การพยากรณ์ทางตลาดการเงิน ด้วยวิธีสับเปลี่ยนสถานะมาร์คอฟและการวิเคราะห์องค์ประกอบหลัก

นายณพรู้ โสภีพันธ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2555

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

FORECASTING FINANCIAL MARKET WITH MARKOV REGIME SWITCHING AND PRINCIPAL COMPONENT ANALYSIS

Nop Sopipan

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Applied Mathematics

Suranaree University of Technology

Academic Year 2012

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

FORECASTING FINANCIAL MARKET WITH MARKOV REGIME SWITCHING AND PRINCIPAL COMPONENT ANALYSIS

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

(Assoc. Prof. Dr. Prapasri Asawakun)

Chairperson

(Prof. Dr. Pairote Sattayatham)

Member (Thesis Advisor)

(Dr. Bhusana Premanode)

Member

(Assoc. Prof. Dr. Samruam Chongcharoen)

Member

(Asst. Prof. Dr. Eckart Schulz)

Member

(Asst. Prof. Dr. Arjuna Chaiyasena)

Member

(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs

(Assoc. Prof. Dr. Prapun Manyum)

Dean of Institute of Science

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

ณพฐ์ โสภีพันธ์ : การพยากรณ์ทางตลาดการเงินด้วยวิธีสับเปลี่ยนสถานะมาร์คอฟและการวิเคราะห์องค์ประกอบหลัก (FORECASTING FINANCIAL MAKET WITH MARKOV REGIME SWITCHING AND PRINCIPAL COMPONENT ANALYSIS) อาจารย์ที่ปรึกษา : ศาสตราจารย์ ดร.ไพโรจน์ สัตยธรรม, 145 หน้า.

วิทยานิพนธ์ฉบับนี้ เสนอการพยากรณ์ของผลตอบแทนรายวันทางการเงิน โดยพัฒนาการวิจัยการพยากรณ์ทางการเงินที่ผ่านมา ซึ่งใช้สมการค่าเฉลี่ยแบบคงที่แต่ไม่สามารถคาดการณ์ได้อย่างแม่นยำ ลักษณะของข้อมูลทางการเงินได้เคลื่อนไหวแบบไดนามิกไปพร้อมกับตัวแปรทางเศรษฐกิจและทางการเงินจำนวนมาก ดังนั้นเราพิจารณาตัวแปรบางตัว ที่เรียกว่า ตัวแปรอธิบาย และตัวแบบอัตตถดถอยเคลื่อนที่เฉลี่ย อันดับ p และ q (ARMA (p, q)) โดยเพิ่มตัวแปร เหล่านั้นในสมการค่าเฉลี่ย สำหรับการพยากรณ์ที่มีความแม่นยำมากขึ้น

การพยากรณ์ในรูปแบบนี้ได้พบปัญหาที่สำคัญเกิดขึ้นสองปัญหา ซึ่งปัญหาแรกคือตัวแปรอธิบายเกิดขึ้นในรูปแบบความสัมพันธ์เชิงพหุ (Multicollinearity) จึงแก้ปัญหาโดยใช้เทคนิคการวิเคราะห์องค์ประกอบหลัก (Principal component analysis: PCA)

ปัญหาที่สอง คือความแปรปรวนของค่าคลาดเคลื่อนไม่คงที่ และอาจจะขึ้นกับเวลา (Heteroskedasticity) ซึ่งได้แก้ปัญหาโดยใช้ตัวแบบความผันผวนไม่คงที่ (Volatility models) เพื่อพยากรณ์ ซึ่งตัวแบบความผันผวนที่ใช้ในวิทยานิพนธ์ฉบับนี้ ได้แก่ ตัวแบบการ์ช (GARCH) ตัวแบบอีการ์ช (EGARCH) ตัวแบบจีเจอาร์การ์ช (GJR-GARCH) และตัวแบบการสับเปลี่ยนสถานะมาร์คอฟแบบการ์ช (Markov Regime Switching GARCH)

สาขาวิชาคณิตศาสตร์ ลายมือชื่อนักศึกษา N.ON.

ปีการศึกษา 2555 ลายมือชื่ออาจารย์ที่ปรึกษา

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

NOP SOPIPAN : FORECASTING FINANCIAL MARKET WITH PRINCIPAL COMPONENT ANALYSIS AND MARKOV REGIME SWITCHING. THESIS ADVISOR : PROF. PAIROTE SATTAYATHAM, Ph.D. 145 PP.

FORECASTING/ TIME SERIES/ HETEROSKEDASTICITY/ MARKOV REGIME SWITCHING/ PRINCIPAL COMPONENT ANALYSIS

This thesis contains the research related with the forecasting of daily returns. It was found that the constant mean equation cannot give accurate forecasts, because the real word financial returns depend concurrently and dynamically on many economic and other financial variables. We therefore introduced some explanatory variables and employed a stationary Autoregressive Moving-average of orders p and q (ARMA (p,q)) for more accuracy.

However, two serious problems arose with this approach. The first problem was multicollinearity in the regression model. We used the Principal Component Analysis (PCA) method to remove possible complications caused by multicollinearity. The second problem was heteroskedasticity. We therefore used several volatility models to forecast volatility. The volatility models we consider were the GARCH, EGARCH, GJR-GARCH and MRS-GARCH (Markov Regime Switching GARCH) models.

School of Mathematics Student's Signature ______

Academic Year 2012 Advisor's Signature P. Sattayatham

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

ACKNOWLEDGEMENTS

I am profoundly grateful to my thesis advisor, Prof. Dr. Pairote Sattayatham, for his support, patient help and offering many useful suggestions and co-advisor Prof.Dr. Bhusana Premanode for inspiration and helpful comments for my thesis.

I would like to acknowledge all the lecturers who taught and helped me during the course of studies at Suranaree University of Technology.

I would like to express my appreciation to Assoc. Prof. Dr. Prapasri Asawakun, Asst. Prof. Dr. Arjuna Chaiyasena, Asst. Prof. Dr. Eckart Schulz, and Assoc. Prof. Dr. Samruam Chongcharoen for their valuable discussions and comments.

I also would like to express my appreciation to all who have helped me. These include all of my friends at Suranaree University of Technology.

I thank the Nakhon Ratchasima Rajabhat University, for grants to support my studies throughout.

Finally, I am deeply grateful to my parents and my family for support, understanding encouragement, and love.

Nop Sopipan

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CONTENTS

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CONTENTS (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CONTENTS (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CONTENTS (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CONTENTS (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

LIST OF TABLES

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

LIST OF TABLES (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

LIST OF TABLES (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

LIST OF TABLES (Continued)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

LIST OF FIGURES

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER I

INTRODUCTION

1.1 Introduction

The characteristic that all stock markets have in common is uncertainty, which is related with their short and long-term future states. This feature is undesirable for investors but it is also unavoidable whenever the stock market is selected as an investment tool. The best that one can do is to reduce this uncertainty by forecasting. Stock market forecasting (or prediction) is one of the instruments for this process.

The financial market forecasting task divides researchers and academics into two groups: those who believe that we can devise mechanisms to predict the market and those who believe that the market is efficient and whenever new information comes up, the market absorbs it by correcting itself, leaving no space for prediction. Furthermore the latter group believes that the financial market follows a random walk, which implies that the best prediction you can have about tomorrow's value is today's value.

There are two main methods of forecasting: qualitative and quantitative. Qualitative forecasting techniques are subjective and based on appropriate opinion and judgment when past data is not available, usually applied to intermediate and long range decisions, such as informed opinion and judgment, Delphi method, Market research, and historical life-cycle analogy. Quantitative forecasting models, on the other hand, are used to estimate future demands as a function of past data, and so are appropriate when past data is available. They are usually applied to

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

short-intermediate range decisions, for example, time series methods and causal econometric forecasting methods.

Most financial returns depend concurrently and dynamically on many economic and financial variables. The fact that the return has a statistically significant autocorrelation indicates that lagged returns might be useful in predicting future returns (Tsay, 2005).

This thesis assumes that returns follow a simple time series model such as a stationary ARMA(p, q) model with some explanatory variables. Nevertheless two serious problems arise with this approach as follows:

1. Heteroskedasticity: the variances of the residuals are not constant and possibly time-dependent, so that we construct models to forecast the returns volatility such as GARCH with Markov Regime Switching (MRS) and compare their performance with GARCH, EGARCH, GJR-GARCH models.

2. Multicolinearity: when the explanatory variables are correlated with each other or there is high correlation between the explanatory variables in a regression equation. We thus use Principal Component Analysis (PCA) to remove possible complications caused by such multicollinearity.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

1.2 Research Objectives

The objectives of the research in this thesis are the following:

1. To forecast return with varying mean equations as follows:

1.1 Constant mean equation.

1.2 Day of the week effect and ARMA process.

1.3 Multiple regression based on PCA and ARMA process.

2. To forecast volatility of return with GARCH, EGARCH, GJR-GARCH, and MRS-GARCH models.

3. To compare models from loss function in forecasting financial return and volatility of return.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

1.3 Literature Review

We now review past research related to forecasting.

Mehmet (2008) and Marcucci (2005) assumed that the mean equation of return is constant, while Easton and Faff (1994), and Kyimaz and Berument (2001) considered a mean equation of return with a one week delay into the regression model. Supoj (2003) considered a mean equation of return with an autoregressive process. Later Tsay (2005) proposed that the constant mean equation could not be forecasted due to inaccuracy of the financial data, since financial returns depend concurrently and dynamically on many economic and financial variables. The fact that returns have a statistically significant autocorrelation themselves indicates that the lagged returns might be useful in predicting future returns. Hence Tsay added some explanatory variables and stationary Autoregressive Moving-average of order p and q (ARMA (p, q)) to replace constant mean equation.

Mehmet (2008) stated that financial returns have three characteristics. The first is volatility clustering, implying that large changes tend to be followed by large changes and small changes tend to be followed by small changes. Second is fat tailedness (excess kurtosis) which means that financial returns often display a fatter tail than a standard normal distribution and the third is the leverage effect i.e. negative returns result in higher volatility than positive returns of the same size.

The generalized autoregressive conditional heteroskedasticity (GARCH) models mainly capture three characteristics of financial returns. The development of GARCH type models was started by Engle (1982). Engle introduced to ARCH model to model the heteroskedasticity by relating the conditional variance of the disturbance term to the linear combination of the squared disturbances in the recent past. Bollerslev (1986) generalized the ARCH (GARCH) model by

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

modeling the conditional variance to depend on its lagged values as well as squared lagged values of disturbance.

The volatility of financial returns is usually affected asymmetrically by positive and negative shocks. The exponential GARCH (EGARCH) of Nelson (1991), the GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993) and Threshold GARCH model of Zakoian (1994) were introduced to account for asymmetric effects of positive and negative shocks on volatility.

In addition, unconditional distribution of financial returns usually have fatter tails than the normal distribution, and standard GARCH or EGARCH models can not often fully capture the excess kurtosis in financial returns with assumption of normality (Pagan, 1996). For that reason, the generalized error distribution (Nelson, 1991) and the student-t distribution (Engle and Bollerslev, 1986) were proposed to overcome the excess kurtosis feature.

There are many extensions and modifications of GARCH type models in the literature. Some of them are long memory GARCH of Ding et al. (1993), Quadratic GARCH of Sentana (1995) and absolute GARCH of Hentschel (1995). Several surveys on those models are available in Bollerslev, Chou and Kroner (1992), Bera, Bollerslev and Higgins (1993), Engle and Nelson (1994), Franses and van Dijk (2000) and Granger and Poon (2003). Although the GARCH type models have proved to be successful in characterizing many features of volatility, they are not problem-free. In empirical studies, parameters of GARCH models are generally assumed to be stable over time. However, conditional distribution of financial returns differs between recession and expansion periods (Perez-Quiros and Timmermann, 2000). Moreover, GARCH models often imply a high volatility persistence of individual shocks. Lamoureux and Lastrapes (1990) argued that high persistence in volatility may be caused by structural changes in the variance

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

process.

Following these ideas, Cai (1994) and Hamilton and Susmel (1994) have independently introduced the Markov Regime Switching ARCH model (MRS-ARCH) which combines the Markov Switching model of Hamilton (1989, 1990) with the ARCH specification. The MRS-ARCH model was designed to capture regime changes in volatility with the unobservable state variable following a first order Markov Chain process. That is, parameters in the ARCH process are allowed to be changed in different states. Although it has been shown that the GARCH specification is better to fit financial data, Cai (1994) and Hamilton and Susmel (1994) used the ARCH specification to overcome the infinite path dependence problem arising in the Markov Regime Switching GARCH model (MRS-GARCH).

On the other hand, Gray (1996) proposed a new approach that allows tractable estimation of the MRS-GARCH model and eliminates the infinite path dependence problem. Also, Dueker (1997) took the same approach as Gray (1996) to overcome the infinite path dependence problem and introduced various alternative MRS-GARCH models. Klaassen (1998) modified Gray's MRS-GARCH model and argued that his specification improves the forecasting performance of the MRS-GARCH models. Recently, Haas, Mittnik, and Paolella (2004) proposed a new method different from Gray's (1996) approach and claimed that analytical tractability of their new model allows derivation of stationarity conditions and dynamic properties.

Hamilton and Susmel (1994) used weekly returns on the New York Stock Exchange Index over the period 1962 to 1987 to test their MRS-ARCH model with two to four regimes. They suggested that the MRS-ARCH specification better fits the data, to forecast volatility and to reduce volatility persistence than the uni-regime GARCH type models. Leon Li and William Lin (1994) and Wai

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Mun Fong (1996) applied the MRS-ARCH model of Hamilton and Susmel (1994) to examine regime shifts and volatility persistence respectively in the weekly the Taiwan Stock Index (TAIEX) and the weekly Japanese Stock Index (TOPIX). They concluded that the MRS-ARCH model provides a better description of the data and a much lower degree of volatility persistence than uni-regime GARCH type models. Moreover, the MRS-ARCH model has been applied to international stock markets by Fornari and Mele (1997), Schaller and Norden (1997), Susmel (1998a, 2000), Bautista (2003), Leon Li (2007) and to exchange rate by Fong (1998).

