta), \end{cases} \tag{3.28}$$

$$\begin{cases} \mu_P(t,\theta) = \tilde{\mu}_B(t,\theta) - r(t,\theta), \\ \sigma_P(t,\theta) = \tilde{\sigma}_B(t,\theta), \end{cases} \tag{3.29}$$

provided that $r(t,\theta)$ and $P(t,\theta)$ are differentiable in θ, and $f(t,T)$ and $B(t,T)$ are differentiable in T.

Proof.

$$\frac{\partial r(t,\theta)}{\partial \theta} = \frac{\partial f}{\partial T}(t, t+\theta)\frac{\partial T}{\partial \theta} = \frac{\partial f}{\partial T}(t, t+\theta),$$
$$\frac{\partial P(t,\theta)}{\partial \theta} = \frac{\partial B}{\partial T}(t, t+\theta)\frac{\partial T}{\partial \theta} = \frac{\partial B}{\partial T}(t, t+\theta)$$

Apply the Itô Lemma (Lemma A.1 in Appendix A) we obtain

$$\begin{aligned} dr(t,\theta) &= df(t,t+\theta) \\ &= \tilde{\mu}_f(t,\theta)dt + \tilde{\sigma}_f(t,\theta)dW_t + \frac{\partial f}{\partial T}(t,t+\theta)dt \\ &= \left(\tilde{\mu}_f(t,\theta) + \frac{\partial r(t,\theta)}{\partial \theta}\right) dt + \tilde{\sigma}_f(t,\theta)dW_t, \end{aligned}$$

and

$$\begin{aligned} dP(t,\theta) &= dB(t,t+\theta) \\ &= \tilde{\mu}_B(t,\theta)B(t,t+\theta)dt + \tilde{\sigma}_B(t,\theta)B(t,t+\theta)dW_t + \frac{\partial B}{\partial T}(t,t+\theta)dt \\ &= \left(\tilde{\mu}_B(t,\theta)B(t,t+\theta)dt + \frac{\partial P(t,\theta)}{\partial \theta}\right) dt + \tilde{\sigma}_B(t,\theta)B(t,t+\theta)dW_t. \end{aligned}$$

Since

$$\frac{\partial P(t,\theta)}{\partial \theta} = -r(t,\theta)P(t,\theta),$$

So

$$dP(t,\theta) = (\tilde{\mu}_B(t,\theta) - r(t,\theta))P(t,\theta)dt + \tilde{\sigma}_B(t,\theta)P(t,\theta)dW_t.$$

■

This lemma shows that the volatility processes of the term structure models under two different parameterisations are identical. Shown by HJM modelling in Theorem 2.3, term structure modelling crucially depends on the volatility modelling, and the drift function of forward rate is determined by the volatility function of the forward rate. Term structure data sets are mostly formatted using the θ-parameterisation in terms of (t,θ). Hence, by the above lemma, if we can model the volatility process of the θ-parameterised term structures accurately, we can obtain the volatility process of the T-parameterised term structures successfully by the simple transformation $\theta = T - t$.

3.5 Martingale Measures and No-Arbitrage Based on θ-parameterised Term Structure of Interest Rates

In this section, we explore the martingale measures and no-arbitrage conditions based on θ-parameterised term structure of interest rates.

3.5.1 Risk-Neutral Martingale Measure P^*

Heath, Jarrow and Morton (HJM, 1992) set up the fundamental martingale theory with T-parameterised term structures. Under the risk-neutral measure P^*, all discounted bond price processes $B(t,T)/B(t)$ are P^* martingales. i.e. numeraire is the risk-free asset $B(t) = \exp\{\int_0^t r(u)du\}$, and $r(t)$ is the spot rate.

In this section, we study the martingale theory with θ-parameterised term structure of interest rates, parallel to martingale theory for the T-parameterised term structures. We find that the discount bond price process $\frac{P(t,\theta)}{B(t)}$ is not a P^* martingale. To understand this claim, let us consider a simple unrealistic special economic market with constant spot rate $r(t) = r$, for all $t \in [0, T^*]$. Under this assumption, the zero coupon bond price is given by,

$$P(t,\theta) = E[e^{-\int_t^{t+\theta} r(u)du}|\mathcal{F}_t] = e^{-r\theta},$$

and

$$r(t,\theta) = -\frac{1}{P(t,\theta)}\frac{\partial P(t,\theta)}{\partial\theta)} = r.$$

It is obvious that $P(t,\theta)$ is a martingale under any measure, since it is independent of time t. But $P(t,\theta)/B(t)$ is not a martingale under any measure, because for any $s < t$,

$$E\left[\frac{P(t,\theta)}{B(t)}|\mathcal{F}_s\right] = E\left[\frac{e^{-t\theta}}{e^{rt}}|\mathcal{F}_s\right] = e^{r(t-\theta)} \neq e^{r(s-\theta)} = \frac{P(s,\theta)}{B(s)}.$$

This simple example illustrates that $P(t,\theta)/B(t)$ is not a martingale under the risk-neutral measure, though the assumption of constant spot rate is, of course, not realistic. Generally, if the spot rate r_t is a stochastic process, we obtain the following result.

Lemma 3.3 *P^* is the risk-neutral measure that ensures that $\frac{B(t,T)}{B(t)}$ is a P^* martingale for every fixed T, if and only if, $\frac{P(t,\theta)}{\tilde{B}(t,\theta)}$ is a P^* martingale for every $\theta > 0$, where*

$$\tilde{B}(t,\theta) = \exp\{\int_0^t (r(u) - r(u,\theta))du$$

that is a risk account accumulating the difference between the spot rate and the forward rates.

Proof. Let

$$Y(t,\theta) = \frac{P(t,\theta)}{\tilde{B}(t,\theta)}.$$

Since

$$d\tilde{B}(t,\theta) = \tilde{B}(t,\theta)(r(t) - r(t,\theta))dt,$$

we obtain

$$\begin{aligned}
dY(t,\theta) &= \frac{1}{\tilde{B}(t,\theta)}dP(t,\theta) + P(t,\theta)d\frac{1}{\tilde{B}(t,\theta)} \\
&= \frac{P(t,\theta)}{\tilde{B}(t,\theta)}(\mu_P(t,\theta)dt - \sigma_P(t,\theta)dW_t) - P(t,\theta)\frac{1}{\tilde{B}^2(t,\theta)}d\tilde{B}(t,\theta) \\
&= Y(t,\theta)\{[\mu_P(t,\theta) - r(t) + r(t,\theta)]dt + \sigma_P(t,\theta)dW_t\}.
\end{aligned}$$

If P^* is the risk -neutral measure equivalent to P, then by Theorem 2.2, there exists an adapted process $\varphi(t)$ satisfying condition (2.14), such that the process $W_t^* = W_t - \int_0^t \varphi(u)du$ is a Brownian motion under the risk-neutral measure P^*. By (2.14), (3.26), (3.27), and Lemma 3.2, we obtain

$$\varphi(t) = -\frac{\tilde{\mu}_B(t,\theta) - r(t)}{\tilde{\sigma}_B(t,\theta)} = -\frac{\mu_P(t,\theta) + r(t,\theta) - r(t)}{\sigma_P(t,\theta)}, \tag{3.30}$$

and

$$dY(t,\theta) = \sigma_P(t,\theta)dW_t^*,$$

So

$$Y(t,\theta) = \frac{P(t,\theta)}{\tilde{B}(t,\theta)}$$

is a P^*-martingale.
Assume now that Y is a P^*-martingale for every $\theta > 0$. Then defining ϕ (3.30) we obtain by similar arguments that that $\frac{B(t,T)}{B(t)}$ is a P^*-martingale for every T. ■

By Lemma 3.2 and Lemma 3.3, under the risk-neutral measure P^*, the stochastic processes $r(t,\theta)$ and $P(t,\theta)$ satisfy the equations

$$dr(t,\theta) = \frac{\partial}{\partial\theta}[r(t,\theta) + \frac{1}{2}(\sigma^*(t,\theta))^2]dt + \sigma(t,\theta)dW_t^*, \tag{3.31}$$

and

$$dP(t,\theta) = [r(t,0) - r(t,\theta)]P(t,\theta)dt - \sigma^*(t,\theta)P(t,\theta)dW_t^*, \tag{3.32}$$

where

$$\sigma^*(t,\theta) = \int_0^\theta \sigma(t,\tau)d\tau, \tag{3.33}$$

and W^* is a Brownian motion under the risk-neutral measure P^*.

3.5.2 T-forward Martingale Measure

Corresponding to Lemma 2.3 of the T-forward martingale measure P^T, we have the following:

Lemma 3.4 *P^T is the T-forward martingale measure that ensures that $\frac{B(t,S)}{B(t,T)}$ is a P^T martingale for any $S \geq 0$, $t \in [0, S \wedge T)]$, if and only if, $\frac{P(t,\tau)/\tilde{B}(t,\tau)}{P(t,\theta)/\tilde{B}(t,\theta)}$ is a P^T martingale for any $\tau \geq 0$, where $\tau = S - t$ and $\theta = T - t$.*

Proof. Let

$$F_P(t,\tau,\theta) = \frac{P(t,\tau)/\tilde{B}(t,\tau)}{P(t,\theta)/\tilde{B}(t,\theta)} = \frac{P(t,\tau)}{P(t,\theta)} e^{\int_0^t (r(u,\tau)-r(u,\theta))du}$$

Following the proof at Lemma 2.3, we find that

$$\begin{aligned} d(\frac{B(t,S)}{B(t,T)}) &= \frac{B(t,S)}{B(t,T)}\{[\mu_B(t,S) - \mu_B(t,T) - \sigma_B(t,T)(\sigma_B(t,S) - \sigma_B(t,T))]dt \\ &\quad +[\sigma_B(t,S) - \sigma_B(t,T)]dW_t\}, \end{aligned}$$

and therefore

$$\begin{aligned} d(\frac{P(t,\tau)}{P(t,\theta)}) &= \frac{P(t,\tau)}{P(t,\theta)}\{[\mu_P(t,\tau) - \mu_P(t,\theta) - \sigma_P(t,\theta)(\sigma_P(t,\tau) - \sigma_P(t,\theta))]dt \\ &\quad +[\sigma_P(t,\tau) - \sigma_P(t,\theta)]dW_t\}. \end{aligned}$$

Then

$$\begin{aligned} dF_P(t,\tau,\theta) &= e^{\int_0^t (r(u,\tau)-r(u,\theta))du} d(\frac{P(t,\tau)}{P(t,\theta)}) + \frac{P(t,\tau)}{P(t,\theta)} e^{\int_0^t (r(u,\tau)-r(u,\theta))du}(r(t,\tau) - r(t,\theta))dt \\ &= F_P\{[\mu_P(t,\tau) - \mu_P(t,\theta) + r(t,\tau) - r(t,\theta) \\ &\quad -\sigma_P(t,\theta)(\sigma_P P(t,\tau) - \sigma_P P(t,\theta))]dt + [\sigma_P(t,\tau) - \sigma_P(t,\theta)]dW_t.\} \end{aligned}$$

The function $\psi(t)$, satisfying equation (2.18), changes the actual measure P to the T-forward martingale measure P^T by

$$\frac{d\hat{P}}{dP} = \exp\left\{\int_0^t \psi(u)dW_u - \frac{1}{2}\int_0^t \psi(u)du\right\},$$

and $W_t^{\hat{P}} = W_t - \int_0^t \psi(u)du$ is a Brownian motion under the T-forward measure P^T.

Let $S = t + \tau$, $T = t + \theta$, changing the parameter of the coefficient functions by equations of (3.26) and (3.27), and by the relations of coefficient functions for the different processes in two different parameterisations (Lemma 3.2), we have

$$\begin{aligned}\psi(t) &= -\frac{\mu_B(t,S)-\mu_B(t,T)-\sigma_B(t,T)(\sigma_B(t,S)-\sigma_B(t,T))}{\sigma_B(t,S)-\sigma_B(t,T)}\\ &= -\frac{\tilde{\mu}_B(t,\tau)-\tilde{\mu}_B(t,\theta)-\tilde{\sigma}_B(t,\theta)(\tilde{\sigma}_B(t,\tau)-\tilde{\sigma}_B(t,\theta))}{\tilde{\sigma}_B(t,\tau)-\tilde{\sigma}_B(t,\theta)}\\ &= -\frac{\mu_P(t,\tau)+r(t,\tau)-\mu_P(t,\theta)-r(t,\theta)-\sigma_P(t,\theta)(\sigma_P(t,\tau)-\sigma_P(t,\theta))}{\sigma_P(t,\tau)-\sigma_P(t,\theta)}.\end{aligned}$$

Hence

$$dF_P(t,\tau,\theta) = F_P[\sigma_P(t,\tau)-\sigma_P(t,\theta)]dW_t^{\hat{P}},$$

and $F_P(t,\tau,\theta)$ is $\hat{P}$-martingale. ∎

By the dynamics of the zero-coupon bond price, it is derived from the above proof that the market is no-arbitrage if the stochastic process $\psi(t)$,

$$\psi(t) = -\frac{\mu_P(t,\tau)+r(t,\tau)-\mu_P(t,\theta)-r(t,\theta)-\sigma_P(t,\theta)(\sigma_P(t,\tau)-\sigma_P(t,\theta))}{\sigma_P(t,\tau)-\sigma_P(t,\theta)} \tag{3.34}$$

is independent of τ, θ.

This condition can be expressed in terms of the parameters defining the dynamics of the forward rates according to the transformation formula of bond and forward rate (3.21). That is,

$$\psi(t) = -\frac{\mu^*(t,\tau)-r(t,\tau)-\mu^*(t,\theta)-r(t,\theta)}{\sigma^*(t,\tau)-\sigma^*(t,\theta)} + \frac{1}{2}(\sigma^*(t,\tau)-\sigma^*(t,\theta)), \tag{3.35}$$

where μ^* and σ^* are the integrals of μ_r and σ_r, respectively, as in (3.20).

3.6 Models of Term Structure of Interest Rates under the θ-parameterisation

The ***Brace-Gatarek-Musiela Model*** (BGM) is a model for the London Interbank Borrowing Rate (LIBOR) forward rates. The BGM model is popular since the LIBOR forward rates are market observable quantities. For a fixed $\delta > 0$, the forward LIBOR $L(t,\theta)$ is defined to be the simple (forward) interest rate for the investment for period $[\theta, \theta+\delta]$. Let P be a risk-neutral measure and let W_t be a Brownian motion under P. Brace, Gatarek and Musiela (BGM) (1997) model the forward rates $r(t,\theta)$ and bond prices $P(t,\theta)$ with the choice of the volatility $\sigma(t,\theta)$ which ensures that the dynamics of the forward LIBOR $L(t,\theta)$, $t,\theta \geq 0$ follows

$$dL(t,\theta) = \left[\frac{\partial}{\partial\theta}L(t,\theta)+\gamma(t,\theta)L(t,\theta)\sigma^*(t,\theta)+\frac{\delta L^2(t,\theta)\gamma^2(t,\theta)}{1+\delta L(t,\theta)}\right]dt + L(t,\theta)\gamma(t,\theta)dW(t),$$

where $\sigma^*(t,\theta) = \int_0^\theta \sigma(t,u)du$, $\gamma(t,\theta)$ is a forward LIBOR volatility function, $t \geq 0$, $\theta \geq 0$. For the relation between the forward rate $r(t,\theta)$ and LIBOR, and the choosing of $\sigma(t,\theta)$ satisfying the above equation, see Brace, Gatarek and Musiela (BGM) (1997). Using the BGM model, pricing of interest rate derivatives were given by BGM (1997). Goldys and Musiela (2001) reformulated the interest rate model in terms of the new parameters (t,θ), as a solution to a stochastic evolution equation in infinite dimensional space. Under the risk-neutral measure P and under the appropriate regularity conditions, the time evolution of the forward rates $r(t,\cdot)$ is completely determined by the initial curve $r(0,\cdot)$ and the volatility structure. Prices of some contingent claims were obtained by solving the related partial differential equation in an infinite number of variables.

Volatility Models. As shown by the HJM (1992) model, the evolution of the instantaneous forward rates depends crucially on the choice of the volatility process and the risk premium. They provided a constant and exponentially decaying volatility structure model given by

$$\sigma_f(t,T) = \sigma \exp(-\lambda(T-t)),$$

which implied

$$\sigma_r(t,\theta) = \tilde{\sigma}_f(t,\theta) = \sigma \exp(-\lambda\theta),$$

hence time independent.

Most empirical studies on interest derivatives have observed a humped shape in the volatility structure of interest rates. Mercurio and Moraleda (1996) proposed a deterministic and humped volatility model of forward rate, given by

$$\sigma_f(t,T) = \sigma(\gamma(T-t)+1)\exp(-\lambda(T-t))$$

that implied

$$\sigma_r(t,\theta) = \widetilde{\sigma}_f(t,\theta) = \sigma(\gamma\theta+1)\exp(-\lambda\theta),$$

where σ, γ and λ are non-negative constants. This volatility structure suggests a humped volatility if $\gamma \geq \lambda$ and a stationary process of the forward rate.

Rama Cont Model. In order to model the stochastic volatility as a process depending on t and θ, Cont (1998) decomposed the variation of the term structure into the variations of the short rate, the long rate and the fluctuation of the curve around its average shape; that is

$$r(t, \theta) = r(t) + s(t)[Y(\theta) + X(t, \theta)],$$

with

$$Y(\theta_{min}) = 0, \quad Y(\theta_{max}) = 1,$$

where Y is a deterministic shape function defining the average profile of the term structure, and

$$X(t, \theta_{min}) = 0, \quad X(t, \theta_{max}) = 0.$$

$X(t, \theta)$ is an adapted process describing random fluctuation of the re-parameterised forward rate around its long term average shape.

A family of models of the term structure dynamics were presented in Cont's paper, describing several statistical features observed in empirical studies of these three processes, $r(t)$, $s(t)$ and $X(t, \theta)$.

For example, we can decompose the term structure as

$$\begin{aligned} r(t, \theta) &= r(t, \theta_{min}) + [r(t, \theta_{max}) - r(t, \theta_{min})] \\ &\cdot \left\{ E[\frac{r(t, \theta) - r(t, \theta_{min})}{r(t, \theta_{max}) - r(t, \theta_{min})}] + \frac{r(t, \theta) - r(t, \theta_{min})}{r(t, \theta_{max}) - r(t, \theta_{min})} - E[\frac{r(t, \theta) - r(t, \theta_{min})}{r(t, \theta_{max}) - r(t, \theta_{min})}] \right\} \end{aligned}$$

where

$$r(t) = r(t, \theta_{min}),$$

$$s(t) = r(t, \theta_{max}) - r(t, \theta_{min}),$$

$$Y(\theta) = E[\frac{r(t, \theta) - r(t, \theta_{min})}{r(t, \theta_{max}) - r(t, \theta_{min})}],$$

$$X(t, \theta) = \frac{r(t, \theta) - r(t, \theta_{min})}{r(t, \theta_{max}) - r(t, \theta_{min})} - E[\frac{r(t, \theta) - r(t, \theta_{min})}{r(t, \theta_{max}) - r(t, \theta_{min})}],$$

and

$$E[X(t, \theta)] = 0,$$

where it is assumed that the above expected values are conditional expected values conditional on pre-information before time t. The volatility of $X(t, \theta)$ can be estimated by the sample observed deviations. Based on Australia Treasury bond yields from 1996 to 2001, maturity in 1 month, 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, and 10 years, this can be estimated using the average (over time t) of the squared deviations $\hat{X}^2(t, \theta)$. Figure 3.1 strongly suggests that the volatility has a maximum around one year maturity. Other authors have made similar observations (Bouchaud et al. 1997, Cont 2001). Moraleda and Vorst (1997) presented a simple humped volatility model.

Discrete-time Approximation. Dietrich-Campbell and Schwartz (1986), Chan (1992) and others estimated the parameters of continuous-time model of interest rates approximately by using a discrete-time specification. This approach is under the assumption that the drift and volatility are constant over each time period interval of time between observations. For the continuous-time model of forward interest rate (3.16), if the constant drift value and volatility value are computed at the beginning of each interval, the discrete-time econometric specification is

$$r(t+1,\theta) - r(t,\theta) = \mu(t,\theta) + \epsilon(t+1,\theta) \tag{3.36}$$

where

$$E[\epsilon(t+1,\theta)|\mathcal{F}_t] = 0, \quad E[\epsilon^2(t+1,\theta)|\mathcal{F}_t] = \sigma^2(t,\theta). \tag{3.37}$$

The discrete-time model of the interest rate increments has the advantage of allowing the mean and variance of the increments to depend directly on the interest rate in a way consistent with the continuous-time model based on diffusion processes.

By HJM theory, modelling the volatility is crucial for term structure modelling. Many researchers have investigated the characteristics of volatility to develop the modelling of volatility. The Autoregressive conditional heteroskedasticity (ARCH) model developed by Engle (1982) and other extended ARCH models study the time varying volatility and capture many observed volatility behaviours in financial time series. Chapter 5 to chapter 7 will develop empirical models for Australian Treasury yields using GARCH technology.

3.7 Summary and Conclusions

This chapter presented the theoretical essentials of the term structure of interest rate under the two parameters, the time evolution t and length of time to maturity θ, which is referred to as the θ-parameterisation of term structures, corresponding to the T-parameterisation of term structures that depend on the time evolution t and time of maturity T. The volatility process under θ-parameterised term structure is identical to the volatility process under T-parameterised term structures. Models of θ-parameterised term structures were proposed corresponding to the T-parameterised term structure models. The discounted bond price $P(t,\theta)/\tilde{B}(t,\theta)$ is a martingale under the risk-neutral measure P^*, with the numeraire $\tilde{B}(t,\theta)$, being a risk account cumulating the difference between the spot rate and forward rate. If $B(t,S)/B(t,T), S,T > 0$ is a P^T-martingale, we derived that $\frac{P(t,\tau)/\tilde{B}(t,\tau)}{P(t,\theta)/\tilde{B}(t,\theta)}$ is a P^T martingale. No-arbitrage conditions are given in terms of the risk-neutral martingale measure P^* of $P(t,\theta)/\tilde{B}(t,\theta)$ and the T-forward martingale P^T of $\frac{P(t,\tau)/\tilde{B}(t,\tau)}{P(t,\theta)/\tilde{B}(t,\theta)}$.

The BGM model, Cont model, the HJM volatility model and the MM volatility model were reviewed using θ-parameterisation. Relying on Cont decomposition approach, our empirical analysis of the Australian Treasury yields confirmed what has been observed in other markets, that the volatility of term structure is humped around 1 year. The discrete-time specification of continuous-time diffusion processes model possesses the advantage that the mean and variance of the increments in discrete-time model depend in a way consistent with these in continuous-time model. By HJM theory, modelling of volatility plays a crucial rule for modelling term structure. Chapter 5 to chapter 7 will develop empirical models for Australian Treasury yields increments using Generalised Autoregressive conditional heteroskedasticity (GARCH) technology.

In this chapter, the stochastic of the term structure is assumed from one dimensional Brownian motion. All theories and modelling can be easily extended to the general case where the Brownian motion is d-dimensional and the volatility process takes values in $\mathbf{R}^d$.

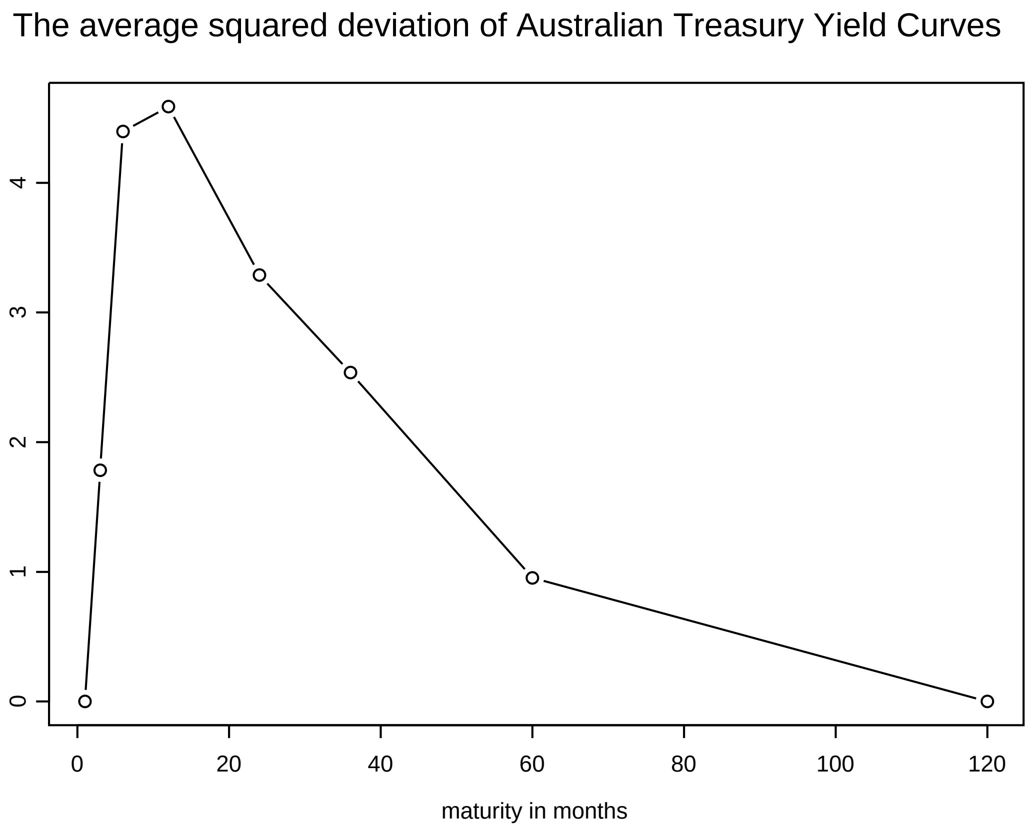


Figure 3.1: Average of squared deviation, $X^2(t, \theta)$, with a maximum at $\theta = 1$ year (maturity in 12 months)

Chapter 4

Australian Yields and Stylised Facts

It is a matter of fact that interest rates cannot be traded in the financial market. Instead, it is only possible to trade related instruments such as bond, options, and swaps. Consequently, interest rates must be derived from the market prices of the associated tradeable products. Methods of estimation vary from institute to institute, and country to country. James and Webber (2000) estimated the yield curves using data from the UK government bond (gilt) market and the US money market.

The author is aware of at least two methods of yield curves estimation used in Australia. One method based on the bond market data with generic bond yields derived from market bond prices, is provided by the Reserve Bank of Australia (RBA). A second method based on the money market data with yield curves constructed from swaps, is provided by the Commonwealth Bank of Australia (CBA).

This chapter introduces the yields data from RBA and analysis of its stylised facts, that will be used in Chapter 5-7 for yield curves modelling. The yield curves constructed from swaps provided by the CBA is introduced in Appendix C and a simple comparison of RBA yields and CBA yields is given.

Section 1 reviews the definition of yield-to-maturity (or yield). Section 2 introduces the generic bond yields derived from market bond prices provided by the RBA. Section 3 presents the statistics of yield curves and stylised facts. Section 4 summarises and concludes the chapter.

4.1 Reviewing of Yield

As addressed in Section 2.2.1, investors who purchased bonds desire to obtain a certain *yield* that is the return provided by the investment. *Yield-to-maturity* represents the percentage rate of return if the security is held until its maturity date. The calculation of yield is based on the coupon rate, the length of time to maturity and the market price of the security. Yield-to-maturity is defined (Musiela and Rutkowski 1997) in a continuous-time framework as follows.

Definition 4.1 *Suppose a coupon bearing bond pays the positive deterministic cash flows* $c_1, \ldots, c_m$ *at dates* $T_1 < \cdots < T_m \leq T^*$. Yield-to-maturity *of the coupon bearing bond at time* t, $R_c(t)$, *is given implicitly by*

$$B_c(t) = \sum_{T_j > t} c_j e^{-R_c(t)(T_j - t)}, \tag{4.1}$$

where $B_c(t)$ *stands for the price at* t *of a coupon bond.*

Let us recall that the *yield-to-maturity* $R(t,T)$ on a zero-coupon bond of maturity T (or called *zero-coupon yield*), given implicitly in Chapter 2 by (2.2), has the form

$$B(t,T) = e^{-R(t,T)(T-t)}, \quad \forall t \in [0,T],$$

where $B(t,T)$ stands for the price at time $t < T$ of a zero coupon bond of maturity T.

Note the difference between coupon bond yield and zero-coupon bond yield. Zero-coupon bond yield is a special case of coupon bond yield if $m = 1$ and $c_m = 1$ in (4.1).

In this thesis, *yield-to-maturity* is simply referred to as *yield.*

4.2 Bond Market: Generic Bond Yields of the Reserve Bank of Australia (RBA)

It is well known that bond yields can be derived from market prices of bonds using (4.1) and (2.2).

Unlike the US Treasury bond market, which has several benchmark bond securities, (such as at two years, three years,. . . , etc), there is only one ten year Treasury bond in Australia, in addition to the short term Treasury Notes whose maturities are one month, three months and six months. Because ‘benchmark’ securities do not exist, other generic

Table 4.1: Treasury Fixed Coupon Bonds

Coupon	Maturity	Yield
9.50%	Aug 2003	4.635
9.00%	Sep 2004	4.580
7.50%	Jul 2005	4.615
10.00%	Feb 2006	4.700
6.75%	Nov 2006	4.780
10.00%	Oct 2007	4.920
8.75%	Aug 2008	5.015
7.50%	Sep 2009	5.155
5.75%	Jun 2011	5.280
6.50%	May 2013	5.375
6.25%	Apr 2015	5.470

bond yields maturities less than ten years are calculated by linear interpolation between the yields of actual government bonds on issue, and by the date of publication of the generic bond data.

The Reserve Bank of Australia presents a daily statistical release of "Indicative Mid Rates of Selected Commonwealth Government Securities", the average of buy or sell rates reported by bond dealers surveyed by the Bank at 4.30 pm AEST. An example of the daily release is shown in Table 4.1, which is the Treasury Fixed Coupon Bonds released on 18 March 2003.

The calculation of generic 5 year bond yields follows the method used by the RBA. From 18 March 2003, the generic 5 year bond yield is calculated for the date 18 March 2008. This date falls between the maturity dates 15 Oct 2007 (10% Coupon of government bonds issued) and 15 Aug 08 (8.75% Coupon of government bonds issued). On 18 March 2003, the 4:30 pm closing yields for these two securities were 4.920% and 5.015% respectively. The number of days between 15 Oct 2007 and 18 Mar 2008 is 155 days, and 150 days for between 18 Mar 2008 and 15 Aug 2008. Using linear interpolation of relevant yields (4.920% and 5.015%) and corresponding numbers of days (155 and 150), gives 4.97%, which is the rate for the generic 5 year bond on 18/3/08.

In the same way, we can calculate all the generic 1 year, 2 years, ..., and 10 years bond yields, and this allows us to construct a yield curve of Treasury bonds prevailing on 18 March 2003.

Using released yields of fixed coupon bonds for each working day, generic yield curves for maturity 1-10 years are derived by linear interpolation as explained above.

4.3 Statistics of Yield and Yield Increments

The objective of this section is to present the statistical features of Australian Treasury yield curves obtained by the RBA maturities in 1 month, 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, and 10 years, from 1996-2001. By *short-term bill yield* we mean the yield of a bill with maturity less than one year and *middle-to-long-term bond yield* if maturity is one year or more.

Autocorrelation function plots of yield series (Figure 4.1) show that each yield series with a fixed maturity are highly autocorrelated. We will analyse the *yield increments* $R(t,\theta) - R(t-1,\theta)$ that will be used for empirical statistical modelling in the following chapters.

Short-term bill yield increments vs. middle-to-long-term bond yield increments. Autocorrelation function plots of yield increments, Figure 4.2, show that the yield increments are not autocorrelated. Autocorrelation function plots of squared yield increments, Figure 4.3, show that the squared middle-to-long-term bond yield increments are autocorrelated, but not for the short-term bill yield increments. The Ljung-Box autocorrelation test (Ljung and Box 1979) is used to test the null hypothesis that all of the autocorrelations are zero. This test is based on the modified Q-statistic which is asymptotically Chi-squared distributed. A simple Lagrange Multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH) effects (Engle 1982) was constructed based on the ARCH regression with the null hypothesis that there are no ARCH effects. Table 4.3 shows the results of Ljung-Box autocorrelation tests (Lag=12) and Lagrange Multiplier (LM) ARCH effects tests, which confirm that squared middle-to-long-term bond yield increments are autocorrelated and demonstrate ARCH effects of the middle-to-long-term bond yield increments, but the squared short-term bill yield increments are not autocorrelated and therefore there are no ARCH effects with the short-term bill yield increments.

Cluster/Persistent. Time series plots (Figure 4.4) of yield increments indicate that

Table 4.2: Tests of the RBA bill and bond yield increments

Test for Autocorrelation: Ljung-Box Null Hypothesis: no autocorrelation

(yield increments)

	1m	3m	6m	1y	2y	3y	5y	10y
Test Stat	37.53	43.54	46.30	19.48	17.14	18.14	17.78	24.21
p.value	0.19	0.06	0.03	0.94	0.97	0.96	0.97	0.80

Test for Autocorrelation: Ljung-Box Null Hypothesis: no autocorrelation

(squared yield increments)

	1m	3m	6m	1y	2y	3y	5y	10y
Test Stat	4.47	1.80	3.94	43.28	61.66	77.57	98.52	132.66
p.value	1.00	1.00	1.00	0.07	0.00	0.00	0.00	0.00

Test for ARCH Effects: LM Test Null Hypothesis: no ARCH effects

(yield increments)

	1m	3m	6m	1y	2y	3y	5y	10y
Test Stat	4.31	1.75	3.62	37.78	57.89	72.19	90.58	118.16
p.value	1.00	1.00	1.00	0.18	0.00	0.00	0.00	0.00

large changes were followed by large changes, and small changes were followed small changes. This means that there are the volatility persistence and clustering. Auto-correlation function plots (Figure 4.2) and (Figure 4.3)show that there is no autocorrelation in the yield increments series themselves for all maturities, while there exists auto-correlation for the squared middle-to-long-term bond yield increments. These indicate that the conditional variance changes over time t for the middle-to-long-term bond yield increments, exhibiting time varying conditional heteroskedasticity and volatility clustering.

Normality. Quantile-quantile normal plots of yield increments, Figure 4.5, showed that the increments of yield are heavy tailed. And the Kolmogorov-Smirnov tests confirm that the yield increments reject that the yield increments are in normal distribution (p-values are 0.000).

Correlation. The bond yields are correlated with RBA cash rate. The plot of correlation between yields in a fixed maturity and cash rates, (top plot on Figure 4.6), shows that the correlation decreases quickly when maturity level increases, and an exponentially decaying correlation structure for yields and cash rates is plausible. The plots of correlation between yield increment and cash rates increment, (bottom plot on Figure 4.6), also show that the correlation decreases quickly when the maturity increases.

Exogenous variables. Figure 4.8 shows the RBA yield curves, along with the target rates set by the RBA and the RBA Board meeting dates at which target rates are adjusted. In Figure 4.8, the arrows indicate the RBA Board meeting dates and the crosses indicate the RBA Board meeting dates at which decisions were made to adjust the target rates. It is obvious that the changes in the yield curves follow the changes in the target cash rate. From the time series plots (Figure 4.7) of yield increments we can see that the larger increments occur right after the RBA cash rate was changed. It appears that the yield increments may have larger variability around the RBA Board meeting dates. This information is taken into account for modelling the yield increments and their volatility. The variables that we use are:

R_t^+: indicator of raising the target cash rate.

R_t^-: indicator of lowering the target cash rate.

M_t: indicator of the RBA Board meeting.

Table 4.3: The Reserve Bank of Australia - Monetary Policy Changes

Cash Rate Target Released	Change in cash rate	New cash rate target
5 June 2002	+0.25	4.75
8 May 2002	+0.25	4.50
5 Dec 2001	-0.25	4.25
3 Oct 2001	-0.25	4.50
5 Sep 2001	-0.25	4.75
4 Apr 2001	-0.50	5.00
7 Mar 2001	-0.25	5.50
7 Feb 2001	-0.50	5.75
2 Aug 2000	+0.25	6.25
3 May 2000	+0.25	6.00
5 Apr 2000	+0.25	5.75
2 Feb 2000	+0.50	5.50
3 Nov 1999	+0.25	5.00
2 Dec 1998	-0.25	4.75
30 Jul 1997	-0.50	5.00
23 May 1997	-0.50	5.50
11 Dec 1996	-0.50	6.00
6 Nov 1996	-0.50	6.50
31 Jul 1996	-0.50	7.00
14 Dec 1994	+1.00	7.50
24 Oct 1994	+1.00	6.50

Since 1996, the Reserve Bank Board has met on the first Tuesday of each month with the exception of 1996, when the "August meeting" was held on 30 July. (Prior to 1997, the "August meetings" were held in late July to enable the Board to sign off the Bank's annual report which was traditionally issued two days after the Budget was brought down in early August; with the Budget moving to May, meetings are now held on the first Tuesday of August.)

For the indicators of raising or lowering the target cash rate, see Table 4.3, which is available on the RBA web site. Monetary policy decisions are expressed in terms of a target cash rate, which is the overnight money market interest rate.