As an alternative estimation technique, Kaufmann and Schnatter (2002) developed Bayesian estimation techniques using Markov Chain Monte Carlo methods (MCMC) for MRS-ARCH models. Also, Kaufmann and Scheicher (2006) applied the MRS-ARCH model performed within the Bayesian framework to describe the daily German Stock Index (DAX). Gray (1996) extended the MRS-ARCH model to the MRS-GARCH case by developing a recombining method that merges conditional variances in different regimes into a single conditional variance. This makes the MRS-GARCH model path independent and allows for constructing a tractable likelihood function. Moreover, a MRS-GARCH model with time varying transition probabilities is proposed in the same study.

To implement his model, Gray (1996) used weekly one-month U.S. Treasury bill rates for the period of 1970 to 1994. He concluded that the MRS-GARCH model outperforms simple uni-regime models in forecasting performance and reduces persistence in volatility more than the MRS-ARCH model of Cai (1994) and Hamilton and Susmel (1994). Dueker (1997) introduced a collapsing procedure based on Kim's (1994) algorithm for the MRS-GARCH and applied it to the daily SP500 index. A modification of Gray's model, which allows multi-step ahead

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

volatility forecasting, was suggested by Klaassen (1998). In addition to the normal distribution, he adopted the student-t distribution for error terms and estimated his MRS-GARCH specification with two regimes using daily U.S. dollar exchange rates. The results show that Klaassen's model improves volatility forecasts and that volatility persistence is time-varying.

Recently, Marcucci (2005) compared a set of GARCH, EGARCH and GJR-GARCH models with a group of MRS-GARCH in terms of their ability to forecast SP100 volatility from one day to one month. Also, he assumed normal, student-t and generalized error distributions for the error terms. The main finding of Marcucci (2005) is that forecasting performance of MRS-GARCH models is significantly better than that of uni-regime GARCH type models at shorter horizons while standard asymmetric GARCH is found better at longer horizon. Daouk and Guo (2004) extended the SWARCH model to Markov Switching Regime Asymmetric GARCH (MRS-Asymmetric GARCH) which allows both regime switching in volatility and asymmetry. Ane and Ureche-Rangau (2006) introduced a Regime Switching Asymmetric Power GARCH model to analyze Asian stock indices. Other studies on the MRS-GARCH model comprise Fong and See (2001, 2002), Yu (2001), Francq and Zakoian (2005), Lee and Yoder (2007), Abramson and Cohen (2007a, 2007b), Brunetti, Mariano, Scotti, and Tan (2007).

This thesis is organized as follows: preliminaries in forecasting the financial market are presented in Chapter II. In Chapter III, we develop Markov Regime Switching GARCH models and discuss them in detail. Moreover, we forecast the price and volatility of two financial assets: the SET50 Index and gold. Afterwards, we apply these models to trading in the futures market. The mean equation model with day of the week effect and ARMA processes are given in Chapter IV. In Chapter V, we use multiple regression based on the PCA and ARMA processes

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

for forecasting the mean equation. Our conclusions are in Chapter VI.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER II

PRELIMINARIES ON FORECASTING FINANCIAL MARKETS

2.1 Financial Returns and Their Characteristics

Financial time series analysis is concerned with the theory and practice of asset valuation over time. Even though a highly empirical discipline, as in other scientific fields, theory forms the foundation for making inferences. There is, however, a key feature that distinguishes financial time series analysis from other time series analyses. Both, financial theory and its empirical time series contain an element of uncertainty. For example, there are several definitions of asset volatility, and for a stock return series, the volatility is not directly observable. As a result of the added uncertainty, statistical theory and methods play an important role in financial time series analysis.

2.1.1 Financial Returns

Most financial studies involve returns, instead of prices, of assets. There are several definitions of an asset return. Let P_t be the price of an asset at time index t. We discuss some definitions of returns. In this thesis, the term *return* mean the continuously compounded return (in percent) with definition as follows.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Definition 2.1. ***Continuously Compounded Return***

The natural logarithm of the simple gross return of an asset is called the continuously compounded return *or* log return *(in percent) as follows [Tsay (2005)]:*

$$r_t = 100 \cdot ln(\frac{P_t}{P_{t-1}}). \qquad (2.1)$$

2.1.2 Properties of Financial Returns

A financial return series, be it stock index returns or exchange rates, often exhibits the following well known properties [Rama, 2001]:

1. The sample mean of the series is close to zero.

2. The maginal distribution is roughly symmetric (or only slightly skewed), has a peak at zero, and is heavy-tailed.

3. - The sample autocorrelations of the series are "small" at almost all lags.
   - The sample autocorrelations of the absolute values and squares of the series are significant for a large number of lags.

4. Volatility is "clustered", i.e. days of either large or small movements are followed by days of similar characteristics.

2.1.3 Distribution Properties of Returns

To study financial returns, it is best to begin with their distributional properties. The objective here is to understand the behavior of the returns across assets and over time. Consider a collection of N assets held for T time periods, say, $t = 1, ..., T$. For each asset i, let r_{it} be its return at time t. The returns under study are $r_{it}; i = 1, ..., N; t = 1, ..., T$ [Tsay, 2005].

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Review of Statistical Distributions and Their Moments

We briefly review some basic properties of statistical distributions as well as moment equations of a random variable. Let R^k be the k-dimensional Euclidean space. A point in R^k is denoted by $x \in R^k$. Consider two random vectors $X = (X_1, ..., X_k)'$ and $Y = (Y_1, ..., Y_q)'$. Let $P(X \in A, Y \in B)$ be the probability that X is in the subspace $A \subset R^k$ and Y is in the subspace $B \subset R^q$.

- **The Joint Distribution**

  The function

$$F_{X,Y}(x, y; \theta) = P(X \leq x, Y \leq y; \theta),$$

  where $x \in R^k, y \in R^q$ and the inequality " $\leq$ " is a component-by-component operation, is a joint distribution function of X and Y with parameter θ. Behavior of X and Y is characterized by $F_{X,Y}(x, y; \theta)$. If the joint probability density function $f_{X,Y}(x, y; \theta)$ of X and Y exists, then

$$F_{X,Y}(x, y; \theta) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x, y; \theta) dy dx.$$

  In this case, X and Y are continuous random vectors.

- **Maginal Distribution**

  The marginal distribution of X is given by

$$F_X(x; \theta) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X,Y}(x, y; \theta) dy dx.$$

  Thus, the marginal distribution of X is obtained by integrating with respect to y. A similar definition applies to the marginal distribution of Y.

  If $k = 1$, X is a scalar random variable and the distribution function becomes

$$F_X(x) = P(X \leq x; \theta),$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

which is known as the *cumulative distribution function (CDF)* of X. The CDF of a random variable is nondecreasing (i.e., $F_X(x_1) \leq F_X(x_2)$ if $x_1 \leq x_2$) and satisfies $F_X(-\infty) = 0$ and $F_X(\infty) = 1$. We use the CDF to compute the p-value for a test statistic in this thesis.

- **Conditional Distribution**

  The *conditional distribution* of X given $Y \leq y$ is given by

$$F_{X|Y \leq y}(x;\theta) = \frac{P(X \leq x, Y \leq y; \theta)}{P(Y \leq y; \theta)}.$$

  If the probability density functions involved do exist, then the conditional density of X given $Y = y$ can be formulated as follows:

$$f_{X|Y}(x;\theta) = \frac{f_{X,Y}(x,y;\theta)}{f_Y(y;\theta)}, \tag{2.2}$$

  where the marginal density function $f_y(y;\theta)$ is obtained by

$$f_Y(y;\theta) = \int_{-\infty}^{\infty} f_{X,Y}(x,y;\theta)dx.$$

  From Eq. (2.2), the relation among joint, marginal, and conditional distributions is

$$f_{X,Y}(x,y;\theta) = f_{X|Y}(x;\theta) \times f_Y(y;\theta). \tag{2.3}$$

  This identity is used extensively in time series analysis (e.g., in maximum likelihood estimation). Finally, X and Y are independent random vectors if and only if $f_{X|Y}(x;\theta) = f_X(x;\theta)$. In this case,

$$f_{X,Y}(x,y;\theta) = f_X(x;\theta) f_Y(y;\theta). \tag{2.4}$$

- **Moments of a Random Variable**

  The lth moment of a continuous random variable X is defined as:

$$m_l = E(X^l) = \int_{-\infty}^{\infty} x^l f(x)dx,$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where E stands for expectation and $f(x)$ is the probability density function of X. The first moment is called the *mean* or *expectation* of X. It measures the central location of the distribution. We denote the mean of X by μ_X. The lth central moment of X is defined as:

$$\bar{m}_l = E[(X - \mu_X)^l] = \int_{-\infty}^{\infty} (x - \mu_X)^l f(x) dx,$$

provided that the integral exists. The second central moment, denoted by σ_X^2, measures the variability of X and is called the *variance* of X. The positive square root, σ_X, of variance is called the *standard deviation* of X. For other distributions, higher order moments are of interest. The third central moment measures the symmetry of X with respect to its mean, whereas the fourth central moment measures the tail behaviour of X. In statistics, *skewness* and *kurtosis*, which are the normalized third and fourth central moments of X, are often used to summarize the extent of asymmetry and tail thickness. In particular, the skewness ($S(X)$) and kurtosis ($K(X)$) of X are defined respectively as:

$$S(X) = E[\frac{(X - \mu_X)^3}{\sigma_X^3}], \tag{2.5}$$

and

$$K(X) = E[\frac{(X - \mu_X)^4}{\sigma_X^4}]. \tag{2.6}$$

The quantity $K(X) - 3$ is called the *excess kurtosis* because $K(X) = 3$ for a normal distribution. Thus, the excess kurtosis of a normal random variable is zero. A distribution with positive excess kurtosis is said to have *heavy tails*, implying that the distribution puts more mass on the tails of its support than a normal distribution does. In practice, this means that a random sample from such a distribution tends to contain more extreme values. Such

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

a distribution is said to be *leptokurtic.* On the other hand, a distribution with negative excess kurtosis has short tails (e.g., a uniform distribution over a finite interval). Such a distribution is said to be *platykurtic.*

In application, skewness and kurtosis can be estimated by their sample counterparts. Let $x_1, ..., x_T$ be a random sample of X with T observations. The sample mean is

$$\hat{\mu}_X = \frac{1}{T}\sum_{i=1}^{T} x_i. \tag{2.7}$$

The sample variance is

$$\hat{\sigma}_X^2 = \frac{1}{T-1}\sum_{i=1}^{T}(x_i - \hat{\mu}_X)^2, \tag{2.8}$$

while the sample skewness is

$$\hat{S}_{(X)} = \frac{1}{(T-1)\hat{\sigma}_X^3}\sum_{i=1}^{T}(x_i - \hat{\mu}_X)^3, \tag{2.9}$$

and the sample kurtosis is

$$\hat{K}_{(X)} = \frac{1}{(T-1)\hat{\sigma}_x^4}\sum_{i=1}^{T}(x_i - \hat{\mu}_X)^4. \tag{2.10}$$

Under the normality assumption, $\hat{S}(X)$ and $\hat{K}(X)-3$ are distributed asymptotically as normal with zero mean and variances $6/T$ and $24/T$, respectively (Snedecor, and Cochran, 1980). These asymptotic properties can be used to test the normality of asset returns. Given an asset return series $\{r_1, ..., r_T\}$, to test the skewness of the returns, we consider the null hypothesis $H_0 : S(r) = 0$ versus the alternative hypothesis $H_a : S(r) \neq 0$. The t-ratio statistic of the sample skewness is

$$t = \frac{\hat{S}(r)}{\sqrt{6/T}}.$$

The decision rule is as follows. Reject the null hypothesis at the α significance level if $|t| > Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the upper $100(\alpha/2)$th quantile of the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

standard normal distribution. Alternatively, one can compute the p-value of the test statistic t and reject H_0 if and only if the p-value is less than α. Similarly, one can test the excess kurtosis of the return series using the hypotheses $H_0 : K(r) - 3 = 0$ versus $H_a : K(r) - 3 \neq 0$. The test statistic is

$$t = \frac{\hat{K}(r) - 3}{\sqrt{24/T}}.$$

which is asymptotically a standard normal random variable. The decision rule is to reject H_0 if and only if the p-value of the test statistic is less than the significance level α. Jarque and Bera (1987) combine the two prior tests and use the test statistic

$$JB = \frac{\hat{S}^2(r)}{6/T} + \frac{(\hat{K}(r) - 3)^2}{24/T},$$

which is asymptotically distributed as a chi-squared random variable with 2 degrees of freedom, to test for the normality of r_t . One rejects H_0 of normality if the p-value of the JB statistic is less than the significance level.

2.2 Financial Volatility and Its Characteristics

Volatility plays a key role in empirical finance so that good forecasts of volatility are crucial for the implementation of derivative pricing, risk management and portfolio selection decisions. Even if any given model outperforms the alternative models in-sample evaluation, it may fail to forecast volatility accurately.

2.2.1 Realized Volatility

In order to assess the forecasting performance of various models, we have first to define a proxy for actual volatility. Since volatility is not directly

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

observable from the market, unlike financial returns, it must be estimated. In the literature, the general approach is to use squared daily (mean adjusted) returns as the measure of actual volatility, that is,

$$h_t = (r_t - \bar{r})^2, \tag{2.11}$$

where $\bar{r}$ is the average daily return at out-of-sample evaluation period.

The squared daily return is an unbiased estimator of actual volatility, but it produces very noisy estimates of unobserved volatility. Andersen and Bollerslev (1998) introduced an alternative volatility measure called *Realized Volatility.* This measure has recently attracted the attention of many researchers as an accurate measure of volatility. If returns are uncorrelated and have zero mean, realized volatility is an unbiased and consistent estimator for actual volatility (Andersen, Bollerslev, Diebold, and Labys, 2003). Realized volatility is obtained by summing the squared intraday returns and the higher frequency intraday data, hence the noise reduction in the volatility estimate. Realized volatility at day t can be formulated as follows

$$h_t = \sum_{d=1}^{D} (r_{t,d})^2, \tag{2.12}$$

where D is the number of intraday returns, such as $D = 24$ for hourly data.