4.4 Conclusions

This chapter explains the derivation of term structure interest rates using the Australian Treasury yield curves of the RBA. Statistics and characteristics of the yield curves are explored based on the RBA yield maturities in 1 month, 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, and 10 years, from 1996-2001.

The analysis shows that the dynamic process of the short-term bill yield increments are not consistent with middle-to-long-term bond yield increments for RBA yield increments. For the short-term bill yield, both the Ljung-Box test and autocorrelation function plots show that there is no autocorrelation left in bill yield increments and squared bill yield increments. The simple Lagrange Multiplier (LM) tests confirm that there are no ARCH effects in the short-term bill yield increments. But, for the middle-to-long-term bond yield, both the Ljung-Box tests and autocorrelation function plots show that there is no autocorrelation left in bond yield increments, while squared bond yield increments are autocorrelated. The simple Lagrange Multiplier (LM) tests confirm that there are ARCH effects in the middle-to-long-term bond yield increments.

The RBA middle-to-long term bond yield increments will used for intensive statistical analysis and modelling in later chapters.

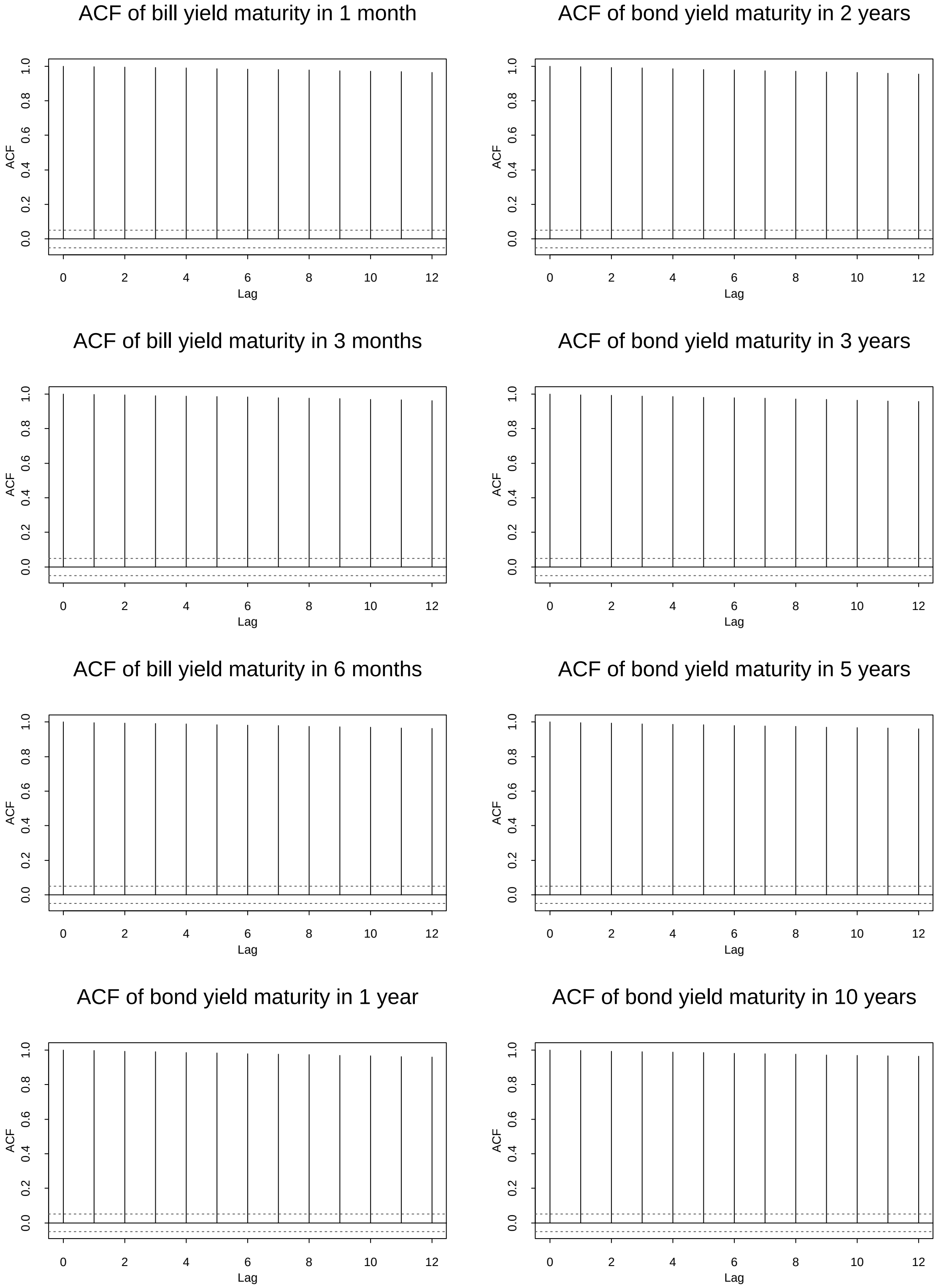


Figure 4.1: ACF of the RBA bill and bond yields

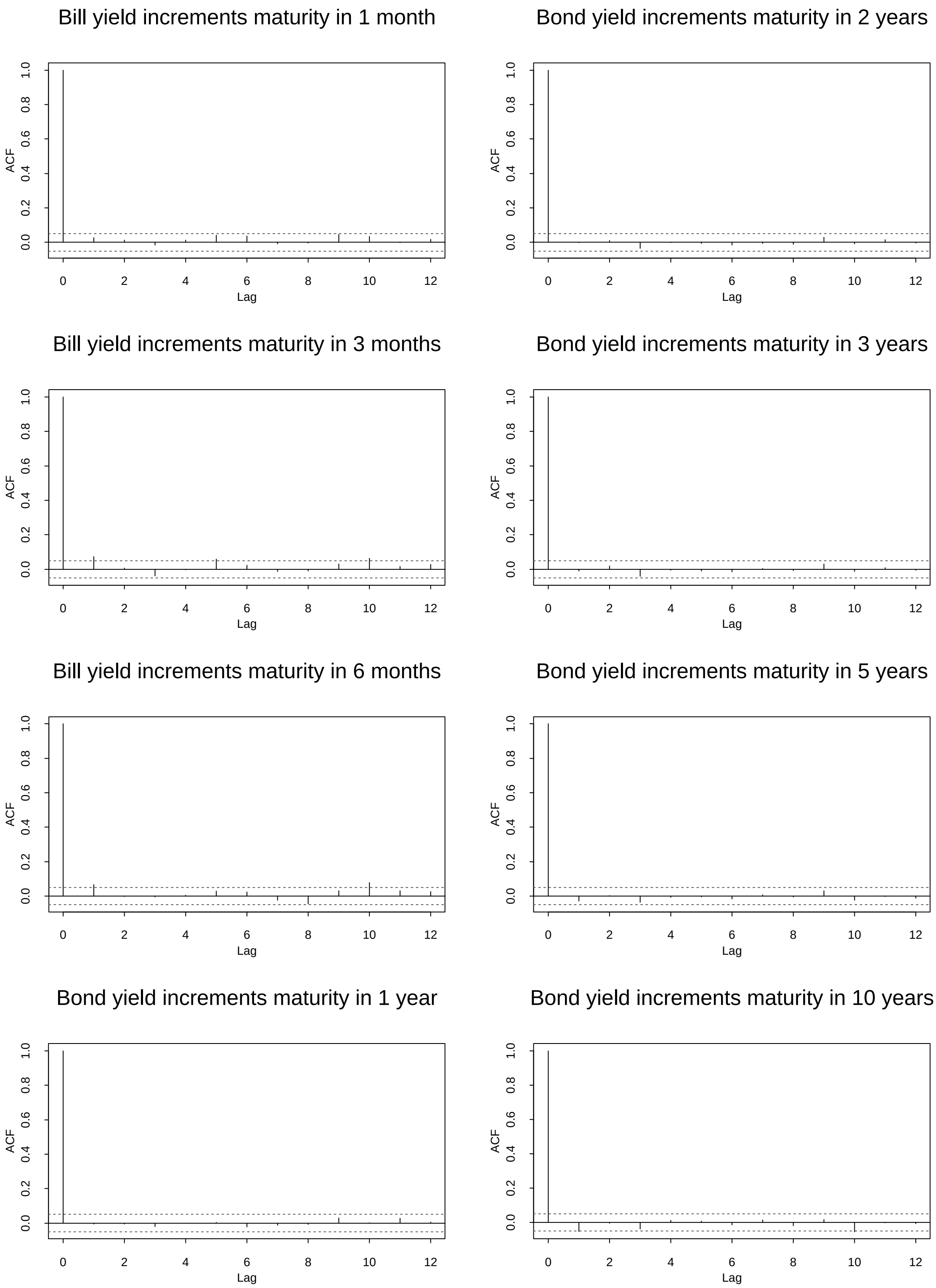


Figure 4.2: ACF of yield increments of the RBA bill and bond

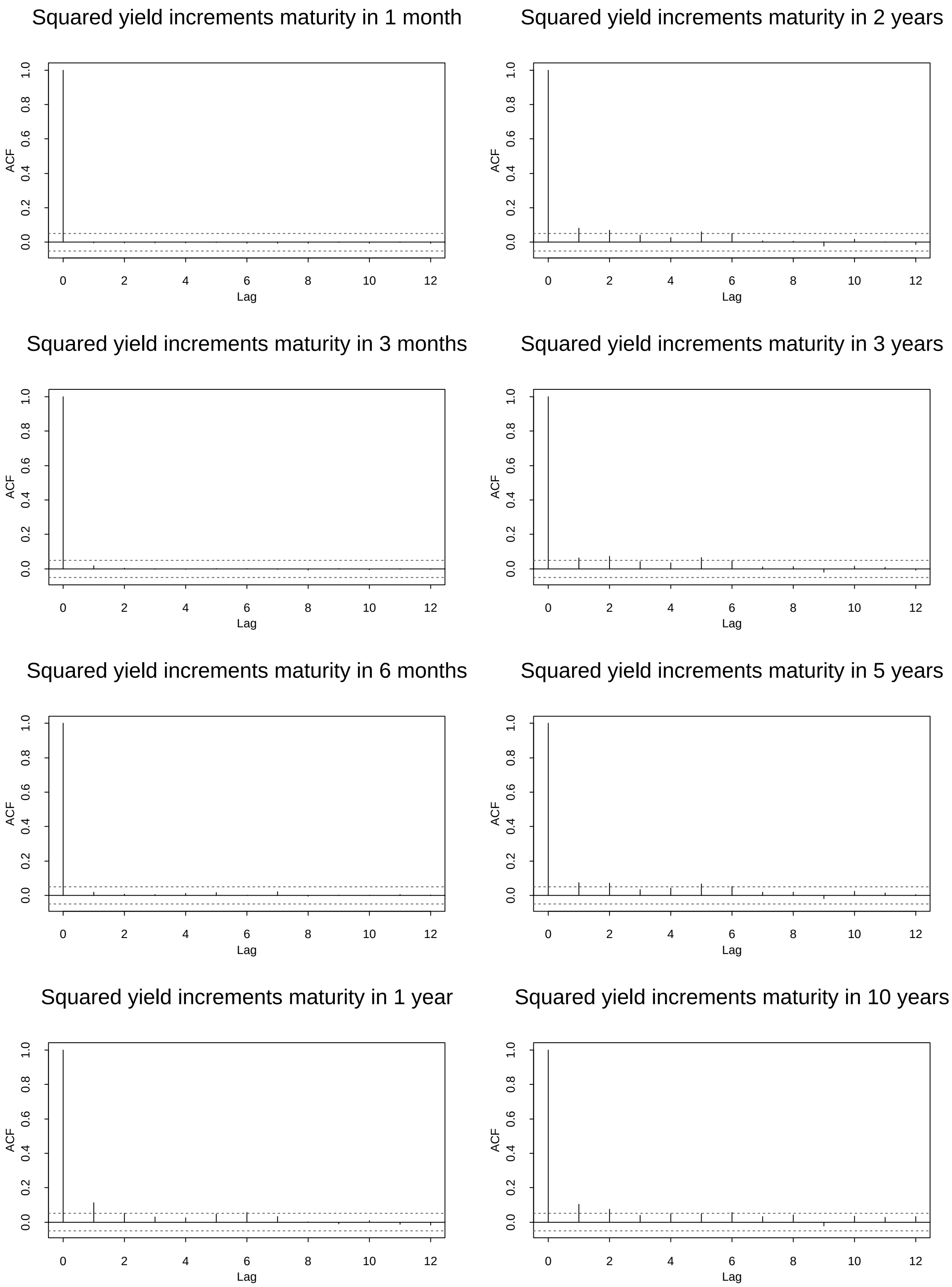


Figure 4.3: ACF of squared yield increments of the RBA bill and bond

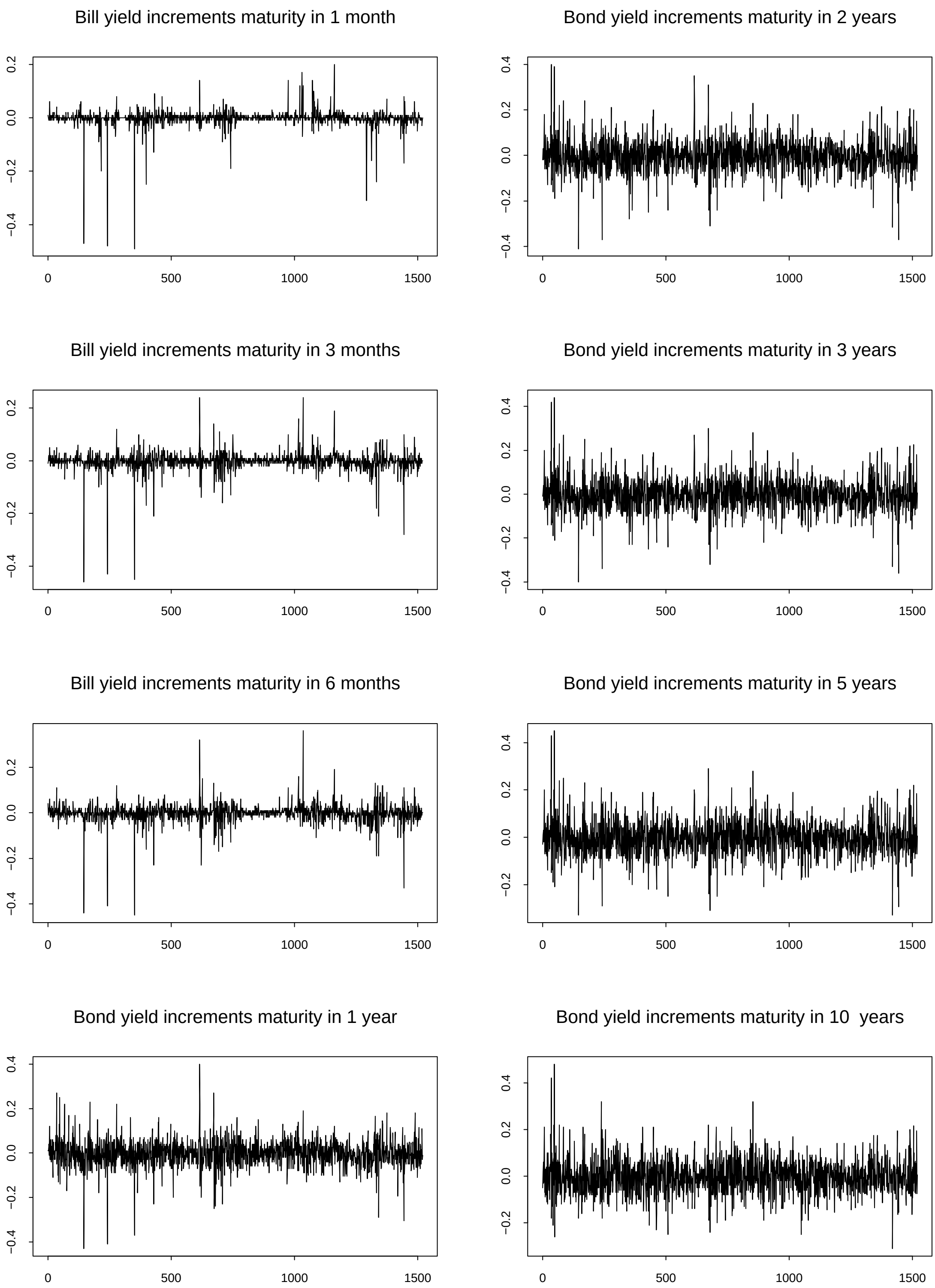


Figure 4.4: Yield increments of the RBA bill and bond

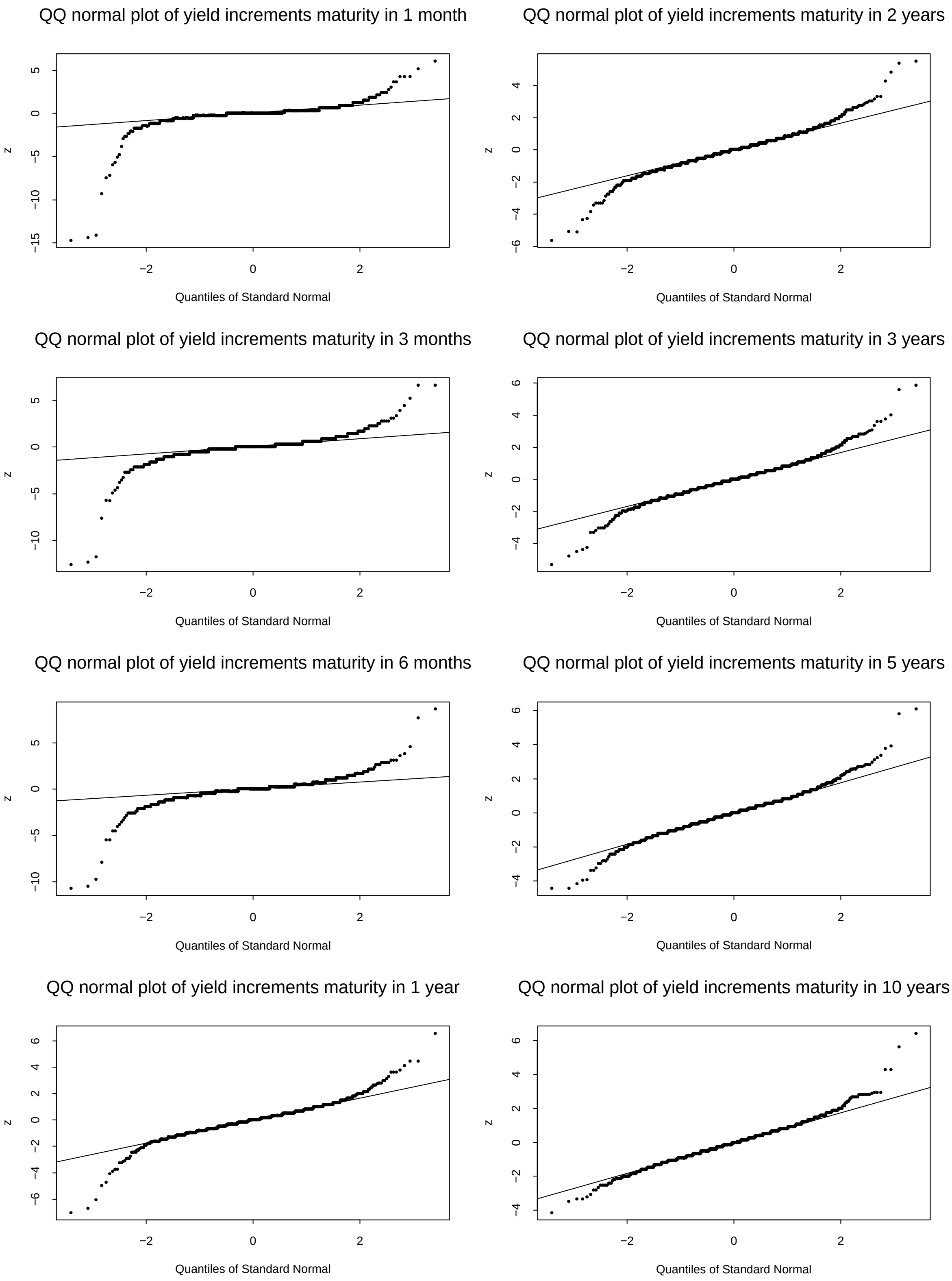
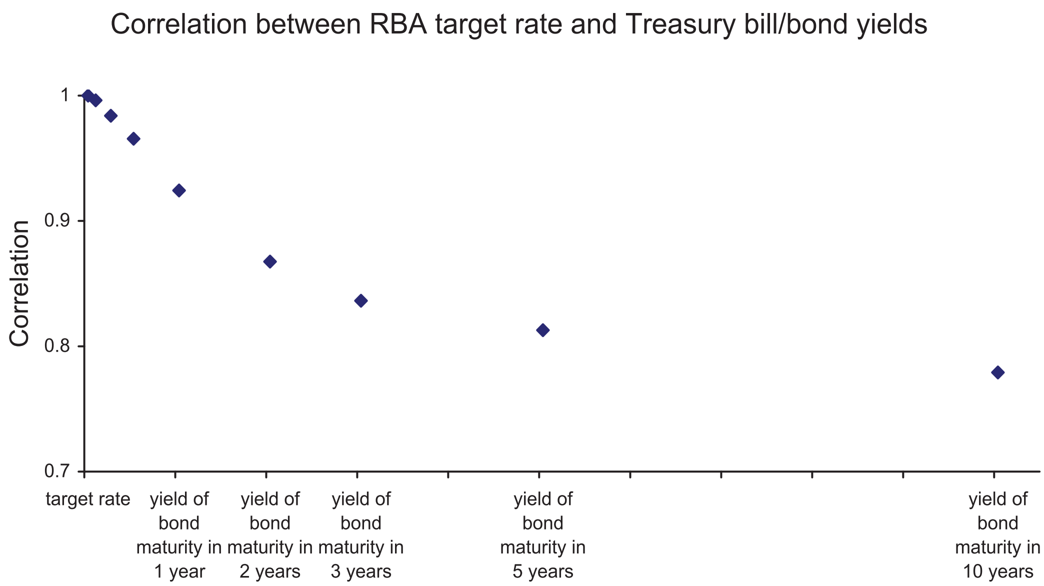


Figure 4.5: QQ normal plots of the RBA yield increments



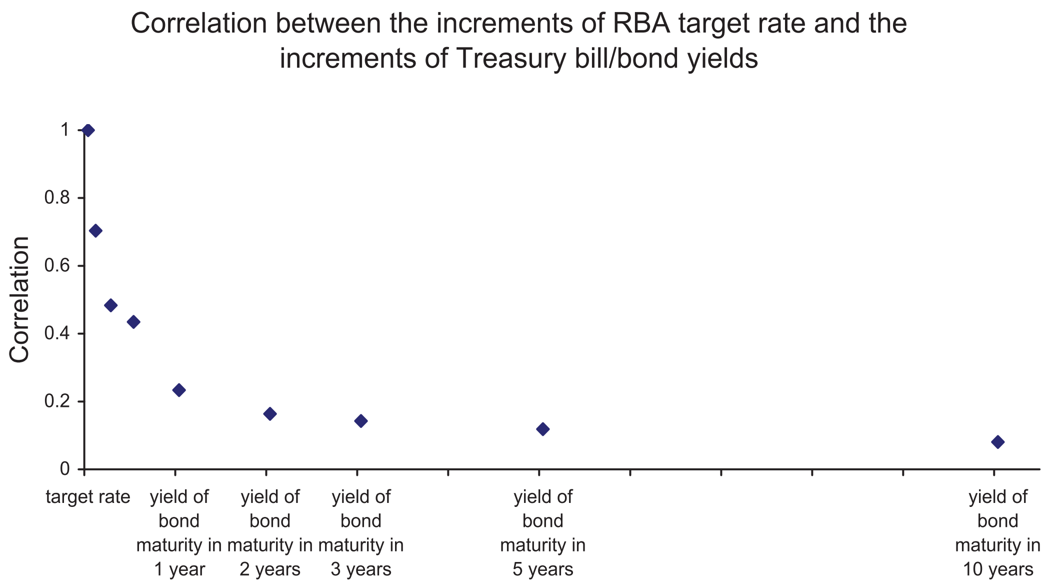


Figure 4.6: Correlation between RBA target rate and Treasury bill/bond yields

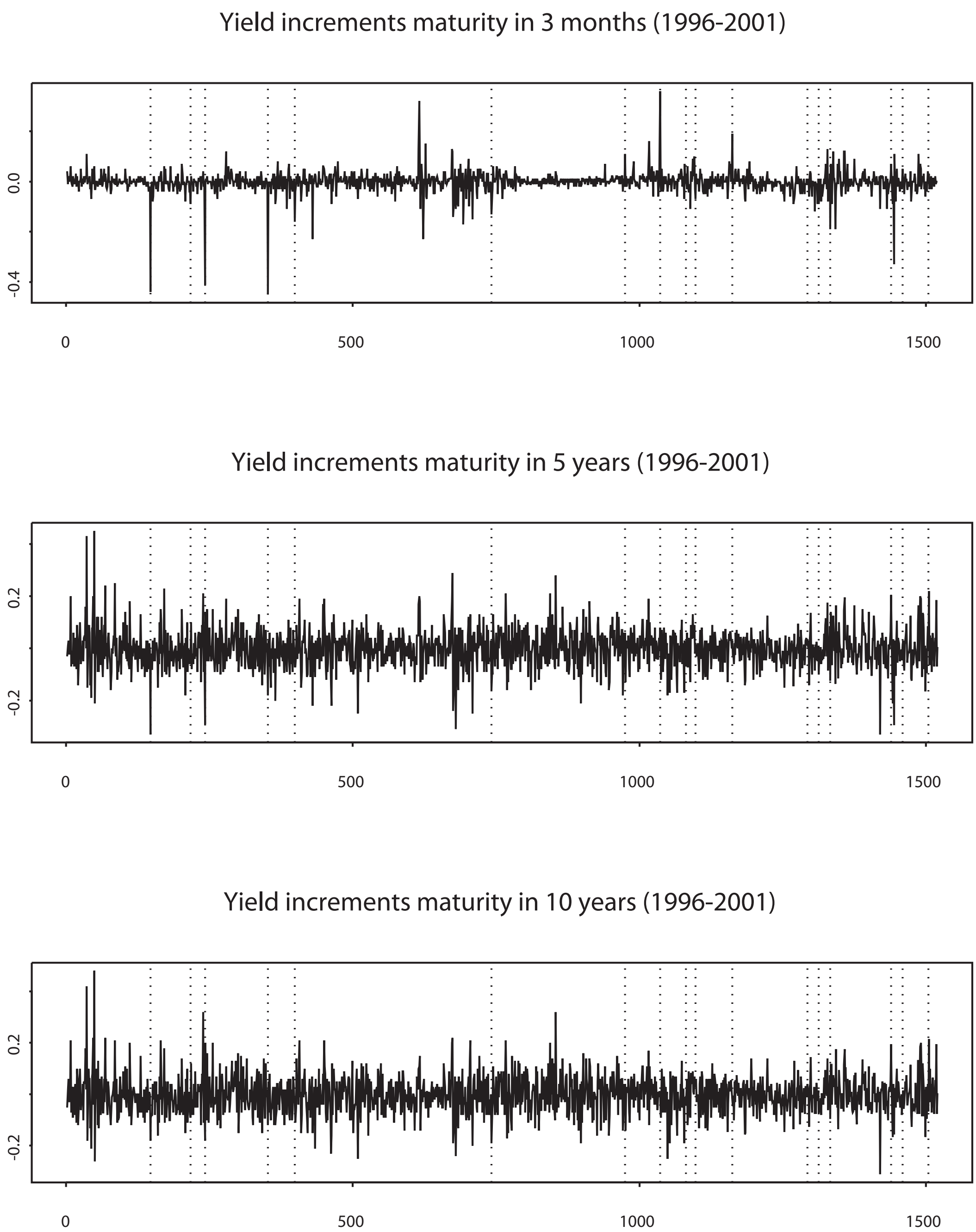


Figure 4.7: RBA yield increments, vertical lines indicate RBA decision making

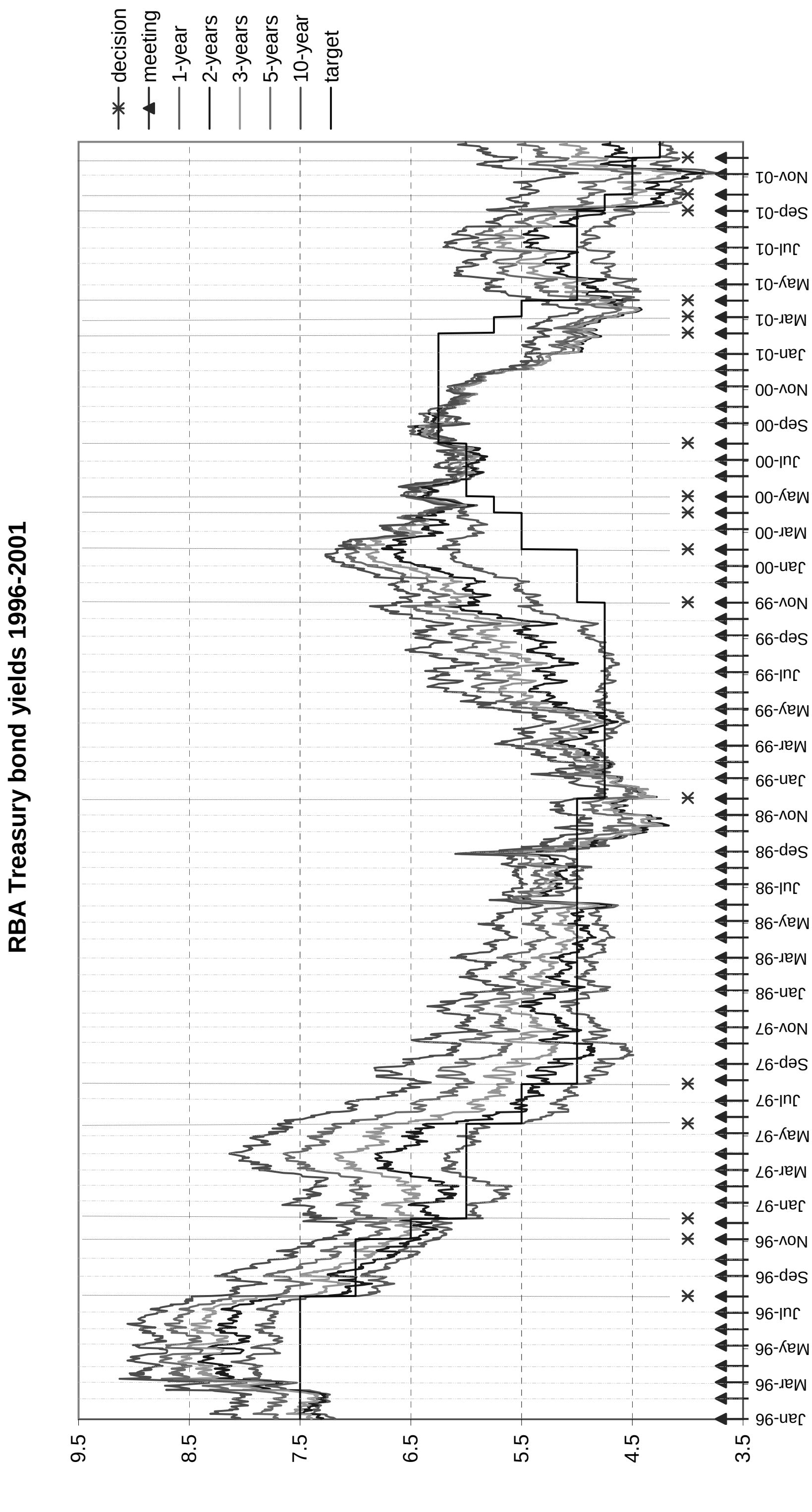


Figure 4.8: RBA yield curves 1996-2001 with indicators of RBA decision making

Chapter 5

Univariate GARCH Modelling of Yield Increments

5.1 Introduction

Chapter 4 reviewed the statistical properties of the Australian yield rates. In particular the volatility properties observed are similar to those observed in other instruments of financial markets such as shares.

Engle (1982) introduced the autoregressive conditional heteroskedasticity (ARCH) model as a way to model the volatility properties commonly encountered in financial time series. Subsequently, numerous extensions of empirical models for stochastic volatility have been applied to financial data series. These include the Generalised autoregressive conditional heteroskedasticity (GARCH) model (Bollerslev 1986), the exponential GARCH (EGARCH) model (Nelson 1991), the integrated GARCH (IGARCH) model (Hamilton 1994), and the fractionally integrated GARCH (FIGARCH) model (Baillie, Bollerslev and Mikkelsen 1996). These models have emphasised the role of persistence of the shocks in the conditional variance (or, *volatility*) process. The objective of this chapter is to develop GARCH type models for the Australian Treasury bond yield increments at each maturity level. The *yield increments* y_t with a fixed maturity θ are the difference between the successive daily yields, i.e. $y_t = R(t,\theta) - R(t-1,\theta)$, where $R(t,\theta)$ is yield of maturity θ (see Section 4.1).

It is worth emphasising that the work in this chapter (and later Chapter 6-7) develops a discrete-time model for the *yield increments* $R(t,\theta) - R(t-1,\theta)$. This approach has strong theoretical and empirical justification. On one hand the analysis of our data set (see Chapter 4) shows that, unlike the yields $R(t, \theta)$, the increments y_t form an approximately stationary sequence (mean stationary, uncorrelated process) thus making statistical modelling easier. On the other hand it provides a convenient starting point for further work on modelling of yield curves in continuous time that describes behaviour of an infinitesimal increment $dR(t, \theta)$ in time dt by multidimensional stochastic differential equations (or stochastic PDEs), as discussed in Section 3.6 under the topic of discrete-time approximation of term structure of interest rates in equations (3.36) and (3.37). The discrete-time specification of a continuous-time diffusion process model possesses the advantage that the mean and variance of the increments in the discrete-time model depend in a way consistent with those in the continuous-time model. Further the yield $R(t, \theta)$ can be derived from the appropriate GARCH model of yield increments $R(t, \theta) - R(t-1, \theta)$ accordingly (See Section 5.3).

In this chapter, the statistical analysis and modelling of interest rates (or yields) are based on yield increments. It is not surprising that the term structure of interest rates are highly related to the economic climate, and are likely to demonstrate volatility properties. The Reserve Bank of Australia target rate is the interest rate at which overnight funds are borrowed and lent in the money market. Because RBA decisions to change the target cash rates are likely to be crucial in modelling the volatility in bond yield series, we incorporate variables to indicate the RBA Board meeting dates and the RBA decisions to raise or lower target rates for yield increments and their volatility in our models. Separate variables to indicate the raising or lowering of target rates allow us to assess potential asymmetry of positive and negative innovations that have been found by other researchers (Engle and Patton 2001, Kearns and Pagan 1993).

Chapter 4 showed that there are ARCH effects in the processes of middle-to-long-term bond yield increments, and that there are no ARCH effects in the processes of short-term bill yield increments. In this chapter, we focus on middle-to-long-term bond yields and model the dynamics of yield increment series using GARCH technology. It has been widely noted in the literature that financial data series are not, in general, normally distributed. Instead, the t-distribution has often been suggested. Kolmogorov - Smirnov tests confirm the non-normality of yield increments of RBA yield from Chapter 4. This chapter develops GARCH(1,1) models, using t-distributions and exogenous variables to indicate the changes in the RBA target rates, for yield increment series maturing in 1, 2, 3, 5, or 10 years. These models can capture many important empirical features of the interest rate increments series we have noted in Chapter 4.

A major finding of this analysis is that the parameters of the GARCH(1,1) models used to specify the mean and variance equations and the degrees of freedom of the t-distribution, are closely related to the length of time to maturity θ. We show that the degrees of freedom of the fitted t-distribution is approximately linearly related to the length of time to maturity, while other parameters depend approximately linearly on the logarithm of maturity. The trends of these increasing or decreasing patterns appear to be plausible and consistent with the financial economy.

The observed functional dependence of the GARCH parameters on time to maturity suggests that the collection of GARCH models of yield increment series for a set of fixed maturities can be linked in a GARCH model having parameters specified as a function of time to maturity specified with new parameters $\boldsymbol{\varphi}$. This model, which is referred to as *GARCH model of term structure of interest rate* (*TS-GARCH* for short), can be used to characterise the bond yield increment series at any middle-to-long-term maturity. To estimate the new parameters $\boldsymbol{\varphi}$ in the TS-GARCH model, we need to use the whole available yield curves data set that is indexed by time t and a discrete set of maturities θ_j. A multivariate GARCH model is built up in Chapter 6 for the Australian Treasury yield increments. The functional dependence of the multivariate GARCH parameters on time to maturity are observed to be consistent with those from univariate GARCH model from this chapter. The extension of the concept of TS-GARCH and estimation of it will be discussed in Chapter 7 using the results from Chapter 6.

Section 2 reviews general concepts of GARCH modelling. Section 3 shows the results of GARCH(1,1) models based on Australian Treasury yield increments over the six year period 1996-2001 with maturities equal to 1, 2, 3, 5, and 10 years. Section 4 explores the functional dependence of the term structured GARCH parameters. Section 5 proposes a GARCH model of term structure of interest rates (TS-GARCH). Finally, Section 6 presents a summary of the results and discussion the future work.

5.2 GARCH Modelling

Engle and Patton (2001) suggested that a good volatility model might be able to forecast volatility as the central requirement in almost all financial applications. They outline the major stylised facts about volatility that should be incorporated in the model; they are:

- Volatility persistence;
- Volatility mean reversion;
- Asymmetry in which the signs of previous innovations affect volatility;
- The possibility of exogenous or pre-determined variables influencing volatility;
- Excess Kurtosis in the distribution of yield increments relative to the Normal distribution (heavy tail).

We now explain each of these concepts in detail.

Volatility clustering/persistence. One feature of volatility is volatility persistence, which means the clustering of large changes and small changes of financial time series. Large changes tend to be followed by large changes, and small changes tend to be followed by small changes. This property implies that the volatility comes and goes, and a volatility shock today will influence the expectation of volatility many periods into the future.

Volatility Mean Reversion. Financial markets may experience excessive volatility from time to time. However, it appears that long run forecasts of volatility will eventually converge to a certain normal level of volatility. Volatility mean reversion implies that the volatility will settle down in the long run forecast that is not affected by the current information. (See Engle and Patton 2000).

Market Asymmetry. An asymmetric effect of market activity upon volatility is that the negative shocks lead to larger volatility than a numerically equivalent positive shocks. Black (1976) found changes in stock return volatility to be negatively correlated, implying that a decrease in return is likely to be accompanied by an increase in volatility and vice versa. Other researchers have confirmed these results. Kearns and Pagan (1993) studied the Australian market, and found that there is weaker evidence for the asymmetric effect in Australian data than in the US data.

Exogenous or pre-determined variables. It is common knowledge that the financial asset prices and their volatility are highly related to the economic climate. The variables containing relevant information affecting the volatility of the financial series are referred to as exogenous variables.

Excess Kurtosis of yield returns (heavy tail). It is well known that the distributions of high frequency financial time series usually have excess kurtosis in the conditional distribution of asset increments than those of the normal distribution. Large changes occur more frequently than a normal distribution would imply. Typically kurtosis estimates larger than 3 indicate non-normality. The volatility model used must be incorporated in capable of allowing for fat tails in the conditional density of asset increments.

Autoregressive conditional heteroskedasticity (ARCH) models developed by Engle (1982) and other extended ARCH models are able to model time varying volatility and capture many observed volatility behaviours in financial time series.

Generalised autoregressive conditional heteroskedasticity GARCH (Bollerslev 1986) models extended the ARCH class of models to allow for both a longer memory and a more flexible lag structure. Let ϵ_t denote a real-valued discrete-time stochastic process, and $\mathcal{F}_t$ the set of all available information through time t. A GARCH(p, q) process is then given by

$$\epsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t), \tag{5.1}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}, \tag{5.2}$$

where

$$p \geq 0, \quad q > 0,$$

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad i = 1, \cdots, q,$$

$$\beta_i \geq 0, \quad i = 1, \cdots, p.$$

Theorem 5.1 *The GARCH(p, q) process defined by (5.1) and (5.2) is wide-sense stationary with $E(\epsilon_t) = 0$, $var(\epsilon_t) = \alpha_0(1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \beta_i)^{-1}$ and $cov(\epsilon_t, \epsilon_s) = 0$ for $t \neq s$ if and only if $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 1$.*

Proof. See Bollereslev (1986). ■

As noted above, financial data series often exhibit non-normal behaviour. For our data series, we have assumed that the distribution of ϵ_t conditional on the past follows a t-distribution with the degrees of freedom denoted by ν and the conditional variance by h_t. The conditional density function of the t-distribution with degrees of freedom ν is

$$f_{\epsilon_t|\mathcal{F}_{t-1}}(x|h_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi(\nu-2)h_t}}(1 + \frac{x^2}{(\nu-2)h_t})^{-(\nu+1)/2}. \tag{5.3}$$

It should be noted that theorem 5.1 also holds if ϵ_t are independently and identically distributed random variables with finite variance and a general distribution. This includes the t-distribution with $\nu > 2$.

The GARCH(p, q) regression model with the exogenous variables in both mean and variance equations is

$$\begin{aligned} y_t &= a + \boldsymbol{b}'\boldsymbol{u}_t + \epsilon_t, \quad \epsilon_t|\mathcal{F}_{t-1} \sim (0, h_t), \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} + \boldsymbol{c}'\boldsymbol{v}_t, \end{aligned}$$

where y_t is the dependent variable, $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$ are vectors of explanatory variables, $\boldsymbol{b}$ and $\boldsymbol{c}$ are unknown vector-valued parameters.

Diagnostic tests to examine the adequacy of the GARCH model explaining the volatility are based on the standardised residuals, $\epsilon_t/\sqrt{h_t}$, where h_t is the estimated conditional variance. To assess the correct distribution of ϵ_t, quantile-quantile plots of the standardised residuals against an appropriate reference distribution (such as t-distribution) can be used. The Ljung-Box test is used to test that there is no autocorrelation left in the standardised residuals and squared standardised residuals series. The simple Lagrange Multiplier (LM) test is used to check that there are no ARCH effects left with the standardised residuals of the model. The S+FinMetrics module is used to fit the models and to conduct these tests.

5.3 A GARCH(1,1)-t Model of Yield Increments

In Chapter 4, we found that the dynamics of the short-term bill yield increments and the middle-to-long-term bond yield increments behave differently. There are autoregressive conditional heteroskedasticity (ARCH) effects with the middle-to-long-term bond yield increment and no ARCH effects with short-term bill yield increments. In this section, we model the increments of middle-to-long-term Australian Treasury bond yields using GARCH modelling applied to the data set for the period 1996-2001 from the Reserve Bank of Australia.

Usually a GARCH(1,1) model is adequate to obtain a good model fit for financial time series. However, Vilasuso (2002), for example, obtained a more accurate forecast of the volatility of the exchange rate using a Fractionally Integrated GARCH (FIGARCH) model (Baillie, Bollerslev, and Mikkelsen 1996) comparing with GARCH or Integrated GARCH (IGARCH) models. We initially compared the fit of simple GARCH(1,1), FIGARCH(1,1), and GARCH(1,1)-t (assuming standardised residuals in t-distribution) models. Then we add the exogenous innovation variables in GARCH(1,1), FIGARCH(1,1), and GARCH(1,1)-t models. At all maturities of 1, 2, 3, 5, and 10 years, we found that the AIC, BIC and log-likelihood of FIGARCH(1,1) and QQ-plots of residuals from FIGARCH were very close to those of GARCH(1,1). Among these three models, the GARCH(1,1)-t gives the smallest AIC and BIC, the biggest likelihood, and QQ-plots shows that the t-distribution is appropriate with the residuals. Because of this, we concentrate on GARCH(1,1)-t for the remainder of this chapter.

The GARCH(1,1)-t model was extended to include exogenous innovation variables incorporating variables to capture the decision making activities of the RBA regarding the target cash rate. The equations for yield increments and for volatility are as follows:

$$\begin{aligned} y_t &= \beta_1 + \beta_2 R_t^- + \beta_3 R_t^+ + \epsilon_t, \quad \epsilon_t|\mathcal{F}_{t-1} \sim t_{\beta_0}(0, h_t), \\ h_t &= \beta_4 + \beta_5 \epsilon_{t-1}^2 + \beta_6 h_{t-1} + \beta_7 R_t^- + \beta_8 R_t^+ + \beta_9 M_t, \end{aligned} \tag{5.4}$$

where R_t^- is an indicator variable taking the value 1 if RBA lowered its target rate and value 0 otherwise, R_t^+ is similar but takes the value 1 when the RBA raised its target rate, and, M_t is an indicator taking the value 1 if time t was a RBA Board meeting date and 0 if it was not. Note that we have used separate variables, R_t^- and R_t^+, to allow for possible asymmetric effects on raising/lowering the target rate.

Fitting of GARCH processes is known to be computationally challenging (See Bollerslev, 1986). S+FinMetrics was used to fit the GARCH(1,1)-t model (5.4). Estimates of this GARCH model are shown in Table 5.1. Exogenous variables are added in the models. The impact of R^- and R^+ on the mean yield increments are both statistically significant, except for the impact on R^+ for yields with a 10 year maturity. A likelihood ratio test (H_0: $B_2 = B_3$ v H_a: $B_2 \neq B_3$) indicates that the coefficients of R^- and R^+ are statistically significantly different (p-values are 0.000, 0.001, 0.002, 0.002 and 0.024 for yield return maturities in 1, 2, 3, 5 and 10 years respectively), while a lowering of the target rate having greater impact on the mean. The coefficients of R^+ in the variance equations have nearly zero negative values and are insignificant except for yields at 5 year maturity. The coefficients of R^- in the variance equation are significantly positive for all maturities. The likelihood ratio test indicates that the coefficients of R^- and R^+ are statistically significantly different (p-values are 0.049, 0.033, 0.026, 0.011 and 0.162

Table 5.1: Output of the initial univariate GARCH(1,1)-t models

	maturity	1 year	2 years	3 years	5 years	10 years
coefficient						
t-distribution	β_0: df	4.92	5.52	5.51	6.34	8.08
Mean	β_1: Constant	-0.0018	-0.0021	-0.0022	-0.0023	-0.0025
	β_2: R_t^-	-0.0983	-0.0926	-0.0926	-0.0840	-0.0742
	β_3: R_t^+	0.0709	0.0622	0.0516	0.0379	0.0073
Variance	β_4: Constant	0.0008	0.0008	0.0008	0.0007	0.0005
	β_5: ARCH(1)	0.0818	0.0604	0.0593	0.0540	0.0527
	β_6: GARCH(1)	0.6649	0.7597	0.7772	0.8049	0.8382
	β_7: R_t^-	0.0100	0.0082	0.0076	0.0050	0.0024
	β_8: R_t^+	0.0016	-0.0007	-0.0014	-0.0028	-0.0013
	β_9: M_t	0.0007	0.0013	0.0014	0.0010	0.0009
p-value						
mean	β_1: Constant	0.075	0.100	0.098	0.093	0.085
	β_2: R_t^-	0.000	0.000	0.000	0.000	0.001
	β_3: R_t^+	0.001	0.003	0.017	0.039	0.432
variance	β_4: Constant	0.000	0.004	0.007	0.010	0.019
	β_5: ARCH(1)	0.001	0.002	0.002	0.003	0.003
	β_6: GARCH(1)	0.000	0.000	0.000	0.000	0.000
	β_7: R_t^-	0.004	0.012	0.017	0.033	0.113
	β_8: R_t^+	0.289	0.381	0.254	0.024	0.249
	β_9: M_t	0.152	0.072	0.069	0.127	0.157

for yield return maturities in 1, 2, 3, 5 and 10 years respectively), while a lowering of the target rate having greater impact on the variance equation. It implies that a lowering of the target rate leads to significant increasing volatility and a raising of the target rate does not change volatility. Thus the response of volatility to RBA changes in target rates are asymmetric. The RBA meeting date variable M_t in the variance equations is weakly significant for maturities of 2 and 3 years, but insignificant for maturities of 1, 5, and 10 years. Also, all coefficients of the RBA Board meeting data variable are very close to zero.

Because the indicator variables of raising cash rate R_t^+ and RBA Board meeting M_t are mostly not significant and very small values of these estimated coefficients, these two variables were excluded from the variance equation in our GARCH model. The final resulting model is

$$\begin{aligned} y_t &= \beta_1 + \beta_2 R_t^- + \beta_3 R_t^+ + \epsilon_t, \quad \epsilon_t | \mathcal{F}_{t-1} \sim t_{\beta_0}(0, h_t), \\ h_t &= \beta_4 + \beta_5 \epsilon_{t-1}^2 + \beta_6 h_{t-1} + \beta_7 R_t^-. \end{aligned} \tag{5.5}$$

The estimated parameters of model (5.5) are shown in Table 5.2. Estimators of the GARCH(1) coefficients β_6 for 1-year to 10-years bond yield increments are in the range of 0.60 to 0.84 implying volatility persistence (See Chapter 7 of Zivot and Wang 2002). The sums of ARCH(1) and GARCH(1) are less than 1 for all maturities, implying that the residual processes ϵ_t are wide-sense stationary by Theorem 5.1 and show mean reversion in volatility. The degrees of freedom for the t-distribution range from 5 to 8, implying different distributions for each individual yield series.

Overall, the final GARCH(1,1)-t model captures the main characteristics of the volatility of yield increments, namely heavy tails, asymmetry with negative or positive innovations, mean reversion in volatility, persistent volatility and wide sense stationary residuals.

Quantile-quantile (QQ-t plots, Figure 5.1) of the standardised residuals from the final GARCH(1,1)-t models show graphically that the t-distributions provide a good model for the distribution of the yield increment residuals. The $QQ - t$ plots in Figure 5.1 shows that the majority of standardised residuals fall on the straight line of QQ-t with a few outliers. Compared to the QQ-normal plots in Figure 5.2, the use of the t-distribution with estimated degree of freedom shows a marked improvement. Even with this improvement, Kolmogorov-Smirnov Tests reject the t-distributions of the residuals (p-values< 0.001),

Table 5.2: Output of the final univariate GARCH(1,1)-t models

	maturity	1 year	2 years	3 years	5 years	10 years
coefficient						
t-distribution	β_0: df	5.0	5.4	5.4	6.3	7.9
mean	β_1: Constant	-0.0017	-0.0021	-0.0022	-0.0022	-0.0024
	β_2: R_t^-	-0.1001	-0.0910	-0.0907	-0.0836	-0.0736
	β_3: R_t^+	0.0728	0.0625	0.0512	0.0326	0.0123
variance	β_4: Constant	0.0010	0.0009	0.0008	0.0006	0.0006
	β_5: ARCH(1)	0.0931	0.0615	0.0608	0.0529	0.0536
	β_6: GARCH(1)	0.6036	0.7537	0.7760	0.8207	0.8400
	β_7: R_t^-	0.0119	0.0088	0.0081	0.0051	0.0026
	ARCH(1)+GARCH(1)	0.6967	0.8152	0.8368	0.8736	0.8937
p-value						
mean	β_1: Constant	0.087	0.099	0.100	0.096	0.087
	β_2: R_t^-	0.000	0.000	0.000	0.000	0.001
	β_3: R_t^+	0.000	0.003	0.024	0.160	0.402
variance	β_4: Constant	0.000	0.003	0.005	0.010	0.014
	β_5: ARCH(1)	0.001	0.002	0.002	0.003	0.002
	β_6: GARCH(1)	0.000	0.000	0.000	0.000	0.000
	β_7: R_t^-	0.005	0.008	0.011	0.021	0.083

Table 5.3: Tests of the RBA bill and bond yield increments

Ljung-Box test for Autocorrelation

Null Hypothesis: no autocorrelation of yield increments

	1y	2y	3y	5y	10y
Test Stat	4.231	3.855	3.981	3.454	7.608
p.value	0.9789	0.9859	0.9838	0.9914	0.815

Ljung-Box test for Autocorrelation

Null Hypothesis: no autocorrelation of squared yield increments

	1y	2y	3y	5y	10y
Test Stat	9.21	5.598	4.819	5.618	6.851
p.value	0.6849	0.935	0.9637	0.9341	0.8673

Lagrange Multiplier test for ARCH Effects

Null Hypothesis: no ARCH effects of yield increments

	1y	2y	3y	5y	10y
Test Stat	9.108	5.776	5.026	5.799	6.613
p.value	0.6937	0.927	0.957	0.9259	0.8821

suggesting the t-distributions may not be sufficiently heavy tailed for these series. Future work could be directed at improving the model by using a different distribution than the t-distribution.

The ACF plots (Figure 5.3) show that the standardised residuals and squared standardised residuals are uncorrelated. The Ljung-Box tests (Brockwell and Davis 2002) confirm the standardised residuals and squared standardised residuals are uncorrelated. The Lagrange Multiplier tests (Lee, J. H. H. 1991) revealed no more ARCH effects in the standardised residuals. See Table 5.3. All of these diagnostics support the conclusion that the GARCH(1,1) model specified by model (5.5) is adequate for the 1-year to 10-years bond yield increments.

As we have emphasised in Section 1, the work in this chapter develops a discrete-time model for the *yield increments* $y_t = R(t,\theta) - R(t-1,\theta)$ at each maturity. The process of the yield $R(t,\theta)$ can be derived from the GARCH model of yield increments y_t by model (5.5) . That is

$$R(t,\theta) = \beta_1 + \beta_2 R_t^- + \beta_3 R_t^+ + R(t-1,\theta) + \epsilon_t, \tag{5.6}$$

or

$$R(t,\theta) = R(0,\theta) + \beta_1 t + \beta_2 \sum_{j=0}^{t} R_j^- + \beta_3 \sum_{j=0}^{t} R_j^+ + \sum_{j=0}^{t} \epsilon_t, \tag{5.7}$$

where $\epsilon_j|\mathcal{F}_{j-1} \sim t_{\beta_0}(0, h_j)$ and $h_j = \beta_4 + \beta_5 \epsilon_{j-1}^2 + \beta_6 h_{j-1} + \beta_7 R_j^-$.

5.4 Term Structured GARCH Parameters

In the last section, we presented a specific GARCH(1,1)-t model and showed that it is adequate for the bond yield increments with maturity of 1, 2, 3, 5, and 10 years. In this section, we will explore the functional patterns of the estimated parameters from the GARCH(1,1)-t model as functions of the maturity levels. The estimates of the parameters are given in Table 5.2.

Figure 5.4 shows plots of the estimated coefficients of GARCH(1,1) model (5.5) versus time to maturity. It can be observed that the GARCH parameters are functionally dependent on the length of time to maturity. Using least squares estimation, linear functions of log time to maturity are obtained for the individual coefficients. In Figure 5.4, solid diamonds represent the estimated coefficients of the individual GARCH(1,1) model (5.5), while the straight line in Panel (β_0) represents the optimal fitting of the linear regression of degree of freedom of t-distribution on length of time to maturity, and other lines in Panel (β_1)-(β_7) represent linear regression of other parameters on the logarithm of length of time to maturity. Regarding Figure 5.4, note that:

Panel (β_0) shows that the degrees of freedom of the t-distribution linearly increases with the length of time to maturity, which implies that the residual distribution is less heavy tailed for longer term bond yield increments.

Panel (β_1) shows that the absolute value of mean yield increments increases as the maturity increases. However, the mean level of the yield increments, ignoring any changes due to cash rates, have estimated values all close to zero (around -0.002, see Table 5.2). These β_1 are not significant from zero at the 5% level, implying that the mean levels of yield increment are nearly zero if there are no RBA cash rates innovations.

Panel (β_2) shows the change in mean yield increments associated with RBA decision to lower the cash rate. Absolute values decrease linearly in mean along the logarithm of length of time to maturity, which implies that the longer term bond average yield increments are less impacted by the RBA's decisions to lower target cash rates.

Panel (β_3) shows the change in mean yield increments associated with a RBA decision to raise the cash rate. Changes decrease linearly as the logarithm of length of time to maturity increases. This implies that the longer term bond average yield increments are less impacted by the RBA's decisions to raise cash rates.

Panel (β_4) shows that the mean levels of the conditional volatility of yield increments decreases linearly with increasing values of the logarithm of length of time to maturity, which implies that the longer term bond yields have smaller average volatility.

Panel (β_5) suggests that the coefficients of the ARCH(1) (component ϵ_{t-1}^2) decrease linearly with the logarithm of length of time to maturity, which implies that the longer term bond yields are less affected by the previous squared residual ϵ_{t-1}^2. This also suggests that as maturity increases, the more rapid is reversion to the mean. Note that the fit of the linear log maturity line does not give as good a representation of ARCH(1) coefficients as it does for other coefficients.

Panel (β_6) gives the coefficients of the GARCH(1) (component h_{t-1}). These increase linearly with the logarithm of length of time to maturity, which imply that the longer term bond yield increments have greater persistent volatility.

Panel (β_7) shows that the change in volatility associated with the RBA's decisions to lower the cash rate decreases linearly with the logarithm of length of time to maturity. This implies that the volatility of longer maturity bond yields is less impacted by decreases in the cash rate.

In summary, the trend of the pattern of GARCH(1,1) coefficients shows that longer maturity yield increment series are less heavy tailed, are more efficient in mean reversion volatility, have greater persistent volatility, and are less impacted by changes in the cash rates in both mean and volatility. Generally speaking, all of these observations appear to be plausible.

Overall, the representation of model parameters (shown in Figure 5.4) by a linear function in maturity (for the degrees of freedom) or by a linear function in the logarithm of maturity (for all other parameters) is reasonable. Exceptions to this appears to be for the ARCH(1) and GARCH(1) parameters for yields in 1 year maturity. However, the sum of these two parameters conforms more closely to the linear log relationship. In view of these observations we will proceed on the assumption that the linear in log representation provides a parsimonious depiction of the yield increments for all maturity.

From the individual estimated GARCH(1,1)-t coefficient, the estimates of the functional patterns by least squares estimation is

$$
\begin{aligned}
\beta_{i0} &= 0.32470\theta_i + 4.6603, \\
\beta_{i1} &= -0.0003\ln(\theta_i) - 0.0018, \\
\beta_{i2} &= 0.0111\ln(\theta_i) - 0.1004, \\
\beta_{i3} &= -0.0272\ln(\theta_i) + 0.0773, \\
\beta_{i4} &= -0.0002\ln(\theta_i) + 0.001, \\
\beta_{i5} &= -0.016\ln(\theta_i) + 0.0826, \\
\beta_{i6} &= 0.0983\ln(\theta_i) + 0.6467, \\
\beta_{i7} &= -0.0041\ln(\theta_i) + 0.0119,
\end{aligned} \tag{5.8}
$$

where i represent i-th yield increment series maturity in θ_i months. These are the equations for the fitted lines on Figure 5.4. Also given in Figure 5.4 are the R^2 measures. These are typically indicative of high levels of fit ($R^2 > 0.9$) except for ARCH β_5 coefficients ($R^2 = 0.72$) and GARCH β_6 coefficients ($R^2 = 0.85$).

5.5 A Proposed GARCH Model of Term Structure of Interest Rates

In the last section we described the patterns of the estimated parameters from the GARCH(1,1)-t model as functions of time to maturity. In this section, we will propose a single GARCH model with functional parameters depending on the maturity level in terms of new parameters that can be applied to all yield increment series in any possible middle-to-long-term maturity.

In Section 3, we found a GARCH(1,1)-t model described by model (5.5), which adequately captures many important aspects of yield increments. The phenomena of functional dependence of term structured GARCH parameters, from Section 4, suggests that the collection of GARCH models for a set of fixed maturities can be linked to a single

GARCH model with functional parameters depending on time to maturity specified with new parameters $\boldsymbol{\varphi}$. That is

$$
\begin{aligned}
& y_{it} = \beta_{i1} + \beta_{i2} R_t^- + \beta_{i3} R_t^+ + \epsilon_{it}, \\
& \epsilon_{it} | \mathcal{F}_{t-1} \sim t_{\beta_{i0}}(0, h_{it}) \\
& h_{it} = \beta_{i4} + \beta_{i5} \epsilon_{i,t-1}^2 + \beta_{i6} h_{i,t-1} + \beta_{i7} R_t^-, \\
& \beta_{i0} = \varphi_{10} + \varphi_{20} \theta_i = (1, \theta_i) \boldsymbol{\varphi}_0, \\
& \beta_{ik} = \varphi_{1k} + \varphi_{2k} \ln(\theta_i) = (1, \ln(\theta_i)) \boldsymbol{\varphi}_k, \quad k = 1, \ldots, 7,
\end{aligned}
\tag{5.9}
$$

where i represents the i-th yield increment series maturity in θ_i month, $\boldsymbol{\varphi}_0 = (\varphi_{10}, \varphi_{20})'$, and $\boldsymbol{\varphi}_k = (\varphi_{1k}, \varphi_{2k})'$, $k = 1, 2, \ldots, 7$, are new parameters. We refer to this model as a *GARCH model of term structure of interest rate* (TS-GARCH) that will be further developed in Chapter 7 .

A major advantage of the TS-GARCH model is that it can be applied to the complete collection of all yield increment series for any possible maturity level, and this should be of considerable practical utility. In particular, we can use data available at a subset of maturities to fit the TS-GARCH model and use it to model interest rate increments at any possible maturity.

5.6 Summary and Conclusions

We have shown that, for Australian Treasury bond yields at each fixed maturity, a GARCH(1,1) model with exogenous variables reflecting the RBA's decisions on cash rate changes and residuals following a t-distribution, captures many important aspects of the yield increments. Model diagnostics showed that GARCH(1,1)-t is adequate for each of the middle-to-long-term bond yield increments series. No ARCH effects remain in the standardised residuals. That is, there remains no autocorrelation in the standardised residuals or their squares. Moreover, it was shown that the parameters of the GARCH(1,1)-t models are functionally dependent on the length of time to maturity. The general shapes of these functional patterns observed appear to be plausible with financial economics, but a theoretical explanation for the linear log relationship requires development. A GARCH model of term structure of interest rates is proposed, referred to as TS-GARCH, which is a GARCH model with parameters expressed as functions of maturity in terms of new parameters.

Section 4 explored the patterns of the functional parameters of GARCH(1,1)-t model (5.5) based on individual yield increment series of Australian Treasury bonds and maturity in 1, 2, 3, 5, or 10 years, and an estimate of TS-GARCH (5.9) was given by (5.8). This estimation used a two step approach based on univariate GARCH models, which firstly estimated the individual univariate GARCH coefficients by maximum log-likelihood estimation (MLE) and then estimated these new parameters describing the patterns of GARCH coefficients by least squares estimation. However, by this approach, the possible correlation between yield increments at different maturities were ignored.

To estimate the new parameters $\boldsymbol{\varphi}$ in the TS-GARCH model (5.9), it is first necessary to develop a multivariate GARCH model which uses the whole yield curve data set indexed by time t and available maturities θ. This is done in Chapter 6. A Generalised TS-GARCH model is proposed in Chapter 7 that expands the ideas of this chapter to the conditional covariance GARCH processes. The estimation of the TS-GARCH parameters is based on the estimation from a multivariate GARCH model developed in Chapter 6.

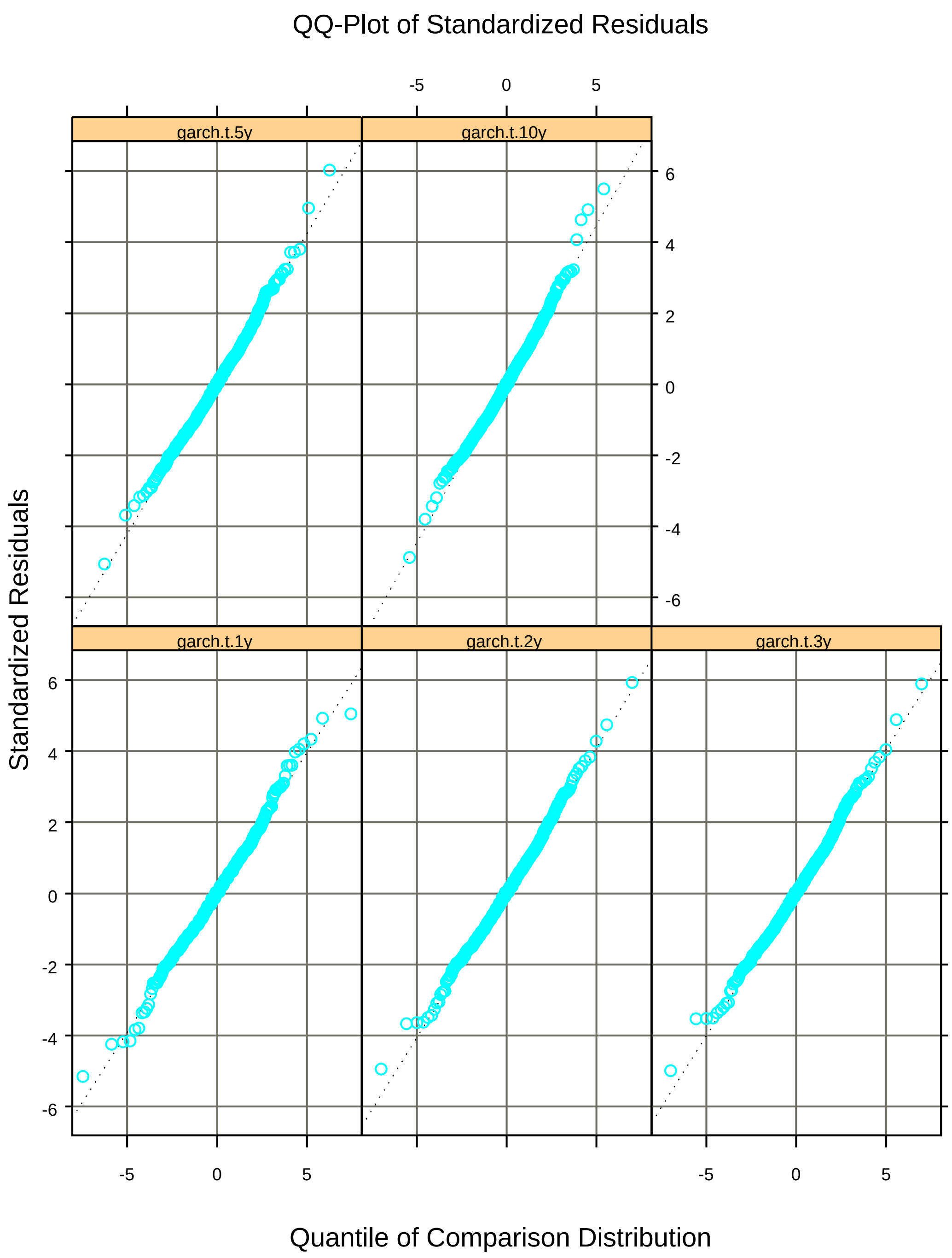


Figure 5.1: QQ plots of GARCH(1,1) in t-distributions

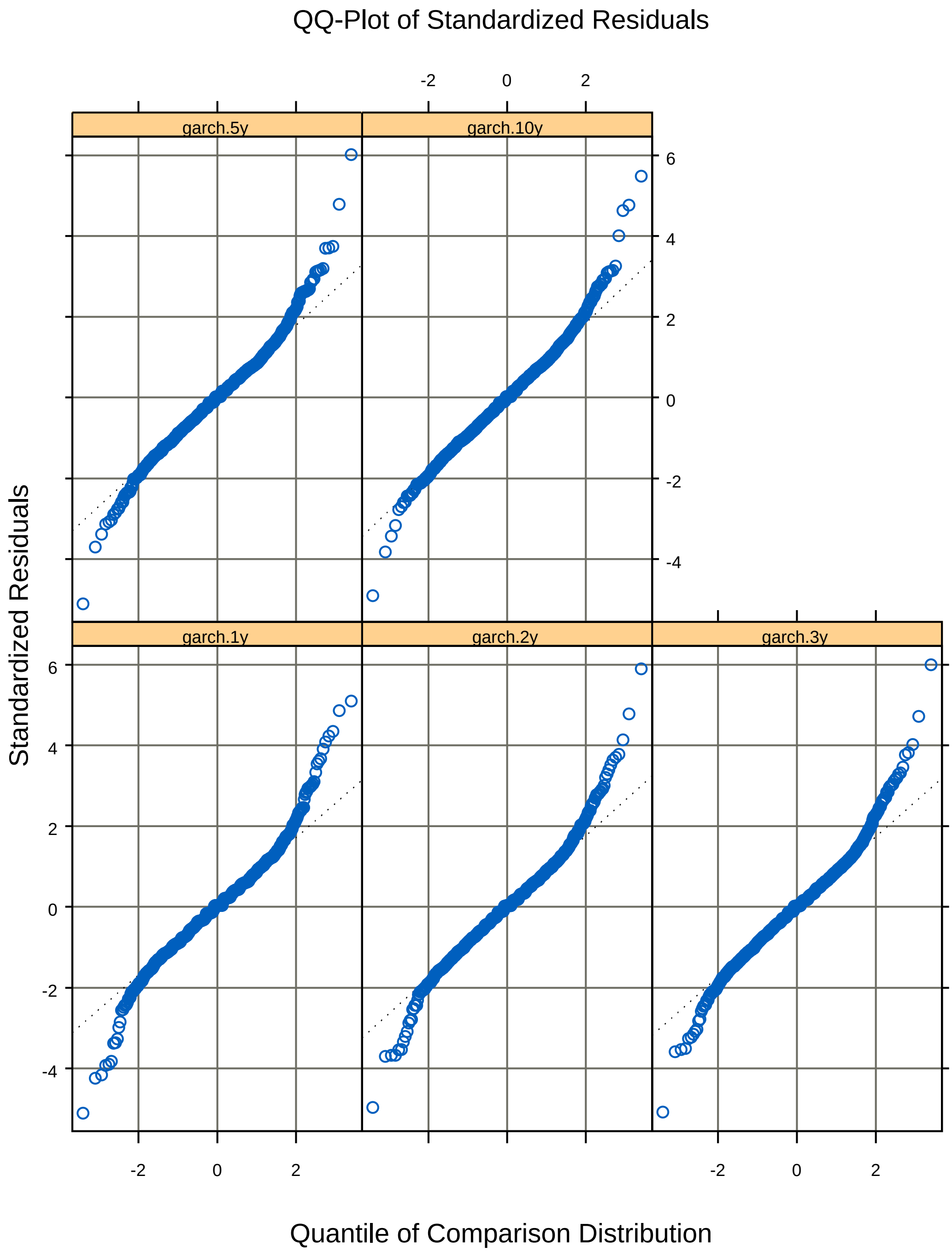


Figure 5.2: QQ plots of GARCH(1,1) in Normal distributions

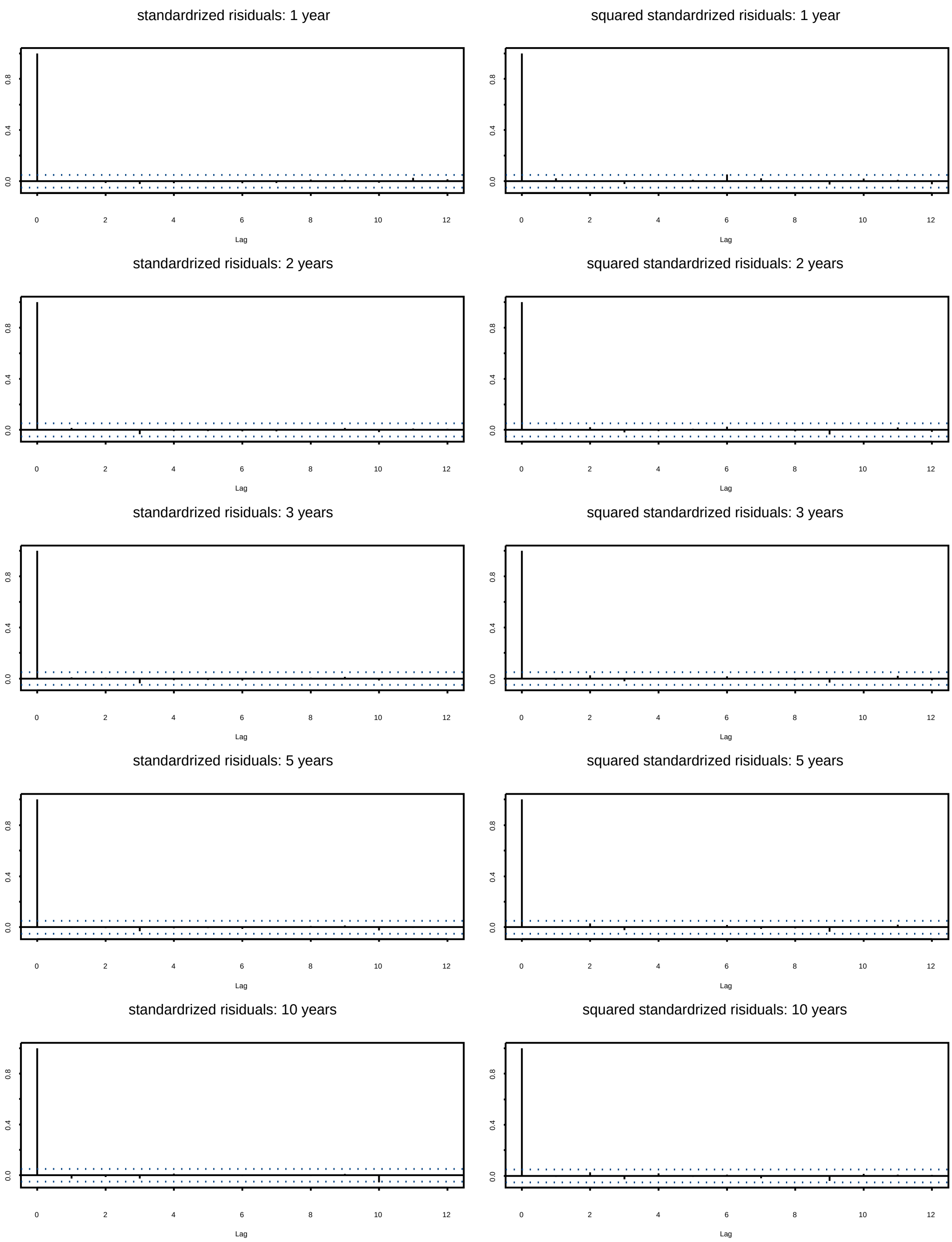


Figure 5.3: ACF of standardised residuals and squared standardised residuals

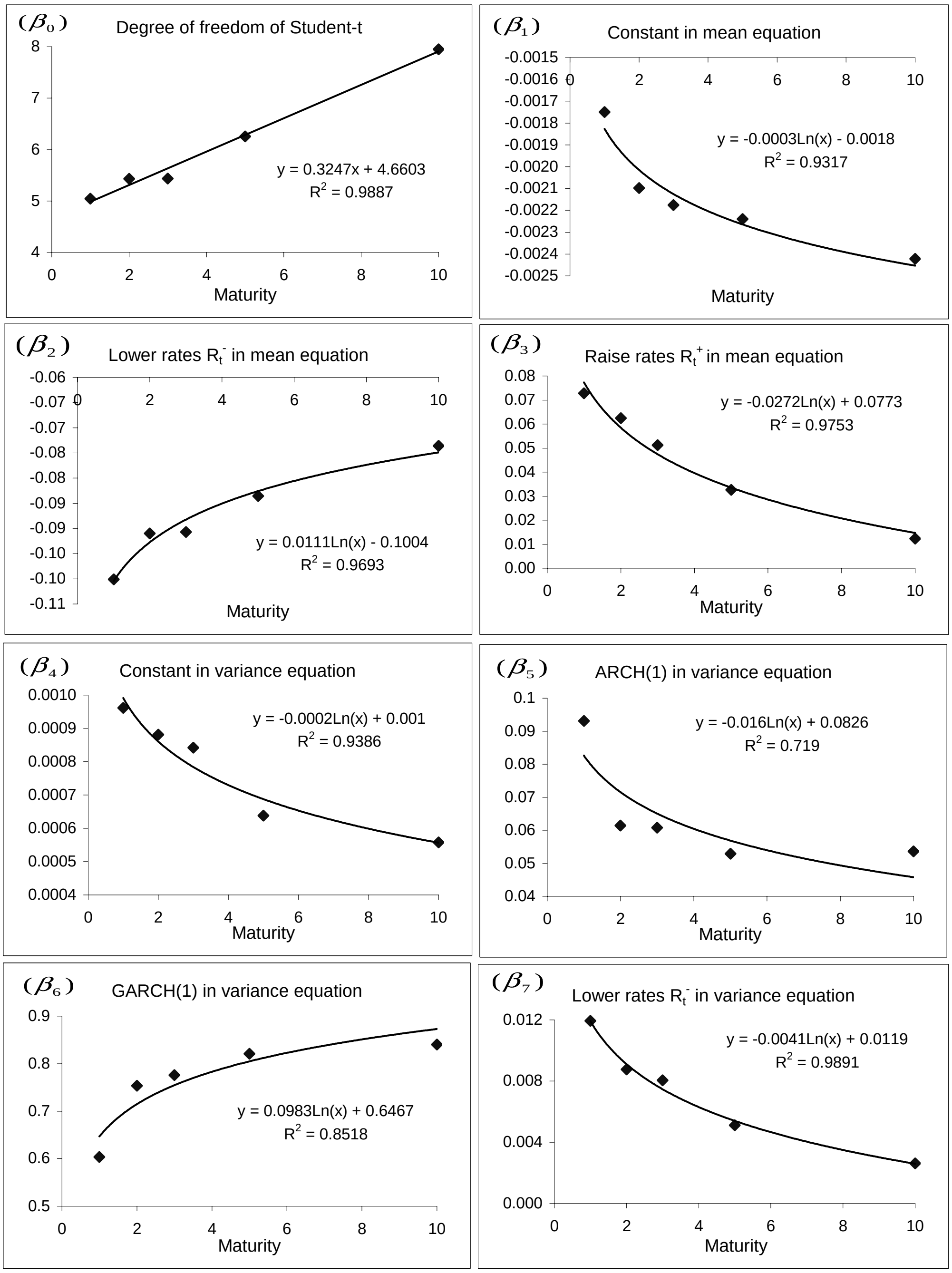


Figure 5.4: Estimated parameters of GARCH(1,1)-t models as a function of time to maturity. Solid diamonds show values for individual estimates. The solid lines are obtained by a least squares fit to log maturity.