However several assets are not traded in the whole day and so any changes during the off-trading hours must be considered. Thus, if this method is applied to the stock market data, realized volatility is defined as sum of squared intraday returns and squared overnight return (Koopman, Jungbacker, and Hol, 2004). That is:

$$h_t = R_{t,0}^2 + \sum_{d=1}^{D} (r_{t,d})^2, \tag{2.13}$$

where $R_{t,0}$ is the overnight return at day t. Hansen, and Lunde (2005) stated that overnight returns relatively large compared to the intraday return lead to a

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

noisy measure and suggested a scaling estimator to obtain a measure for whole day volatility. So,

$$h_t = \hat{c}\sum_{d=1}^{D}(r_{t,d})^2, \tag{2.14}$$

where

$$\hat{c} = \frac{T^{-1}\sum_{t=1}^{T}}{T^{-1}\sum_{t=1}^{T}\sum_{d=1}^{D} r_{t,d}^2}.$$

Other studies on realized volatility include Martens (2002), Barucci and Reno (2002), Areal and Taylor (2002).

Since intraday data of financial market data was not available to this thesis, we used squared daily returns as actual volatility for the forecast horizon one. In order to calculate volatility over the k days, following Klaassen (2002), we sum squared daily returns over the relevant (5, 10 and 22 days) horizons. This method is unbiased and more accurate than the traditional method which is the squared return of the forecast horizon. We can define the actual volatility over the k days $t, ..., t+k-1$ as

$$h_{t,k} = \sum_{i=t}^{t+k-1}(r_i - \bar{r})^2, \tag{2.15}$$

where $\bar{r}$ is the average daily return from an out-of-sample evaluation period. In practice, an investor, who has a one month investment horizon, generally is concerned with the volatility forecast over the next 22 days rather than volatility forecast for day $t+22$ made on day t. So, we focus on the volatility forecast over the next k days instead of the $k-step-ahead$ forecasts. In order to compute the volatility forecast over the next k days, we aggregate $k-step-ahead$ forecasts. The $h_{t,k}$ denoting the volatility forecast over next k days, can be formulated as follows:

$$h_{t,k} = \sum_{i=1}^{k} h_{t+i}, \tag{2.16}$$

where h_{t+i} denotes the $i-step-ahead$ forecast made at time t.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

2.2.2 Stylized Facts about Volatility in Financial Time Series

Financial time series such as stock returns, exchange rates etc. exhibit certain patterns which are crucial for correct model specification, estimation and forecasting. The stylized facts about volatility follow:

- **Fat tails:** When the distribution of financial time series such as stock returns are compared with the normal distribution, fatter tails are observed. This observation is also referred to as excess kurtosis. The standardized fourth moment for a normal distribution is 3 whereas for many financial time series, a value well above 3 is observed (Mandelbrot, 1963 and Fama, 1963; 1965) are the first studies to report this feature. For modeling excess kurtosis, distributions that have fatter tails than normal are proposed in the literature such as Paretian and Levy.

- **Volatility clustering:** The second stylized fact is the periods of volatility clustering which refers to the observation of large movements being followed by large movements. This is an indication of the persistence in shocks. Correlograms and corresponding Box-Ljung statistics show significant correlations which exist at extended lag lengths.

- **Leverage effects:** This refers to the idea that price movements are negatively correlated with volatility, first suggested by Black (1976) for stock returns. Black argued, however, that the measured effect of stock price changes on volatility was too large to be explained solely by leverage effects. Empirical evidence on leverage effects can be found in Nelson (1991), Gallant, Rossi, and Tauchen (1992, 1993), Campbell and Kyle (1993) and Engle and Ng (1993).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

- **Long memory:** Especially for high-frequency data, volatility is highly persistent and there exists evidence of near unit root behavior of the conditional variance process. This observation led to two propositions for modeling persistence: unit root and the long memory process. The Autoregressive Conditional Heteroscedasticitiy (ARCH) models use the second idea for modeling persistence.

- **Co-movement in volatility:** When we look at financial time series across different markets, such as looking at exchange rate returns for different currencies, we observe big movements in one currency being matched by big movements in another. This suggests the importance of multivariate models in modeling cross-correlations in difference markets.

These observations about volatility led many researchers to focus on the cause of these stylized facts. The fact of information arrivals was a prominent one in the literature as many studies link asset returns to the information flow. Asset returns are observed and measured at fixed time intervals such as daily, weekly or monthly. Much more frequent observations such as tick-by-tick data are also available. The rate of information arrival is non-uniform and not directly observable. Mandelbrot and Taylor (1967) use the idea of time deformation to explain the observed fat tails in the data. The same idea is used by Clark (1973) to explain volatility. Easley and O'Hara (1992) try to link market volatility with trading volume; quote arrivals, forecastable events such as dividend announcements, and market closures.

To get reliable forecasts of future volatilities, it is crucial to account for the stylized facts. There are various approaches in the literature for volatility modeling that try to capture these stylized facts which we will discuss in the following sections.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

2.3 Stationarity

The foundation of time series analysis is the concept of *stationarity.* There are two definitions as follow:

Definition 2.2. *A time series* $\{r_t\}$ *is said to be* ***strictly stationary*** *if the joint distribution of* $(r_{t_1}, ..., r_{t_k})$ *is identical to that of* $(r_{t_1+t}, ..., r_{t_k+t})$ *for all* t*, where* k *is an arbitrary positive integer and* $(t_1, ..., t_k)$ *is a collection of* k *positive integers. In other words, strict stationarity requires that the joint distribution of* $(r_{t_1}, ..., r_{t_k})$ *is invariant under time shift.*

This is a very strong condition that is hard to verify empirically, hence a weaker version of stationarity is often assumed:

Definition 2.3. *A time series* $\{r_t\}$ *is* ***weakly stationary*** *if both the mean of* r_t *and the covariance between* r_t *and* r_{t-l} *are time-invariant, where* l *is an arbitrary integer. More specifically,* $\{r_t\}$ *is weakly stationary if* $E(r_t) = \mu$*, a constant, and* $Cov(r_t, r_{t-l}) = \gamma_l$.

In practice, suppose that we have observed T data points $\{r_t | t = 1, ..., T\}$. The weak stationarity implies that the time plot of the data would show that the T values fluctuate with constant variation around a fixed level. In applications, weak stationarity enables one to make inferences concerning future observations (e.g., predictions).

Implicitly, in the condition of weak stationarity, we assume that the first two moments of r_t are finite. From the definitions, if r_t is strictly stationary and its first two moments are finite, then r_t is also weakly stationary. The converse is not true in general. However, if the time series r_t is normally distributed, then weak stationarity is equivalent to strict stationarity.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In this thesis, we are mainly concerned with weakly stationary series. The covariance $\gamma_l = Cov(r_t, r_{t-l})$ is called the lag-l autocovariance of r_t . It has two important properties: $\gamma_0 = Var(r_t)$ and $\gamma_{-l} = \gamma_l$.

The second property holds because

$$\begin{aligned} Cov(r_t, r_{t-(-l)}) &= Cov(r_{t-(-l)}, r_t) \\ &= Cov(r_{t+l}, r_t) \\ &= Cov(r_{t_1}, r_{t_1-l}), \end{aligned} \tag{2.17}$$

where $t_1 = t + l$.

In the finance literature, it is common to assume that an asset return series is weakly stationary. This assumption can be checked empirically provided that a sufficient number of historical returns are available. For example, one can divide the data into subsamples and check the consistency of the results obtained across the subsamples.

2.3.1 Correlation and Autocorrelation Function

The correlation coefficient between two random variables X and Y is defined as:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sqrt{E(X-\mu_X)^2 E(Y-\mu_Y)^2}},$$

where μ_X and μ_Y are the means of X and Y, respectively, and it is assumed that the variances exist. This coefficient measures the strength of linear dependence between X and Y, and it can be shown that $-1 \leq \rho_{X,Y} \leq 1$ and $\rho_{X,Y} = \rho_{Y,X}$. The two random variables are uncorrelated if $\rho_{X,Y} = 0$. In addition, if both X and Y are normal random variables, then $\rho_{X,Y} = 0$ if and only if X and Y are independent. When the sample$(X_t, Y_t)_{t=1}^{T}$ is available, the correlation can be

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

consistently estimated by its sample counterpart

$$\hat{\rho}_{X,Y} = \frac{\sum_{t=1}^{T}(X_t - \bar{X})(Y_t - \bar{Y})}{\sqrt{\sum_{t=1}^{T}(X_t - \bar{X})^2(Y_t - \bar{Y})^2}}.$$

- **Autocorrelation Function (ACF)**

Consider a weakly stationary return series r_t . When the linear dependence between r_t and its past values r_{t-l} is of interest, the concept of correlation is generalized to autocorrelation. The correlation coefficient between r_t and r_{t-l} is called the lag-l autocorrelation of r_t and is commonly denoted by ρ_l, which under the weak stationarity assumption is a function of l only. Specifically, we define as:

$$\rho_l = \frac{Cov(r_t, r_{t-l})}{\sqrt{Var(r_t)Var(r_{t-l})}} = \frac{Cov(r_t, r_{t-l})}{Var(r_t)} = \frac{\gamma_l}{\gamma_0}, \tag{2.18}$$

where the property $Var(r_t) = Var(r_{t-l})$ for a weakly stationary series is used.

From the definition, we have $\rho_0 = 1, \rho_l = \rho_{-l}$, and $-1 \leq \rho_l \leq 1$. In addition, a weakly stationary series r_t is not serially correlated if and only if $\rho_l = 0$ for all $l > 0$.

In general, the lag-l sample autocorrelation of r_t is defined as:

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{T}(r_t - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=1}^{T}(r_t - \bar{r})^2}, 0 \leq l \leq T - 1. \tag{2.19}$$

If r_t is an $i.i.d.$ sequence satisfying $E(r_t^2) < \infty$, then $\hat{\rho}_l$ is asymptotically normal with mean zero and variance $1/T$ for any fixed positive integer l. More generally, if r_t is a weakly stationary time series satisfying $r_t = \mu + \Sigma_{i=0}^{q}\psi_i a_{t-i}$, where $\psi_0 = 1$, q is a non-negative integer, and a_j is a Gaussian white noise series, then $\hat{\rho}_l$ is asymptotically normal with mean zero and variance $(1 + 2\Sigma_{i=1}^{q}\rho_i^2)/T$ for $l > q$. This is referred to as Bartletts formula in the time series.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

- **Testing Individual ACF**

For a given positive integer l, the previous result can be used to test H_0 : $\rho_l = 0$ versus $H_a : \rho_l \neq 0$. The test statistic is

$$t - ratio = \frac{\hat{\rho}_l}{\sqrt{(1 + 2\sum_{i=1}^{l-1} \hat{\rho}_i^2)/T}}.$$

If r_t is a stationary Gaussian series satisfying $\rho_j = 0$ for $j > l$, the t-ratio is asymptotically distributed as a standard normal random variable. Hence, the decision rule of the test is to reject H_0 if $|t - ratio| > Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the $100(1-\alpha/2)th$ percentile of the standard normal distribution.

In finite samples, $\hat{\rho}_l$ is a biased estimator of ρ_l. The bias is on the order of $1/T$, which can be substantial when the sample size T is small. In most financial applications, T is relatively large so that the bias is not serious.

- **Portmanteau Test**

Financial applications often require us to test jointly that several autocorrelations of r_t are zero. Box and Pierce (1970) propose the Portmanteau statistic

$$Q^*(m) = T\sum_{l=1}^{m} \hat{\rho}_l^2,$$

as a test statistic for the null hypothesis $H_0 : \rho_1 = \ldots = \rho_m = 0$ against the alternative hypothesis $H_a : \rho_i \neq 0$ for some $i \in [1, \ldots, m]$. Under the assumption that r_t is an $i.i.d.$ sequence with certain moment conditions, $Q^*(m)$ is asymptotically a chi-squared random variable with m degrees of freedom. Ljung and Box (1978) modify the $Q^*(m)$ statistic as below to increase the power of the test in finite samples:

$$Q(m) = T(T+2)\sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{T-l}. \tag{2.20}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The decision rule is to reject H_0 if $Q(m) > \chi^2_\alpha$, where χ^2_α denotes the $100(1-\alpha)th$ percentile of a chi-squared distribution with m degrees of freedom. Most software packages will provide the p-value of $Q(m)$. The decision rule is then to reject H_0 if the p-value is less than or equal to α, the significance level.

2.4 General Autoregressive Moving-average (ARMA) Models

Basically, an ARMA model combines the ideas of AR (autoregressive) and MA (moving-average) models into a compact form so that the number of parameters used is kept small. For the return series in finance, the chance of using ARMA models is low. However, the concept of ARMA models is highly relevant in volatility modeling. As a matter of fact, the generalized autoregressive conditional heteroscedastic (GARCH) model can be regarded as an ARMA model, albeit nonstandard, for the ε_t.

A general ARMA(p,q) model is in the form:

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}, \tag{2.21}$$

where ε_t is a white noise series, and p and q are non-negative integers. The AR and MA models are special cases of the ARMA(p,q) model. Using the back-shift operator, the model can be written as

$$(1 - \phi_1 B - ... - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - ... - \theta_q B^q)\varepsilon_t. \tag{2.22}$$

The polynomial $1 - \phi_1 B - ... - \phi_p B^p$ is the AR polynomial of the model. Similarly, $1 - \theta_1 B - ... - \theta_q B^q$ is the MA polynomial. We require that there are no common factors between the AR and MA polynomials; otherwise the order (p,q) of the model would be reducible. Like a pure AR model, the AR polynomial

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

introduces the characteristic equation of an ARMA model. If all of the solutions of the characteristic equation are less than 1 in absolute value, then the ARMA model is weakly stationary.

2.4.1 Identifying ARMA Models

The techniques of model identification which are most commonly used were propounded originally by Box and Jenkins (1994). Their basic tools were the sample autocorrelation function and the partial autocorrelation function. We shall describe these functions and their use separately from the spectral density function which ought, perhaps, to be used more often in selecting models. The fact that spectral density function is often overlooked is probably due to an unfamiliarity with frequency-domain analysis on the part of many model builders.

Identifying the order of differencing and the constant:

- Rule 1: If the series has positive autocorrelations out to a high number of lags, then it probably needs a higher order of differencing.

- Rule 2: If the lag-1 autocorrelation is zero or negative, or the autocorrelations are all small and patternless, then the series does not need a higher order of differencing. If the lag-1 autocorrelation is -0.5 or more negative, the series may be overdifferenced.

- Rule 3: The optimal order of differencing is often the order of differencing at which the standard deviation is lowest.

- Rule 4: A model with no orders of differencing assumes that the original series is stationary (among other things, mean-reverting). A model with one order of differencing assumes that the original series has a constant average trend (e.g. a random walk). A model with two orders of total differencing

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

assumes that the original series has a time-varying trend (e.g. a random trend).

- Rule 5: A model with no orders of differencing normally includes a constant term (which represents the mean of the series). A model with two orders of total differencing normally does not include a constant term. In a model with one order of total differencing, a constant term should be included if the series has a non-zero average trend.

Identifying the numbers of AR and MA terms:

- Rule 6: If the partial autocorrelation function (PACF) of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is positive i.e., if the series appears slightly *underdifferenced* then consider adding one or more AR terms to the model. The lag beyond which the PACF cuts off is the indicated number of AR terms.

- Rule 7: If the autocorrelation function (ACF) of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is negative i.e., if the series appears slightly *overdifferenced* then consider adding an MA term to the model. The lag beyond which the ACF cuts off is the indicated number of MA terms.

- Rule 8: It is possible for an AR term and an MA term to cancel each other's effects, so if a mixed ARMA model seems to fit the data, also try a model with one fewer AR term and one fewer MA term particularly if the parameter estimates in the original model require more than 10 iterations to converge.

- Rule 9: If there is a unit root in the AR part of the model i.e., if the sum of the AR coefficients is almost exactly 1 you should reduce the number of AR terms by one and increase the order of differencing by one.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

- Rule 10: If there is a unit root in the MA part of the model i.e., if the sum of the MA coefficients is almost exactly 1 you should reduce the number of MA terms by one and reduce the order of differencing by one.

- Rule 11: If the long-term forecasts appear erratic or unstable, there may be a unit root in the AR or MA coefficients.

2.4.2 Forecasting Using an ARMA Model

Forecasts of an ARMA(p, q) model have characteristics similar to those of an AR(p) model after adjusting for the impacts of the MA(q) component on the lower horizon forecasts. Denote the forecast origin by h and the available information by F_h. The 1-step ahead forecast of r_{h+1} can easily be obtained from the model as

$$\begin{aligned} \hat{r}_h(1) &= E(r_{h+1}|F_h) \\ &= \hat{\phi}_0 + \sum_{i=1}^{p} \hat{\phi}_i r_{h+1-i} - \sum_{i=1}^{q} \hat{\theta}_i \varepsilon_{h+1-i}, \end{aligned}$$

and the associated forecast error is $\varepsilon_{h+1} = r_{h+1} - \hat{r}_h(1)$. The variance of 1-step ahead forecast error is $Var[\varepsilon_{h+1}] = h_t$. For the ℓ-step ahead forecast, we have

$$\begin{aligned} \hat{r}_h(\ell) &= E(r_{h+\ell}|F_h) \\ &= \hat{\phi}_0 + \sum_{i=1}^{p} \hat{\phi}_i \hat{r}_h(\ell - i) - \sum_{i=1}^{q} \hat{\theta}_i \varepsilon_h(\ell - i), \end{aligned}$$

where it is understood that $\hat{r}_h(\ell - i) = r_{h+\ell-i}$ if $\ell - i \leqslant 0$ and $\varepsilon_h(\ell - i) = 0$ if $\ell - i > 0$ and $\varepsilon_h(\ell - i) = \varepsilon_{h+\ell-i}$ if $\ell - i \leqslant 0$.

Thus, the multistep ahead forecasts of an ARMA model can be computed recursively. The associated forecast error is $\varepsilon_{h+\ell} = r_{h+\ell} - \hat{r}_h(\ell)$ which can be computed easily.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

2.5 The Multiple Regression Model

Multiple Linear Regression (MLR) attempts to model the relationship between two or more explanatory variables and a response variable, by fitting a linear equation to the observed data. The dependent variable (Y) is given by:

$$Y = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \varepsilon, \tag{2.23}$$

where $X_i, i = 1, \ldots, p$ are the explanatory independent variables, $\beta_i, i = 1, \ldots, p$ are the regression coefficients, and ε is the error associated with the regression and assumed to be normally distributed with both expectation value zero and constant variance (J.C.M Pires et al. (2007)).

2.5.1 Assumptions in Multiple Regression Model

Not surprisingly, the assumptions for a multiple regression analysis are very similar to those required for a simple linear regression.

- **Linearity:** Because of the multiple X variables, the assumption of linearity is not as straightforward as for simple linear regression. Multiple regression analysis assumes that the relationship between Y and each X is linear. This means that if all other X variables are held constant, then changes in the particular X variable lead to a linear change in the Y variable. Because this is a relation, simple plots of Y versus each X variable may not be linear, since pairwise plots can not hold the other variables fixed.

- **Correct sampling scheme:** The Y must be a random sample from the population of Y values for every set of X value in the sample. Fortunately, it is not necessary to have a completely random sample from the population as the regression line is valid even if the X values are deliberately chosen.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

However, for a given set of X, the values from the population must be a simple random sample.

- **No outliers or influential points:** All the points must belong to the relationship and there should be no unusual points. The residual plot of the residual against the row number or against the predicted value should be investigated to see if there are indeed any unusual points.

- **Equal variation along the line:** The variability about the regression plane requirement is similar for all sets of X, i.e. the scatter of the points above and below the fitted surface should be roughly constant over the entire line. This is assessed by looking at the plots of the residuals against each X variable to see if the scatter is roughly uniformly scattered around zero with no increase and no decrease in spread over the entire line.

- **Independence:** Each value of Y is independent of any other value of Y. The most common cases where this fails are time series data.

- **Normality of errors:** The difference between the value of Y and the expected value of Y is assumed to be normally distributed. This is one of the most misunderstood assumptions. Many people erroneously assume that the distribution of Y over all X values must be normally distributed, i.e. they look simply at the distribution of the Ys ignoring the Xs. The assumption states that only the residuals, the difference between the value of Y and the point on the line must be normally distributed.

- **X variables measured without error:**It sometimes turns out that the X variables are not known precisely.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The predicted value given by the regression model ($\hat{Y}$) is calculated by:

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{p} \hat{\beta}_i X_i. \tag{2.24}$$

The most common method to estimate the regression parameters $\hat{\beta}_i, i = 1, \ldots, p$ is use of the Ordinary Least square Estimator (OLS).

MLR is one of the most frequently used methods for forecasting, to fit observed data and to create models that can be used for prediction in many research fields such as biology, medicine, psychology, economic and environment.

Definition 2.4. ***Multicollinearity*** *is a statistical phenomenon in which two or more explanatory (or predictor) variables in a multiple regression model are highly correlated.*

In this situation the coefficient estimates may change erratically in response to small changes in the model or the data. Multicollinearity does not reduce the predictive power or reliability of the model as a whole, at least within the sample data themselves; it only affects calculations regarding individual predictors. That is, a multiple regression model with correlated predictors can indicate how well the entire bundle of predictors predicts the outcome variable, but it may not give valid results about any individual predictor, or about which predictors are redundant with respect to others.

A high degree of multicollinearity can also prevent computer software packages from performing the matrix inversion required for computing the regression coefficients, or it may render inaccuarate the results of that inversion.

Detection of multicollinearity

Indicators of the possible presence of multicollinearity in a model:

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

- Large changes in the estimated regression coefficients when a predictor variable is added or deleted.

- Insignificant regression coefficients for the affected variables in the multiple regression, but a rejection of the joint hypothesis that those coefficients are all zero (using an F-test).

- Some have suggested a formal detection-tolerance or the variance inflation factor (VIF) for multicollinearity: the variance inflation factor (VIF) quantifies the severity of multicollinearity in an ordinary least squares regression analysis, providing an index that measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of collinearity.

Calculation and analysis

The VIF can be calculated and analyzed in three steps:

Step one: Calculate k different VIFs, one for each X_i by first running an ordinary least square regression that has X_i as a function of all the other explanatory variables in the first equation.

If $i = 1$, for example, the equation would be

$$X_1 = \alpha_2 X_2 + \alpha_3 X_3 + ... + \alpha_k X_k + c_0 + e,$$

where c_0 is a constant and e is the error term.

Step two: Then, calculate the VIF factor for $\hat{\beta}_i$ in equation 2.24 with the following formula:

$$VIF = \frac{1}{1 - R_i^2},$$

where R_i^2 is the coefficient of determination of the regression equation in step one, but with X_i on the left hand side, and all other predictor variables (all

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the other X variables) on the right hand side.

Step three: Analyze the magnitude of multicollinearity by considering the size of the VIF. A common rule of thumb is that if $VIF > 5$ then multicollinearity is high. Also 10 has been proposed as a cut off value.

- **Condition Number Test:** The standard measure of ill-conditioning in a matrix is the condition index. It will indicate that the inversion of the matrix is numerically unstable with finite-precision numbers (standard computer floats and doubles). This indicates the potential sensitivity of the computed inverse to small changes in the original matrix. The Condition Number is computed by finding the square root of the quotient of the maximum eigenvalue divided by the minimum eigenvalue. If the Condition Number is above 30, the regression is said to have significant multicollinearity.

- **Farrar-Glauber Test:** If the variables are found to be orthogonal, there is no multicollinearity; if the variables are not orthogonal, then multicollinearity is present.

- Construction of a pair-wise correlation matrix will yield indications as to the likelihood that any given couplet of right-hand-side variables are multicollinear.

In this thesis, we use VIF as the indicator for multicollinearity.

2.6 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (i.e., uncorrelated with) the preceding components.

PCA can be mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.

Consider a random variable $X = (X_1, \ldots, X_p)'$ with mean $\mu = (\mu_1, \ldots, \mu_p)'$, $(.)'$ denoting its transpose, $\mu_i < \infty; i = 1, \ldots, p$ and the random variables X having a known variance matrix $\Sigma = (\sigma_{ij}); \sigma_{ij} < \infty; i, j = 1, \ldots, p$. The more realistic case, where Σ is unknown, follows by replacing Σ by a sample covariance matrix $\mathbf{S}$. Assume that the rank of Σ is p and

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0, \tag{2.25}$$

are the p eigenvalues of Σ .

With PCA we desire an uncorrelated linear function of $X_1, \ldots, X_p$, say, $Z_1, \ldots, Z_m, m \leq p$, such that the variances $V(Z_1), \ldots, V(Z_m)$ account for most of the total variances among $X_1, \ldots, X_p$. Also, we require
$V(Z_1) > V(Z_2) > \ldots > V(Z_m)$. Algebraically, principal components are particular linear combinations of $X_1, \ldots, X_p$. Geometrically, the principal components represent a new coordinate system obtained by rotating the original axes $X_1, \ldots, X_p$. The new axes represent the directs with maximum variability.

Let $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{ip})', i = 1, \ldots, p$ be $p \times 1$ a vector of weights for the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

respective components of X. Consider the linear function

$$Z_1 = \alpha_1' X = \sum_{i=1}^{p} \alpha_{i1} X_i. \tag{2.26}$$

Our aim is to find α_1 such that $V(Z_1)$ is maximum subject to the condition $\alpha_1'\alpha_1 = 1$. It is clear that $V(Z_1)$ can be increased by multiplying α_1 by some constant. To eliminate this arbitrariness we restrict our attention to coefficient vectors of unit lengths.

Now, $V(Z_1) = \alpha_1'\Sigma\alpha_1$. Hence, we are required to find a vector α_1 such that

$$\alpha_1'\Sigma\alpha_1, \tag{2.27}$$

is maximum subject condition $\alpha_1'\alpha_1 = 1$.

To maximize $\alpha_1'\Sigma\alpha_1$ subject to $\alpha_1'\alpha_1 = 1$, the standard approach is to use the technique of Lagrange multipliers. Maximize $\alpha_1'\Sigma\alpha_1 - \lambda(\alpha_1'\alpha_1 - 1)$, where λ is a Lagrange multiplier.

Differentiation with respect to α_1 gives

$$\begin{aligned} \Sigma\alpha_1 - \lambda\alpha_1 &= 0, \\ (\Sigma - \lambda I)\alpha_1 &= 0, \end{aligned} \tag{2.28}$$

where I is the $p \times p$ identity matrix.

Since, $\alpha_1 \neq 0$, there can be a solution only if $\Sigma - \lambda I$ is singular, i.e. if $|\Sigma - \lambda I| = 0$ such that if λ is a eigenvalue of Σ and α_1 is its corresponding normalized eigenvector of Σ.

Thus, λ is an eigenvalue of Σ and α_1 is the corresponding eigenvector. To decide which of the p eigenvectors gives $\alpha_1'X$ with maximum variance, note that the quantity to be maximized is

$$\alpha_1'\Sigma\alpha_1 = \alpha_1'\lambda\alpha_1 = \lambda\alpha_1'\alpha_1 = \lambda$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

so λ must be as large as possible. Thus, α_1 is the eigenvector corresponding to the largest eigenvalue of Σ, and $Var[Z_1] = Var[\alpha_1' X] = \alpha_1' \Sigma \alpha_1 = \lambda = \lambda_1$, the largest eigenvalue.