Chapter 6

Multivariate GARCH Modelling of Yield Increments

Chapter 5 showed that the yield increment series for middle-to-long-term bonds are conditionally heteroskedastic, and that univariate GARCH(1,1)-t models could capture most of the empirical properties of these series at each maturity level. Moreover, the estimated coefficients of the GARCH(1,1)-t models, for each individual series at a fixed maturity level, conform to simple functional patterns depending on the maturity levels. In particular, the degrees of freedom for the t-distributions varied with maturity level. However, the modelling in Chapter 5 ignores possible correlation between the squared residuals of the series for different maturities. It is very likely that, when treated as a multiple time series, the collection of squared yield increments at all observed maturities will be cross correlated.

In this chapter, we extend the univariate GARCH(1,1)-t models of Chapter 5 to a multivariate GARCH(1,1)-$\boldsymbol{t}$ model. The current literature and statistical software for fitting the multivariate GARCH models cover the multivariate Normal distribution and the multivariate t-distribution with the same degrees of freedom for each component series. However, as Chapter 5 has demonstrated, this assumption is not appropriate for analysis of the term structure of Australian Treasury yield increments. The yield increments at different maturity levels require use of the t-distribution with different degrees of freedom. In this chapter, we will develop a multivariate GARCH(1,1) model for the term structure of yield increments in which the marginals have t-distributions with different degrees of freedom. This extension of existing GARCH models is based on the concepts of copulas of elliptically contoured distributions and the meta-elliptical t-distribution which have recently been developed for applications to financial studies of complex multivariate system, see Kotz, S. and Seeger, J. P. (1991) and Fang, Fang and Kotz (2002). We modify the definition of the multivariate asymmetric t-distribution for our GARCH modelling purpose. The Matrix-Diagonal GARCH(1,1) model with multivariate asymmetric t-distribution allows for different degrees of freedom in each margin of t-distribution. We refer to this new model as the *Matrix-Diagonal GARCH(1,1)-AMt model*. The effectiveness of the Matrix-Diagonal GARCH(1,1)-AMt is demonstrated by using standard statistical diagnostics on the standardised residuals. The stylised facts about volatility such as volatility mean reversion and volatility persistence are discussed.

Section 1 presents the definition of multivariate asymmetric t-distribution. Section 2 presents the Matrix-Diagonal GARCH(1,1)-AMt model of yield increments for middle-to-long-term Australian bonds. Section 3 provides the estimation. Section 4 provides the estimated results and diagnostics of the model. Section 5 presents a summary.

6.1 The Multivariate Asymmetric t-distribution

Fang, Fang and Kotz (2002) presented the standard multivariate asymmetric t-distribution as an application of meta-elliptical distribution, with zero mean and a dispersion matrix specified as a correlation matrix without loss of generality. The fact that the transformation functions are the distribution functions of Student's t-distribution implies that both the original random variables and the transformed random variables have the same unit dispersions. This section will present the definition of the multivariate asymmetric t-distribution that generalises the multivariate asymmetric t-distribution given by Fang, Fang and Kotz (2002). For our purpose of modelling the volatility of yield curves, we define the multivariate asymmetric t-distribution with general mean and general covariance matrix, and moreover, the transformations ensure that the constructed random variables have the same variances as original random variables .

The multivariate t-distribution

The d-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)'$ is said to have a multivariate t-distribution with m degree of freedom, mean vector $\boldsymbol{\mu}$ and positive-definite dispersion S, denoted $\boldsymbol{X} \sim \boldsymbol{t}_d(m, \boldsymbol{\mu}, S)$, if its density is given by

$$f(\boldsymbol{x}) = \frac{\Gamma(\frac{m+d}{2})}{\Gamma(\frac{m}{2})\sqrt{(\pi m)^d |S|}}\left(1 + \frac{(\boldsymbol{x}-\boldsymbol{\mu})' S^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}{m}\right)^{-\frac{m+d}{2}} \quad , \tag{6.1}$$

where $m > 2$. Note that in this standard parameterisation $cov(\boldsymbol{X}) = \frac{m}{m-2}S$.

This definition reduces to univariate t-distribution when $d = 1$. The univariate random variable X with a t-distribution with m degree of freedom, mean value μ and dispersion value s is denoted $X \sim t_m(\mu, s)$ where $var(X) = \frac{m}{m-2}s$. Specifically, $X \sim t_m(0, 1)$ is a Student's t-distribution with degree of freedom m.

The multivariate asymmetric t-distribution

Based on the definition of a meta-elliptical distribution (see Definition 1.2 in Fang, Fang and Kotz 2002), we define the multivariate asymmetric t-distribution with general mean vector and covariance matrix as follows, where the transformations ensure that each component of the constructed random vector $\boldsymbol{Z}$ has the same variance as that of the original random vector $\boldsymbol{X}$.

Let $\boldsymbol{X} = (X_1, \ldots, X_d)'$ be a d-dimensional random vector with each component X_i having t-distribution with degree of freedom m_i, mean μ_i and variance h_{ii}, ie. $X_i \sim t_{m_i}(\mu_i, h_{ii}), m_i > 2$. The d-dimensional random vector $\boldsymbol{X}$ is said to have a multivariate asymmetric t-distribution with $(m; m_1, \ldots, m_d)$ degrees of freedom, mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)'$ and covariance matrix $H = (h_{ij})_{d\times d}$, denoted $\boldsymbol{X} \sim AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}, S)$, $S = \frac{m-2}{m}H$, if its density is given by

$$f(x_1, \ldots, x_d) = \tilde{\boldsymbol{q}}_m(Q_{m,1}^{-1}(Q_{m_1}(x_1)), \ldots, Q_{m,d}^{-1}(Q_{m_d}(x_d))) \cdot \prod_{i=1}^{d} \frac{q_{m_i}(x_i)}{q_{m,i}(Q_{m,i}^{-1}(Q_{m_i}(x_i)))} \quad , \tag{6.2}$$

where $m > 2$, $\tilde{\boldsymbol{q}}_m$ is the d-variate density function of multivariate t-distribution $\mathbf{t}_d(m, \boldsymbol{\mu}, S)$, $q_{m,i}$, q_{m_i} and $Q_{m,i}$, Q_{m_i} are density functions and distribution functions of $t_m(\mu_i, \frac{m-2}{m}h_{ii})$ and $t_{m_i}(\mu_i, \frac{m_i-2}{m_i}h_{ii})$, respectively, $i = 1, ..., d$.

Note that the marginals of the multivariate asymmetric t-distribution $\boldsymbol{X}$ have t-distributions with different degrees of freedom $m_1, \ldots, m_d$. The multivariate asymmetric t-distribution $AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}, S)$ reduces to a multivariate t-distribution $t_d(m, \boldsymbol{\mu}, S)$ when $m = m_i, i = 1, \ldots, d$.

If $\boldsymbol{X} \sim AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}, S)$, let

$$\boldsymbol{Z} = (Z_1, \ldots, Z_d)' = (Q_{m,1}^{-1}(Q_{m_1}(X_1)), \ldots, Q_{m,d}^{-1}(Q_{m_d}(X_d)))' \quad . \tag{6.3}$$

Then it can be noted that

- S is the dispersion matrix of $\boldsymbol{Z}$, the transformed vector of $\boldsymbol{X}$ given by (6.3).
- $\boldsymbol{Z}$ has a multivariate t-distribution with m degree of freedom, mean $\boldsymbol{\mu}$ and dispersion S, ie. $\boldsymbol{Z} \sim \boldsymbol{t}_d(m, \boldsymbol{\mu}, S)$, $S = \frac{m-2}{m}H$ and $H = cov(\boldsymbol{Z})$.
- X_i and Z_i have the same means μ_i and same variances h_{ii}, $i = 1, \ldots, d$, and degrees of freedom m_i and m respectively, ie. $X_i \sim t_{m_i}(\mu_i, \frac{m_i-2}{m_i}h_{ii})$ and $Z_i \sim t_m(\mu_i, \frac{m-2}{m}h_{ii})$, $i = 1, ..., d$. One of the main features of our construction of the multivariate asymmetric t-distribution is that the original random variable X_i and transformed variable Z_i have the same variances for each i, $i = 1, \ldots, d$. As a result, the GARCH process describing the temporal evaluation of the conditional variance of Z_i represents that of the original yield increment series X_i, $i = 1, \ldots, d$. (Note: The X_i and Z_i have the same variances as discussed above, however, the dispersions of X_i and Z_i are different due to the different degrees of freedom of t-distributions, m_i and m respectively, of X_i and Z_i, $i = 1, ..., d$. It is a feature different from the multivariate asymmetric t-distribution constructed by Fang, Fang and Kotz 2002 that allows the same dispersions but different variances of X_i and Z_i, $i = 1, \ldots, d$.)

6.2 Matrix-Diagonal GARCH(1,1)-AMt Model

In Appendix A.1, we review the most popular multivariate GARCH models. They are the diagonal VEC model (Bollerslev, Engle and Wooldridge, 1988), the Matrix-Diagonal model (Ding 1994, Bollerslev, Engle and Nelson 1994), the BEKK model (Baba, Engle, Kraft and Kroner 1991, Engle and Kroner 1995), the Constant Conditional Correlation (CCC) model (Bollerslev 1990) and the Dynamical Conditional Correlation (DCC) model (Engle and Sheppard 2001). Any one of these models could be generalised to allow multivariate asymmetric t-distributions. For definiteness and to illustrate how this can be done we selected the Matrix-Diagonal model to analyse the term structure of Australian yield increments. The Matrix-Diagonal model allows the volatility process of each series to follow a univariate GARCH process, while the covariance process follows a GARCH model in terms of the product of two errors. The interpretation of the GARCH coefficient matrices is clear and simple. Also, the Matrix-Diagonal model guarantees that the time varying covariance matrixes H_t are positive-definite over all time t.

It is assumed that the term structure of Australian yield increments follows a multivariate asymmetric t-distribution with given marginals in different degrees of freedom of t-distributions. In this section we develop a Matrix-Diagonal GARCH(1,1) model of d-dimensional Australian yield increment series $\boldsymbol{y}_t$, with a multivariate asymmetric t-distribution, and we refer to it as the Matrix-Diagonal GARCH(1,1)-AMt model.

In view of the results presented in Chapter 5 for each individual maturity yield increment series, based on the multivariate Matrix-Diagonal GARCH-$\boldsymbol{t}$ model with exogenous variables (Zivot and Wang 2003), we take as the starting point for the multivariate extension the general Matrix-Diagonal GARCH(1,1)-AMt model specified as

$$\begin{aligned} &\boldsymbol{y}_t|\mathcal{F}_{t-1} \sim AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}_t, S_t), \quad S_t = \tfrac{m-2}{m} H_t, \\ &\boldsymbol{\mu}_t = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \cdot R_t^- + \boldsymbol{\beta}_3 \cdot R_t^+, \\ &H_t = A_0 A_0' + (AA') \otimes (\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + (BB') \otimes H_{t-1} + \boldsymbol{c}\boldsymbol{c}' \cdot R_t^-, \end{aligned} \tag{6.4}$$

where $\boldsymbol{\epsilon}_t = \boldsymbol{z}_t - \boldsymbol{\mu}_t$ and $\boldsymbol{z}_t = (Q_{m,1,t}^{-1}(Q_{m_1,t}(y_{1t})), \ldots, Q_{m,d,t}^{-1}(Q_{m_d,t}(y_{dt})))'$, $Q_{m,i,t}$ and $Q_{m_i,t}$ are the cumulative distribution functions of $t_m(\mu_{it}, \frac{m-2}{m} h_{ii,t})$ and $t_{m_i}(\mu_{it}, \frac{m_i-2}{m_i} h_{ii,t})$, $i = 1, ..., d$; $\boldsymbol{\epsilon}_t|\mathcal{F}_{t-1} \sim \boldsymbol{t}_d(m, \boldsymbol{0}, S_t)$; And R^+ and R^- are indicator variables of raising and lowering the target cash rate respectively (See Section 4.4) and the symbol $\otimes$ denotes the Hadamard product, i.e. element-by-element multiplication. In (6.4), $m_1, m_1, \ldots, m_d$ and m are scalar parameters; $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$ and $\boldsymbol{c}$ are d-dimensional vectors; A_0, A and B are $d \times d$ lower triangular matrices.

We simplified the above model using the following steps:

Step 1. Assuming that the degrees of freedom $m_i \equiv m$ (the symmetric t-distribution), the Matrix-Diagonal GARCH(1,1)-AMt model (6.4) reduces to a simple multivariate Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model. We used S+Finmetrics to estimate the parameters (scalars, vectors and matrices).

Step 2. The model in sStep 1 was simplified using the following principles:

- all non-significant coefficients removed,
- structural parameters retained.

In Step 1, we estimated the multivariate Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model using function *mgarch* in S+FinMetrics. Starting from the standard multivariate Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.7) in Appendix A.2 where parameter matrices A_0, A

and B in the covariance equation are square lower triangular $d \times d$ matrices. The log likelihood value of the estimated model is 19644. It is observed that all elements except column 1 of B are not significant. By reducing the parameter matrix B to a 1-dimensional vector $\boldsymbol{b}$, the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.8) showed a marked improvement that with a larger log likelihood value (19748) and with 10 fewer parameters. Also the estimated coefficients still posses a functional pattern of the maturity θ consistent with that observed in Chapter 5 (β_6 and β_7) and also each component is significant. However, if we further reduce the $d \times d$ matrix A to a 1-dimensional vector, there is no clear pattern of the coefficients along the maturity θ that was observed in Chapter 5. The matrix A_0 in S+Finmetrics must be a $d \times d$ matrix so that the H_t are positive-definite matrices. The estimation results of the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.8) are displayed in Appendix A.2.

According to the results of simplifying the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model, the Matrix-Diagonal GARCH(1,1)-AMt model of d-dimensional Australian yield increment series $\boldsymbol{y}_t$ (6.4) is simplified as

$$\begin{aligned}
&\boldsymbol{y}_t|\mathcal{F}_{t-1} \sim AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}_t, S_t), \quad S_t = \tfrac{m-2}{m} H_t, \\
&\boldsymbol{\mu}_t = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \cdot R_t^- + \boldsymbol{\beta}_3 \cdot R_t^+, \\
&H_t = A_0 A_0' + (AA') \otimes (\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + (\boldsymbol{b}\boldsymbol{b}') \otimes H_{t-1} + \boldsymbol{c}\boldsymbol{c}' \cdot R_t^-,
\end{aligned} \tag{6.5}$$

where $\boldsymbol{\epsilon}_t$, $\boldsymbol{z}_t$, R^+ and R_- are the same as those described in model (6.4). In Matrix-Diagonal GARCH(1,1)-AMt model (6.5), $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are d-dimensional parameter vectors; A_0 and A are lower triangular $d \times d$ parameter matrices.

In the rest of the Chapter, we use the Matrix-Diagonal GARCH(1,1)-AMt model (6.5).

6.3 Estimations

The preliminary analysis used to arrive at the model (6.5) used *mgarch* function in S+FinMetrics. However, there is no currently available statistical software for estimating the GARCH model with a multivariate asymmetric t-distribution. This section describes an estimation approach for the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) which combines use of the Matlab function $fmincon$, to find the minimum of a constrained nonlinear multivariate function, and S+FinMetrics function *mgarch* for estimation of the Matrix-Diagonal GARCH model.

Let $\boldsymbol{\nu} = \{m_1, \ldots, m_d\}$, $\boldsymbol{\beta} = \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{b}, \boldsymbol{c}, A_0, A\}$ and assume that m is fixed. Conditional on initial conditions and values of R^- and R^+, the log likelihood function of the Matrix-Diagonal GARCH(1,1)-AMt Model (6.5) is

$$L(\boldsymbol{\nu}, \boldsymbol{\beta}) = L_1(\boldsymbol{\nu}, \boldsymbol{\beta}) + L_2(\boldsymbol{\nu}, \boldsymbol{\beta}) \quad , \tag{6.6}$$

where

$$\begin{aligned} L_1(\boldsymbol{\nu}, \boldsymbol{\beta}) &= \textstyle\sum_{t=1}^{N} \log \tilde{\mathbf{q}}_{m,t}(z_{t1}, \ldots, z_{td}; m, \boldsymbol{\nu}, \boldsymbol{\beta}, R_t^-, R_t^+) \quad , \\ L_2(\boldsymbol{\nu}, \boldsymbol{\beta}) &= \textstyle\sum_{t=1}^{N} \log q_{m_i,t}(y_{ti}; m_i, \boldsymbol{\beta}, R_t^-, R_t^+) - \sum_{t=1}^{N} \log q_{m,i,t}(z_{ti}; m, m_i, \boldsymbol{\beta}, R_t^-, R_t^+) \quad , \end{aligned} \tag{6.7}$$

and, at time t,

$$\begin{aligned} \tilde{\mathbf{q}}_m(z_1, \ldots, z_d) &= \frac{\Gamma(\frac{m+d}{2})}{\Gamma(\frac{m}{2})\sqrt{(\pi m)^d |S|}}(1 + \frac{(\boldsymbol{z}-\boldsymbol{\mu})' S^{-1} (\boldsymbol{z}-\boldsymbol{\mu})}{m})^{-\frac{m+d}{2}} \quad , \\ q_{m,i}(z_i) &= \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})\sqrt{\pi(m-2)h_{ii}}}(1 + \frac{(z_i-\mu_i)^2}{(m-2)h_{ii}})^{-\frac{m+1}{2}} \quad , \\ q_{m_i}(y_i) &= \frac{\Gamma(\frac{m_i+1}{2})}{\Gamma(\frac{m_i}{2})\sqrt{\pi(m_i-2)h_{ii}}}(1 + \frac{(y_i-\mu_i)^2}{(m_i-2)h_{ii}})^{-\frac{m_i+1}{2}} \quad . \end{aligned} \tag{6.8}$$

It is implicit in Fang, Fang and Kotz (2002) that the parameter m in asymmetric multivariate $\boldsymbol{t}$-distribution $AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}, S)$ can be estimated from the data. Our attempts to estimate m as well as $m_1, \ldots, m_d$ resulted in the estimate values did not converge and the log likelihood values did not converge. This may suggest that m is not identifiable when $m_1, \ldots, m_d$ are also free parameters. To overcome this difficulty, we fixed m at 6.3, the estimate of the degrees of freedom of $\boldsymbol{t}$-distribution of the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.8) from S+FinMetrics (SE=0.429236). The estimate for this is that the transformed random vector $\boldsymbol{Z}$ in (6.3) has degrees of freedom typical of the original random vector $\boldsymbol{X}$; m is set as the 'average' of $m_1, \ldots, m_d$.

Altogether there are 60 parameters in the Matrix-Diagonal GARCH(1,1)-AMt model (6.5). It is hard to estimate the 60 parameters by minimizing the likelihood function over all parameters simultaneously. This was attempted by minimizing the log-likelihood function (6.6) of Matrix-Diagonal GARCH(1,1)-AMt Model (6.5) using $fmincon$ in Matlab, but the computer algorithms failed to converge. Because of this, the estimation of Matrix-Diagonal GARCH(1,1)-AMt model (6.5) implemented here uses an iteration method. The steps are:

Step A: For fixed values of the degrees of freedom of marginals $\boldsymbol{\nu}$, estimate GARCH coefficients $\boldsymbol{\beta}$ in S+FinMetrics by optimising $L_1(\boldsymbol{\nu}, \boldsymbol{\beta})$ based on Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ of $\boldsymbol{Z}$. Call the result $(\boldsymbol{\nu}, \hat{\boldsymbol{\beta}}(\boldsymbol{\nu}))$.

Step B: Fixing $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}(\boldsymbol{\nu})$, optimise the combined likelihood $L(\boldsymbol{\nu}, \boldsymbol{\beta}) = L_1(\boldsymbol{\nu}, \boldsymbol{\beta}) + L_2(\boldsymbol{\nu}, \boldsymbol{\beta})$ over $\boldsymbol{\nu}$. Optimisation uses Matlab. Result is $\hat{\boldsymbol{\nu}}$.

Step C: Repeat Step 1 and Step 2 until convergence. Call the converged result $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\beta}})$.

Step A estimates GARCH coefficients $\boldsymbol{\beta}$ in S+FinMetrics based on Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model of $\boldsymbol{Z}$ with common degree of freedom m, given a fixed $\boldsymbol{\nu} = \{m_1, \ldots, m_d\}$. However, $\boldsymbol{Z}$ is a transformed vector from vector $\boldsymbol{y}$ depending on the parameters $\boldsymbol{\beta}$. The estimation of Step A used an sub-iteration method in S+FinMetrics, that is,

Step A1: Get initial values of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$, and variances matrix $(h_{ii,t}), i = 1, \ldots, d$ from univariate GARCH(1,1)-$\boldsymbol{t}$ models of each yield increment series with given degrees of freedom $m_1, \ldots, m_d$, respectively.

Step A2: Calculate the mean vector $\boldsymbol{\mu}_t = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \cdot R_t^- + \boldsymbol{\beta}_3 \cdot R_t^+$. Calculate $\boldsymbol{z_t} = (Q_{m,1,t}^{-1}(Q_{m_1,t}(y_{1t})), \ldots, Q_{m,d,t}^{-1}(Q_{m_d,t}(y_{dt})))'$ where $Q_{m,i,t}$ and $Q_{m_i,t}$ are the cumulative distribution functions of $t_m(\mu_{it}, \frac{m-2}{m} h_{ii,t})$ and $t_{m_i}(\mu_{it}, \frac{m_i-2}{m_i} h_{ii,t})$, $i = 1, ..., d$;

Step A3: Estimate $\boldsymbol{\beta}$ and H from Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model of $\boldsymbol{Z}$ using function *mgarch* in S+FinMetrics. Output the log-likelihood value of the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model of $\boldsymbol{Z}$.

Step A4: Iterate Steps A2 and A3 until the log-likelihood value of the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model of $\boldsymbol{Z}$ converge and estimated values $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$,$\boldsymbol{\beta}_3$, A_0, A, $\boldsymbol{b}$, $\boldsymbol{c}$ converge.

Note that the converged result $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\beta}})$ from the above iteration method may not be the MLE of $L(\boldsymbol{\nu}, \boldsymbol{\beta})$ in (6.6), because the estimate GARCH coefficients $\boldsymbol{\beta}$, in Step A, in S+FinMetrics, is optimizing $L_1(\boldsymbol{\nu}, \boldsymbol{\beta})$ instead of optimizing $L(\boldsymbol{\nu}, \boldsymbol{\beta})$. In this sample, $L_1(\boldsymbol{\nu}, \boldsymbol{\beta}) = 19738$ (99.924%) and $L_2(\boldsymbol{\nu}, \boldsymbol{\beta}) = 15$ (0.076%), and hence $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\beta}})$ from the above iteration method should be approximately the MLE maximazing $L(\boldsymbol{\nu}, \boldsymbol{\beta})$ in (6.6). $L_2(\boldsymbol{\nu}, \boldsymbol{\beta})$ presented in (6.6) is a very small value. The ratio $\frac{q_{m_i,t}(y_{ti}; m_i, \boldsymbol{\beta}, R_t^-, R_t^+)}{q_{m,i,t}(z_{ti}; m, m_i, \boldsymbol{\beta}, R_t^-, R_t^+)}$ is equal to 1, greater than 1 and less than 1 when $m_i = m$, $m_i > m$ and $m_i < m$ respectively, and is very close to 1 when m_i is close to m. The product of the 5 ratios $\frac{q_{m_i,t}(y_{ti};m_i,\boldsymbol{\beta},R_t^-,R_t^+)}{q_{m,i,t}(z_{ti};m,m_i,\boldsymbol{\beta},R_t^-,R_t^+)}$ $(i = 1,\ldots,5)$ is close to 1 when m is the 'average' of $m_1,\ldots,$ and m_5, where $m_1,\ldots,$ and m_5 are not ver different. So $L_2(\boldsymbol{\nu},\boldsymbol{\beta})$ is close to zero, which makes $L(\boldsymbol{\nu},\boldsymbol{\beta})$ nearly equal to $L_1(\boldsymbol{\nu},\boldsymbol{\beta})$. Hence $(\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\beta}})$ from the above iteration method should be an approximately efficient estimate of Matrix-Diagonal GARCH(1,1)-AMt model (6.5). Comte and Lieberman (2003) established the asymptotic normality for the general multivariate GARCH model under certain regularly conditions. In particular, $\boldsymbol{\epsilon}_t$ would be required to have finite 8^{th} moment. We have observed that a t-distribution with between 5 to 8 degrees of freedom provides a good model. These do not satisfy the condition in Comte and Lieberman (2003). As the result, the asymptotic normality of the estimation is not guaranteed.

6.4 Results and Diagnostics of the Matrix-Diagonal GARCH-AMt Model

This section presents the results of estimation of the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) using 5-dimensional Australian Treasury yield increments with maturity in 1, 2, 3, 5 and 10 years for the period from 1996 to 2001. A diagnostic assessment of this model is provided by testing for no autocorrelation in the standardised residuals and the squared standardised residuals, and testing no ARCH effect in the standardised residuals. The stylised facts about volatility such as volatility asymmetry, volatility mean reversion and volatility persistence are discussed.

The results of the estimation of the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) are $\hat{m}_1 = 5.1782$ (SE=0.4666), $\hat{m}_2 = 5.9367$ (SE=0.6565), $\hat{m}_3 = 5.9921$ (SE=0.6464), $\hat{m}_4 = 6.8607$ (SE=0.8955) and $\hat{m}_5 = 8.4470$ (SE=1.3523). Note that the estimates of $(m_1,\ldots,m_d)$ and their standard errors are obtained at the convergence of Step B using Matlab and are therefore not the MLE SEs. The standard errors derived in this way may underestimate the SE's from MLE but at this stage we can't derive the MLE SE's. The estimated GARCH coefficients are displayed in Table 6.1.

The Ljung-Box tests for squared standardised residuals show that, for yield increment series with maturity in 1, 2, 3, 5 and 10 years, the squared standardised residuals are not autocorrelated, and the Lagrange multiplier tests confirm there are no ARCH effects

Table 6.1: Outputs of the Matrix-Diagonal GARCH(1,1)-AMt model

	Value	Std.Error	t value	Pr(> $\|t\|$)
$\boldsymbol{\beta}_1(1)$	-0.00224	0.0012165	-1.8414	0.0658
$\boldsymbol{\beta}_1(2)$	-0.00273	0.001483	-1.8412	0.0658
$\boldsymbol{\beta}_1(3)$	-0.002731	0.0015414	-1.7719	0.0766
$\boldsymbol{\beta}_1(4)$	-0.002819	0.0015517	-1.817	0.0694
$\boldsymbol{\beta}_1(5)$	-0.002939	0.0016127	-1.8225	0.0686
$\boldsymbol{\beta}_2(1)$	-0.077303	0.0274658	-2.8145	0.0049
$\boldsymbol{\beta}_2(2)$	-0.061312	0.0311766	-1.9666	0.0494
$\boldsymbol{\beta}_2(3)$	-0.057955	0.03029	-1.9133	0.0559
$\boldsymbol{\beta}_2(4)$	-0.056295	0.0313548	-1.7954	0.0728
$\boldsymbol{\beta}_2(5)$	-0.051871	0.0251595	-2.0617	0.0394
$\boldsymbol{\beta}_3(1)$	0.079483	0.0515967	1.5405	0.1237
$\boldsymbol{\beta}_3(2)$	0.070406	0.0626582	1.1236	0.2613
$\boldsymbol{\beta}_3(3)$	0.059669	0.0633119	0.9425	0.3461
$\boldsymbol{\beta}_3(4)$	0.043983	0.0688903	0.6385	0.5233
$\boldsymbol{\beta}_3(5)$	0.023504	0.0901604	0.2607	0.7944
$A_0(1,1)$	0.02108	0.0013057	16.1444	0.0000
$A_0(2,1)$	0.022855	0.0014414	15.8562	0.0000
$A_0(3,1)$	0.02338	0.0014605	16.0075	0.0000
$A_0(4,1)$	0.023035	0.0015494	14.8676	0.0000
$A_0(5,1)$	0.021327	0.0019111	11.1596	0.0000
$A_0(2,2)$	0.004983	0.0006957	7.1623	0.0000
$A_0(3,2)$	0.004853	0.0008221	5.9025	0.0000
$A_0(4,2)$	0.005366	0.0010004	5.3636	0.0000
$A_0(5,2)$	0.005531	0.0016636	3.324	0.0009
$A_0(3,3)$	0.002557	0.0002995	8.54	0.0000
$A_0(4,3)$	0.00312	0.0004451	7.01	0.0000
$A_0(5,3)$	0.004103	0.0011408	3.597	0.0003
$A_0(4,4)$	0.003598	0.0003603	9.984	0.0000
$A_0(5,4)$	0.006771	0.0010337	6.551	0.0000
$A_0(5,5)$	0.006611	0.0008771	7.537	0.0000
A(1,1)	0.259417	0.0167842	15.456	0.0000
A(2,1)	0.238178	0.015222	15.647	0.0000
A(3,1)	0.233945	0.0152581	15.333	0.0000
A(4,1)	0.221518	0.0151084	14.662	0.0000
A(5,1)	0.192439	0.0179142	10.742	0.0000
A(2,2)	0.052654	0.0078784	6.683	0.0000
A(3,2)	0.057371	0.009564	5.999	0.0000
A(4,2)	0.049485	0.0119074	4.156	0.0000
A(5,2)	0.041191	0.0199011	2.07	0.0386
A(3,3)	0.030198	0.0042291	7.14	0.0000
A(4,3)	0.018743	0.0071562	2.619	0.0089
A(5,3)	0.023614	0.0188112	1.255	0.2096
A(4,4)	0.021275	0.0076403	2.785	0.0054
A(5,4)	0.037484	0.0242057	1.549	0.1217
A(5,5)	0.04297	0.0103649	4.146	0.0000
$\boldsymbol{b}(1)$	0.884575	0.0117526	75.267	0.0000
$\boldsymbol{b}(2)$	0.90209	0.01022	88.244	0.0000
$\boldsymbol{b}(3)$	0.90363	0.01005	89.908	0.0000
$\boldsymbol{b}(4)$	0.90728	0.01096	82.772	0.0000
$\boldsymbol{b}(5)$	0.91641	0.01495	61.294	0.0000
$\boldsymbol{c}(1)$	0.06692	0.013	5.146	0.0000
$\boldsymbol{c}(2)$	0.06336	0.0157	4.037	0.0001
$\boldsymbol{c}(3)$	0.06055	0.01634	3.706	0.0002
$\boldsymbol{c}(4)$	0.04841	0.01621	2.987	0.0029
$\boldsymbol{c}(5)$	0.02584	0.01763	1.465	0.1430

left with the standardised residuals. The estimators of the coefficients of the model are displayed in Table 6.1. Most of the coefficients of the model (6.5) are significantly different from zero with small standard errors.

The coefficient matrices of the multivariate Matrix-Diagonal GARCH(1,1)-AMt model (6.5), are given below.

The constant coefficient matrix in covariance equation is close to zero, that is

$$A_0A_0' = \begin{bmatrix} 0.00044 & 0.00048 & 0.00049 & 0.00049 & 0.00045 \\ 0.00048 & 0.00055 & 0.00056 & 0.00055 & 0.00051 \\ 0.00049 & 0.00056 & 0.00058 & 0.00057 & 0.00054 \\ 0.00049 & 0.00055 & 0.00057 & 0.00058 & 0.00056 \\ 0.00045 & 0.00051 & 0.00054 & 0.00056 & 0.00059 \end{bmatrix}.$$

The ARCH(1) coefficient matrix is

$$AA' = \begin{bmatrix} 0.06730 & 0.06179 & 0.06069 & 0.05747 & 0.04992 \\ 0.06179 & 0.05950 & 0.05874 & 0.05537 & 0.04800 \\ 0.06069 & 0.05874 & 0.05893 & 0.05523 & 0.04810 \\ 0.05747 & 0.05537 & 0.05523 & 0.05232 & 0.04591 \\ 0.04992 & 0.04800 & 0.04810 & 0.04591 & 0.04254 \end{bmatrix},$$

where all elements of AA' are very close to zero, implying that the variance process h_{it} and covariance process h_{ijt} are slightly affected by lagged residuals given that most of all elements of A are significantly different from zero.

The GARCH(1) coefficient matrix is

$$\boldsymbol{bb}' = \begin{bmatrix} 0.78247 & 0.79797 & 0.79933 & 0.80255 & 0.81063 \\ 0.79797 & 0.81377 & 0.81515 & 0.81845 & 0.82669 \\ 0.79933 & 0.81515 & 0.81654 & 0.81984 & 0.82809 \\ 0.80255 & 0.81845 & 0.81984 & 0.82315 & 0.83144 \\ 0.81063 & 0.82669 & 0.82809 & 0.83144 & 0.83981 \end{bmatrix},$$

where all elements, GARCH(1,1) coefficients, are significantly different from zero and close to one (all are more than 0.8), implying that the conditional covariances are persistent.