The second principal component is given by the linear function

$$Z_2 = \alpha_2' X$$

where α_2 is such that $Var[Z_2]$ is maximum subject to the condition $\alpha_2' \alpha_2 = 1$ and Z_2 is orthogonal to Z_1 (i.e. $\alpha_2' \alpha_1 = 0$, since we require Z_1, Z_2 to be stochastically independent). It is known from matrix algebra that

$$max_{(\alpha_2 : \alpha_2' \alpha_2 = 1, \alpha_2' \alpha_1 = 0)} \alpha_2' \Sigma \alpha_2 = \lambda_2, \tag{2.29}$$

and the maximum is attained, the normalized eigenvector corresponding to λ_2

Alternatively, consider the Lagrangian

$$\phi = \alpha_2' \Sigma \alpha_2 - \theta_1 (\alpha_2' \alpha_2 - 1) - \theta_2 (\alpha_2' \alpha_1 - 0)$$

where θ_1, θ_2 are Lagrange multipliers. Differentiating both sides with respect to α_2 and equating the result to zero,

$$\frac{\partial \phi}{\partial \alpha_2} = 2(\Sigma - \theta_1 I)\alpha_2 - \theta_2 \alpha_1 = 0. \tag{2.30}$$

Premultiplying by α_1'

$$2\alpha_1' \Sigma \alpha_2 - \theta_2 = 0. \tag{2.31}$$

Again, premultiplying the relation

$$(\Sigma - \lambda_1 I)\alpha_1 = 0$$

by α_2' ,

$$\alpha_2' \Sigma \alpha_1 = 0.$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Hence, from Eq. (2.31), $\theta_2 = 0$. Therefore, by Eq. (2.30),

$$(\Sigma - \theta_1 I)\alpha_2 = 0, \tag{2.32}$$

and it follows similarly (as in Eq. (2.28) and subsequent statements) that $\theta_1 = \lambda_2$, the corresponding eigenvector. The second principal component is, therefore,

$$Z_2 = \alpha_2' X$$

with

$$Var(Z_2) = \lambda_2 \tag{2.33}$$

Clearly,

$$\begin{aligned} Cov(Z_1, Z_2) &= Cov(\alpha_1' X, \alpha_2' X) \\ &= \alpha_1' \Sigma \alpha_2 \\ &= \alpha_1' \lambda_2 \alpha_2 \\ &= 0. \end{aligned}$$

We continue by induction to find the kth principal component, $Z_k = \alpha_k' X$, we are to find α such that $Var(Z_k)$ is maximum subject to the conditions

$$\begin{aligned} \alpha_k' \alpha_k &= 1 \\ \alpha_k' \alpha_i &= 0 \ (i = 1, \ldots, k-1). \end{aligned}$$

We thus obtain p random variables $Z_k, (k = 1, ..., p)$,

$$Z_k = \alpha_k' X$$

with

$$Var(Z_k) = \lambda_k, \tag{2.34}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where α_k is the normalized eigenvector corresponding to λ_k. Clearly,

$$\begin{aligned} Cov(Z_k, Z_{k'}) &= Cov(\alpha_k' X, \alpha_{k'}' X) \\ &= \alpha_k' \Sigma \alpha_{k'} \\ &= \alpha_k' \lambda_{k'} \alpha_{k'} \\ &= 0 \ (k \neq k'). \end{aligned} \tag{2.35}$$

Theorem 2.1. *Let Σ be a variance matrix. Σ has the eigenvalue-eigenvector pair $(\lambda_1, \alpha_1), \ldots, (\lambda_p, \alpha_p)$ where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$. The kth principal component of X is given by*

$$Z_k = \alpha_k' X$$

for $k = 1, 2, ..., p$ with

$$Var(Z_k) = \lambda_k.$$

Also, $Cov(Z_k, Z_{k'}) = 0 (k \neq k')$. If some $\lambda_k's$ are equal, the corresponding principal components may not be uniquely chosen.

Proof: see Jolliffe (2002).

Theorem 2.2. ***Spectral Decomposition (or Jordan Decomposition)***
If $A = A'$, there exists an orthogonal matrix Γ such that

$$A = \Gamma \Lambda \Gamma' \tag{2.36}$$

where $\Lambda = Diag(\lambda_1, \ldots, \lambda_p)$ and $\lambda's$ are latent roots of A (some of which may be zero and some of which may be identical). Here $\Gamma = (\alpha_1, \ldots, \alpha_p)$ where α_i $(p \times 1)$ is the normalized eigenvector corresponding to λ_i, $\alpha_i' \alpha_j = \delta_{i,j}$ where $\delta_{i,j}$ is Dirac function. Hence

$$A = \Sigma_{i=1}^{p} \lambda_i \alpha_i \alpha_i'. \tag{2.37}$$

The matrix $\gamma_i \gamma_i'$ is called the $i - th$ spectral decomposition of A and the decomposition Eq. (2.37) the spectral decomposition (or Jordan Decomposition) of A.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Proof: see Jolliffe (2002). The following result holds:

- **(a)** If A is a non-singular symmetric matrix, then for any integer k, positive or negative,
$$A^k = \sum_{i=1}^{p} \lambda_i^k \gamma_i \gamma_i' = \Gamma \Lambda^k \Gamma' \tag{2.38}$$
where $\Lambda^k = Diag(\lambda_1^k, \ldots, \lambda_p^k)$.

- **(b)** If all the eigenvalues of (a symmetric matrix) A are positive, we can define the rational powers
$$A^{r/s} = \Gamma \Lambda^{r/s} \Gamma' \tag{2.39}$$
for any $s > 0$ and r are integers, where $\Lambda^{r/s} = Diag(\lambda_1^{r/s}, \ldots, \lambda_p^{r/s})$.

- **(c)** If A is a symmetric singular matrix (at least one eigenvalue of A is zero). then Eq. (2.38) and (2.39) hold if the exponents (k or r/s) are restricted to be non-negative.

The result (2.38) follows easily, since

$$A^2 = A \cdot A = (\Gamma \Lambda \Gamma') \cdot (\Gamma \Lambda \Gamma') = (\Gamma \Lambda^2 \Gamma),$$

and the proof follows by induction. Thus ,

$$A^{-2} = \Gamma \Lambda^{-2} \Gamma'. \tag{2.40}$$

If $\lambda_i > 0$, $\forall i$,

$$A^{-1/2} = \Gamma \Lambda^{-1/2} \Gamma'. \tag{2.41}$$

If $\lambda_i \geq 0$, $\forall i$,

$$A^{1/2} = \Gamma \Lambda^{1/2} \Gamma'. \tag{2.42}$$

The decomposition (2.42) is called *the symmetric square root decomposition* of A.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

By the Spectral Decomposition Theorem, we can write

$$\Sigma = A\Lambda A'$$

where $A = (\alpha_1, \ldots, \alpha_p), \Lambda = Diag.(\lambda_1, \ldots, \lambda_p)$.

Note that some of the λ_i's may be zero. Therefore, the total population variance among $X_1, \ldots, X_p$ is

$$\begin{aligned} \sum_{i=1}^{p} Var(X_i) &= tr(\Sigma) \\ &= tr(A\Lambda A') \\ &= tr(\Lambda A A') \\ &= tr(\Lambda) \\ &= \sum_{i=1}^{p} \lambda_i \\ &= \sum_{i=1}^{p} Var(Z_i) \end{aligned} \tag{2.43}$$

2.7 Statistical Loss Functions

After making forecasts and choosing a proxy for actual data, the researchers should choose a statistical loss function to see how close the forecasts are to their target and compare forecasting performance of challenging models. In the literature, various loss functions have been used to evaluate forecast errors.

Let y_{t+k} be actual data and $\hat{y}_{t+k}$ be forecasting data. Popular measures for forecasting performance are given by the Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), QLIKE Loss Function, R2LOG Loss Function, Mean Absolute Deviation (MAD) and Heteroscedasticity-adjusted Mean Square

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Error (HMSE);

$$\begin{aligned}
MSE &= \frac{1}{n}\sum_{t=1}^{n}(y_{t+K}-\hat{y}_{t+K})^2, \\
QLIKE &= \frac{1}{n}\sum_{t=1}^{n}(ln(\hat{y}_{t+K})-\frac{y_{t+K}}{\hat{y}_{t+K}}), \\
R2LOG &= \frac{1}{n}\sum_{t=1}^{n}[ln(\frac{y_{t+K}}{\hat{y}_{t+K}})]^2, \\
MAD &= \frac{1}{n}\sum_{t=1}^{n}|y_{t+K}-\hat{y}_{t+K}|, \\
HMSE &= \frac{1}{n}\sum_{t=1}^{n}(\frac{y_{t+K}}{\hat{y}_{t+K}}-1)^2.
\end{aligned} \tag{2.44}$$

Since there is no unique criterion indicating the best forecasting model, we have used all of them rather than choosing a single loss function, following Marcucci (2005), as this should provide more comprehensive forecast evolution.

The MSE is the most widely used measure in forecast accuracy. The main drawback of these loss functions is that they penalize both over forecasting and under forecasting equally. Bollerslev and Ghysels (1996) argued that MSE may be unreliable in the presence of heteroscedasticity and proposed HMSE. Moreover, Bollerslev et al. (1994) introduced QLIKE loss function which corresponds to the loss implied by a Gaussian likelihood. The loss function R2LOG is similar to the R^2 of logarithmic version of Mincer-Zarnowitz (1969) (MZ) regression, where $log(\sigma^2_{t,K})$ is regressed on $log(h_{t,K})$ and a constant. More detailed analysis of all these loss functions are provided in Patton and Sheppard (2007). The lower the loss function values, the better forecasting performance.

As well as forecasting volatility accurately, predicting the direction of volatility may be helpful for practitioners while constructing their investment strategies. For this purpose, we also evaluate out-of-sample forecasts by comparing the fractions of the volatility forecast that have same sign of change as the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

actual volatility. We consider the Success Ratio (SR) as follows:

$$SR = \frac{1}{n}\sum_{j=1}^{n} I_{\hat{\sigma}^2_{t+j,k}\cdot\hat{h}_{t+j,k}>0}, \tag{2.45}$$

where I is indicator function ; $k = 1, 2, \cdots, 22$; $\hat{\sigma}^2_{t+j,k} = \sigma^2_{t+j,k} - \bar{\sigma}^2_{t,k}$ and $\hat{h}_{t+j,k} = h_{t+j,k} - \bar{h}_{t,k}$.

2.8 Forecasting Financial Returns

2.8.1 Forecasting Financial Returns Using Mean Equation

Assumption 2.1. *The financial return series are (weakly) stationary processes.*

This assumption can be checked empirically provided that a sufficient number of historical returns are available. For example, one can divide the data into subsamples and check the consistency of the results obtained across the subsamples.

In time series, let P_t denote the series of the financial price at time t on a probability space (Ω, F, P) and r_t be the log return of an asset at time index t (in percent), i.e.

$$\begin{aligned} r_t &= 100 \cdot ln(\frac{P_t}{P_{t-1}}) \\ &= \mu_t + \varepsilon_t, \\ \varepsilon_t &= \eta_t\sqrt{h_t}, \end{aligned} \tag{2.46}$$

where μ_t is mean equation, h_t is conditional variance of errors ε_t, η_t is *i.i.d.* with $D(0, 1)$. The distribution D is generally assumed to be normal, student-t or GED distribution.

The basic idea behind volatility study is that the series $\{r_t\}$ is either serially uncorrelated or with minor lower order serial correlations, but it is a dependent series.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

To put the volatility models in proper perspective, it is informative to consider the conditional mean and variance of r_t given F_{t-1}; that is,

$$\begin{aligned} \mu_t &= E(r_t|F_{t-1}), \\ h_t &= Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}], \end{aligned} \tag{2.47}$$

where F_{t-1} is sigma algebra. As shown, serial dependence of a stock return series r_t is weak if it exists at all. Therefore, the equation for μ_t in (2.46) should be simple, and we assume that r_t follows a simple time series model such as a stationary ARMA(p, q) model with some explanatory variables. In other words, we entertain the model

$$\begin{aligned} r_t &= \mu_t + \varepsilon_t, \\ \mu_t &= \mu_0 + \sum_{i=1}^{n} \beta_i X_{it} + \sum_{s=1}^{p} \phi_s r_{t-s} - \sum_{m=1}^{q} \theta_m \varepsilon_{t-m}, \end{aligned} \tag{2.48}$$

where μ_0, β_i, for $i = 1, ..., n$, ϕ_s, for $s = 1, ..., p$, θ_m, for $m = 1, ..., q$ are constants, n, p, and q are non-negative integers, and X_{it} are explanatory variables.

2.9 Forecasting Volatility of Financial Returns

In time-series, a financial price is transformed to log return series for stationary processes which look like white noise. Mehmet (2008) said financial returns have three characteristics. First is volatility clustering: large changes tend to be followed by large changes and small changes tend to be followed by small changes. Second is fat tailedness (excess kurtosis): financial returns often display a fatter tail than a standard normal distribution and the third is the leverage effect which means that negative returns result in higher volatility than positive returns of the same size. The generalized autoregressive conditional heteroskedasticity (GARCH) models mainly capture these three characteristics of financial returns.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

2.9.1 Model Building

Building a volatility model for an asset return series consists of four steps:

1. Specify a mean equation by testing for serial dependence in the data and, if necessary, building an econometric model (e.g., an ARMA model) for the return series to remove any linear dependence.

2. Use the residuals of the mean equation to test for ARCH effects.

3. Specify a volatility model if ARCH effects are statistically significant and perform a joint estimation of the mean and volatility equations.

4. Check the fitted model carefully and refine it if necessary.

For most asset return series, the serial correlations are weak, if any. Thus, building a mean equation involves removing the sample mean from the data if the sample mean is significantly different from zero. For some daily return series, a simple AR model might be needed. In some cases, the mean equation may employ some explanatory variables.

Testing for ARCH Effect

For ease in notation, let $\varepsilon_t = r_t - \mu_t$ be the residuals of the mean equation. The squared series ε_t^2 is then used to check for conditional heteroscedasticity, which is also known as the ARCH effects. Two tests are available. The first test is to apply the usual Ljung-Box statistics $Q(m)$ to the $\{\varepsilon_t^2\}$ series; see McLeod and Li (1983). The null hypothesis is that the first m lags of ACF of the ε_t^2 series are zero. The second test for conditional heteroscedasticity is the Lagrange multiplier test of Engle (1982). This test is equivalent to the usual F statistic for testing

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\alpha_i = 0 (i = 1, ..., m)$ in the linear regression

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_m \varepsilon_{t-m}^2 + e_t, \ t = m+1, ..., T,$$

where e_t denotes the error term, m is a prespecified positive integer, and T is the sample size.

Specifically, the null hypothesis is $H_0 : \alpha_1 = ... = \alpha_m = 0$. Let $SSR_0 = \sum_{t=m+1}^{T} (\varepsilon_t^2 - \bar{\omega})^2$ where $\bar{\omega} = (1/T) \sum_{t=1}^{T} \varepsilon_t^2$ is the sample mean of ε_t^2 and $SSR_1 = \sum_{t=m+1}^{T} \hat{e}_t^2$, where $\hat{e}_t$ is the least squares residual of the prior linear regression. Then we have

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - 2m - 1)},$$

which is asymptotically distributed as a chi-squared distribution with m degrees of freedom under the null hypothesis. The decision rule is to reject the null hypothesis if $F > \chi_m^2(\alpha)$, where $\chi_m^2(\alpha)$ is the upper $100(1-\alpha)th$ percentile of χ_m^2, or the $p - value$ of F is less than α.