Also

$$AA' + \boldsymbol{bb}' = \begin{bmatrix} 0.84977 & 0.85976 & 0.86002 & 0.86002 & 0.86056 \\ 0.85976 & 0.87327 & 0.87390 & 0.87381 & 0.87469 \\ 0.86002 & 0.87390 & 0.87548 & 0.87507 & 0.87619 \\ 0.86002 & 0.87381 & 0.87507 & 0.87547 & 0.87735 \\ 0.86056 & 0.87469 & 0.87619 & 0.87735 & 0.88235 \end{bmatrix}. \tag{6.9}$$

The diagonal elements of $AA' + \boldsymbol{bb}'$ are less than 1, implying that the individual GARCH(1,1) process $\epsilon_{it}, i = 1, \ldots, d$ are wide-sense variance stationary by Theorem 5.1 (Bollerslev 1986) and the volatility mean reverts to its long run level (see Zivot and Wang 2003). We can also claim that d-dimensional process of $\boldsymbol{\epsilon}_t$ is stationary by the multivariate Matrix-Diagonal GARCH model by the theorem from Boussama (1998) and Theorem 1 of Comte and Lieberman (2003). That is,
let *vech* the vector-half operator that stacks the lower triangular portion of a matrix into a vector, and $\rho(A)$ the spectral radius of A, i.e., the largest modulus of eigenvalues of A.

Theorem 6.1 *$\boldsymbol{\epsilon}_t$ is a d-dimensional random vector, $E(\boldsymbol{\epsilon}_t|\mathcal{F}_{t-1}) = 0$ a.s. and $E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'|\mathcal{F}_{t-1}) = H_t$. For the model given by*

$$\epsilon_t = H^{1/2}\eta_t, \quad \eta_t \sim iid(0, I_d) \tag{6.10}$$

where I_d is $d \times d$ identity matrix, and

$$vech(H_t) = vech(C) + \sum_{s=1}^{q} \tilde{A} vech(\boldsymbol{\epsilon}_{t-s}\boldsymbol{\epsilon}_{t-s}') + \sum_{s=1}^{q} \tilde{B} vech(H_{t-s}), \tag{6.11}$$

assume that the $\boldsymbol{\epsilon}_t$ admit a density absolutely continuous with respect to the Lebesgue measure, positive in a neighbourhood of the origin. Assume moreover that

$$\rho(\sum_{s=1}^{q} \tilde{A} + \sum_{s=1}^{q} \tilde{B}) < 1,$$

and let Y be defined by

$$Y_t = (vech(H_{t+1})', vech(H_t)', \ldots, vech(H_{t-p+2})', \boldsymbol{\epsilon}_t', \boldsymbol{\epsilon}_{t-1}', \boldsymbol{\epsilon}_{t-q+1})'.$$

Then the recurrence relations (6.10) and (6.11) for Y have an almost unique strictly stationary causal solution which constitutes a positive Harris recurrent Markov chain which is geometrically ergodic and β-mixing.

The multivariate matrix diagonal GARCH model of Z in (6.5) can be rewritten as

$$vech(H_t) = vech(A_0 A_0') + \tilde{A} vech(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + \tilde{B} vech(H_{t-1}) + vech(\boldsymbol{cc}')R_t^-, \qquad (6.12)$$

where $\tilde{A} = diag(vech(AA'))$ and $\tilde{B} = diag(vech(\boldsymbol{bb}'))$. Note that $\rho(\tilde{A} + \tilde{B}) < 1$, because $\tilde{A} + \tilde{B} = diag(vech(AA' + \boldsymbol{bb}'))$ and the eigenvalues are the elements of the lower triangular matrix $AA' + \boldsymbol{bb}'$ which has all elements less than 1, see equation (6.9). By Theorem 6.1, the d-dimensional process of $\boldsymbol{\epsilon}_t$ is stationary by multivariate Matrix-Diagonal GARCH(1,1)-AMt model (6.5).

The expression of unconditional covariance of ϵ_{it} and ϵ_{jt} can be given as follows. Let $A_0 A_0' = (a_{ij})$, $AA' = (\alpha_{ij})$ and $\boldsymbol{bb}' = (\beta_{ij})$ from Matrix-Diagonal GARCH(1,1)-AMt model (6.5). For $i, j = 1, \ldots, d$, $E[\epsilon_{it}|\mathcal{F}_{t-1}] = 0$, and $E[\epsilon_{it}\epsilon_{it}|\mathcal{F}_{t-1}] = h_{ijt}$, and let $u_{ijt} = \epsilon_{it}\epsilon_{jt} - E(\epsilon_{it}\epsilon_{it}|\mathcal{F}_{t-1})$ be a white noise, the above equation of Matrix-Diagonal model can be

$$\epsilon_{it}\epsilon_{jt} = a_{ij} + (\alpha_{ij} + \beta_{ij})\epsilon_{i(t-1)}\epsilon_{j(t-1)} + u_{ijt} - \beta_{ij} u_{ij(t-1)},$$

so

$$E[\epsilon_{it}\epsilon_{jt}] = a_{ij} + (\alpha_{ij} + \beta_{ij})E[\epsilon_{i(t-1)}\epsilon_{j(t-1)}].$$

Assuming $\epsilon_{it}\epsilon_{jt}$ is stationary, then

$$E[\epsilon_{it}\epsilon_{jt}] = a_{ij} + (\alpha_{ij} + \beta_{ij})E[\epsilon_{it}\epsilon_{jt}].$$

Thus

$$cov(\epsilon_{it}, \epsilon_{jt}) = E[\epsilon_{it}\epsilon_{jt}] = a_{ij}/(1 - (\alpha_{ij} + \beta_{ij})).$$

Finally, the coefficient matrix of the indicator variable for lowering the cash rate in the covariance equation is

$$\boldsymbol{cc}' = \begin{bmatrix} 0.00448 & 0.00424 & 0.00405 & 0.00324 & 0.00173 \\ 0.00424 & 0.00401 & 0.00384 & 0.00307 & 0.00164 \\ 0.00405 & 0.00384 & 0.00367 & 0.00293 & 0.00156 \\ 0.00324 & 0.00307 & 0.00293 & 0.00234 & 0.00125 \\ 0.00173 & 0.00164 & 0.00156 & 0.00125 & 0.00067 \end{bmatrix}.$$

The estimated coefficients of the mean equation and coefficient matrices of covariance of the multivariate Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model are shown in Figure 6.1. The trends of coefficients of mean, panel (β_1), (β_2), and (β_3), are very similar to those of the univariate GARCH(1,1) models in Chapter 5 (Figure 5.4).

Panels (β_4), (β_5), and (β_6), and (β_7) of Figure 6.1 are rows of the coefficient matrices of $A_0A_0', AA', \boldsymbol{bb}'$, and $\boldsymbol{cc}'$ along the maturity level. Figure 6.2 shows contour plots of coefficient matrices of $A_0A_0', AA', \boldsymbol{bb}'$, and $\boldsymbol{cc}'$ on 2-dimensions of time to maturity. From these figures, patterns of the coefficient matrices of covariance are observed. When the length of time to maturity increases, the elements of AA' (ARCH(1)) decrease, the elements of $\boldsymbol{bb}'$ (GARCH(1)) slightly increase and the elements $\boldsymbol{cc}'$ (lowering rates in covariance) decrease.

Overall, the trends of the coefficients of mean and covariance of Matrix-Diagonal GARCH(1,1) are consistent with the patterns of univariate GARCH(1,1) models in Chapter 5. That is, both mean level and conditional variance of the increments of bond yield with longer maturity are less impacted by the changing of RBA cash rates and are more efficient in persistence of conditional covariances.

It needs to be emphasised that the GARCH processes of individual yield increments at different maturities obtained by the univariate GARCH(1,1)-t models (5.4) use $\boldsymbol{\epsilon}_t = \boldsymbol{y}_t - \boldsymbol{\mu}_t$ and these are modelled using t-distributions with different degrees of freedom. These are not exactly the same as the diagonal conditional variance GARCH processes in Matrix-Diagonal GARCH(1,1)-AMt model (6.5) which uses $\boldsymbol{\epsilon}_t = \boldsymbol{z}_t - \boldsymbol{\mu}_t$, in which the degrees of freedom for each component are the same.

An alternative specification of the Matrix-Diagonal GARCH(1,1)-AMt model would use $\boldsymbol{\epsilon}_t = \boldsymbol{y}_t - \boldsymbol{\mu}_t$ in equation (6.5). This alternative would provide equations for the conditional variance of each component series that have a similar interpretation to those former from univariate models (5.4). However, the disadvantage of this alternative (with $\boldsymbol{\epsilon}_t = \boldsymbol{y}_t - \boldsymbol{\mu}_t$) is that the S+FinMetrics *mgarch* function can't be used. Software for fitting this alternative would need to be developed and that is a topic for further research.

Figure 6.4 presents the comparison of standard deviations h_t from the univariate GARCH-t (dot lines) and Matrix-Diagonal GARCH-AMt (solid lines), maturities in 1 year, 5 years and 10 years. The trends of the two time series are consistent, and the univariate GARCH models provide larger standard deviations than the Matrix-Diagonal GARCH-AMt model.

6.5 Model Tests and Selection

6.5.1 Likelihood Ratio Tests for Degrees of Freedom Parameters

Note that the estimation of the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) displayed in Table 6.1 are very close to those estimated by the simple Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (6.5) displayed in Table A.2 in Appendix A.2, although the distributions are different. See Figure 6.3. The simple Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (6.5) is nested within the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) under the assumption of the same degrees of freedom of the t-distribution of marginals. The Matrix-Diagonal GARCH(1,1)-AMt model (6.5) gives a larger log-likelihood value of 19753 with four additional parameters, comparing to the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.8) that gives a log-likelihood value of 19748. The resulting likelihood ratio test statistic for testing the null hypothesis H_0: $m_1 = \cdots = m_d = m$ is equal to 10 which, using the $\chi^2_{(4)}$ distribution, has an approximate p-value of 0.04. Hence the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) is significantly different from the Matrix-Diagonal GARCH(1,1)-t (A.8) at 0.05 level.

6.5.2 Akaike Information Criterion (AIC) for Comparing the Univariate Models with Multivariate Models

This section is to compare the Matrix-Diagonal GARCH(1,1)-AMt model (6.5) with a set of univariate GARCH(1,1)-t models in Chapter5, marginals of both models are in t-distributions with different degrees of freedom. Because the two models are not nested, the Akaike Information Criterion (AIC) can be used as a criterion for selecting models. The AIC is: $AIC = -2L_q + 2q$ where L_q is the maximised log-likelihood and q is the number of parameters in the model. A smaller value of AIC indicates the preferred model. The Matrix-Diagonal GARCH(1,1)-AMt model which takes into account correlation between increments volatilities and has more parameters has an AIC value: AIC=-39385. This model is preferred to the set of univariate GARCH models which ignores the correlation between increments but uses less parameters with an AIC=-19517.

6.6 Summary

In this chapter, we defined the multivariate asymmetric t-distribution using meta-elliptical distribution concepts (Fang, Fang and Kotz 2002) which allows different marginals. The univariate GARCH(1,1)-t model of Australian Treasury bond yield increments developed in Chapter 5 is extended to a multivariate GARCH(1,1) model with multivariate asymmetric t-distribution. We demonstrated that the Matrix-Diagonal GARCH(1,1)-AMt model is appropriate for the term structure of yield increments of Australian Treasury bonds. The estimation of the Matrix-Diagonal GARCH(1,1)-AMt model has been implemented using a 2-stage method within which iterations are required. The estimated results show that the Matrix-Diagonal GARCH(1,1)-AMt model captures the properties of volatility mean reversion, volatility persistence and a stationary GARCH process. Also, the likelihood ratio test shows that the Matrix-Diagonal GARCH(1,1)-AMt (A.8) is a significant improvement to the simple Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.8). Based on Akaike information criterion (AIC), the Matrix-Diagonal GARCH(1,1) with multivariate asymmetric t-distribution is the preferred model to a set of univariate GARCH models.

The trends of the coefficients of mean and conditional variances GARCH process over the maturity levels from the Matrix-Diagonal GARCH(1,1)-AMt model are consistent with the results from individual univariate GARCH(1,1)-t models discussed in Chapter 5. We will further investigate the patterns of coefficients of the Matrix-Diagonal GARCH(1,1)-AMt in Chapter 7 with the aim of developing a covariance extension to the TS-GARCH model proposed in Chapter 5.

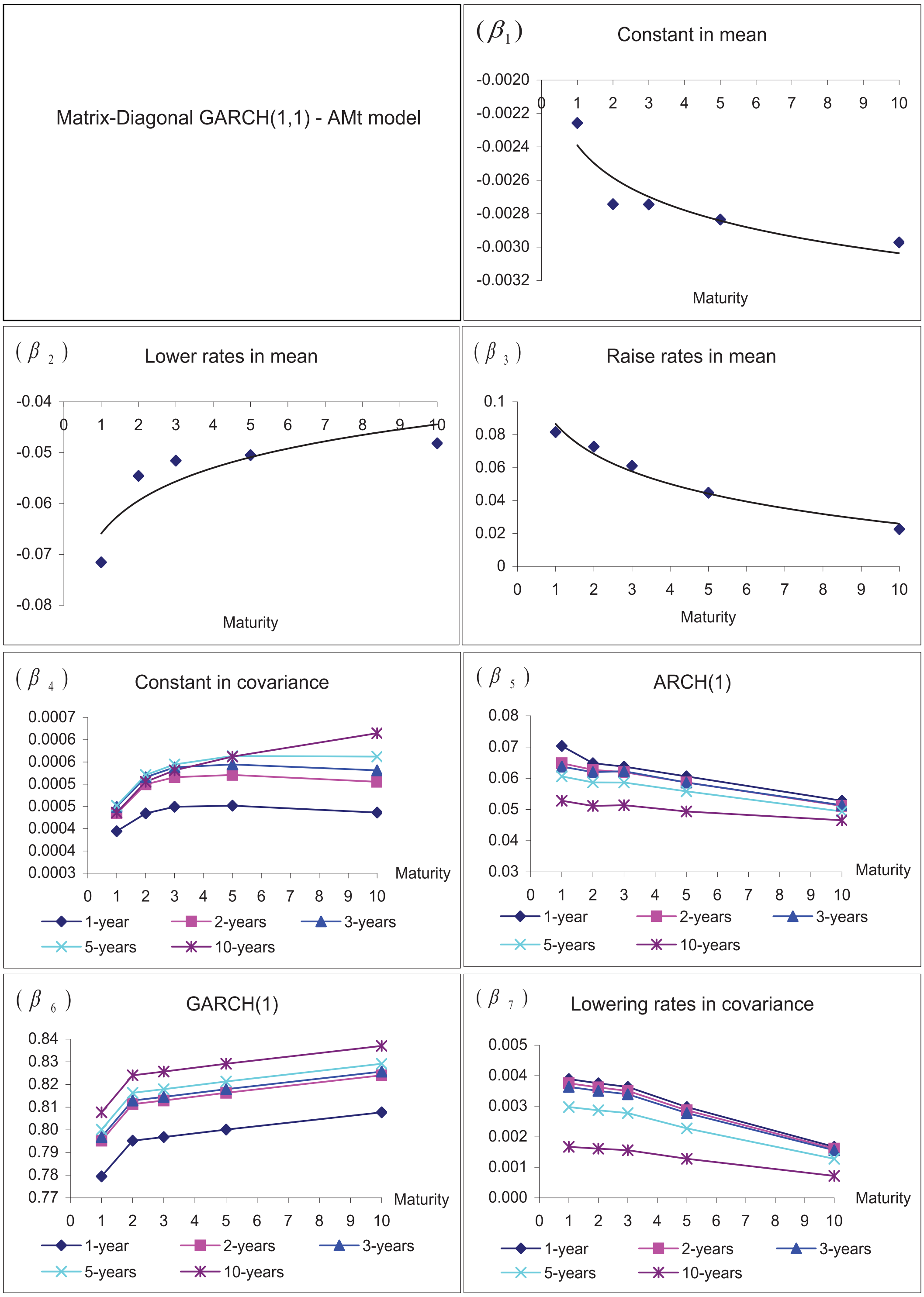


Figure 6.1: Coefficients of mean and variance: Matrix-Diagonal GARCH(1,1)-AMt

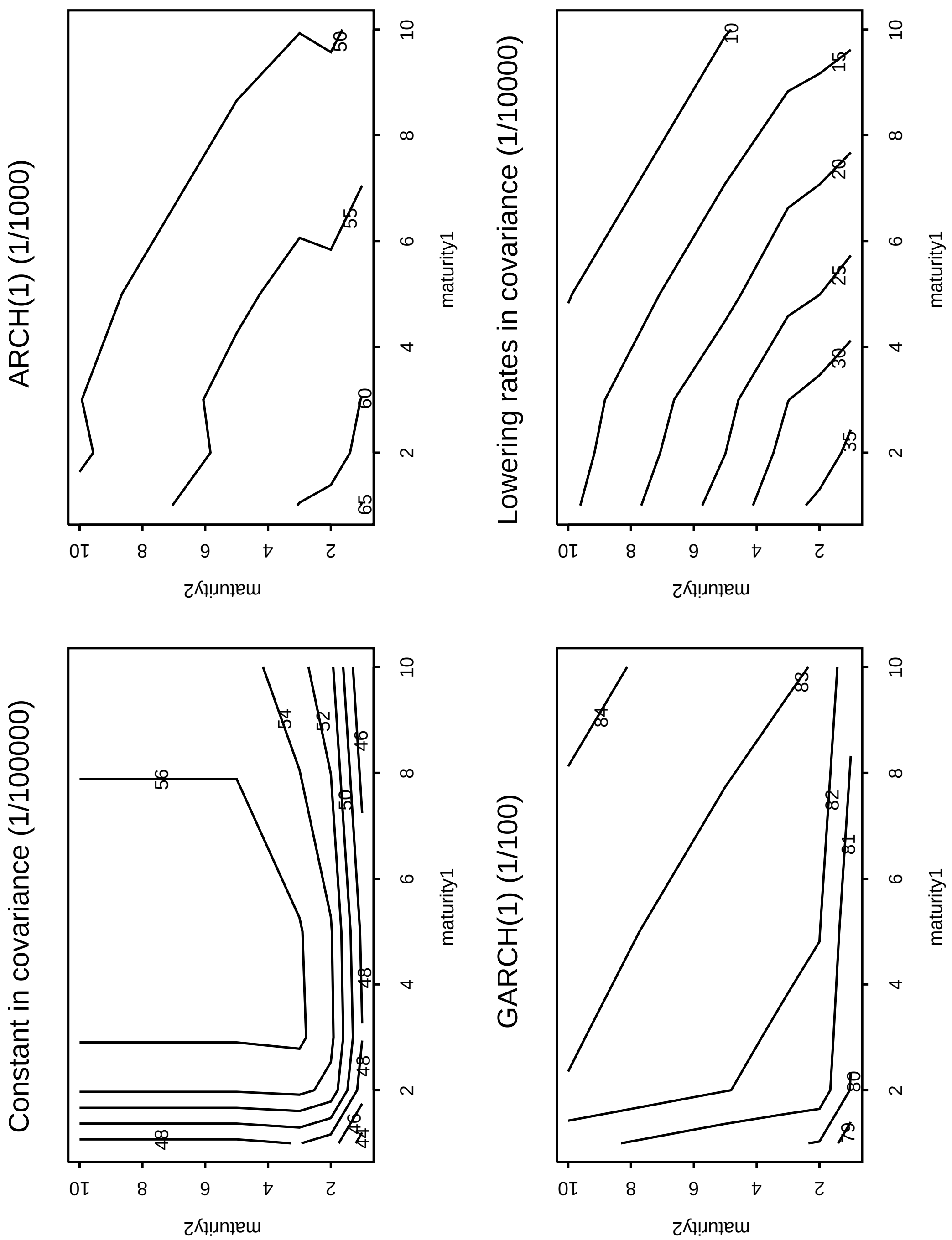


Figure 6.2: Contour plots of coefficient matrices of covariance from Matrix-Diagonal GARCH(1,1)-t

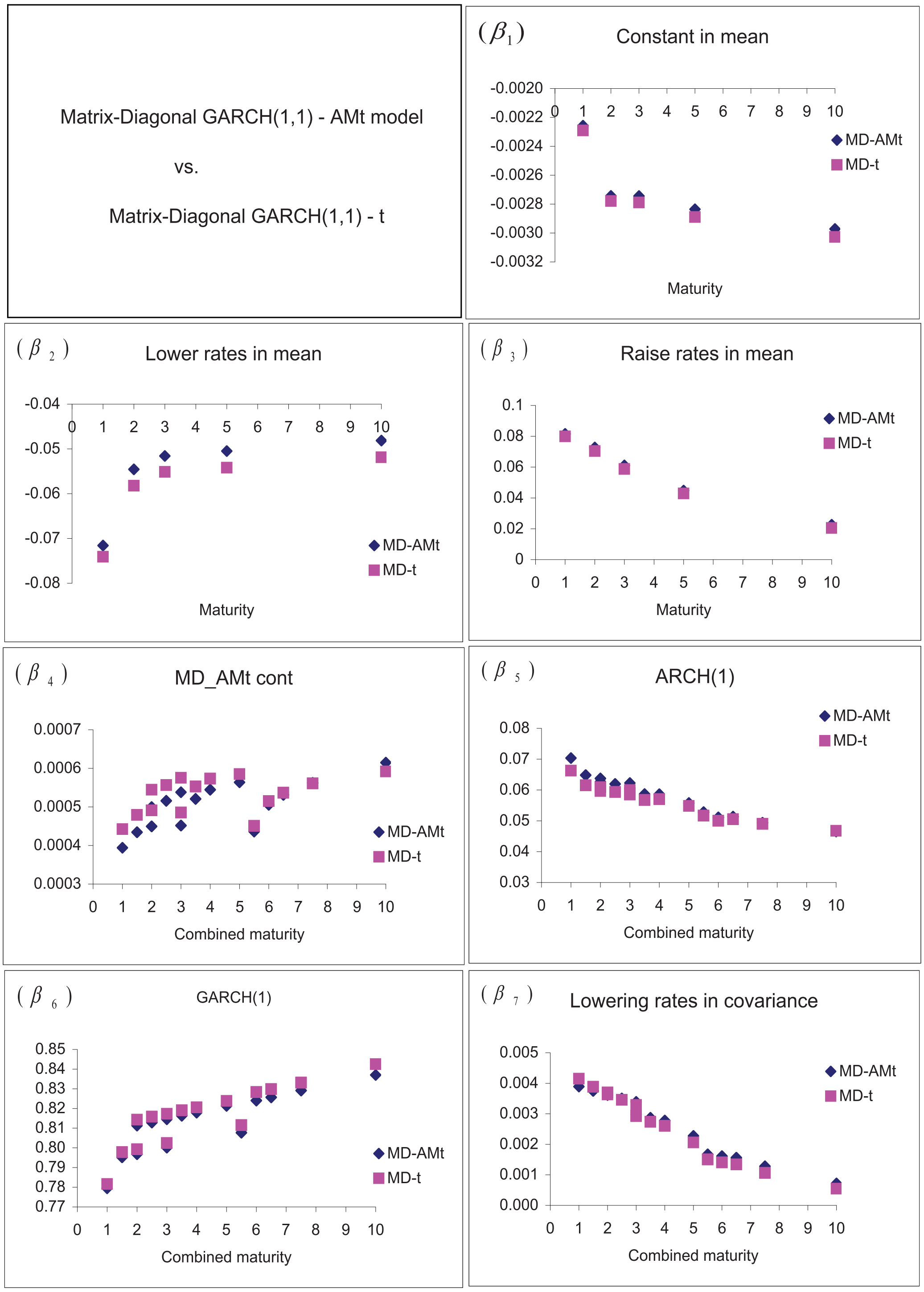


Figure 6.3: Estimations: Matrix-Diagonal GARCH(1,1)-AMt vs. Matrix-Diagonal GARCH(1,1)-t, combined-maturity$=(\theta_i + \theta_j)/2$

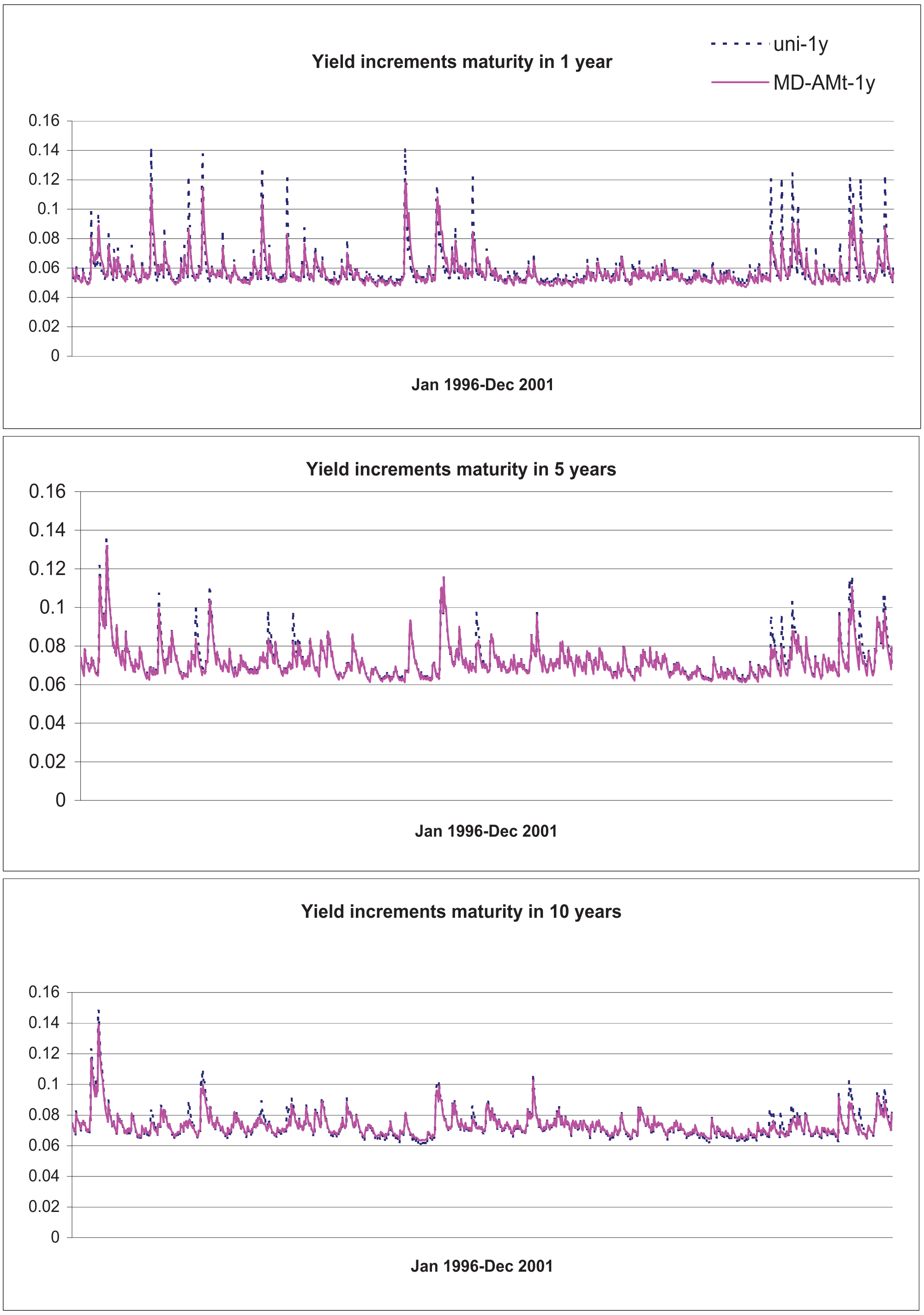


Figure 6.4: Comparison of standard deviations from univariate GARCH-t (dot lines) and Matrix-Diagonal GARCH-AMt (solid lines).

Chapter 7

A GARCH Model of Term Structures

7.1 Introduction

In Chapter 6, a multivariate GARCH model was developed for the collection of yield increment series at the maturities available for model building purposes. This multivariate model allowed for the conditional covariances between the individual series to be modelled. We observed in Chapter 6 that the parameters of the multivariate Matrix-Diagonal GARCH(1,1)-AMt model, characterising the GARCH process of conditional covariance, varied as smooth patterns in time to maturity. The trends of these patterns (Figure 6.1), except for the parameters of the constant term for the covariance, are similar to those observed for the univariate GARCH models of Chapter 5 (Figure 5.3).

In Chapter 5, we developed univariate GARCH type models of the interest rate (or, bond yield) increment time series for each fixed maturity. A major finding of this modelling is that the collection of univariate GARCH models for a set of yield increment series, each with fixed maturities, can be linked in a single GARCH model with parameters $\boldsymbol{\beta}$ expressed as functions of maturity in terms of new parameters $\boldsymbol{\varphi}$. We proposed a single GARCH model (5.9) in Chapter 5, and referred to it as a *GARCH model of term structure of interest rates* (TS-GARCH). The TS-GARCH model of Chapter 5 considers the conditional variance GARCH process only. The purpose of this chapter is to extend the TS-GARCH model proposed in Chapter 5 to the conditional covariances GARCH processes based on the multivariate GARCH(1,1)-AMt modelling of Chapter 6.

The conditional covariance TS-GARCH is presented in Section 2. The estimation based on the multivariate Matrix-Diagonal GARCH(1,1)-AMt modelling is given in Section 3. Section 4 presents a generalised concept of the TS-GARCH model and Section 5 gives a summary.

7.2 TS-GARCH Model

Chapter 6 developed a multivariate Matrix-Diagonal GARCH(1,1)-AMt model of d-dimensional yield increment series $\boldsymbol{y}_t$ described by Equation (6.5). Let

$$\begin{aligned}
\boldsymbol{y}_t &= (y_{1t}, \ldots, y_{dt})' \quad ,\\
\boldsymbol{\mu}_t &= (\mu_{1t}, \ldots, \mu_{dt})' \quad ,\\
\boldsymbol{\epsilon}_t &= (\epsilon_{1t}, \ldots, \epsilon_{dt})' \quad ,\\
\boldsymbol{z}_t &= (z_{1t}, \ldots, z_{dt})' \quad ,\\
H_t &= (h_{ijt})_{i,j=1,\ldots,d} \quad ,\\
\boldsymbol{\beta}_0 &= (m_1, \ldots, m_d)' \quad ,\\
\boldsymbol{\beta}_1 &= (\beta_{11}, \ldots, \beta_{d1})' \quad ,\\
\boldsymbol{\beta}_2 &= (\beta_{12}, \ldots, \beta_{d2})' \quad ,\\
\boldsymbol{\beta}_3 &= (\beta_{13}, \ldots, \beta_{d3})' \quad ,\\
A_0 A_0' &= (\beta_{ij4})_{i,j=1,\ldots,d} \quad ,\\
AA' &= (\beta_{ij5})_{i,j=1,\ldots,d} \quad ,\\
\boldsymbol{bb}' &= (b_i b_j) = (\beta_{ij6})_{i,j=1,\ldots,d} \quad ,\\
\boldsymbol{cc}' &= (c_i c_j) = (\beta_{ij7})_{i,j=1,\ldots,d} \quad .
\end{aligned} \tag{7.1}$$

According to the similarity of trends of parameters from multivariate Matrix-Diagonal GARCH(1,1)-AMt model (Figure 6.1) and those from the univariate GARCH models (Figure 5.3), the TS-GARCH model (5.9) proposed in Chapter 5 is extended to the conditional covariance TS-GARCH as follow:

$$\begin{aligned}
&\boldsymbol{y}_t|\mathcal{F}_{t-1} \sim AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}_t, S_t), \quad S_t = \tfrac{m-2}{m} H_t,\\
&\mu_{it} = \beta_{i1} + \beta_{i2} R_t^- + \beta_{i3} R_t^+, \quad i = 1, \ldots, d\\
&h_{ijt} = \beta_{ij4} + \beta_{ij5}\epsilon_{i(t-1)}\epsilon_{j(t-1)} + \beta_{ij6} h_{ij(t-1)} + \beta_{ij7} R_t^-, \quad i,j = 1, \ldots, d, \quad i \le j,
\end{aligned} \tag{7.2}$$

where

$$\begin{aligned}
m_i &= \varphi_{10} + \varphi_{20}\theta_i, \quad i = 1, \ldots, d,\\
\beta_{ik} &= \varphi_{1k} + \varphi_{2k}\ln(\theta_i), \quad i = 1, \ldots, d, k = 1, \ldots, 3,\\
\beta_{ijk} &= \varphi_{1k} + \varphi_{2k}\ln(\theta_{ij}), \quad \theta_{ij} = \tfrac{\theta_i + \theta_j}{2}, \quad i,j = 1, \ldots, d, i \le j, k = 4, \ldots, 7,
\end{aligned}$$

and $\epsilon_{it} = z_{it} - \mu_{it}$, $z_{it} = Q^{-1}_{m,i,t}(Q_{m_i,t}(y_{it}))$, $Q_{m,i,t}$ and $Q_{m_i,t}$ are the cumulative distribution functions of $t_m(\mu_{it}, \frac{m-2}{m} h_{ii,t})$ and $t_{m_i}(\mu_{it}, \frac{m_i-2}{m_i} h_{ii,t})$ respectively; and θ_i is the number of years to maturity of i-th yields; $i = 1, \ldots, d$.

The TS-GARCH model (7.2) implies that term structure interest rates increments are described by a single GARCH model with functional parameters. The functional parameters of degrees of freedom of t-distribution of marginals $\{m_1, \ldots, m_d\}$ depend linearly on maturity, the functional parameters of the mean equation $\{\beta_{i1}, \beta_{i2}, \beta_{i3}\}$ depend linearly on log maturity, and the functional parameters of conditional covariance equation $\{\beta_{ij4}, \beta_{ij5}, \beta_{ij6}, \beta_{ij7}\}$ depend linearly on the log *combined-maturity* of θ_i and θ_j denoted as θ_{ij}, while the average of maturities θ_i and θ_j is used as the combined-maturity of θ_i and θ_j in the model. The choice of the combined-maturity satisfies the following criteria. Firstly, the combined-maturity has to be a bivariate symmetric function of θ_i and θ_j because of the equal contributions of the two maturities θ_i and θ_j in covariance of ϵ_{it} and ϵ_{jt}. Secondly, the combined-maturity has to satisfy that $\theta_{ii} = \theta_i$ because the variance is a special covariance, $var(\epsilon_{it}) = cov(\epsilon_{it}, \epsilon_{it})$. The average of θ_i and θ_j is the most simple form of combined-maturity satisfying these two criteria. Other possible forms of combined-maturity will be discussed in Section 4.

7.3 Estimations

This section presents a two step procedure for estimation of the TS-GARCH model (7.2) based on the multivariate Matrix-Diagonal GARCH(1,1)-AMt model. We fixed $m = 6.3$ as explained in Chapter 6. The first step is, using the observed yield increment series at available maturities, to estimate the parameters $\{m_1, \ldots, m_d$; $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{b}, \boldsymbol{c}$, A_0, and $A\}$ of multivariate Matrix-Diagonal GARCH(1,1)-AMt (6.5) in Chapter 6, and calculate the TS-GARCH coefficients $m_1, \ldots, m_s$, $\{\beta_{ik} | k = 1, 2, 3; i = 1, \ldots, d\}$ and $\{\beta_{ijk} | k = 4, 5, 6, 7; i, j = 1, \ldots, d; i \leq j\}$ in (7.2) by equations in (7.1). Secondly, the new parameters $\boldsymbol{\varphi}$ of TS-GARCH model (7.2) are estimated by least squares estimation (LSE).

In Figure 7.1, the estimates of the parameters of the degrees of freedom of the marginals and the parameters of the mean equation $\{m_1, \ldots, m_d; \beta_{ik} | k = 1, 2, 3; i = 1, \ldots, d\}$ are plotted using solid dots against the maturity θ_i in Panels labelled (β_0) to(β_3). Panels labelled (β_0) to(β_3) give plots of the estimates of the parameters in the covariance equation, $\{\beta_{ijk} | k = 4, 5, 6, 7; i, j = 1, \ldots, d; i \leq j\ \}$, against the combined-maturity $\{\theta_{ij}|i,j = 1,\dots,d; i \leq j\}$. It is observed that log-linear trends exist clearly for most estimated β. Least Squares Estimation (LSE) of the trend lines are added on Figure 7.1. The estimated new parameters $\{\varphi_{1k}, \varphi_{2k}|k = 1,\dots,7\}$ of TS-GARCH model (7.2) explaining the trends are the coefficients of the equations displayed on Figure 7.1. The R^2 values are greater than 0.75, except for the term of the constant in variance equation, showing that the log linear regressions are quite good for describing the smooth patterns of the GARCH coefficients.

It is worth pointing out that the estimated trends of GARCH coefficients, displayed in Figure 7.1, are consistent with the results obtained in Chapter 5. Panel (β_2) and Panel (β_3) suggest that longer-term maturity yield increments are less impacted by RBA decisions of lowering or raising cash rates. Panel (β_5) suggests that the conditional covariance of yield increments with longer-term combined-maturity are less affected by the previous cross residuals $\epsilon_{i(t-1)}\epsilon_{j(t-1)}$, that is more rapid in mean reversion of the conditional covariances. Panel (β_6) suggests that the conditional covariances of yield increments with longer-term combined-maturity have greater persistence. Panel (β_7) suggests that conditional covariance of yield increments with longer-term combined-maturity is less impacted by RBA decisions of lowering the cash rate.

7.3.1 Using the Model for the Out of Sample Testing and Interpolates

The estimation of the TS-GARCH model (7.2) displayed in Figure 7.1 is based on RBA Treasury bond yields with maturities of 1, 2, 3, 5 and 10 years and for the period from January 1996 to December 2001. We now assess the performance of the model in forecasting volatility into the future and interpolating the volatility for a finer grid of maturities in the future.