2.9.2 The ARCH Model

The first model that provides a systematic framework for volatility modeling is the ARCH model of Engle (1982). The basic idea of ARCH models is that (a) the shock ε_t of an asset return is serially uncorrelated, but dependent, and (b) the dependence of ε_t can be described by a simple quadratic function of its lagged values. Specifically, an ARCH(m) model assumes that

$$\begin{aligned} \varepsilon_t &= \eta_t \sqrt{h_t}, \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_m \varepsilon_{t-m}^2, \end{aligned} \tag{2.49}$$

where $\{\eta_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and variance equal to one, $\alpha_0 > 0$, and $\alpha_i \geq 0$ for

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$i > 0$. The coefficients α_i must satisfy some regularity conditions to ensure that the unconditional variance of ε_t is finite. In practice, η_t is often assumed to follow the standard normal or a standardized Student-t distribution or a generalized error distribution.

2.9.3 The GARCH Model

Although the ARCH model is simple, it often requires many parameters to adequately describe the volatility process of an asset return. For instance, an ARCH model is needed for the volatility process. Some alternative models must be sought. Bollerslev (1986) proposes a useful extension known as the generalized ARCH (GARCH) model. For a log return series r_t, let $\varepsilon_t = r_t - \mu_t$ be the innovation at time t . Then ε_t follows a GARCH(p, q) model if

$$\begin{aligned} \varepsilon_t &= \eta_t \sqrt{h_t}, \\ h_t &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}, \end{aligned} \tag{2.50}$$

where again η_t is a sequence of i.i.d. random variables with mean zero and variance equal to one, $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$, and $\sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i) < 1$. Here it is understood that $\alpha_i = 0$ for $i > p$ and $\beta_j = 0$ for $j > q$. The latter constraint on $\alpha_i + \beta_i$ implies that the unconditional variance of ε_t is finite, whereas its conditional variance σ_t^2 evolves over time. As before, η_t is often assumed to be a standard normal or standardized student-t distribution or generalized error distribution. Equation (2.50) reduces to a pure ARCH(p) model if $q = 0$. The α_i and β_j are referred to as ARCH and GARCH parameters, respectively.

The strengths and weaknesses of GARCH models can easily be seen by focusing on the simplest GARCH(1,1) model with

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, \tag{2.51}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $0 \leq \alpha_1, \beta_1 \leq 1, (\alpha_1 + \beta_1) < 1$.

First, a large ε_{t-1}^2 or h_{t-1} gives rise to a large h_t . This means that a large ε_{t-1}^2 tends to be followed by another large ε_t^2 , generating, again, the well-known behavior of volatility clustering in financial time series.

Second, it can be shown that if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, then

$$\frac{E(\varepsilon_t^4)}{[E(\varepsilon_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3. \tag{2.52}$$

Consequently, similar to ARCH models, the tail distribution of a GARCH(1,1) process is heavier than that of a normal distribution.

Third, the model provides a simple parametric function that can be used to describe the volatility evolution.

For one-step-ahead, volatility forecasting from GARCH(1,1) model is shown in equation (2.50),

$$\hat{h}_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 h_t. \tag{2.53}$$

In order to forecast volatility for $2-step-ahead$, the fact $E\left[\varepsilon_{t+1}^2 | F_t\right] = \hat{h}_{t+1}$ Then,

$$\begin{aligned} \hat{h}_{t+2} &= \alpha_0 + \alpha_1 \varepsilon_{t+1}^2 + \beta_1 \hat{h}_{t+1}, \\ &= \alpha_0 + \{\alpha_1 + \beta_1\} \hat{h}_{t+1}. \end{aligned} \tag{2.54}$$

Therefore, the forecasting formula can be generalized for k-step-ahead forecast as follows,

$$\hat{h}_{t+k} = \alpha_0 \Sigma_{i=1}^{k-1} (\alpha_1 + \beta_1)^{i-1} + (\alpha_1 + \beta_1)^{k-1} \hat{h}_{t+1}. \tag{2.55}$$

2.9.4 The EGARCH(1,1) Model

The main problem of the standard GARCH model is that positive and negative shocks have the same effects on volatility. However, impacts of positive

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and negative shocks on the volatility may be asymmetric (Black, 1976). Several alternative GARCH models have been proposed to capture the asymmetric nature of volatility responses. One of them is the exponential GARCH (EGARCH) model of Nelson (1991). In this specification, conditional variance is modeled in logarithmic form, which means that there is no restriction on parameters in the model to avoid negative variances. The conditional variance equation of EGARCH(1,1) is defined as

$$ln(h_t) = \alpha_0 + \alpha_1 |\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}| + \beta_1 ln(h_{t-1}) + \xi \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}, \tag{2.56}$$

where ξ is the asymmetry parameter to capture the leverage effect. The EGARCH process is covariance stationary if the condition $\beta_1 < 1$ is satisfied.

One-step-ahead volatility forecast is computed as

$$ln(\hat{h}_{t+1}) = \alpha_0 + \alpha_1 |\frac{\varepsilon_t}{\sqrt{h_t}}| + \beta_1 ln(h_t) + \xi \frac{\varepsilon_t}{\sqrt{h_t}}. \tag{2.57}$$

Then, the multi-step-ahead volatility forecast is computed as

$$ln(\hat{h}_{t+k}) = \alpha_0 + \beta_1 ln(\hat{h}_{t+k-1}). \tag{2.58}$$

2.9.5 GJR-GARCH(1,1) Model

Another model that allows for different impacts of positive and negative shocks on volatility is the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). The GJR-GARCH(1,1) model takes the following form,

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 (1 - I_{\{\varepsilon_{t-1} > 0\}}) + \beta_1 h_{t-1} + \xi \epsilon_{t-1}^2 (I_{\{\varepsilon_{t-1} > 0\}}), \tag{2.59}$$

where $I_{\{\varepsilon_{t-1} > 0\}}$ is equal to one when ε_{t-1} is greater than zero. The conditions $\alpha_0 > 0$, $(\alpha_1 + \xi)/2 > 0$ and $\beta_1 > 0$ must be satisfied to ensure positive conditional variance. Also, the process is covariance-stationary if $[(\alpha_1 + \xi)/2] + \beta_1 < 1$. Then,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

unconditional variance is defined as

$$\sigma^2 = \frac{\alpha_0}{1-(\alpha_1+\xi)/2+\beta_1}. \tag{2.60}$$

One-step-ahead volatility forecast is computed as

$$\hat{h}_{t+1} = \alpha_0 + \alpha_1\varepsilon_t^2(1 - I_{\{\varepsilon_t>0\}}) + \beta_1 h_t + \xi\epsilon_t^2(I_{\{\varepsilon_t>0\}}). \tag{2.61}$$

Then, multi-step-ahead volatility forecast is computed as

$$\hat{h}_{t+k} = \alpha_0 + \left(\frac{\alpha_1+\xi}{2} + \beta_1\right)\hat{h}_{t+k-1}. \tag{2.62}$$

2.9.6 Distributions for Standardized Errors

The standard normal distribution sometimes may not be enough to describe the fat-tail feature of the financial returns. In order to capture the fat-tail feature in the data, Bollerslev (1987) and Nelson (1991) proposed the student-t and generalized error distributions (GED), respectively. Although these two distributions are also symmetric, they have fatter tails than the normal distribution captures. In this study, we assume that standardized errors follow the student-t and GED distributions as well as the normal distribution.

In the case of normal distribution, the conditional probability density function of errors is defined as:

$$f(\varepsilon_t\,|\varepsilon_{t-1},\ \varepsilon_{t-2}, ...) = \frac{1}{\sqrt{2\pi h_t}}\exp(-\frac{1}{2}\cdot\frac{\varepsilon_t^2}{h_t}). \tag{2.63}$$

When errors are assumed to follow student-t distribution, the conditional probability density function of errors is defined as:

$$f(\varepsilon_t\,|\varepsilon_{t-1},\ \varepsilon_{t-2}, ...) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\pi(\nu-2)}\Gamma(\nu/2)}\frac{1}{\sqrt{h_t}}[1+\frac{\varepsilon_t^2}{h_t(\nu-2)}]^{\frac{-(\nu+1)}{2}}, \tag{2.64}$$

where $\Gamma(.)$ is the Gamma function, and ν is the degree of freedom which must be greater than 2. When $\nu \to \infty$, the student-t distribution becomes the normal distribution. So, lower values of ν imply fatter tails.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

If GED is considered as a distribution assumption, the conditional probability density function of errors is defined as:

$$f(\varepsilon_t \,|\varepsilon_{t-1},\ \varepsilon_{t-2}, ...) = \frac{\nu \exp\left[\left(\frac{-1}{2}\right)\left|\frac{\varepsilon_t}{\delta\sqrt{h_t}}\right|^{\nu}\right]}{\delta 2^{\left(\frac{\nu+1}{\nu}\right)}\Gamma(1/\nu)\sqrt{h_t}}, \tag{2.65}$$

where $\delta = \sqrt{\left(\frac{2^{(-2/\nu)}\Gamma(1/\nu)}{\Gamma(3/\nu)}\right)}$, Γ is the Gamma function and ν is the tail thickness parameter. When $\nu = 2$, GED becomes a standard normal distribution. It has fatter tails than the normal distribution in the case of $\nu < 2$, whereas the normal distribution has fatter tails than GED in the case of $\nu > 2$.

The parameters in GARCH type models are generally estimated by the Maximum Likelihood Estimation (MLE) method. The idea behind this method is to determine the set of parameters that maximize the likelihood (probability) function of the sample data under assumption about standardized residuals. This is done by forming the likelihood function. Since the maximum of likelihood function can not be obtained analytically for GARCH type models, numerical optimization techniques are used to find the set of parameters that maximize likelihood function.

The log-likelihood functions for a sample with T observations are as follows.

For the normal distribution,

$$L = -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(2\pi) + \ln(h_t) + \frac{\varepsilon_t^2}{h_t}\right]. \tag{2.66}$$

For the student-t distribution,

$$\begin{aligned} L &= T\cdot\ln[\Gamma((\nu+1)/2)] - \ln[\Gamma(\nu/2)] - \frac{1}{2}\ln[\pi(\nu-2)] \\ &\quad - \frac{1}{2}\sum_{t=1}^{T}[\ln(h_t^2) + (\nu+1)\ln(1+\frac{\varepsilon_t^2}{h_t^2(\nu-2)})]. \end{aligned} \tag{2.67}$$

For the GED distribution,

$$L = \sum_{t=1}^{T}\left[\ln(\nu/\delta) - \frac{1}{2}\left|\frac{\varepsilon_t}{\delta\sqrt{h_t}}\right|^{\nu} - \left(\frac{\nu+1}{\nu}\right)\ln(2) - \ln\left[\Gamma(1/\nu)\right] - \frac{1}{2}\ln(h_t)\right]. \tag{2.68}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

2.10 Markov Regime Switching

The Markov switching model of Hamilton (1989), also known as the regime switching model, is one of the most popular nonlinear time series models in the literature. This model involves multiple structures (equations) that can characterize the time series behaviors in different regimes. By permitting switching between these structures, this model is able to capture more complex dynamic patterns.

A novel feature of the Markov switching model is that the switching mechanism is controlled by an unobservable state variable that follows a first-order Markov chain. In particular, the Markovian property regulates that the current value of the state variable depends on its immediate past value. As such, a structure may prevail for a random period of time, and it will be replaced by another structure when a switching takes place. This is in sharp contrast with the random switching model of Quandt (1972) in which the events of switching are independent over time.

The Markov switching model is also different from the models of structural changes. While the former allows for frequent changes at random time points, the latter admits only occasion and exogenous changes. The Markov switching model is therefore suitable for describing correlated data that exhibit distinct dynamic patterns during different time periods.

Numerous empirical evidences suggest that the time series behaviours of economic and financial variables may exhibit different patterns over time. Instead of using one model for the conditional mean of a variable, it is natural to employ several models to represent these patterns. A Markov switching model is constructed by combining two or more dynamic models via a Markovian switching mechanism. Following Hamilton (1989, 1994), we shall focus on the Markov

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

switching AR model. In this section, we first illustrate the features of Markovian switching using a simple model and then discuss more general model specifications.

2.10.1 A Simple Model

Let S_t denote an unobservable state variable assuming the value one or zero. A simple switching model for the variable z_t involves two AR specifications:

$$z_t = \begin{cases} \alpha_0 + \beta z_{t-1} + \varepsilon_t, S_t = 0 \ , \\ \alpha_0 + \alpha_1 + \beta z_{t-1} + \varepsilon_t, S_t = 1. \end{cases} \tag{2.69}$$

where $|\beta| < 1$ and ε_t are *i.i.d.* random variables with mean zero and variance σ_ε^2.

This is a stationary AR(1) process with mean $\alpha_0/(1-\beta)$ when $S_t = 0$, and it switches to another stationary AR(1) process with mean $(\alpha_0 + \alpha_1)/(1-\beta)$ when S_t changes from 0 to 1. Then provided that $\alpha_1 \neq 0$, this model admits two dynamic structures at different levels, depending on the value of the state variable S_t. In this case, z_t are governed by two distributions with distinct means, and S_t determines the switching between these two distributions (regimes).

When $S_t = 0$ for $t = 1, ..., \tau_0$ and $S_t = 1$ for $t = \tau_0 + 1, .., T$, the model (2.69) is the model with a single structural change in which the model parameter experiences one (and only one) abrupt change after $t = \tau_0$. When S_t are independent Bernoulli random variables, it is the random switching model of Quandt (1972). In the random switching model, the realization of S_t is independent of the previous and future states so that z_t may *jump* (switching back and forth between different states). If S_t is postulated as the indicator variable $1_{\lambda_t \leqslant c}$ such that $S_t = 0$ or 1 depending on whether the value of λ_t is greater than the cut-off (threshold) value c, equation (2.69) becomes a threshold model. It is quite common to choose a lagged dependent variable (say, z_{t-d}) as the variable t.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

While these models are all capable of characterizing the time series behaviours in two regimes, each of them has its own limitations. For the model with a single structural change, it is very restrictive because only one change is admitted. Although extending this model to allow for multiple changes is straightforward, the resulting model estimation and hypothesis testing are typically cumbersome; see e.g., Bai and Perron (1998) and Bai (1999). Moreover, changes in such models are solely determined by time which is exogenous to the model. The random switching model, by contrast, permits multiple changes, yet its state variables are still exogenous to the dynamic structures in the model. This model also suffers from the drawback that the state variables are independent over time and hence may not be applicable to time series data. On the other hand, switching in the threshold model is dependent and endogenous and results in multiple changes. Choosing a suitable variable λ_t and the threshold value c for this model is usually a difficult task, however.

One approach to circumventing the aforementioned problems is to consider a different specification for S_t. In particular, suppose that S_t follows a first order Markov chain with the following transition matrix:

$$P = \begin{pmatrix} Pr(S_t = 0|S_{t-1} = 0) & Pr(S_t = 1|S_{t-1} = 0) \\ Pr(S_t = 0|S_{t-1} = 1) & Pr(S_t = 1|S_{t-1} = 1) \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix},$$

where $p_{ij}(i,j = 0,1)$ denote the transition probabilities of $S_t = j$ given that $S_{t-1} = i$. Clearly, the transition probabilities satisfy $p_{i0} + p_{i1} = 1$. The transition matrix governs the random behavior of the state variable, and it contains only two parameters (p_{00} and p_{11}). The model (2.69) with the Markovian state variable is known as a Markov switching model. The Markovian switching mechanism was first considered by Goldfeld and Quandt (1973). Hamilton (1989) presents a thorough analysis of the Markov switching model and its estimation method; see

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

also Hamilton (1994) and Kim and Nelson (1999).

In the Markov switching model, the properties of z_t are jointly determined by the random characteristics of the driving innovations ε_t and the state variable S_t. In particular, the Markovian state variable yields random and frequent changes of model structures, and its transition probabilities determine the persistence of each regime. While the threshold model also possesses similar features, the Markov switching model is relatively easy to implement because it does not require choosing a priori the threshold variable λ_t. Instead, the regime classification in this model is probabilistic and determined by data. A difficultly with the Markov switching model is that it may not be easy to interpret because the state variables are unobservable.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER III

FORECASTING WITH CONSTANT MEAN EQUATION AND VOLATILITY MODELS

In this chapter, we use the mean equation with constant. We consider two types of equations; the first constant mean equation employ the same constant for all regimes. In the second constant mean equation, the constant depends on the regime. Nevertheless, we forecast the volatility of the financial return with heteroskedasticity i.e. GARCH, EGARCH, GJR-GARCH. Next we describe models for forecasting financial return with constant mean equations. After that we use the Markov Regime Switching include GARCH model, called Markov Regime Switching GARCH (MRS-GARCH). Finally, we apply these models to some financial instruments i.e. the gold price and the SET50 Index.

3.1 Forecasting Financial Returns with Constant Mean Equation

In time series, let P_t denote the series of the financial price at time t on a probability space (Ω, F, P) and r_t be the log return of an asset at time index t (in percent) in (2.46), i.e.

$$\begin{aligned} r_t &= 100 \cdot ln(\frac{P_t}{P_{t-1}}) \\ &= \mu_t + \varepsilon_t. \end{aligned}$$

We use the two type of constant mean equations in (2.46).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The first constant mean equation uses the same constant for all regimes, i.e.

$$r_t = \delta + \varepsilon_t. \tag{3.1}$$

The second constant mean equation uses a constant depending on the regime, i.e.

$$r_t = \delta_{S_t} + \varepsilon_t, \tag{3.2}$$

where S_t is a regime, as defined before in chapter 2.

3.2 Markov Regime Switching GARCH Models

An important technique for analyzing structural breaks in financial returns is the Markov Switching model of Hamilton (1989, 1990). In his study, Hamilton extended the Markov switching regression model of Goldfeld and Quandt (1973) to the time series framework and analyzed the growth rate of U.S. real GNP. In Hamilton's model, the process is allowed to switch stochastically between different regimes. Also, regimes are usually governed by a first order Markov Chain process. In our study, for simplicity, we assume that there are two unobservable regimes.

Let $\{P_t\}$ denote the series of the financial price at time t and $\{r_t\}_{t>0}$ be a sequence of random variables on a probability space (Ω, F, P) . The index t denotes the daily closing observations and $t = -R+1, ..., n$. The sample period consist of an estimation (or in-sample) period with R observations $(t = -R+1, ..., 0)$, and an evolution (or out-of-sample) period with n observations $(t = 1, ..., n)$; r_t denotes the logarithmic return (in percent) on the financial price at time t in equation (2.46)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The GARCH(1,1) model for the series of the returns can be written as:

$$\begin{aligned} r_t &= \delta + \varepsilon_t \\ &= \delta + \eta_t \sqrt{h_t}, \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, \end{aligned} \tag{3.3}$$

where $\alpha_0 > 0, \alpha_1 \geq 0$ and $\beta_1 \geq 0$ are assumed to be non-negative real constants, to ensure that $h_t \geq 0$. We assume η_t is an i.i.d. process with zero mean and unit variance.

The parameters of the GARCH model are generally considered as constants. But the movement of financial returns between recession and expansion is different, and may result in differences in volatility. Gray (1996) extended the GARCH model to the MRS-GARCH model in order to capture regime changes in volatility with unobservable state variables. It was assumed that those unobservable state variables satisfy the first order Markov Chain process.

The MRS-GARCH model with only two regimes can be represented as follows:

$$\begin{aligned} r_t &= \delta_{S_t} + \varepsilon_t \\ &= \delta_{S_t} + \eta_t \sqrt{h_{t,S_t}}, \end{aligned} \tag{3.4}$$

and

$$h_{t,S_t} = \alpha_{0,S_t} + \alpha_{1,S_t} \varepsilon_{t-1}^2 + \beta_{1,S_t} h_{t-1}, \tag{3.5}$$

where $S_t = 1$ or 2 , δ_{S_t} is the mean and h_{t,S_t} is the volatility under regime S_t on F_{t-1}, both are measurable function of $F_{t-\tau}$ for $\tau \leq t-1$. In order to ensure easily the positivity of the conditional variance, we impose the restrictions $\alpha_{0,S_t} > 0, \alpha_{1,S_t} \geq 0$ and $\beta_{1,S_t} \geq 0$. The sum $\alpha_{1,S_t} + \beta_{1,S_t}$ measures the persistence of a shock to the conditional variance.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The unobserved regime variable S_t is governed by a first order Markov Chain with constant transition probabilities given by

$$Pr(S_t = i|S_{t-1} = j) = p_{ji}, \tag{3.6}$$

for $i, j = 1, 2$.

In matrix notation

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}, \tag{3.7}$$

for $p = p_{11}$ and $q = p_{22}$.

3.2.1 Forecasting Volatility of Financial Returns with MRS-GARCH

In the MRS-GARCH model with two regimes, Klaassen (2002) forecast volatility for *k-step-ahead* by using the recursive method as in the standard GARCH model where $k = 1, 2, ..., n$. In order to compute the multi-step-ahead volatility forecasts, we firstly compute a weighted average of the multi-step-ahead volatility forecasts in each regime where the weights are the *prediction probability* $(Pr(S_{t+\tau} = i|F_{t-1}))$ for $i = 1, 2$.

Since there is no serial correlation in the returns, the *k-step-ahead* volatility forecast at time t depends on the information at time $t-1$. Let $\hat{h}_{t,t+k}$ denotes the time t aggregated volatility forecasts for the next k steps. It can be calculated as follows: (See, for example Marcucci (2005) page 8)

$$\begin{aligned} \hat{h}_{t,t+k} &= \sum_{\tau=1}^{k} \hat{h}_{t,t+\tau} \\ &= \sum_{\tau=1}^{k} [\sum_{i=1}^{2} Pr(S_{t+\tau} = i|F_{t-1}) \hat{h}_{t,t+\tau,S_{t+\tau}=i}], \end{aligned} \tag{3.8}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\hat{h}_{t,t+\tau,S_{t+\tau}=i}$ indicates the $\tau - step - ahead$ volatility forecast in regime i made at time t and can be calculated recursively as follows:

$$\begin{aligned}
\hat{h}_{t,t+\tau,S_{t+\tau}=i} &= E_{t-1}[h_{t+\tau}|S_{t+\tau}=i] \\
&= E_{t-1}[\alpha_0 + \alpha_1\varepsilon^2_{t+\tau-1} + \beta_1 h_{t+\tau-1}|S_{t+\tau}=i] \\
&= \alpha_{0,S_{t+\tau}=i} + \alpha_{1,S_{t+\tau}=i}E_{t-1}[\varepsilon^2_{t+\tau-1}|S_{t+\tau}=i] + \beta_{1,S_{t+\tau}=i}E_{t-1}[h_{t+\tau-1}|S_{t+\tau}=i] \\
&= \alpha_{0,S_{t+\tau}=i} + \alpha_{1,S_{t+\tau}=i}E_{t-1}[E_{t-1}[\varepsilon^2_{t+\tau-1}|S_{t+\tau-1}=j]|S_{t+\tau}=i] \\
&\quad + \beta_{1,S_{t+\tau}=i}E_{t-1}[h_{t+\tau-1}|S_{t+\tau}=i] \\
&= \alpha_{0,S_{t+\tau}=i} + (\alpha_{1,S_{t+\tau}=i} + \beta_{1,S_{t+\tau}=i})E_{t-1}[h_{t+\tau-1}|S_{t+\tau}=i]. \qquad (3.9)
\end{aligned}$$

Also, in generally the prediction probability in (3.8) is computed as:

$$\begin{pmatrix} Pr(S_{t+\tau}=1|F_{t-1}) \\ Pr(S_{t+\tau}=2|F_{t-1}) \end{pmatrix} = P^{\tau+1} \begin{pmatrix} Pr(S_{t-1}=1|F_{t-1}) \\ Pr(S_{t-1}=2|F_{t-1}) \end{pmatrix},$$

where P defined in (3.7) and $Pr(S_{t-1}=i|F_{t-1})$ will be discussed in (3.14).

Lastly, we compute the expectation part $E_{t-1}[h_{t+\tau-1}|S_{t+\tau}=i]$ in (3.9) as follows:

$$\begin{aligned}
E_{t-1}[h_{t+\tau-1}|S_{t+\tau}=i] &= E_{t-1}[[E_{t-1}[r^2_{t+\tau-1}|S_{t+\tau-1}=j] \\
&\quad -[E_{t-1}[r_{t+\tau-1}|S_{t+\tau-1}=j]]^2]|S_{t+\tau}=i] \\
&= E_{t-1}[E_{t-1}[r^2_{t+\tau-1}|S_{t+\tau-1}=j]|S_{t+\tau}=i] \\
&\quad -E_{t-1}[[E_{t-1}[r_{t+\tau-1}|S_{t+\tau-1}=j]]^2|S_{t+\tau}=i] \qquad (3.10)
\end{aligned}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where

$$\begin{aligned}
& E_{t-1}[E_{t-1}[r^2_{t+\tau-1}|S_{t+\tau-1}=j]|S_{t+\tau}=i] \\
= & \sum_{j=1}^{2} E_{t-1}[r^2_{t+\tau-1}|S_{t+\tau-1}]Pr(S_{t+\tau-1}=j|S_{t+\tau}=i,F_{t-1}) \\
= & \sum_{j=1}^{2} E_{t-1}[(\delta+\varepsilon_{t+\tau-1})^2|S_{t+\tau-1}]Pr(S_{t+\tau-1}=j|S_{t+\tau}=i,F_{t-1}) \\
= & \sum_{j=1}^{2} E_{t-1}[\delta^2+2\delta\varepsilon_{t+\tau-1}+\varepsilon^2_{t+\tau-1}|S_{t+\tau-1}]Pr(S_{t+\tau-1}=j|S_{t+\tau}=i,F_{t-1}) \\
= & \sum_{j=1}^{2} \tilde{p}_{ji,t-1}[\delta^2_{S_{t+\tau-1}=j}+h_{t+\tau-1,S_{t+\tau-1}=j}], \qquad (3.11)
\end{aligned}$$

and

$$\begin{aligned}
\tilde{p}_{ji,t-1} &= Pr\left(S_{t+\tau-1}|S_{t+\tau},F_{t-1}\right) \\
&= \frac{p_{ji}Pr\left(S_{t+\tau-1}=j|F_{t-1}\right)}{Pr\left(S_{t+\tau}=i\right)|F_{t-1})}.
\end{aligned}$$

Similarly , we compute the second term of the right hand side in (3.10) by

$$E_{t-1}[[E_{t-1}[r_{t+\tau-1}|S_{t+\tau-1}=j]]^2|S_{t+\tau}=i]=\sum_{j=1}^{2}\tilde{p}_{ji,t-1}[\delta_{S_{t+\tau-1}=j}]^2. \qquad (3.12)$$

Substituting both (3.11) and (3.12) into (3.10) we obtain

$$\begin{aligned}
& E_{t-1}[h_{t+\tau-1}|S_{t+\tau}=i] \\
= & \sum_{j=1}^{2}\tilde{p}_{ji,t-1}[\delta^2_{S_{t+\tau-1}=j}+h_{t+\tau-1,S_{t+\tau-1}=j}]-\sum_{j=1}^{2}\tilde{p}_{ji,t-1}[\delta_{S_{t+\tau-1}=j}]^2. \qquad (3.13)
\end{aligned}$$

In the next step, we will compute those regime probabilities

$$p_{it}=Pr(S_t=i|F_{t-1}),$$

for $i=1,2$. Note that when the regime probabilities are based on information up to time t, we describe this as *filtered probability* $(Pr(S_t=i|F_t)$.

In order to compute the regime probabilities, we denote

$$f_{1t}:=f(r_t|S_t=1,F_{t-1}),$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$f_{2t} := f(r_t|S_t = 2, F_{t-1}).$$

Then, the conditional distribution of the return series r_t becomes a mixture-of-distribution model in which the mixing variable is regime probability p_{it}. That is

$$r_t|F_{t-1} \sim \begin{cases} f(r_t|S_t = 1, F_{t-1}) & \text{with probability } p_{1t}, \\ f(r_t|S_t = 2, F_{t-1}) & \text{with probability } p_{2t} = 1 - p_{1t}, \end{cases}$$

where $f(r_t|S_t, F_{t-1})$ denotes one of the assumed conditional distributions for errors: Normal distribution (N), Student-t distribution with only single degree of freedom (t) or double degree of freedom ($2t$) and Generalized error distributions (GED).