Using the recently updated RBA Treasury bond yields from July 2000 to April 2004 maturity in yearly (from 1 year to 10 years), see Figure 7.4, the goodness of interpolation and forecasting by the TS-GARCH model (7.2) are examined. The interpolation is for the volatility of yield increments with additional maturities of 4, 6, 7, 8 and 9 years for the period 3/07/2000-31/12/2001, and the forecasting is for yield increments maturity in 1, 2, 3, 5 and 10 years period January 2002 - April 2004. The standardised residuals for the interpolated series and forecasted series are computed using the variances obtained from the TS-GARCH model. There is no ARCH effects left in the standardised residuals, and no correlation in the standardised residuals and squared standardised residuals for the interpolated series and forecasted series with a fixed maturity. QQ plots showed that t-distribution is adequate with the standardised residuals.

These results demonstrate that the TS-GARCH model is adequate for fitting, interpolation and forecasting of yield increments with any possible middle-to-long-term maturity.

7.3.2 Other Estimations of TS-GARCH Model

For the TS-GARCH model (7.2), denote

$$\boldsymbol{\beta} = (\underbrace{m_1, \ldots, m_d}_{\{\beta_{i0}, i=1,\ldots,d\}}, \underbrace{\beta_{11}, \ldots, \beta_{d1}}_{\{\beta_{i1}, i=1,\ldots,d\}}, \ldots, \underbrace{\beta_{13}, \ldots, \beta_{d3}}_{\{\beta_{i3}, i=1,\ldots,d\}}, \underbrace{\beta_{114}, \ldots, \beta_{dd4}}_{\{\beta_{ij4}, i,j=1,\ldots,d, i\leq j\}}, \ldots, \underbrace{\beta_{117}, \ldots, \beta_{dd7}}_{\{\beta_{ij7}, i,j=1,\ldots,d, i\leq j\}})', \tag{7.3}$$

$$\boldsymbol{\varphi} = (\varphi_{10}, \varphi_{20}, \varphi_{11}, \varphi_{21}, \ldots, \varphi_{17}, \varphi_{27})',$$

and denote $l = dim(\boldsymbol{\beta}) = 60$, and $s = dim(\boldsymbol{\varphi}) = 16$. The functional dependence of GARCH coefficients $\boldsymbol{\beta}$ in terms of new TS-GARCH parameters $\boldsymbol{\varphi}$ can then be expressed as

$$\boldsymbol{\beta} = A\boldsymbol{\varphi}, \tag{7.4}$$

where A is a $l \times s$ full column rank matrix. Let $\hat{\boldsymbol{\beta}}$ is the estimate of $\boldsymbol{\beta}$ from the first step, and assume $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \boldsymbol{\eta}$. Then the LSE of $\boldsymbol{\varphi}$ at the second step is

$$\hat{\boldsymbol{\varphi}} = (A'A)^{-1}A'\hat{\boldsymbol{\beta}}. \tag{7.5}$$

To discuss a consistent, asymptotically normal and efficient estimation of TS-GARCH model, a multivariate diagonal VEC (DVEC) GARCH(1,1)-AMt model is described by equations,

$$\begin{aligned} &\boldsymbol{y}_t|\mathcal{F}_{t-1} \sim AMt_d(m; m_1, \ldots, m_d; \boldsymbol{\mu}_t, S_t), \quad S_t = \tfrac{m-2}{m}H_t, \\ &\boldsymbol{\mu}_t = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \cdot R_t^- + \boldsymbol{\beta}_3 \cdot R_t^+, \\ &H_t = \boldsymbol{\beta}_4 + \boldsymbol{\beta}_5 \otimes (\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + \boldsymbol{\beta}_6 \otimes H_{t-1} + \boldsymbol{\beta}_7 \cdot R_t^-, \end{aligned} \tag{7.6}$$

where $\boldsymbol{\epsilon}_t$ is same as defined in (6.5), $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, and $\boldsymbol{\beta}_3$ are d-dimensional vectors; $\boldsymbol{\beta}_4 = (\beta_{ij4}), \boldsymbol{\beta}_5 = (\beta_{ij5}), \boldsymbol{\beta}_6 = (\beta_{ij6})$ and $\boldsymbol{\beta}_7 = (\beta_{ij7})$ are $d \times d$ symmetric matrices.

The estimation of $\{m_1, \ldots, m_d,\ \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{b}, \boldsymbol{c}, A_0, A\}$ of the multivariate Matrix-Diagonal GARCH(1,1)-AMt (6.5) from Section 6.3 is not guaranteed to be asymptotically normal. However, given a consistent, asymptotically normal estimation of $\{m_1, \ldots, m_d$, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{b}, \boldsymbol{c}, A_0, A\}$ of the multivariate Matrix-Diagonal GARCH(1,1)-AMt (6.5), by

δ-method (see Rao 1973), $\hat{\boldsymbol{\beta}}$ derived from the estimation by equations (7.1) and (7.3) is a consistent, asymptotically normal estimation of the multivariate diagonal VEC GARCH(1,1)-AMt model (7.6). Then the LSE of $\boldsymbol{\varphi}$ by (7.5) is a consistent, asymptotically normal estimation of TS-GARCH model (7.2). However, the two step procedure estimation method is generally not efficient.

We now describe an asymptotically normal and efficient estimation method for the TS-GARCH model.

Comte and Lieberman (2003) established the consistent and asymptotic normality for the general multivariate GARCH model under certain regularly conditions. Let $\hat{\boldsymbol{\beta}}$ be the MLE of $\boldsymbol{\beta}_0$ of the above DVEC model (7.6), where $\boldsymbol{\beta}_0$ denotes the true value of the parameters and $\boldsymbol{\beta}_0 = A\boldsymbol{\varphi}_0$ where φ_0 are the true values of φ in TS-GARCH model (7.2). The asymptotic consistency and asymptotic normality for $\hat{\boldsymbol{\beta}}$ are expressed as

$$\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$$

and

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{\mathcal{D}} N(0, J^{-1})$$

where $J = E[\nabla_\beta \ln f(\boldsymbol{y}|\boldsymbol{\beta}_0)\{\nabla_\beta \ln f(\boldsymbol{y}|\boldsymbol{\beta}_0)\}']$ exists and is nonsingular.

Let $\hat{\boldsymbol{\varphi}} = (A'A)^{-1}A'\hat{\boldsymbol{\beta}}$ be the LSE of $\boldsymbol{\varphi}$ for the relation equation (7.4) at $\hat{\boldsymbol{\beta}}$. If the true value of $\boldsymbol{\varphi}$ is $\boldsymbol{\varphi}_0$, $\boldsymbol{\varphi}_0 = (A'A)^{-1}A'\boldsymbol{\beta}_0$. Then $\hat{\boldsymbol{\varphi}}$ is a consistent and asymptotically normal estimator of covariance version TS-GARCH model (7.2), because

$$\hat{\boldsymbol{\varphi}} = (A'A)^{-1}A'\hat{\boldsymbol{\beta}} \xrightarrow{p} (A'A)^{-1}A'\boldsymbol{\beta}_0 = \boldsymbol{\varphi}_0,$$

and

$$\sqrt{n}(\hat{\boldsymbol{\varphi}_W} - \boldsymbol{\varphi}_0) \xrightarrow{\mathcal{D}} N(0, (A'A)^{-1}A'J^{-1}A(A'A)^{-1}).$$

Unfortunately, this estimation is not asymptotically efficient because

$$\begin{aligned} E[\nabla_\varphi \ln f(\boldsymbol{y}|\boldsymbol{\varphi}_0)\{\nabla_\varphi \ln f(\boldsymbol{y}|\boldsymbol{\varphi}_0)\}'] &= A'E[\nabla_\beta \ln f(\boldsymbol{y}|\boldsymbol{\beta}_0)\{\nabla_\beta \ln f(\boldsymbol{y}|\boldsymbol{\beta}_0)\}']A \\ &= A'JA \\ &\neq [(A'A)^{-1}A'J^{-1}A(A'A)^{-1}]^{-1}, \end{aligned}$$

in general, since $J \neq \sigma^2 I$ (I is a unit matrix) and A is not square.

However, adjusting the LSE $\hat{\boldsymbol{\varphi}}$ to generalised (weighted) least squares estimation (WLSE) $\hat{\boldsymbol{\varphi}}_W$, it can provide an asymptotically normal and efficient estimation. Because

$$\hat{\boldsymbol{\varphi}}_W = (A'JA)^{-1}A'J\hat{\boldsymbol{\beta}}$$

and

$$
\begin{aligned}
var(\sqrt{n}(\hat{\boldsymbol{\varphi}} - \boldsymbol{\varphi}_0)) &= (A'JA)^{-1}A'Jvar(\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0))JA(A'JA)^{-1} \\
&\to (A'JA)^{-1}A'JJ^{-1}JA(A'JA)^{-1} \\
&= (A'JA)^{-1} \\
&= (E[\nabla_{\varphi} \ln f(\boldsymbol{y}|\boldsymbol{\varphi}_0)\{\nabla_{\varphi} \ln f(\boldsymbol{y}|\boldsymbol{\varphi}_0)\}'])^{-1}.
\end{aligned}
$$

Thus, first estimating $\boldsymbol{\beta}$ of the DVEC model (7.6), by MLE and then estimating $\boldsymbol{\varphi}$ by WLSE, will provide a consistent, asymptotically normal and asymptotically efficient estimation of TS-GARCH.

However, there is difficulty to implement this method. Because,

- the estimation of the multivariate diagonal VEC GARCH(1,1)-AMt model (7.6) derived from the estimation of the multivariate Matrix-Diagonal GARCH(1,1)-AMt (6.5) in Section 6.3 by equations (7.1) and (7.3) is not guaranteed to be a MLE or approximately MLE, although we have claimed that the estimation from Section 6.3 should be approximately the MLE;
- one of the certain regularly conditions of the asymptotic normality (Comte and Lieberman, 2003) requires $\boldsymbol{\epsilon}_t$ have finite 8^{th} moment. It was observed that a t-distribution with between 5 to 8 degrees of freedom provides a good model. So, the condition of $\boldsymbol{\epsilon}_t$ having finite 8^{th} moment is not satisfied; and,
- the estimation of the covariance matrix of estimates J is not available from the multivariate GARCH model using S+FinMetrics.

To get an efficient estimation of TS-GARCH, we attempted to developed a MATLAB program to get a maximum likelihood estimation (MLE) of the TS-GARCH model directly for the parameters of $m_1, \ldots, m_d$; $\varphi_{10}, \varphi_{20}, \varphi_{11}, \varphi_{21}, \ldots, \varphi_{17}, \varphi_{27}$. Unfortunately, the algorithm did not converge. It is a further study to develop an algorithm for efficient estimation of TS-GARCH parameters using MLE on the constrained parameterisation of (7.4).

7.4 Further Generalisation of the TS-GARCH Model

In this section, we discuss additional generalisations to the TS-GARCH model. The first provides alternative specifications of the combined maturity for the covariance parameters functions and finds the functional patterns of TS-GARCH model that assure positive-definite H_t. The second extension presents the general functions of TS-GARCH model. The third extension discussed is use of other meta-elliptical distributions with given marginals.

7.4.1 The First Extension: Alternative Specifications of the Combined Maturity

Section 2 proposed a TS-GARCH(1,1) model (7.2) for covariances processes of RBA Treasury bond yield increments under the assumption of multivariate asymmetric t-distribution. In Figure 7.1, the log-linear functions for all TS-GARCH coefficients in model (7.2) depend on the average combined-maturity of θ_i and θ_j given by $\theta_{ij} = (\theta_i + \theta_j)/2$ $(i, j = 1, \ldots, d, i \leq j)$, and $\theta_{ii} = \theta_i$. In some Panels of Figure 7.1, the individual parameter estimates represented by the solid dots are not well represented by the fitted line. See for example, the parameters for the constant in the covariance equation in Panel (β_4) and parameters for the GARCH(1) in the covariance equation in Panel (β_6). The TS-GARCH model (7.2) is implied by the multivariate diagonal VEC GARCH model (7.6). However, the H_t in the diagonal VEC model (7.6) cannot be guaranteed to be positive-definite if the coefficient matrices $\boldsymbol{\beta}_4, \ldots,$ $\boldsymbol{\beta}_6$ are not positive-definite. The estimated $\boldsymbol{\beta}_4$, $\boldsymbol{\beta}_6$ and $\boldsymbol{\beta}_7$, calculated by the estimated coefficients in equations of the TS-GARCH model displayed in Figure 7.1 and the formula in (7.2), are not positive-definite matrices. Hence the estimated H_t are not positive-definite for many t. Alternative definitions of combined-maturity for functional patterns are defined for describing the coefficient matrices in covariance equation of the Matrix-Diagonal GARCH(1,1)-AMt that guarantees the H_t are positive-definite. The alternative combined-maturity has to retain the basic properties of (i) being a bivariate symmetric function of θ_i and θ_j, and (ii) the combined-maturity is equal to the maturity when i=j, as $\theta_{ii} = \theta_i$, $i, j = 1, \ldots, d$. One possibility is to include the difference of the two-maturity levels in a combined-maturity as well as the average of the two-maturity levels. After some experiments, an alternative combined-maturity $\theta_{ij} = \frac{1}{2}(\theta_i + \theta_j - \frac{2}{3}|\theta_i + \theta_j|)$, $i, j = 1, \ldots, d$ is proposed for $\boldsymbol{\beta}_4$, $\boldsymbol{\beta}_6$, an alternative combined-maturity $\theta_{ij} = \frac{1}{2}(\theta_i + \theta_j - \frac{1}{2}|\theta_i + \theta_j|)$, $i, j = 1, \ldots, d$ is proposed for $\boldsymbol{\beta}_7$ and the average combined-maturity $\theta_{ij} = \frac{1}{2}(\theta_i + \theta_j)$ is still used for $\boldsymbol{\beta}_5$. These new choices make the fit displayed on Figure 7.2 better than Figure 7.1 and produces estimated coefficients matrices being positive-definite, ensuring that H_t will be positive-definite for all time t.

7.4.2 The Second Extension: General Functions of TS-GARCH Model

An alternative approach to estimate TS-GARCH that directly guarantees the positive-definite H_t is to find the pattern of $A_0, A, \boldsymbol{b}$ and $\boldsymbol{c}$ of Matrix-diagonal GARCH(1,1)-AMt model as functions of maturity and combined-maturity. The estimated $\boldsymbol{b}$ and $\boldsymbol{c}$ were displayed on Figure 7.3 as functions of maturity by solid points. The trend lines of $\boldsymbol{b}$ and $\boldsymbol{c}$ depending on the maturity are clear and displayed. $R^2(> 0.9)$ confirm that the trends are significant and the trend lines fit very well. So the alternative proposed functions in TS-GARCH can be

$$\beta_{ij6} = (\varphi_{16} + \varphi_{26}\ln(\theta_i))(\varphi_{16} + \varphi_{26}\ln(\theta_j)),$$

and

$$\beta_{ij7} = (\varphi_{17} + \varphi_{27}\theta_i)(\varphi_{17} + \varphi_{27}\theta_j).$$

However, it is difficult to find the patterns in the elements of A_0 and A which are appropriate functions of combined-maturity. If the matrix A is reduced to a d-dimensional vector, the diagonal of AA' loses the pattern as observed in univariate GARCH(1,1)-t models in Chapter 5. The A_0 is restricted to be a square lower triangular matrix in S+FinMetrics.

In order to cover all theses variants, for a d-dimensional random vector $\boldsymbol{y}_t$, a generalised TS-GARCH model can be described by, assuming that there are in total s parameters in distribution, m parameters in the mean equation and n parameters in the covariance GARCH equation,

$$\begin{aligned} &(y_{it}, h_{ijt}) \sim GARCH(\beta_i^{(1)}, \ldots, \beta_i^{(m)}; \beta_{ij}^{(1)}, \ldots, \beta_{ij}^{(n)}) \\ &\beta_i^{(k)} = f_k(\theta_i, \boldsymbol{\phi}_k), \quad k = 1, \ldots, m, \\ &\beta_{ij}^{(k)} = g_k(\theta_i, \theta_j, \boldsymbol{\varphi}_k), \quad k = 1, \ldots, n, \end{aligned} \tag{7.7}$$

where $i, j = 1, \ldots, d; i \leq j$. $f_k, (k = 1, \ldots, m)$, are smooth functions describing the trend of the k-th GARCH parameters $\{\beta_1^{(k)}, \ldots, \beta_d^{(k)}\}$ in mean equation, and $g_k, (k = 1, \ldots, m)$, are functions describing the trend of k-th GARCH parameters $\{\beta_{ij}^{(k)}, i \neq j\}$ in covariance equation. A reasonable assumption is that $dim(\boldsymbol{\phi}_k) < d, k = 1, \ldots, m$ and $dim(\boldsymbol{\varphi}_k) < d(1 + d)/2, k = 1, \ldots, n$.

The dimension of all β parameters is $dim(\boldsymbol{\beta}) = md + nd(1+d)/2$, and the dimension of all new (ϕ, φ) parameters of TS-GARCH is

$$dim(\boldsymbol{\phi}, \boldsymbol{\varphi}) = \sum_{1}^{m} dim(\boldsymbol{\phi}_k) + \sum_{1}^{n} dim(\boldsymbol{\varphi}_k) < md + nd(1+d)/2 = dim(\boldsymbol{\beta}).$$

Thus, the TS-GARCH not only smoothes the GARCH coefficients that provides the possible interpolation and forecasting for any middle-to-long-term maturities, but also reduces the number of parameters to be estimated.

For the generalised TS-GARCH model by equation (7.7), the MLE is to find $(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\varphi}})$ which maximise the log likelihood function

$$L(\boldsymbol{\phi}, \boldsymbol{\varphi} | y_{11}, \ldots, y_{1n}, \ldots, y_{d1}, \ldots, y_{dn}), \tag{7.8}$$

subject to the constraint that the covariance matrices H_t are positive-definite.

Further research is required to develop a reliable and convergent algorithm for the MLE estimation of TS-GARCH.

7.4.3 The Third Extension: Other Meta-elliptical Distributions with Given Marginals

It needs to be noted that the t-distribution was rejected by Kolmogorov-Smirnov Test for the RBA Treasury bond yield increments, although the $QQ-t$ plots indicate that the t-distribution shows a marked improvement over normal distribution – See Chapter 5. Further work could be to find an appropriate distribution for each individual yield increment series in univariate GARCH modelling, and replacing the multivariate asymmetric t-distribution by a meta-elliptical distribution with given distributions of marginals for multivariate GARCH modelling.

7.5 Conclusions

In this chapter, we developed the concept of GARCH model of term structure of interest rates (TS-GARCH) initially proposed in Chapter 5. TS-GARCH is a parsimonious specification for dealing with the GARCH parameters in relation to maturity. The patterns exhibited in Figure 5.3 and Figure 7.1 justify such specification for both univariate GARCH models and multivariate models. With such a specification, one can have a predicted specification for yield increments under a different maturity even if the data on that particular maturity are absent. Thus, TS-GARCH is a single GARCH equation that characterises all covariance processes for any possible selection of maturities. An estimation of the TS-GARCH was provided using a two-step method based on the estimation of the multivariate Matrix-Diagonal GARCH(1,1)-AMt. This estimation is generally not efficient. Using the WLSE, instead of the LSE, we have discussed a consistent, asymptotic normal and efficient estimate of TS-GARCH by a two-step approach. However, it is difficult to implement the estimation. It is worth developing a reliable algorithm that pools the available yield curves data in a single maximum likelihood fit to obtain a converging and efficient MLE of the TS-GARCH.

Based on sample and out-of-sample assessment of Australian Treasury bond yield data, the TS-GARCH model provides an appropriate model for yield increment series in any possible middle-to-long-term maturity.

As it stands, the TS-GARCH model we are proposing is empirically derived and is not supported by financial theory as far as we can tell. Derivation of the log-linear (or similar) relation from financial theory could be an interesting future research challenge.

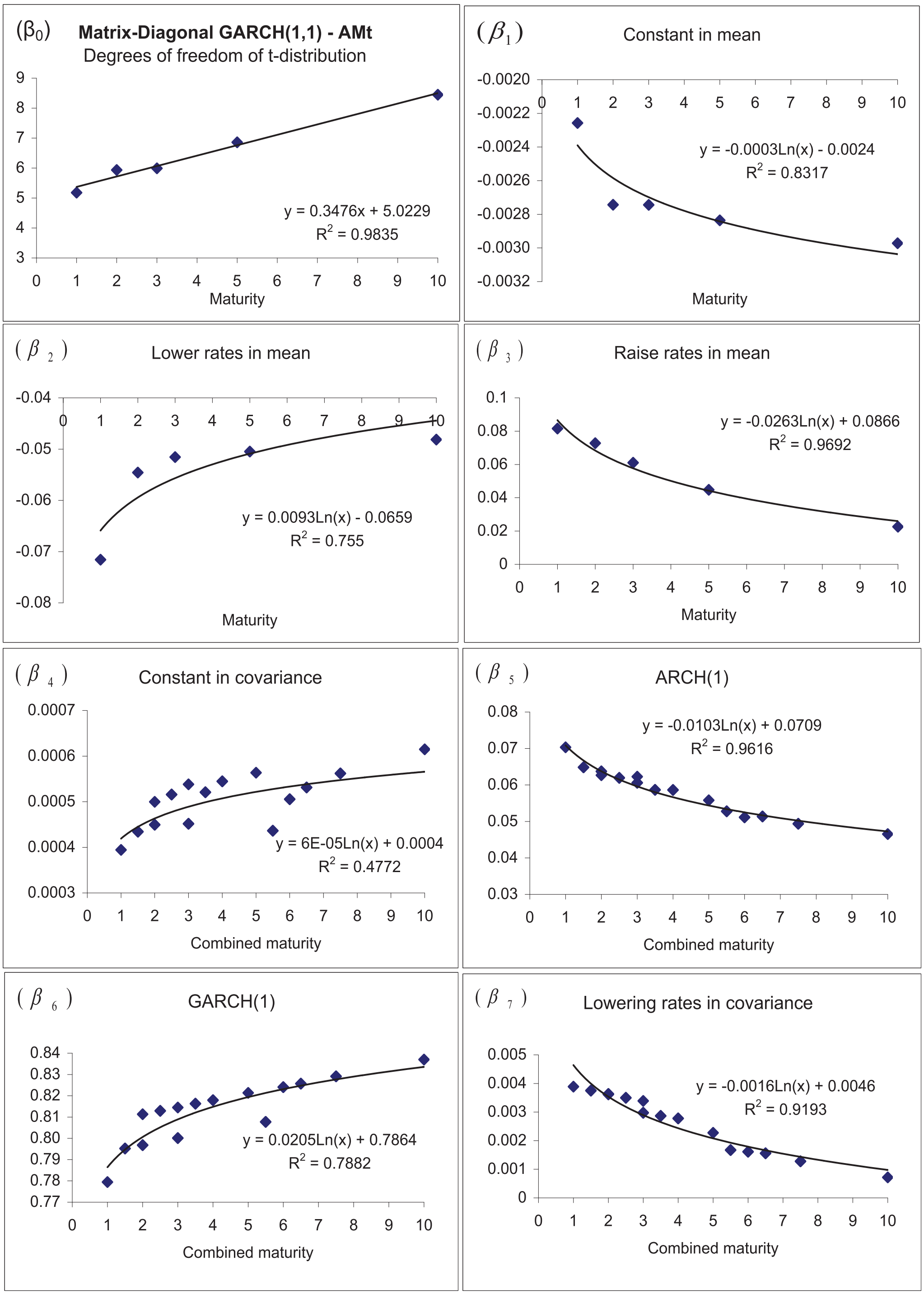


Figure 7.1: An estimate of TS-GARCH based on multivariate Matrix-Diagonal GARCH(1,1)-AMt, combined-maturity=$(\theta_i + \theta_j)/2$

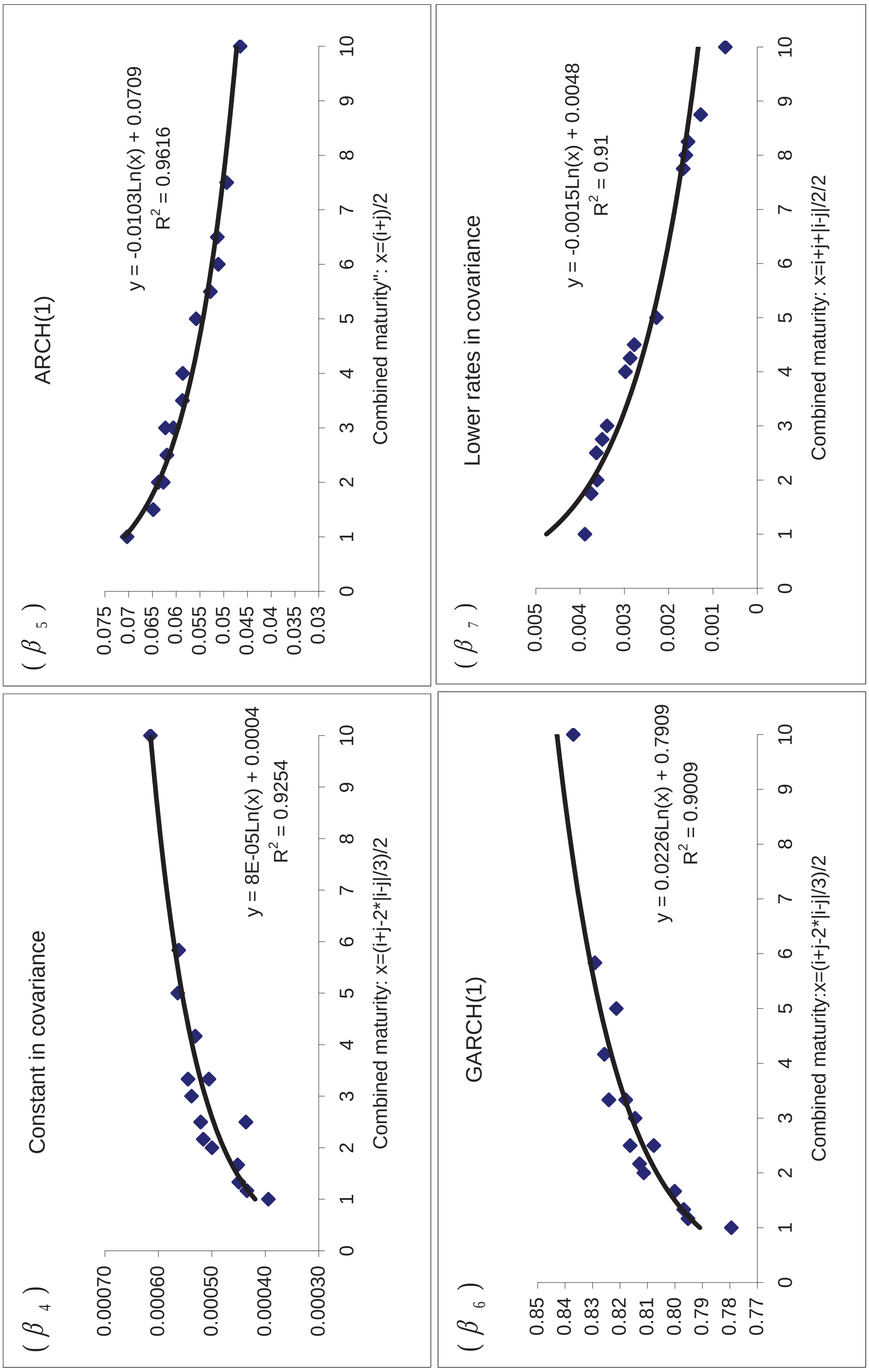


Figure 7.2: Another estimates of TS-GARCH: covariance process

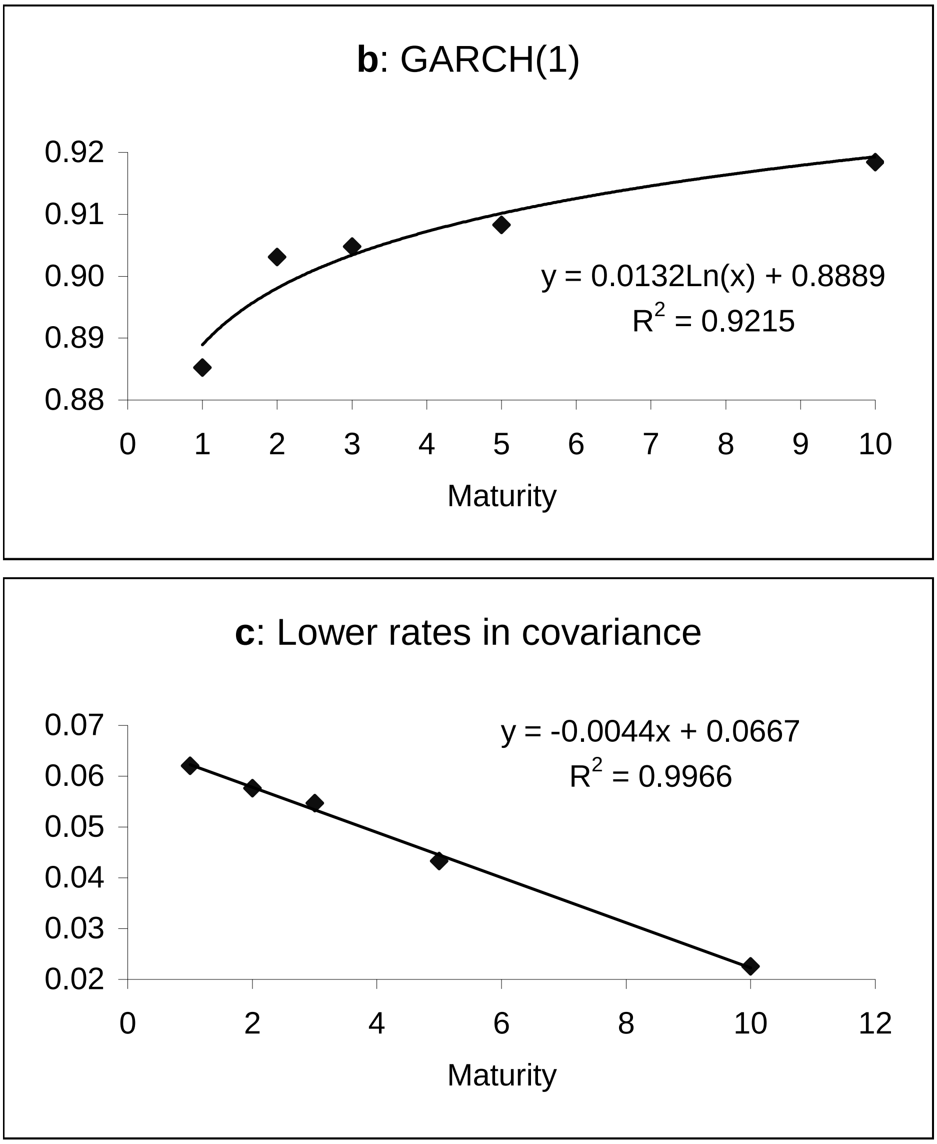


Figure 7.3: Alternative TS-GARCH functions

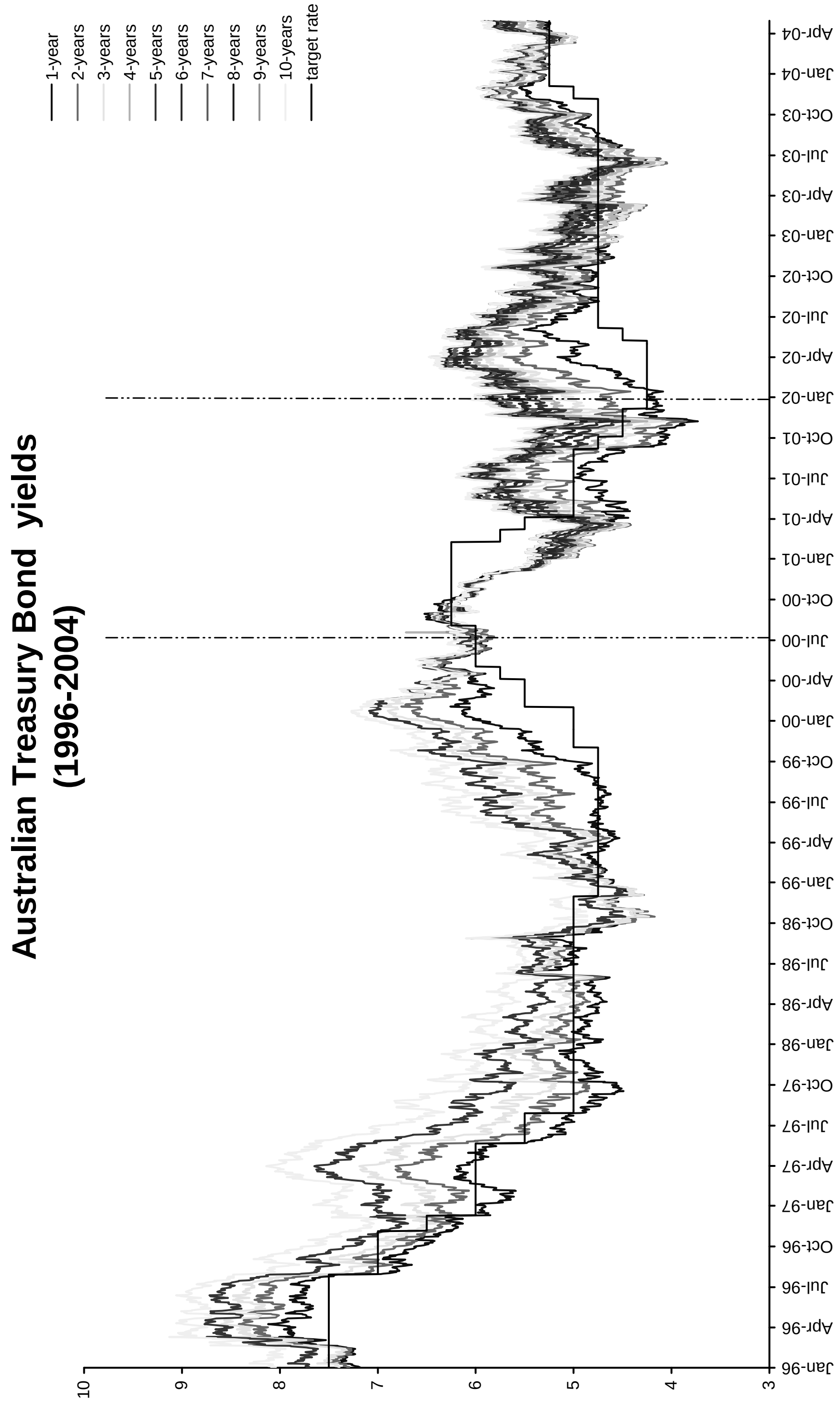


Figure 7.4: RBA yield curves 1996-2004

Chapter 8

Summary and Conclusions

This thesis has made contributions in the following areas

- Theoretical investigation of θ-parameterised term structure of interest rates;
- Derivations of yield curves;
- Univariate GARCH modelling and the specification of GARCH parameters to propose TS-GARCH;
- Definition of the multivariate asymmetric t-distribution using meta-elliptical distribution concepts (Fang, Fang and Kotz 2002);
- Development of Matrix-Diagonal GARCH(1,1) with multivariate asymmetric t-distribution for yield increments;
- The estimation of Matrix-Diagonal GARCH(1,1) with multivariate asymmetric t-distribution using an iterative approach which uses available software.
- Development of the concept of term structured GARCH models and a method of estimation for it with application to RBA term structure of interest rates.

We now expand on each of these contributions.

The θ-parameterisation of term structure of interest rates provides a convenient mathematical concept, as well as a convenient language for statistical reporting, analysis and modelling of yield curves. The volatility functions of the term structure in θ-parameterisation and T-parameterisation are identified and made to correspond to each other by a simple transformation $\theta = T - t$. Systematically investigating the term structure of interest rates in the θ-parameterisation, under certain conditions, we find a martingale process of the relative bond price and a martingale process of bond price ratio. Consequently, we get a no-arbitrage condition in terms of the drift process and volatility process of term structure of interest rates $r(t, \theta)$. Also, using a model of θ-parameterised term structure of interest rates, we find that the volatility of Australian Treasury bond yield is humped around the maturity in one year, a fact which has been found in other literature.

The two types of yields data sets available in Australia are the generic yield curves produced by the Reserve Bank of Australia (RBA) based on the Treasury bonds on issue and the constructed yield curves of the Commonwealth Bank of Australia (CBA) derived from swap rates. We use the RBA yield increments for statistical analysis and empirical modelling. It is found that the short-term (having maturity less than one year) bill yield increments have different volatility dynamics than do middle-to-long-term (having maturity in one year or over one year) bond yield increments. The short-term bill yield increment has no ARCH effects, while both yield increments and squared yield increments are not autocorrelated. The middle-to-long term bond yield increment has ARCH effects, while yield increments are not autocorrelated and squared yield increments are autocorrelated. The distributions for all yield and yield increments are non-normal.

Using available middle-to-long-term Treasury bond yield data from the RBA we develop individual GARCH models for yield increments for 1, 2, 3, 5, 10 year maturities. These are GARCH(1,1) models with residuals in a t-distribution and exogenous variables indicating RBA decisions of lowering or raising target cash rate. We found, from the process of GARCH modelling based on RBA yield data, that: a) The indicator of RBA monthly Board meetings is not significant in both the mean and variance equation; b) RBA decisions of lowering target rates is significant in both the mean and variance equation; c) RBA decision of raising target rate only impacts on the mean equation; and d) The effects, on both the mean increments and volatility, of RBA raising rates are significant from lowering rates.

The GARCH(1,1) model effectively captures many important phenomena of the financial series considered here such as heavy tails, volatility persistency, volatility mean reversion, asymmetric impact of positive or negative innovations in mean and volatility, and the wide stationarity of the residuals process. It is an important finding in this thesis that the GARCH coefficients of individual yield increment series with a fixed maturity time θ are functionally dependent on the time to maturity θ. The functional patterns of the parameter estimates can be interpreted as:

a) The degrees of freedom of t-distribution of longer-term bond yield increments is greater, implying that longer-term bond yield increments are closer to being normally distributed;

b) A change in RBA cash target rate has less impact on longer-term yield increments;

c) The longer-term bond yield has smaller volatility;

d) The longer-term bond yields are less affected by the previous residuals in the yield increments. i.e. are more efficient in mean reversion;

e) The longer-term bond yields are more persistent with the previous volatility;

f) RBA decisions to lower target rates have less impact on longer-term bond yields volatility;

g) The residuals of GARCH(1,1) are wide-sense stationary;

h) The asymmetric impact, on both the mean increments and volatility, of the RBA decisions of raising and lowering the RBA target cash rate. The RBA decision of lowering the RBA target cash rate has significantly stronger impact on the mean of yield increment then the decision of raising the RBA target cash rate. The RBA decision of lowering the RBA target cash rate has significant impact on the volatility of yield increment, while the RBA decision of raising the RBA target cash rate has an insignificant impact on the volatility of yield increment.

Our investigation of the patterns of the GARCH parameters as functions of maturity shows that the degrees of freedom depend linearly on the time to maturity, and other GARCH parameters depend linearly on the logarithm of time to maturity. These empirical observations lead us to propose a specification for a new model. This GARCH model, having coefficients expressed as parametric functions of the time to maturity, can characterise the volatility in a collection of yield increment series for any possible maturity times. The model that was initially proposed in Chapter 5, for the conditional variance process only is referred to as GARCH model of term structure of interest rates (*the variance version* of the TS-GARCH).

Chapter 6 studied the multivariate GARCH models of the dynamic conditional covariance of term structure of yield increments. Available literature in multivariate GARCH modelling and statistical software for computing the estimation of the multivariate GARCH models are based on the assumption of multivariate Normal distribution or multivariate t-distribution with the same degrees of freedom for all marginals. However, the univariate GARCH models developed in Chapter 5 for individual yield increment time series at each maturity found that the degrees of freedom depend linearly on the time to maturity, which implies the different marginal distributions of the multivariate yield increments.

The multivariate asymmetric t-distribution using meta-elliptical distributions concepts (Fang, Fang and Kotz 2002) is defined in Section 6.1, which allows different marginals. It extends and modifies the multivariate asymmetric t-distribution presented in Fang, Fang and Kotz (2002). The multivariate asymmetric t-distribution presented in Fang, Fang and Kotz (2002) had zero mean and a dispersion matrix specified as a correlation matrix without loss of generality. Also the transformation functions are the distribution functions of Student's t-distribution that implies the same dispersions for the original random variables and for the constructed random variables, dispersions being 1. For the purpose of modelling the volatility of yield curves, we define the multivariate asymmetric t-distribution with general mean and covariance matrix, and moreover, the original random vector and the constructed random vector have the same variances.

With the general multivariate asymmetric t-distribution, a Matrix-Diagonal GARCH(1,1) model for RBA Treasury yields is developed that allows the different degrees of freedom of t-distributions in Section 6.2. The estimation of the model using MLE is computationally challenging. We have, alternatively, successfully implemented an iterative method in Section 6.3. It is concluded in Section 6.4 that the Matrix-Diagonal GARCH (1,1) model with multivariate asymmetric t-distribution and including exogenous variables indicating RBA raising and lowing target cash rate (6.5) is an appropriate model. This model captured the main characteristics of financial data series such as mean reversion, persistency and stationarity of the conditional variances of the residual processes.

The likelihood ratio test shows that the Matrix-diagonal GARCH(1,1) with multivariate asymmetric t-distribution (different t-df of marginals) are significantly different from the simple Matrix-diagonal GARCH(1,1) with multivariate t-distribution (same t-df of marginals). The Akaike Information Criterion (AIC) shows that the Matrix-diagonal GARCH(1,1) model with multivariate asymmetric t-distribution (The correlation of yield curves is taken into account, and having more parameters) fits better than the univariate GARCH(1,1) models (The correlation of yield curves is ignored, and having less parameters). The estimates of the parameters in the mean equation and covariance equation of multivariate Matrix-Diagonal GARCH model also follow functional patterns depending on the maturities, and the interpretation of these patterns is mostly consistent with those of the univariate GARCH models. Combining the variance version TS-GARCH modelling ideas of Chapter 5 with the conditional covariance GARCH process from multivariate Matrix-Diagonal GARCH modelling of Chapter 6, Chapter 7 proposed the *covariance version* TS-GARCH model which is a single GARCH model characterising a collection of conditional covariance processes of term structure of yield increments.

TS-GARCH is a parsimonious specification dealing with the GARCH parameters in relation to maturity. The patterns exhibited in Figure 5.3 and Figure 7.1 justify such specification for both univariate GARCH models and multivariate models. With such a specification, one can have a predicted specification for yield increments under a different maturity even if the data on that particular maturity is absent. The main benefit of the TS-GARCH modelling is, using the limited available observed yield data, to model the dynamic conditional covariance processes of term structure of yield increments in any possible maturities based on the smoothing functional GARCH parameters depending on the maturities. Thus, TS-GARCH is a single GARCH equation that characterises all covariance processes for any possible selection of maturities. Estimations of the TS-GARCH was provided by a two-step approach based on the estimation of the multivariate Matrix-Diagonal GARCH(1,1)-AMt model. We have tested the appropriateness of TS-GARCH by diagnosis the model fitting, interpolation and forecasting. Based on sample (1996-2001, maturities in 1, 2, 3, 5, 10 years) and out-of-sample (2000-2004, maturity in yearly 1, 2, ..., 10 years) assessment of RBA Treasury bond yield data, the TS-GARCH model provides an appropriate model for yield increment series in any possible middle-to-long-term maturity. It is worth developing an algorithm that pools the available yield curves data in a single maximum likelihood fit to obtain a converging and efficient MLE of the TS-GARCH.

It is interesting and challenging future work to further investigate of the derivation of the smoothing functions of multivariate GARCH parameters in relation to maturity and combined maturity, to describe and support from financial theory.

Appendix A

Multivariate GARCH Models

A.1 Review of Multivariate GARCH Models

In this section, we review the most popular multivariate GARCH models. They are the Diagonal VEC model (Bollerslev, Engle and Wooldridge, 1988), the Matrix-Diagonal model (Ding 1994, Bollerslev, Engle and Nelson 1994), the BEKK model (BEKK 1991, Engle and Kroner 1995), the CCC model (Bollerslev 1990) and the DCC model (Engle and Sheppard 2001). This review is based substantially on Zivot and Wang (2003), as well as the references cited.

Consider a multivariate (d-dimensional) time series $\boldsymbol{\epsilon}_t$, $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \cdots, \epsilon_{dt})'$, $t = 1, \cdots, T$. Assume that the $\boldsymbol{\epsilon}_t$ have zero mean vector. The conditional covariance matrix of $\boldsymbol{\epsilon}_t$ is assumed to follow the time-varying structure given by

$$cov(\boldsymbol{\epsilon}_t|\mathcal{F}_{t-1}) = H_t, \tag{A.1}$$

where $\mathcal{F}_t$ is the information set at time t. The diagonal elements of H_t, $h_{iit}(= h_{it})$, are the conditional variances of ϵ_{it}; and the non-diagonal elements, h_{ijt}, are the conditional covariance of ϵ_{it} and ϵ_{jt}, where $t = 1, \cdots, T$, $i, j = 1, \cdots, d, i \neq j$.

Diagonal VEC Model (DVEC). It was first proposed by Bollerslev, Engle and Wooldridge (1988). It allows for a flexible and stationary model for time varying covariance in the form

$$H_t = A_0 + \sum_{i=1}^{p} A_i \otimes (\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}'_{t-i}) + \sum_{i=1}^{q} B_j \otimes H_{t-j}, \tag{A.2}$$

where the symbol $\otimes$ stands for the Hadamard product, i.e., element-by-element multiplication. All the coefficient matrices are $d \times d$ symmetric matrices. In this model, the volatility process h_{it} of each series follows a univariate GARCH process, while the covariance process h_{ijt} follows a GARCH model in terms of the product of errors $\epsilon_{it}\epsilon'_{jt}$. The coefficients of these GARCH processes are stored in matrices A_0, A_i and B_j in the corresponding position. Hence each element of the coefficient matrices can be easily interpreted for the GARCH process. A covariance matrix must be positive semi-definite (PSD). However, the H_t in the DVEC model cannot be guaranteed to be PSD even if the lagged H_{t-j} are, which is a weakness of the DVEC model.

Matrix-Diagonal Model. To overcome the weakness of the DVEC model, Ding (1994), and Bollerslev, Engle and Nelson (1994) proposed the *Matrix-Diagonal model* by letting

$$H_t = A_0 A'_0 + \sum_{i=1}^{p} (A_i A'_i) \otimes (\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}'_{t-i}) + \sum_{i=1}^{q} (B_j B'_j) \otimes H_{t-j}, \tag{A.3}$$

where A_0, $A_i(i = 1, \cdots, p)$ and $B_j(j = 1, \cdots, q)$ are all lower triangular matrices. Then the coefficient matrices $A_0 A_0$, $A_i A'_i(i = 1, \cdots, p)$ and $B_j B'_j(j = 1, \cdots, q)$ are positive-definite matrices, and, as a result, the time varying covariance matrix H_t is positive-definite (Zivot and Wang, 2003). Obviously, the Matrix-Diagonal model can be simplified by restricting $A_i(i = 1, \cdots, p)$ and $B_j(j = 1, \cdots, q)$ to be vectors or positive scalars. The Matrix-Diagonal model keeps the DVEC model's main features that the volatility process h_{it} of each series follows a univariate GARCH process, while the covariance process h_{ijt} follows a GARCH model in terms of the product of errors $\epsilon_{it}\epsilon'_{jt}$. The coefficients of the GARCH processes are stored in matrices $A_0 A_0$, $A_i A'_i$ and $B_j B'_j$ in corresponding positions, and the interpretation of the GARCH coefficient matrices, $A_0 A_0$, $A_i A'_i$ and $B_j B'_j$, is clear and simple.

BEKK Model. With the DVEC and Matrix-Diagonal models, the conditional variance and covariance are only dependent on their own lagged element and the corresponding cross-product terms. To allow for inclusion of the variances and covariances corresponding to other component series, Engle and Kroner (1995) formalised an alternative model of the conditional covariance process, denoted the BEKK (Baba, Engle, Kraft and Kroner) model:

$$H_t = A_0 A_0' + \sum_{i-1}^{p} A_i(\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}_{t-i}')A_i' + \sum_{i-1}^{q} B_j H_{t-j} B_j', \tag{A.4}$$

where A_0 is a lower triangular matrix, and $A_i(i = 1, \cdots, p)$ and $B_j(j = 1, \cdots, q)$ are unrestricted square matrices. It can be shown that H_t is guaranteed to be symmetric and positive-definite matrix. The dynamics allowed in the BEKK model are richer than that of the DVEC and Matrix-Diagonal models, and the conditional variance and covariance implied by BEKK are more volatile then that implied by the DVEC and Matrix-Diagonal models. The conditional variance and conditional covariance process are impacted by their lagged process and cross-moments of errors, and other series' lagged process and cross-moments of errors as well. The weakness of the BEKK model is that there are a large number of parameters to be estimated and interpretation of the individual coefficients is more difficult than the DVEC and Matrix-Diagonal models.

Univariate GARCH-based Model: CCC and DCC. The DVEC, Matrix-Diagonal and BEKK models describe the conditional covariance directly, with a large number of parameters to model the variances and covariances together. Another approach is to separate the likelihood function of multivariate GARCH into the components of uncorrelated GARCH (univariate based GARCH) and correlated GARCH, and then apply the univariate GARCH models to each of those uncorrelated series.

A simple model of this approach is the ***Constant Conditional Correlation model*** **(CCC)**, proposed by Bollerslev (1990). It assumes that the conditional correlation matrix is constant over time. The $k \times k$ covariance matrix H_t is decomposed according to the equation:

$$H_t = D_t R D_t,$$

where R is the constant conditional correlation matrix, and D_t is the diagonal matrix with vector $(\sqrt{h_{1t}}, \cdots, \sqrt{h_{dt}})$ on the diagonal, and h_{it} is the conditional variance of the i-th series at time t:

$$D_t = \begin{bmatrix} \sqrt{h_{1t}} & & \\ & \ddots & \\ & & \sqrt{h_{dt}} \end{bmatrix} \tag{A.5}$$

with h_{it} following a univariate GARCH process, for $i = 1, \cdots, d$.

Simply by requiring each univariate conditional variance to be positive and the constant correlation matrix to be of full rank, the CCC model with the assumption of constant conditional correlation makes the estimation of a large model feasible and guarantees the covariance matrix to be positive-definite. However, some literature has shown that observed financial time series do not always exhibit constant conditional correlation. For example, Tsui and Yu (1999) rejected the constant correlation assumption for stock market increments. Fitting the CCC model to Australian yield curves, we have found that the squared standardised residuals of the CCC model are still correlated, and the corresponding ARCH effects remain in the standardised residuals. This implies that the CCC model is not appropriate for Australian Treasury bond yield increments.

Engle and Sheppard (2001) proposed a ***Dynamical Conditional Correlation model*****(DCC)** which preserves the estimation based on univariate GARCH and allows for the correlation matrix to change over time. The problem of multivariate conditional covariance estimation can be achieved by using a two step method. Step one, estimate univariate GARCH models for each asset; and step two, using transformed residuals (referred to as standardised residuals) obtained in step one to estimate a conditional correlation estimator. Let $\boldsymbol{\epsilon}_t$ be the residual process of a d-dimensional process $\boldsymbol{y}_t$, then Engle and Sheppard's DCC model is formulated as follows:

$$\begin{aligned}
&\boldsymbol{\epsilon}_t|\mathcal{F}_t \sim N(0, H_t),\\
&H_t = D_t R_t D_t,\\
&D_t^2 = diag\{\omega_i\} + diag\{\kappa_i\} \otimes \boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' + diag\{\lambda_i\} \otimes D_{t-1}^2,\\
&\boldsymbol{r}_t = D_t^{-1}\boldsymbol{\epsilon}_t,\\
&Q_t = S \otimes (ll' - A - B) + A \otimes \boldsymbol{r}_t\boldsymbol{r}_t' + B \otimes Q_{t-1},\\
&R_t = diag\{Q_t\}^{-1/2} Q_t diag\{Q_t\}^{-1/2},
\end{aligned} \tag{A.6}$$

where D_t is the $d \times d$ diagonal matrix (A.5), with i-th diagonal element $\sqrt{h_{it}}$ being the time varying conditional standard deviation of the i-th series. R_t is the time varying correlation matrix, with diagonal elements being ones and (i, j)-th element being the time varying conditional correlation of i-th series and j-th series. The third equation of (A.6) simply expresses the assumption that conditional variance of each series follows a univariate GARCH process. The vector of standardised residuals, denoted by $\boldsymbol{r}_t$, is given by the fourth equation of (A.6). The fifth equation of (A.6) expresses the dynamics of correlation that depend on its own lagged element and the corresponding cross-product of errors only, where matrix S is the unconditional covariance of the standardised residuals and l is a vector of ones. The last equation is used to ensure that R_t is a correlation matrix. In theory, the covariance matrix of standardised residuals should be equal to their correlation matrix, however this is normally not satisfied on a sample basis estimation process. It is claimed that Q_t will be PSD if A, B and (ll'-A-B) are PSD (Engle 2002). The parameters of D_t, ω_i, κ_i and λ_i, specify the time evolution of the univariate GARCH process of the i-th series. The matrices A and B are the parameters describing the correlation process R_t.

Engle (2002) developed estimation of the DCC model in the special case where the matrices A and B are simply a scalar or diagonal rather than a whole matrix. This assumes that the conditional correlations all follow the same dynamic structure, which may be inappropriate in cases where there are many asset increments. Hafner and Franses (2003) Generalised the estimation of the DCC model to the case where A equals $\alpha\alpha'$ (α is a d-dimensional vector) and B is a scalar.

A.2 Matrix-Diagonal GARCH-t Model of Australian Treasury Yield Increment

The Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model for Australian bond yield increments is

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \cdot R_t^- + \boldsymbol{\beta}_3 \cdot R_t^+ + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \boldsymbol{t}_m(0, H_t), \\ H_t &= A_0A_0' + (AA') \otimes (\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + (BB') \otimes H_{t-1} + \boldsymbol{c}\boldsymbol{c}' \cdot R_t^-, \end{aligned} \tag{A.7}$$

where $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$ and $\boldsymbol{c}$ are parameter vectors. A_0, A and B are lower triangular parameter matrices.

The result of estimation from S+FinMetrics are: the estimated parameter $m = 5.99$ with standard error 0.429, and the log likelihood value is 19644. It is observed that all elements except column 1 of B are not significant. By reducing the parameter matrix B to a 1-dimensional vector $\boldsymbol{b}$, the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model for Australian bond yield increments is

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \cdot R_t^- + \boldsymbol{\beta}_3 \cdot R_t^+ + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \boldsymbol{t}_m(0, H_t), \\ H_t &= A_0A_0' + (AA') \otimes (\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + (\boldsymbol{b}\boldsymbol{b}') \otimes H_{t-1} + \boldsymbol{c}\boldsymbol{c}' \cdot R_t^-, \end{aligned} \tag{A.8}$$

where $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, $\boldsymbol{c}$ and $\boldsymbol{b}$ are parameter vectors. A_0 and A are lower triangular parameter matrices.

The Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model A.8 is the same as the Matrix-Diagonal GARCH(1,1)-AMt (6.5) under the assumption of the same degrees of freedom of marginals, $m = m_1 = m_2 = \cdots = m_d$. Estimate the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model A.8, the results are: the estimated parameter $m = 6.233582$ with standard error 0.429236, and the log likelihood value is 19748. The estimators of the coefficients are displayed in Table A.1. Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model (A.8) showed a marked improvement that has less parameters and larger log likelihood value.

Table A.1: Outputs of the Matrix-Diagonal GARCH(1,1)-$\boldsymbol{t}$ model

	Value	Std.Error	t value	Pr($>|t|$)
$\boldsymbol{\beta}_1(1)$	-0.002304	0.0012048	-1.9123	0.028
$\boldsymbol{\beta}_1(2)$	-0.002795	0.0014744	-1.8954	0.029
$\boldsymbol{\beta}_1(3)$	-0.002805	0.0015361	-1.8260	0.034
$\boldsymbol{\beta}_1(4)$	-0.002906	0.0015501	-1.8748	0.030
$\boldsymbol{\beta}_1(5)$	-0.003047	0.0016175	-1.8840	0.029
$\boldsymbol{\beta}_2(1)$	-0.077925	0.0253436	-3.0747	0.001
$\boldsymbol{\beta}_2(2)$	-0.061143	0.0295956	-2.0660	0.019
$\boldsymbol{\beta}_2(3)$	-0.058135	0.0285033	-2.0396	0.020
$\boldsymbol{\beta}_2(4)$	-0.055957	0.0287820	-1.9442	0.026
$\boldsymbol{\beta}_2(5)$	-0.052374	0.0251526	-2.0823	0.018
$\boldsymbol{\beta}_3(1)$	0.081739	0.0483859	1.6893	0.045
$\boldsymbol{\beta}_3(2)$	0.072150	0.0596670	1.2092	0.011
$\boldsymbol{\beta}_3(3)$	0.059969	0.0604605	0.9919	0.160
$\boldsymbol{\beta}_3(4)$	0.043436	0.0658144	0.6600	0.254
$\boldsymbol{\beta}_3(5)$	0.020096	0.0879205	0.2286	0.409
$A_0(1,1)$	0.020800	0.0012636	16.4610	0.000
$A_0(2,1)$	0.022560	0.0013982	16.1351	0.000
$A_0(3,1)$	0.023111	0.0014374	16.0781	0.000
$A_0(4,1)$	0.022852	0.0015433	14.8067	0.000
$A_0(5,1)$	0.021207	0.0019560	10.8421	0.000
$A_0(2,2)$	0.004980	0.0006785	7.3399	0.000
$A_0(3,2)$	0.004842	0.0008204	5.9014	0.000
$A_0(4,2)$	0.005329	0.0010295	5.1761	0.000
$A_0(5,2)$	0.005339	0.0017764	3.0054	0.001
$A_0(3,3)$	0.002536	0.0003002	8.4487	0.000
$A_0(4,3)$	0.003117	0.0004558	6.8386	0.000
$A_0(5,3)$	0.004003	0.0012285	3.2582	0.000
$A_0(4,4)$	0.003639	0.0003639	10.0014	0.000
$A_0(5,4)$	0.006721	0.0011159	6.0231	0.000
$A_0(5,5)$	0.006368	0.0009837	6.4738	0.000
$A(1,1)$	0.255469	0.0164353	15.5440	0.000
$A(2,1)$	0.237497	0.0150649	15.7649	0.000
$A(3,1)$	0.235170	0.0152702	15.4006	0.000
$A(4,1)$	0.226176	0.0154043	14.6826	0.000
$A(5,1)$	0.199620	0.0187390	10.6526	0.000
$A(2,2)$	0.050221	0.0078864	6.3681	0.000
$A(3,2)$	0.055298	0.0096606	5.7240	0.000
$A(4,2)$	0.047822	0.0119927	3.9876	0.000
$A(5,2)$	0.040133	0.0208356	1.9262	0.027
$A(3,3)$	0.029558	0.0042247	6.9966	0.000
$A(4,3)$	0.01816	0.007189	2.527	0.005
$A(5,3)$	0.02516	0.020096	1.252	0.105
$A(4,4)$	0.02062	0.008247	2.500	0.006
$A(5,4)$	0.04534	0.024759	1.831	0.033
$A(5,5)$	0.04573	0.010981	4.164	0.000
$\boldsymbol{b}(1)$	0.88523	0.011351	77.985	0.000
$\boldsymbol{b}(2)$	0.90313	0.009863	91.566	0.000
$\boldsymbol{b}(3)$	0.90481	0.009784	92.481	0.000
$\boldsymbol{b}(4)$	0.90828	0.010821	83.935	0.000
$\boldsymbol{b}(5)$	0.91842	0.015056	61.001	0.000
$\boldsymbol{c}(1)$	0.06209	0.011622	5.342	0.000
$\boldsymbol{c}(2)$	0.05762	0.014147	4.073	0.000
$\boldsymbol{c}(3)$	0.05472	0.014941	3.662	0.000
$\boldsymbol{c}(4)$	0.04329	0.015313	2.827	0.002
$\boldsymbol{c}(5)$	0.02257	0.017463	1.292	0.098

Appendix B

Girsanov Theorem and Itô Lemma

We recall the Girsanov's Theorem, which allows changing the probability measure of a given Itô process by changing its drift.

Theorem B.1 *(Girsanov) Let W_t be a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, P)$. Suppose that $\varphi(t)$ is an adapted real-valued stochastic process such that*

$$E_P[\rho(t)] = 1$$

where

$$\rho(t) = exp\left\{\int_0^t \varphi(s)dW_s - \frac{1}{2}\int_0^t \varphi(s)^2 ds\right\}$$

Define a probability measure $\tilde{P}$ on $(\Omega, \mathcal{F})$ equivalent to P by means of the Radon-Nikodym derivative

$$\frac{d\tilde{P}}{dP} = \rho(t), \quad P - as$$

Then the process $\tilde{W}$, which is given by the formula

$$\tilde{W}_t = W_t - \int_0^t \varphi(s)ds, \quad \forall t \in [0, T]$$

follows a standard Bronwian motion on the space $(\Omega, \mathcal{F}, \tilde{P})$.

See Dothan (199) or Musiela and Rutkowski (1997).

The Itô Lemma is a very useful tool and widely used in finance pricing. It shows how to derive the stochastic dynamic process $f(t,x)$, while x described by a stochastic differential equation.

Lemma B.1 *(Itô) Suppose a stochastic process x follows an Itô process*

$$dx = \mu(t,\omega)dt + \sigma(t,\omega)dW$$

Then any function $f(\in C^2)$ of the process x and time t satisfies

$$df = \left[\frac{\partial f(t,x)}{\partial t} + \frac{\partial f(t,x)}{\partial x}\mu(t,\omega) + \frac{1}{2}\frac{\partial f^2(t,x)}{\partial x^2}\sigma^2(t,\omega)\right]dt + \frac{\partial f(t,x)}{\partial x}\sigma(t,\omega)dW$$

provided the function is sufficiently differentiable.

See Dothan (1990).

Appendix C

Swap and Constructed Yield Curve of the Commonwealth Bank Australia

C.1 Swap, Swap Rates and Valuation of Swaps

Swaps are popular interest rate derivatives that are important for the modelling of interest rates, for the estimates of interest rates (or yield) can be derived from the market prices of swaps. A method to construct a yield curve is introduced in Chapter 4.

A *Swap* is an agreement between two companies to exchange cash flows in the future. The agreement defines the dates when the cash flows are to be paid and the way in which they are to be calculated. The most common type of swap is a *'plain vanilla' interest rate swap*. In this contract, a company agrees to pay cash flows equal to interest at a predetermined *fixed rate* on a *notional principal* for a number of years. In return, it receives interest at a *floating rate* on the same notional principal for the same period of time. The principal itself is not exchanged, so it is termed the notional principal. In Europe, the floating rate in many interest rate swap agreements is the *London Interbank Offer Rate* (LIBOR). The floating rate is usually set at the beginning of each period to which it will apply and is paid at the end of the period. We say a swap is settled *in arrears* if the payment is made at the end of each period. A swap is settled *in advance* if payments are made at the beginning of each period. In that case, the payments are discounted to the beginning of each period to correspond to a swap settled in arrears.

A payer swap is the *fixed-for-floating swap* when the fixed rate is paid and the floating rate is received. Similarly, *a receiver swap* is the *floating-for-fixed swap* when the floating rate is paid and the fixed rate is received.

Let us consider a collection of future dates: $T_0 = T < T_1 < \ldots < T_n$. A *forward interest rate swap* is a swap agreement entered at trade date t , $t \leq T_0 = T < T_1 < \ldots < T_n$. T_0 is the start date of the swap, and $T_1, \ldots, T_n$ are settlement (payment) dates. Let δ_j , $j = 1, \ldots, n$ be the number of years of the $j-th$ *accrual period* $[T_{j-1}, T_j]$. Consider the *forward payer swap settled in arrears.* The floating rate $R(T_j, T_{j+1})$ received at time T_{j+1} is set at time T_j by reference to the price of zero-coupon bond over that period, so $R(T_j, T_{j+1})$ is treated as the zero rate (or yield-to maturity) over the time interval $[T_j, T_{j+1}]$, and satisfies

$$B(T_j, T_{j+1})^{-1} = 1 + \delta_j R(T_j, T_{j+1}),$$

which agrees with the market LIBOR.

More generally, the forward rate $F(t, T_j, T_{j+1})$ satisfies

$$1 + \delta_{j+1} \cdot F(t, T_j, T_{j+1}) = \frac{B(t, T_j)}{B(t, T_{j+1})},$$
$$R(T_j, T_{j+1}) = F(T_j, T_j, T_{j+1}).$$

Stipulated by the no-arbitrage theory, the value of a swap is zero at the initial time t. But it may become positive or negative after it has been in existence for some time. A swap contract is equivalent to a long position in one bond combined with a short position in another bond or to a portfolio of forward rate agreements. In either case we use the zero rate for discounting. Define:

- B_{fix}: value of fixed-rate bond underlying the swap,
- B_{fl}: value of floating-rate bond underlying the swap.

The value of the payer swap is

$$V_{swap} = B_{fl} - B_{fix}. \tag{C.1}$$

Without loss of generality, assume that the principal is $N = 1$. At date T_j, $j = 1, \ldots, n$, the cash flows of a payer swap are $F(t, T_{j-1}, T_j)\delta_j$ and $-K\delta_j$. The value at time

t of a forward payer swap with discounting factors $B(t, T_j)$, $j = 1, \ldots, n$, is given by

$$\begin{aligned}
V_{FS} &= \sum_{j=1}^{n} B(t, T_j) \{F(t, T_{j-1}, T_j)\delta_j - K\delta_j\} \\
&= \sum_{j=1}^{n} B(t, T_j) \left\{ \frac{B(t, T_{j-1})}{B(t, T_j)} - 1 - K\delta_j \right\} \\
&= \sum_{j=1}^{n} \{B(t, T_{j-1}) - B(t, T_j) - K\delta_j B(t, T_j)\} \qquad \text{(C.2)} \\
&= B(t, T_0) - B(t, T_n) - \sum_{j=1}^{n} K\delta_j B(t, T_j),
\end{aligned}$$

and after rearranging,

$$V_{FS} = B(t, T_0) - \sum_{j=1}^{n-1} K\delta_j B(t, T_j) - (1 + K\delta_n)B(t, T_n). \qquad \text{(C.3)}$$

Equation (C.3) shows that a forward payer swap settled in arrears is a contract to deliver a coupon-bearing bond and to receive in the same time t a zero-coupon bond. It can be described by a replicating strategy (portfolio) as follows: at time t, $t \leq T_0 = T < T_1 < \ldots, T_n$,

- Sell 1 zero-coupon bond with maturity T_0,
- Buy a coupon-bearing bond which pays holder the amount $K\delta_j$ (fixed interest) at $T_1, \ldots, T_{n-1}$ and (interest plus principal) $1 + K\delta_n$at date T_n, where K is the fixed rate which makes the value of the portfolio zero at time t.

Similarly, for any notional principal N, the value of the forward swap is

$$V_{FS} = N \cdot \left\{ B(t, T_0) - \sum_{j=1}^{n-1} K\delta_j B(t, T_j) - (1 + K\delta_n B(t, T_n)) \right\}.$$

For forward swaps settled in advance, payments will be made at the beginning of each period; these payments are discounted back to the beginning of each period corresponding the swap settled in arrears. The discount methods vary from country to country. The discounting factors, for the cash flow at each payment date T_j, for both fixed or floating parts are the same, which is $1/[1 + F(t, T_j, T_{j+1})\delta_{j+1}]$. The value at time t of the forward swap pay in advance is

$$\begin{aligned}
V_{FS}^{*} &= \sum_{j=0}^{n-1} B(t,T_j)\frac{F(t,T_j,T_{j+1})\delta_{j+1} - K\delta_{j+1}}{1+F(t,T_j,T_{j+1})\delta_{j+1}} \\
&= \sum_{j=0}^{n-1} B(t,T_j)\frac{B(t,T_j)/B(t,T_{j+1}) - 1 - K\delta_{j+1}}{B(t,T_j)/B(t,T_{j+1})} \\
&= \sum_{j=0}^{n-1} \{B(t,T_j) - B(t,T_{j+1}) - K\delta_{j+1}B(t,T_{j+1}\} \\
&= \sum_{j=1}^{n} \{B(t,T_{j-1}) - B(t,T_j) - K\delta_j B(t,T_j)\},
\end{aligned} \tag{C.4}$$

which exactly agrees with the value of the forward swap settled in arrears, in (C.2).

In the Australian market, the floating part is discounted by the floating rate and the fixed part is discounted by the fixed rate K. The cash flows at each payment date T_j are

$$\frac{F(t,T_j,T_{j+1})\delta_{j+1}}{1+F(t,T_j,T_{j+1})\delta_{j+1}}$$

and

$$-\frac{K\delta_{j+1}}{1+K\delta_{j+1}}.$$

The value at time t of such a forward swap settled in advance is

$$\begin{aligned}
V_{FS}^{**} &= \sum_{j=0}^{n-1} B(t,T_j)\left\{\frac{F(t,T_j,T_{j+1})\delta_{j+1}}{1+F(t,T_j,T_{j+1})\delta_{j+1}} - \frac{K\delta_{j+1}}{1+K\delta_{j+1}}\right\} \\
&= \sum_{j=0}^{n-1} B(t,T_j)\left\{\frac{B(t,T_j)/B(t,T_{j+1}) - 1}{B(t,T_j)/B(t,T_{j+1})} - \frac{K\delta_{j+1}}{1+K\delta_{j+1}}\right\} \\
&= \sum_{j=0}^{n-1} \left\{B(t,T_j) - B(t,T_{j+1}) - \frac{K\delta_{j+1}}{1+K\delta_{j+1}}B(t,T_j)\right\} \\
&= \sum_{j=0}^{n-1} \left\{\frac{1}{1+K\delta_{j+1}}B(t,T_j) - B(t,T_{j+1})\right\} \\
&= \sum_{j=1}^{n} \left\{\frac{1}{1+K\delta_j}B(t,T_{j-1}) - B(t,T_j)\right\},
\end{aligned} \tag{C.5}$$

which is the value of the forward payer swap settled in arrears discounting at the fixed rate K, because

$$V_{FS}^{**} = \sum_{j=1}^{n} \frac{B(t,T_j)}{1+K\delta_j} \left\{ \frac{B(t,T_{j-1})}{B(t,T_j)} - (1+K\delta_j) \right\}. \tag{C.6}$$

From now on, we restrict our attention to the forward interest rate swap settled in arrears.

The *forward swap rate* $K(t,T,T_n)$ is the value of the fixed rate K which makes the value of the forward swap at initiation zero. i.e. the value of K for which $V_{FS}(t) = 0$. By (C.3),

$$K(t,T,T_n) = \frac{B(t,T) - B(t,T_n)}{\sum_{j=1}^{n} \delta_j B(t,T_j)}. \tag{C.7}$$

A *swap* is the forward swap with $t = T$. The *swap rate* is the forward swap rate with $t = T$, and

$$K(T,T,T_n) = \frac{1 - B(T,T_n)}{\sum_{j=1}^{n} \delta_j B(T,T_j)}. \tag{C.8}$$

Forward swap rates and swap rates are quoted daily by the financial institutions that offer the interest rate swap contract. Practically, the average of bid and offer fixed rates in the forward swap (swap) market is called the forward swap rate (swap rate), respectively.

C.2 Money Market: Constructed Yield Curve of the Commonwealth Bank Australia (CBA)

Term structures of interest rates are usually derived from bond prices. However, the bond market in Australia is underdeveloped, because only 10 year 'benchmark' exists in the Australian bond market. An alternative approach of estimating yield curves is used by the Commonwealth Bank Australia, which uses bank notes to derive yield curves. Under the generally accepted assumption of no-arbitrage, the yield curves obtained at CBA should be consistent with those produced by RBA as introduced in the last section.

The method used in this section is based on the valuation of swaps discussed in Section 2.2.2. We will show how the yield curve is derived from those bank's notes, in particular, how forward rates are deduced from swap rates. We will consider the interest rate swap settled in arrears.

From (C.8), we have

$$K(T,T,T_n) = \frac{1 - B(T,T_n)}{\sum_{j=1}^{n} \delta_j B(T,T_j)},$$

and therefore solving for bond price $B(T, T_n)$ (or called *discount factor*)

$$B(T, T_n) = \frac{1 - K \sum_{j=1}^{n-1} \delta_j B(T, T_j)}{1 + K\delta_n}, \tag{C.9}$$

the yield-to-maturity $R(T, T_n)$ is obtained by

$$R(T, T_n) = \frac{B^{-1}(T, T_n) - 1}{T_n - T}. \tag{C.10}$$

The procedure to construct the yield curve from an interest rate swap settled in arrears uses the following steps.

Step a) Calculate the discount factor using the available market short-term bill rates.

Step b) Price the synthetic swap rates by formula (C.8).

Step c) Linearly interpolate intermediate swap rates using the synthetic swap rates and already known long term swap rates.

Step d) Calculate corresponded discounting factors from swap rates by formula (C.9).

Step e) Find the yield-to-maturity from formula (C.10).

We test this approach on small data sets and show how it works in an Excel template in Appendix B, which is called *Yield Curve Builder* by the RBA.

The forward interest rate $F(t, T, T_n)$ can be similarly constructed by using the forward swap rate $K(t, T, T_n)$. The discounting factors $B(t, T_n)$ can be solved from (C.7) to give

$$B(t, T_n) = \frac{B(t, T) - K \sum_{j=1}^{n-1} \delta_j B(t, T_j)}{1 + K\delta_n}, \tag{C.11}$$

and forward interest rate is obtained by

$$F(t, T, T_n) = \frac{1}{T_n - t} \left\{ \frac{B(t, T)}{B(t, T_n)} - 1 \right\}.$$

C.3 An Example: "Yield Curve Builder"

As an example, we show a yield curve builder playing with inputs using the Excel spread sheet.

INPUTS

Bill

	Dates	Bill Rates
reference date	24-Jan-02	
1 day	25-Jan-02	4.3100 %
1 month	24-Feb-02	4.3300 %
2 months	24-Mar-02	4.3583 %
3 months	24-Apr-02	4.3617 %
6 months	24-Jul-02	4.3915 %

Swaps

	Frequency	Swap Rates
1 year	4	5.220 %
3 years	4	5.595 %
4 years	2	5.835 %
5 years	2	5.970 %

OUTPUTS: Yield curve builder

Step a) discounting factors $B(t, T_i)$ (from the bill rates)

	Dates	Bill Rates	discounting factors
reference date	24-Jan-02		1
1 day	25-Jan-02	4.3100 %	**0.99988193**
1 month	24-Feb-02	4.3300 %	**0.99633594**
2 months	24-Mar-02	4.3583 %	**0.99300436**
3 months	24-Apr-02	4.3617 %	**0.98935956**
6 months	24-Jul-02	4.3915 %	**0.97868710**

step a)

Step b) Pricing of 6 months Synthetic Swap by (C.8)

Principal	1,000,000.00
Freq	2
Fixed Rate	**4.367816 %**

step b)

Step c) Linear interpolate swap rates

Swaps	Maturity Date	Swap Fixed Rates	Linear Interpolated Fixed Swap rate	Swap Rates
6 months	24-Jul-02	4.367816 %		4.3678 %
9 months	24-Oct-02	N/A	**4.79391 %**	4.7939 %
1 year	24-Jan-03	5.22000 %		5.2200 %

step c)

Step d, e) calculate discounting factors by (C.9), and yield by (C.10)

Bill Rates	Dates	Bill Rates	Length of period	Discount factor
reference date	24-Jan-02		0	1
3 months	24-Apr-02	4.3617 %	0.246575342	0.98935956
6 months	24-Jul-02	4.3915 %	0.495890411	0.97868710
9 months	24-Oct-02	**5.6711** %	0.747945205	**0.95930943**
1year	24-Jan-03	**6.0570** %	1	**0.94289155**
		step e)		step d)

Results: Yield Curve Builded

	Dates to Maturities	Yield Rates
1 day	25-Jan-02	4.3100 %
1 month	24-Feb-02	4.3300 %
2 months	24-Mar-02	4.3583 %
3 months	24-Apr-02	4.3617 %
6 months	24-Jul-02	4.3915 %
9 months	24-Oct-02	5.6710 %
1year	24-Jan-03	6.0567 %

This sample is to construct the yield curve, input 3 months, 6 months zero rates and 1 year fixed swap rate, output 9 month, 12 month zero rates. By the same procedure, we can construct 9 months, 12 months, 15 months, 18 months, 24 months zero rates if input 3 months, 6 months zero rates and 2 years swap rate. For the accuracy of the construction, we normally try to use short-term swap rate and linear interpolation for other swap rates.

C.4 CBA Yields–Compared to the RBA Yields

This section is based on CBA yields and the RBA yields maturities in 1 month, 3 months, 3 years, 5 years and 10 years, from 02 Jan 1996 to 13 Jan 1999. Time series plots, see Figure C.1, show that the yield curves of CBA have very similar trends to the RBA's. Typically, CBA yields are slightly higher than the RBA yields. Autocorrelation function plots of yield series (Figure C.2) show that each yield series with a fixed maturity are highly autocorrelated.

Figure C.3 shows short-term bill yield increments of CBA vs. RBA, and Figure C.4 shows middle-to-long-term bond yield increments of CBA vs. RBA.

Figure C.5 shows that the autocorrelation functions (ACF) of the CBA yield increments are similar to those of the RBA yield increments. Both for the CBA and the RBA series, yield increments and squared yield increments of short-term bill are uncorrelated. For the middle-to-long-term bonds, yield increments are serially un-autocorrelated, but squared yield increments are autocorrelated. This means that the middle-to-long-term bond yield increments exhibit volatility.

Figure C.6 shows that the quantile-quantile (QQ) normal plots (Lee 1995) of CBA yield increments are similar to those for RBA yield increments at each individual maturity level. It is also clear from this figure that the yield increments are not normally distributed.

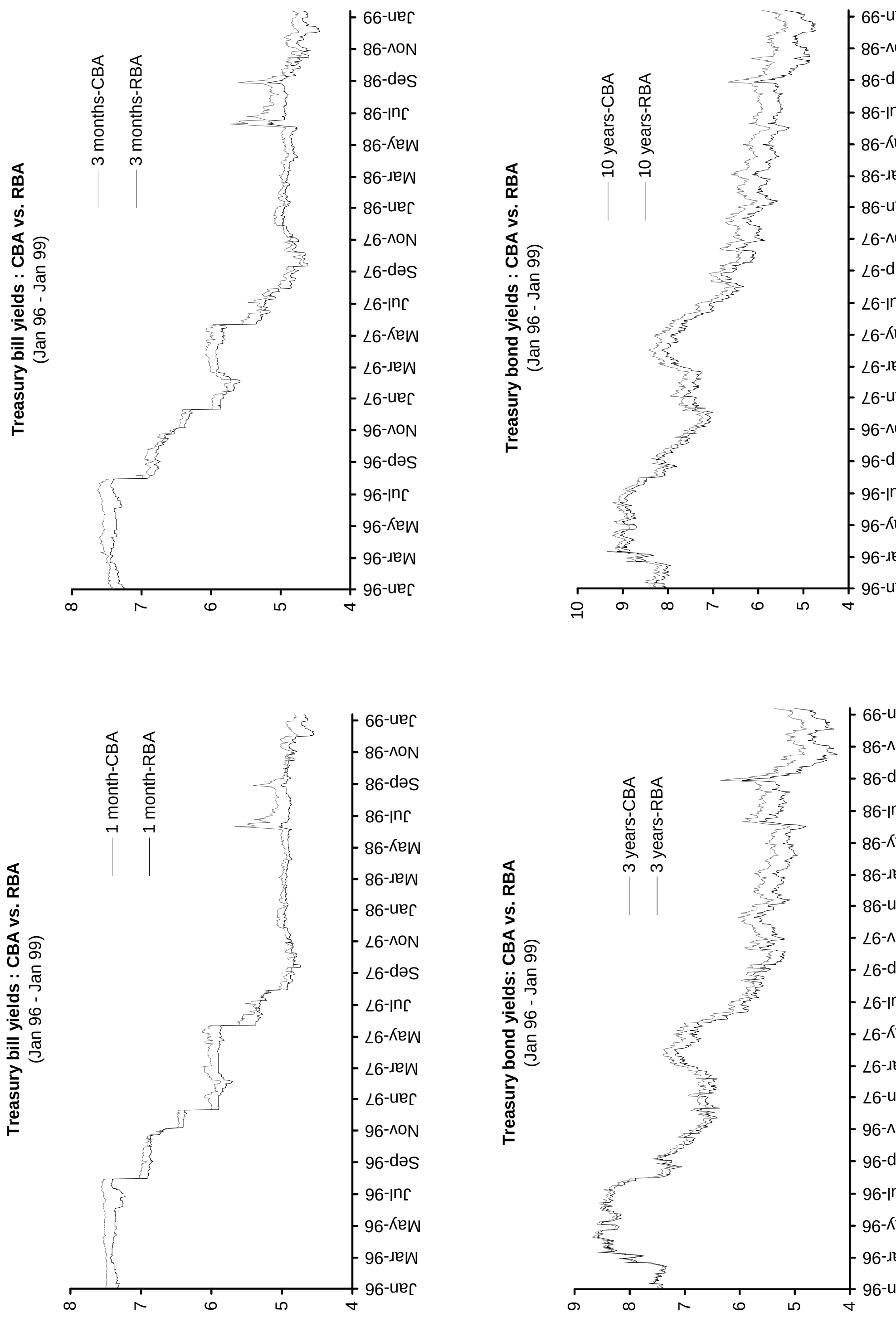


Figure C.1: CBA/RBA yield curves 1996-2004

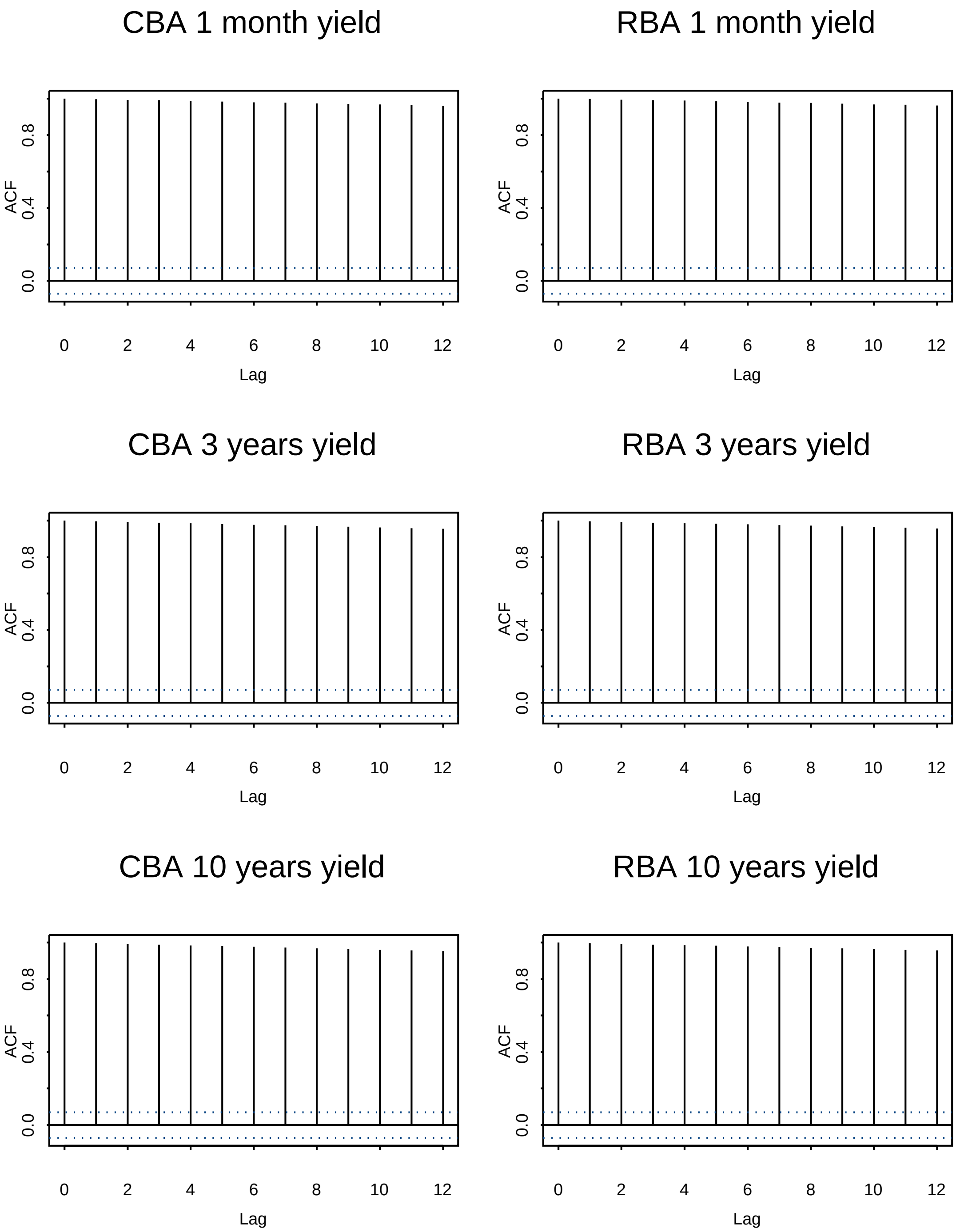


Figure C.2: ACF of CBA/RBA yield (1 month, 3 years and 10 years).

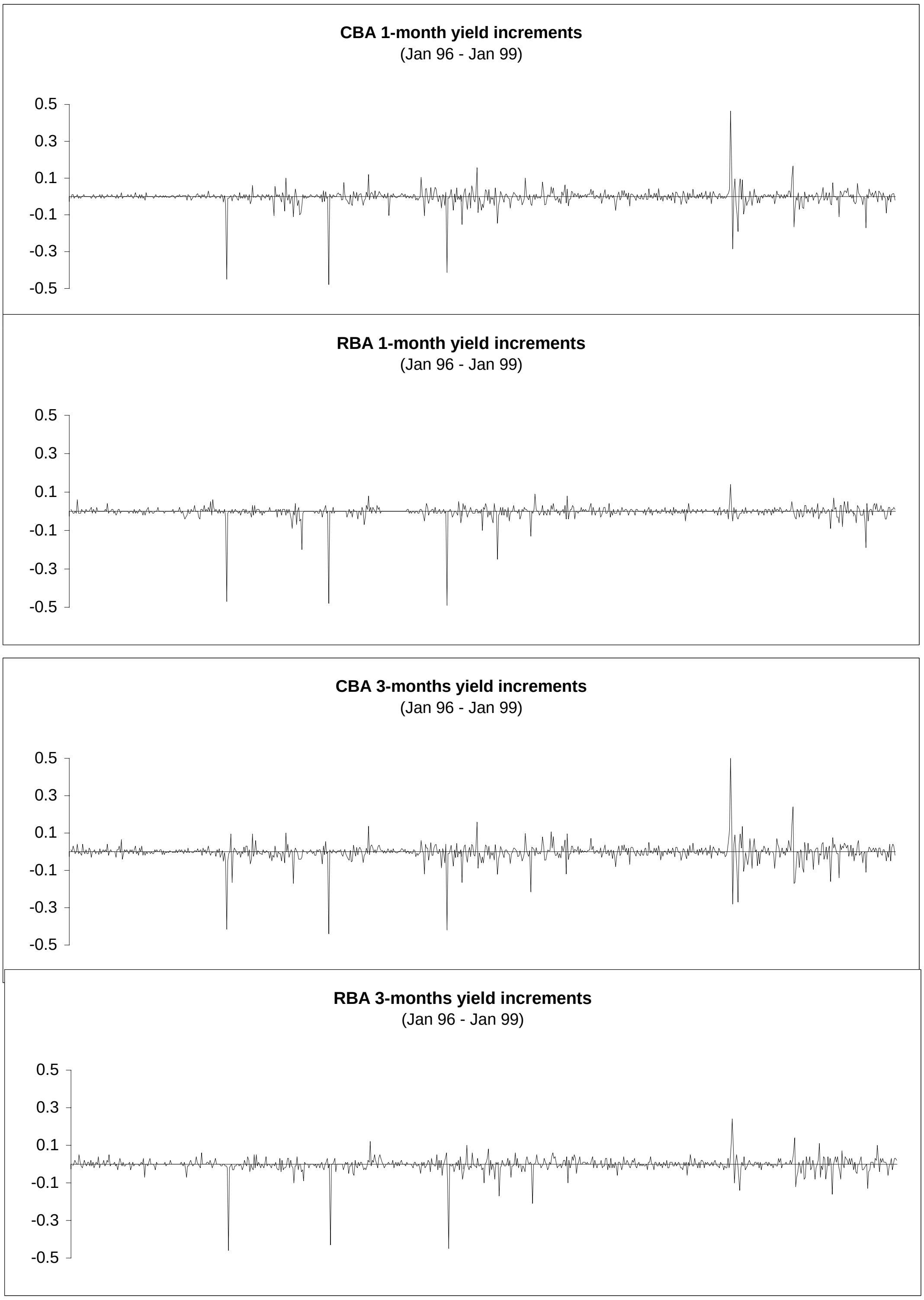


Figure C.3: CBA/RBA short-term bill yield increments

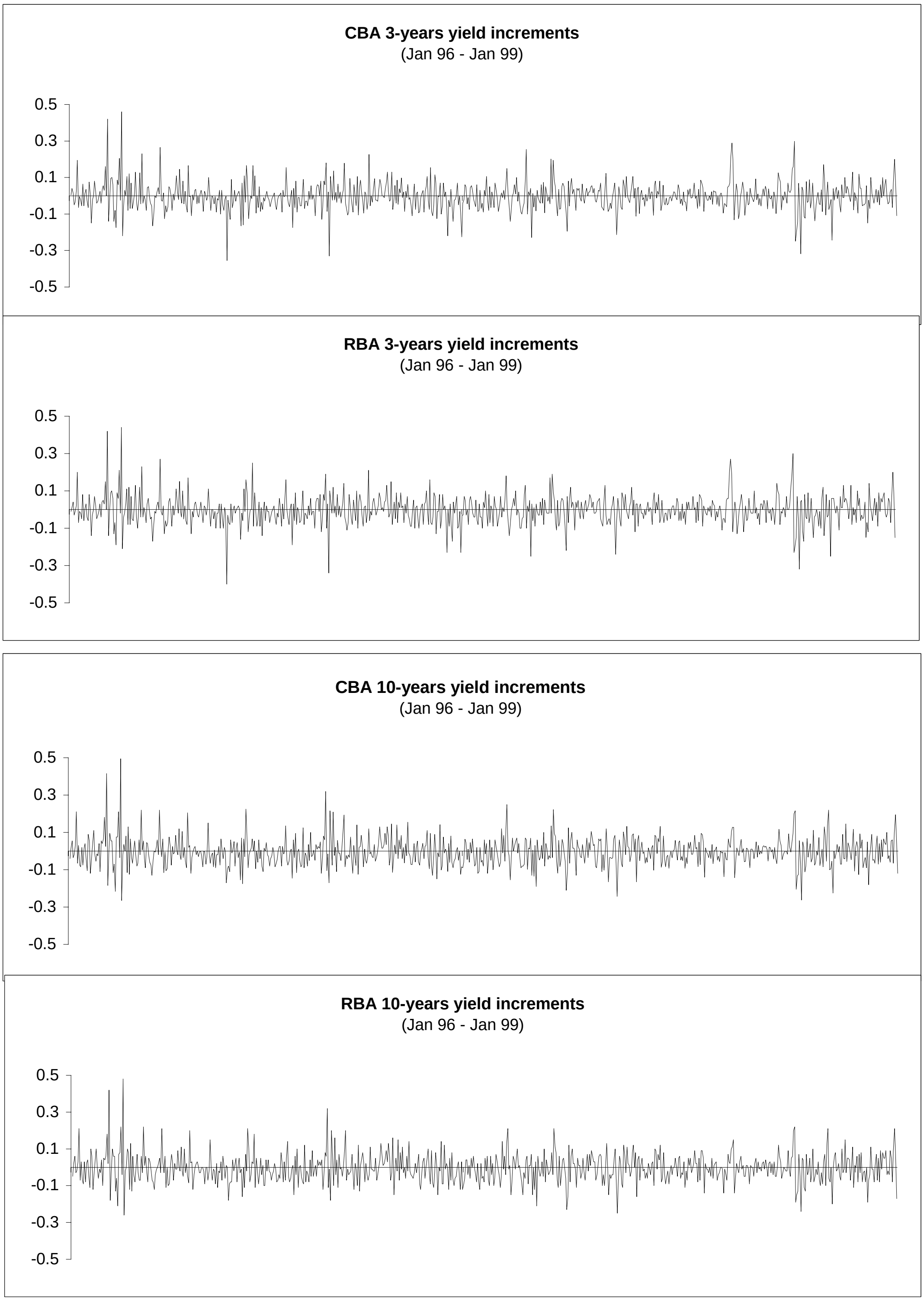


Figure C.4: CBA/RBA middle-to-long-term bond yield increments

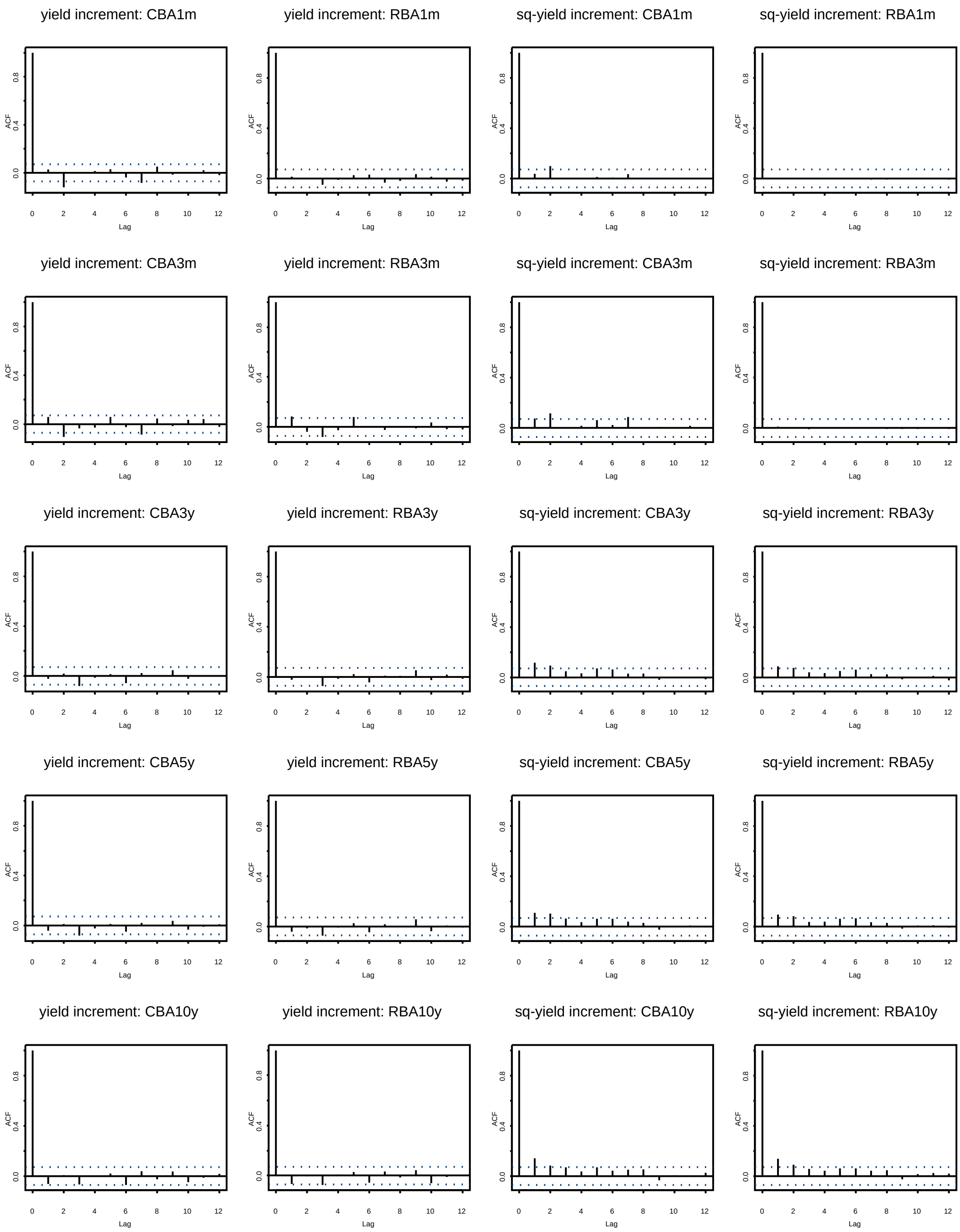


Figure C.5: ACF of CBA/RBA yield increments and squared yield increments.

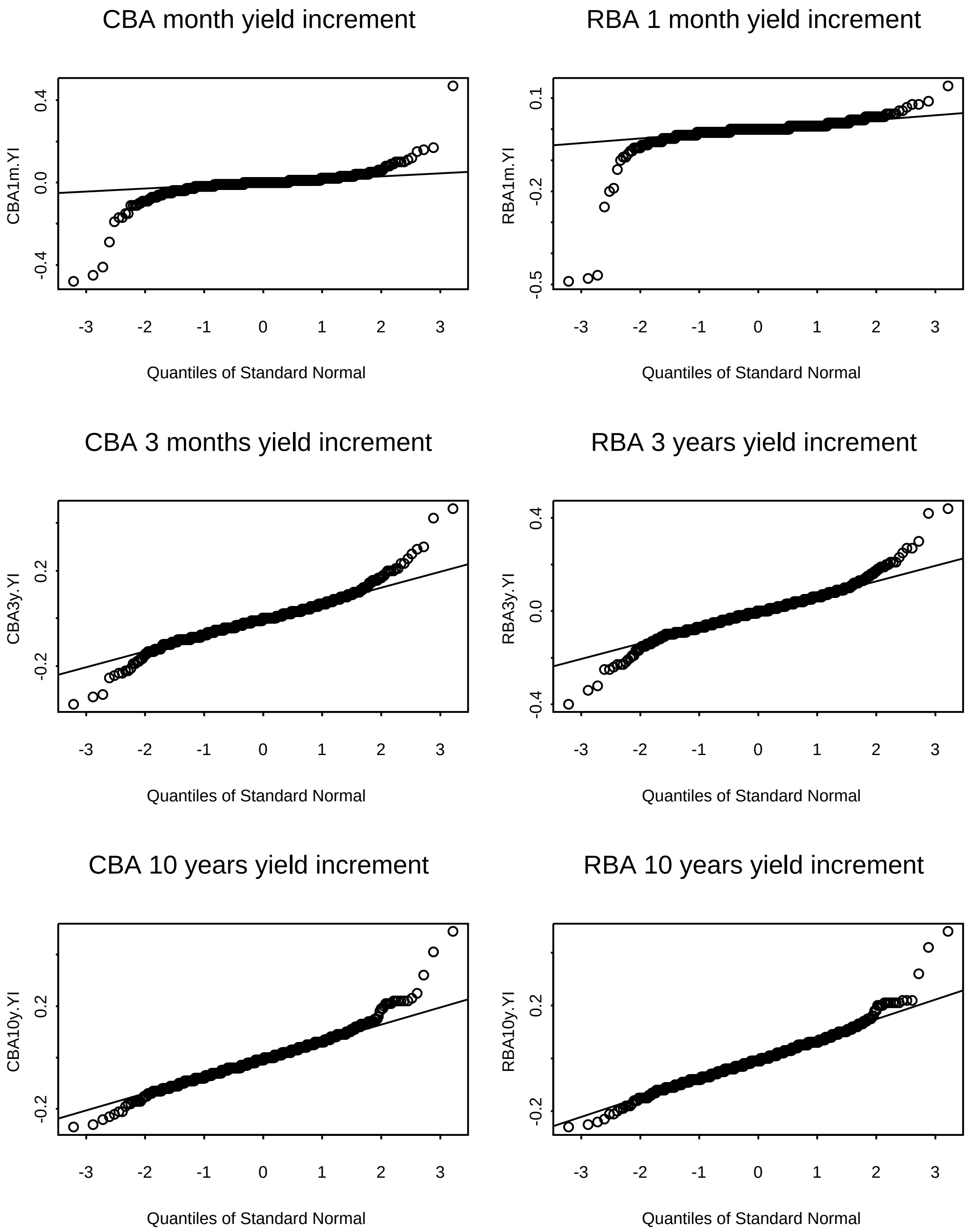


Figure C.6: QQ normal plots of CBA/RBA yield increments

Appendix D

Acronyms

ARCH	AutoRegressive Conditional Heteroskedasticity
BGM	Brace, Gatarek and Musiela
CBA	The Commonwealth Bank of Australia
CCC	Constant Conditional Correlation
CIR	Cox, Ingersoll and Ross
DCC	Dynamic Conditional Correlation
DVEC	Diagonal VEC
EBKK	Baba, Y., Engle, R. F., Kraft, D. F. and K. F. Kroner
FIGARCH	Fractionally Integrated GARCH
GARCH	Generalised AutoRegressive Conditional Heteroskedasticity
HJM	Heath-Jarrow-Morton, 1992
IGARCH	Integrated GARCH
RBA	The Reserve Bank of Australia

Bibliography

[1] Ait-Sahalia, Y. (1996a). "Nonparametric Pricing of Interest Rate Derivative Securities", *Econometrica*, vol. 64, No.3, 527-560.

[2] Ait-Sahalia, Y. (1996b). "Testing Continuous-Time Models of the Spot Interest Rate", *The Review of Financial Studies*, Vol.9, No.2, 385-426.

[3] Baba, Y., Engle, R. F., Kraft, D. F. and Kroner, K. F. (1991). "Multivariate Simultaneous Generalized ARCH", University of California, San Diego: Department of Economics, Discussion Paper No. 89-57.

[4] Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). "Fractionally Intergrated Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometerics* 74, 3-30.

[5] Black, F. and Scholes, M. (1973). "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-659.

[6] Black, F. (1976) "Studies of stock price volatility changes", *Proceedings of the 1976 Meetings of the Business and Economics Statistics Section*, 177-181, American Statistical association.

[7] Bollerslev, T., Engle, R. and Wooldridge, J. M., (1988). "A Capital Asset Pricing Model with Time Varying Covariances", *Journal of Political Economy*, 96, 116-131

[8] Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, 307-327.

[9] Bollerslev, T. (1990). "Modelling the Coherence in Short Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model", *1992 Review of Economics and Statistics* 72, 3, 498 - 505.

[10] Bollerslev, T., Engle, R., and Nelson, D. (1994). “ARCH Models”, in R. Eggle and D. McFadden (eds) *Handbook of Econometrics*, Vol. 4.

[11] Bouchaud, J-P., Cont, R., El-Karoui, N., Potters, M. and Sagna, N (1997). “Phenomenology of the Interest Rate Curve: a statistical analysis of term structure deformations”, Working papar.

[12] Boussama, F. (1998). “Ergodicity, Mixing and Estimation in GARCH Models”, Ph.D Dissertation, University of Paris 7.

[13] Brace, A., Gatarek, D. and Musiela, M. (1997). “The Market Model of Interest Rate Dynamics”, *Mathematical Finance* 7, No. 2, 127-147.

[14] Brace, A. and Musiela, M. (1994). “A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton”, *Mathematical Finance* 4(3), 259-283.

[15] Brennan, M. J. and Schwartz E. S. (1980). “Analyzing Convertible Bonds”, *Journal of Financial and Quantitative Analysis* 15: 907-929.

[16] Brockwell, P. and Davis, R. (2002). “Introduction to Time Series and Forecasting”, 2nd. Ed., *Springer*, 36.

[17] Chan, K. C., Karolyi, A., Longstaff, F. A. and Sanders B. S. (1992). ”An Empirical Comparison of Alternative Models of Short-Term Interest Rate”, *Journal of Finance*, 47, 1209-1227

[18] Comte, F. and Lieberman, O. (2003). “Asymptotic theory for multivariate GARCH processes”, *Journal of Multivariate Analysis*, 84, 61-68

[19] Cont, R. (1998). “Modeling Term Structure Dynamics: an infinite dimensional approach”, *International Journal of Theoretical and Applied Finance*, Vol 8, No 3, 357-380, 2005.

[20] Cont, R. (2001). “Empirical Properties of Asset Returns: Stylised facts and Statistical Issues”, *Quantitive Finance Volume*, 1, 223-236.

[21] Courtadon, G. (1982). “The Pricing of Options on Default Free Bonds”, *Journal of Financial and Quantitative Analysis* 17, 75-100

[22] Cox, J. and Ross, S. (1976). "The Valuation of Options for Alternative Stochastic Processes", *Journal of Financial Economics.*

[23] Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1980). "An Analysis of Variable Rate Contracts". J*ournal of Finance*, 35, 389-403

[24] Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). "An Interemporal General Equilibrium Model of Asset Prices", *Econometrica*, 53, 363-384.

[25] Dietrich-Campbell,B. and Schwartz, E. (1986). "Valuing debt options: empirical evidence", *Journal of Financial Economics* 13, 529-546.

[26] Ding,Z. (1994). "Time series Analysis of Speculative Returns", *Ph.D Thesis*, Department of Economics, University of California, San Diego.

[27] Dothan, L. (1978). "On the Term Structure of Interest Rates", *Journal of Financial Economics* 6, 59-69

[28] Engle, R. (1982). "Autoregressive Conditional Heteroscedasty with Estimates of the Variance of the United Kingdom", *Econometrica* 50, 987-1008.

[29] Engle, R. and Kroner, K. F.(1995). "Multivariate Simultaneous Generalized ARCH", *Econometric Theory*, 11, 122-150.

[30] Engle, R. and Patton, A. (2001). "What good is a Volatility Model?" *Quantivative Finance*, Volume 1, 237-245.

[31] Engle, R. and Sheppard, K. (2001). "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH", No 8554 in NBER Working Papers from National Bureau of Economic Research, Inc.

[32] Engle, R. (2002). "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models", *Journal of Business & Economic Statistics*, Vol. 20, No. 3, 339-350.

[33] Fang, H., Fang, K. and Kotz, S. (2002). "The Meta-elliptical Distributions with Given Marginals", *Journal of Multivariate Analysis* 82, 1-26.

[34] Goldys, B. and Musiela, M. (2001). "Infinite Dimensional Diffusions, Kolmogorov Equations and Interest Rate Models", Option Pricing, Interest Rates and Risk Management, 9.

[35] Hafner, C. and Franses, P. H. (2003). "A Generalized Dynamic Conditional Correlation Model for Many Asset Returns", Econometric Institute Report EI 2003-18.

[36] Hamilton, J. D. (1994). "Time Series Analysis", Princeton University Press, Princeton.

[37] Harrison, J. M. and Pliska, S. R. (1981). "Martingales and stochastic integrals in the theory of continuous trading", Stochastic Processes and their Applications 11, 215-260.

[38] Harrison, J. M. and Pliska, S. R. (1983). "A stochastic calculus model of continuous trading: Complete markets", Stochastic Processes and their Applications 15, 313-316.

[39] Heath, D., Jarrow, R. and Morton, A. (1992). "Bond Pricing and Term Structure of Interest Rates: A new methodology for contingent claim valuation", *Econometrica* 60(1), 77-105.

[40] Hull, J. (2003). "Options, Futures and Other Derivative Securities", third edition, Prentice-Hall, Upper Saddle River, New Jersey.

[41] Hull, J. and White, A. (1990). "Pricing Interest Rate and Other Derivative Securities", *Review of Financial Studies* 3, 573-592.

[42] James, J. and Webber, N. (2000). "Interest Rate Modelling", Wiley.

[43] Johnson, N. L. and Kotz, S. (1972) "Distributions in Statistics: Continuous Multivariate Distributions", John Wiley & Sons, Inc.

[44] Kearns, P. and Pagan, A. R. (1993). "Australian Stock Market Volatility: 1875-1987", *The Economic Record* Vol 69, No. 205, 163-78.

[45] Kotz, S. and Seeger, J. P. (1991). "A new approach to dependence in multivariate distributions", Advance in Probability Distributions with Given Marginals (G. Dall'Aglio, S. Kotz, and G. Salinetti, Eds.), 13-50, Kluwer Academic, Dordrecht.

[46] Lee, A. (1995). "Data Analysis, an introduction based on R", Dept. stat., University of Auckland.

[47] Lee, J. H.H.(1991). "A Lagrange multiplier test for GARCH models", *Economics Letters*, 37, 3 (November),265-271

[48] Ljung, T. and Box, G. E. P. (1979). "The Likelihood Function for a Stationary Autoregressive Moving Average Process", *Biometrika* 66, 265-270.

[49] Mercurio, F. and Moraleda, J. M. (1996). "An Analytically Tractable Interest Rate Model with Humped Volatility", Discussion paper 96-116/2, Tinbergen Institute, Erasmus University Rotterdam, the Netherlands.

[50] Merton, R. C. (1973). "Theory of rational option pricing", *Bell Journal of Economics and Management Science*, 4 (1), 141-183.

[51] Moraleda, J. M. (1997). "On the Pricing of Interest-Rate Options", No. 55, Tinbergen Institute Research Series.

[52] Moraleda, J. M. and Pelsser, A. (1997). "Forward versus Spot Interest Rate Models of the Term Structure: An Empirical Comparison", *The Journal of Derivatives*, 9-21.

[53] Moraleda, J. M. and Vorst, A. C. F. (1997). "Pricing American Interest Rate Claims with Humped Volatility Models", *Journal of Banking and Finance* 21, 1131-1157.

[54] Musiela, M. (1993): "Stochastic PDEs and Term Structure Models", *Journees Internationales de Finance*, IGR-AFFI, La Baule

[55] Musiela, M. and Rutkowski, M. (1997). "Martingale methods in financial modelling", Springer Verlag, New York. Application of Mathematics 36.

[56] Newey, W. K. and McFadden, D. (1994). "Large Sample Estimation and Hypothesis Testing", *Handbook of econometrics* Volume IV, Chapter 36, 2113-2245.

[57] Nelson, D.B. (1991). "Conditional heteroskedasticity in asset pricing: A new approach", *Econometrica* 59, 347-370.

[58] Rao, C. R. (1973). "Linear Statistical Inference and Its Application", John Wiley & Sons.

[59] Theodossiou, P. (1998). "Financial Data and the Skewed Generalized T Distribution", *Management Science*, Vol. 44, No. 12, Part 1 of 2. (Dec., 1998), pp. 1650-1661.

[60] Tse, Y. K. and Tsui, A. K. (2002). "A Multivariate Generalized Autoregressive Conditional Heteroskedasticity Model with Time-Varying Correlations", *Journal of Business & economic Statastics*, Vol. 20. No. 3.

[61] Tusi, A. K. and Yu, Q. (1999). "Constant Conditional Correlation in a Bivariate GARCH model: evidence from the stock market in China. *Mathematics and Computers in Simulation* 48, 503-509.

[62] Vasicek, O. A. (1977). "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics* 5, 177-188.

[63] Vilasuso, J. (2002). "Forecasting Exchange Rate Volatility", *Economics Letter* 76, 59-64.

[64] Zivot, E. and Wang, J. (2003). "Modelling Financial Time Series with S-plus", Springer.