We shall compute regime probabilities recursively by following two steps (Kim, and Nelson (1999), page 63):

Step 1: Given $Pr(S_{t-1} = j|F_{t-1})$ at the end of the time $t-1$ the regime probabilities $p_{it} = Pr(S_t = i|F_{t-1})$ are computed as:

$$Pr(S_t = i|F_{t-1}) = \sum_{j=1}^{2} Pr(S_t = i, S_{t-1} = j|F_{t-1}).$$

Since the current regime (S_t) only depends on the regime one period ago (S_{t-1}), then

$$\begin{aligned} Pr(S_t = i|F_{t-1}) &= \sum_{j=1}^{2} Pr(S_t = i, S_{t-1} = j|F_{t-1}) \\ &= \sum_{j=1}^{2} Pr(S_t = i|S_{t-1} = j) Pr(S_{t-1} = j|F_{t-1}) \\ &= \sum_{j=1}^{2} p_{ji} Pr(S_{t-1} = j|F_{t-1}). \end{aligned}$$

Step 2: At the end of the time t, when observed return at time t (r_t) the information at time t set $F_t = \{F_{t-1}, r_t\}$, the $Pr(S_t = i|F_t)$ is calculated as follows:

$$\begin{aligned} Pr(S_t = i|F_t) &= Pr(S_t = i|r_t, F_{t-1}) \\ &= \frac{f(r_t, S_t = i|F_{t-1})}{f(r_t|F_{t-1})}, \end{aligned}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $f(r_t, S_t = i|F_{t-1})$ is the joint density of returns, and the unobserved regime at state i for $i = 1, 2$ variables which can be written as follows:

$$\begin{aligned} f(r_t, S_t = i|F_{t-1}) &= f(r_t|S_t = i, F_{t-1})f(S_t = i|F_{t-1}) \\ &= f(r_t|S_t = i, F_{t-1})Pr(S_t = i|F_{t-1}), \end{aligned}$$

and $f(r_t|F_{t-1})$ is the marginal density function of returns which can be constructed as follows:

$$\begin{aligned} f(r_t|F_{t-1}) &= \sum_{i=1}^{2} f(r_t, S_t = i|F_{t-1}) \\ &= \sum_{i=1}^{2} f(r_t|S_t = i, F_{t-1})Pr(S_t = i|F_{t-1}). \end{aligned}$$

We use Bayesian arguments

$$\begin{aligned} Pr(S_t = i|F_t) &= \frac{f(r_t, S_t = i|F_{t-1})}{f(r_t|F_{t-1})} \\ &= \frac{f(r_t|S_t = i, F_{t-1})Pr(S_t = i|F_{t-1})}{\sum_{j=1}^{2} f(r_t|S_t = i, F_{t-1})Pr(S_t = i|F_{t-1})} \\ &= \frac{f_{it}p_{it}}{\sum_{i=1}^{2} f_{it}p_{it}}. \end{aligned} \tag{3.14}$$

Then, all regime probabilities p_{it} can be computed by iterating these two steps. However, at the beginning of iteration, $Pr(S_0 = i|F_0)$ for $i = 1, 2$ are necessary to start the iteration. Hamilton (1989, 1990) suggest we should use unconditional regime probabilities of S_t, i.e.

$$\begin{aligned} \pi_1 &= Pr(S_0 = 1|F_0) \\ &= \frac{1-q}{2-p-q}, \end{aligned}$$

and

$$\begin{aligned} \pi_2 &= Pr(S_0 = 2|F_0) \\ &= \frac{1-p}{2-p-q}. \end{aligned}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Given initial values for regime probabilities, conditional mean and conditional variance in each regime, the parameters of the MRS-GARCH model can be obtained by maximizing numerically the log-likelihood function. The log-likelihood function is constructed recursively similar to that in GARCH models.

3.2.2 Forecasting Price with MRS-GARCH

We forecast financial price at *k-step-ahead* with MRS-GARCH models. Denote by $\hat{r}_{t,t+k}$ the *k-step-ahead* forecasting logarithm return of the financial price at time t depending on F_{t-1}.

We compute as:

$$\begin{aligned} \hat{r}_{t,t+k} &= E_{t-1}[r_{t+k}] \\ &= \sum_{i=1}^{2} Pr(S_{t+k} = i|F_{t-1})\hat{r}_{t,t+k,S_{t+k}=i}, \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} \hat{r}_{t,t+k,S_{t+k}=i} &= E_{t-1}[r_{t+k}|S_{t+k} = i] \\ &= E_{t-1}[\delta + \varepsilon_{t+k}|S_{t+k} = i] \\ &= E_{t-1}[\delta|S_{t+k} = i] + E_{t-1}[\varepsilon_{t+k}|S_{t+k} = i] \\ &= \sum_{j=1}^{2} Pr(S_{t+k-1} = j|S_{t+k} = i, F_{t-1})\delta_{S_{t+k}=i} \\ &= \sum_{j=1}^{2} \tilde{p}_{ji,t-1}\delta_{S_{t+k}=i}. \end{aligned}$$

Applied to forecasting financial price *one-step-ahead*, we combine the log returns (3.15) and (2.46) of the financial price to obtain;

$$\hat{P}_{t+1} = P_t \cdot exp\left[\frac{\sum_{i=1}^{2} Pr(S_{t+1} = i|F_{t-1}) \sum_{j=1}^{2} \tilde{p}_{ji,t-1}\delta_{S_{t+k}=i}}{100}\right]. \tag{3.16}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

3.3 Forecasting Volatility of Gold Price using Markov Regime Switching and Trading Strategy

Gold is a precious metal which is also classed as a commodity and a monetary asset. Gold has acted as a multifaceted metal through the centuries, possessing similar characteristics to money in that it acts as a store of wealth, a medium of exchange and a unit of value. Gold has also played an important role as a precious metal with significant portfolio diversification properties. Gold is used in industrial components, jewellery and as an investment asset. The quantity of gold required is determined by the quantity demanded for industry investment and jewellery use. Therefore an increase in the quantity demanded by the industry will lead to an increase in the price of the metal.

The changing price of gold can also be the result of a change in the Central Banks holding of these precious metals. In addition, changes in the rate of inflation, currency markets, political harmony, equity markets, and producer and supplier hedging, all affect the price equilibrium.

Gold futures are an alternative investment tool which rely on the gold price movement. The investors can benefit from the gold futures investment by making profit from both directions, either up or down, which is like stock index futures trading. In addition, gold futures can also hedge against gold price fluctuations or stock market volatility due to the negative correlation to the stock market. This will provide a greater opportunity to make profit when the stock market declines during an economic downturn.

Gold futures is a futures contract with gold (96.5 percent purity) as an underlying asset. Gold is the oldest precious metal known to man and for thousands of years it has been valued as a global currency, a commodity, an investment and

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

simply an object of beauty. The characteristic of gold price movement that do not correlate with the equity market makes gold futures a very interesting investment option. Holding precious metals such as gold in a portfolio can give apparent benefits in the form of speculative gains, investment gains, hedging against macroeconomic and geopolitical risk and or wealth preservation.

Edel Tully et al., (2005) have investigated how the Asymmetric power GARCH model can capture the dynamics of the gold market. Results suggest that the APGARCH model provides the most adequate description for the gold price.

In this chapter, we use GARCH, EGARCH, GJR-GARCH and MRS-GARCH models to forecast the volatility of gold prices and to compare their performance. Moreover we shall use this estimated volatility to forecast the closing price of gold. Finally, we apply the forecasting price of gold to trading in gold future contracts with a maturity date of August 2011 (GF10Q11).

3.3.1 Empirical Methodology and Model Estimation Results of Gold Price

The data set used in this study are the daily closing prices P_t over the period 4/01/2007 through 31/08/2011 ($t = 1, \ldots, 1,213$ observations). The data set was obtained from the basis of the London Gold Market Fixing Limited on day and the foreign exchange rate for Baht to US dollars announced by TFEX (The Thailand Futures Exchange) on day, after conversion for weight and fineness. The data set is divided into in-sample (R =1,192 observations) and out-of-sample (n =21 observations). The plot of P_t and log returns series (r_t) are given in Figure 3.1. Plot P_t and r_t displays usual properties of financial data series. As expected, volatility is not constant over time and exhibits volatility clustering with large

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

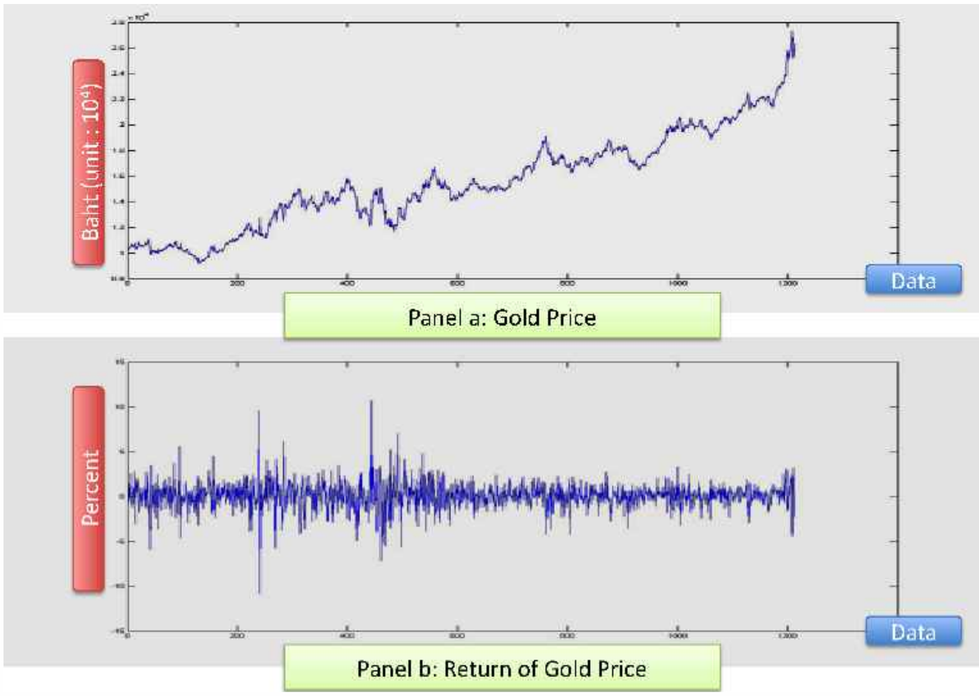


Figure 3.1 Graph of (a) closing gold price and (b) log returns series for the period 4/01/2007 through 30/08/2011.

changes in the indices often followed by large changes, and small changes often followed by small changes.

Descriptive statistics of r_t are presented in Table 3.1. As Table 3.1 shows, r_t has a positive average return of 0.074 percent. The daily standard deviation is 1.537 percent. The series also displays a negative skewness of -0.102 and an excess kurtosis of 9.457. These values indicate that the returns are not normally distributed, namely they have fatter tails because skewness does not equal zero and kurtosis is greater than 3. Also, the Jarque-Bera test statistic of 2,107.620 confirms the non-normality of r_t . And the Augmented Dickey-Fuller test of -35.873 indicates that r_t is stationary. The autocorrelation functions (ACF) tests the significance level of autocorrelation in Table 3.2, when we apply Ljung and Box

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.1 Summary statistics of gold price log returns series(r_t).

Statistic	Return
Min	-10.823
Max	10.71
Mean	0.074
Standard deviation	1.537
Skewness	-0.102
Kurtosis	9.457
Jarque-Bera Normality test	2,107.620 (P-value= 0.000)
Augmented Dickey-Fuller test	-35.873 (P-value= 0.000)

Q-test. The null hypothesis of the test is that there is no serial correlation in the return series up to the specified lag. Serial correlation in the P_t is confirmed as non-stationary but r_t is stationary. Because the serial correlation in the squared returns is non-stationary this suggests conditional heteroskedasticity. Therefore, we analyze the significance of autocorrelation in the squared mean adjusted return $(r_t - \delta)^2$ series by Ljung-Box Q-test and apply Engle's ARCH test.

3.3.2 Empirical Methodology of Gold Price

This empirical part adopts GARCH type and MRS-GARCH(1,1) models to estimate the volatility of the P_t. The GARCH type models considered are GARCH (1,1), EGARCH (1,1) and GJR-GARCH (1,1). In order to account for the fat tail feature of financial returns, we consider three different distributions for the innovations: Normal (N), Student-t (t) and Generalized Error Distributions (GED).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.2 ACF of gold closing price, log returns series, square return and results for Engle's ARCH Test.

Lags		1	5	10	22
	ACF	0.99	0.97	0.94	0.89
ACF of Gold Price	LBQ Test	1202.13	5885.60	11458.61	23713.86
	P-value	0.00	0.00	0.00	0.00
	ACF	−0.04	0.02	0.01	−0.05
ACF of Gold return	LBQ Test	1.47	6.95	19.03	34.41
	P-value	0.23	0.23	0.04	0.05
	ACF	0.24	0.03	0.01	0.06
ACF of Gold square return	LBQ Test	67.59	77.34	90.10	178.44
	P-value	0.00	0.00	0.00	0.00
	ARCH test	67.68	70.70	79.01	126.93
Result for Engle test	P-value	0.00	0.00	0.00	0.00

GARCH type Models

Table 3.3 presents an estimation of the results for GARCH type models. It is clear from the table that almost all parameter estimates including δ in GARCH type models are highly significant at $\alpha = 0.01$. However, the asymmetry effect term ξ in the EGARCH models is significantly different from zero, which indicates unexpected negative returns implying higher conditional variance as compared to same size positive returns. All models display strong persistence in volatility ranging from 0.9654 to 0.9741 unless EGARCH models are very low, that is, volatility is likely to remain high over several price periods once it increases.

Markov Regime Switching GARCH Models

Estimation results and summary statistics of MRS-GARCH models are presented in Table 3.4. Most parameter estimates in MRS-GARCH are significantly

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

different from zero at least 95 percent confidence level. But α_0 and β_1 are insignificant in some states. All models display strong persistence in volatility, that is, volatility is likely to remain high over several price periods, once it increases.

3.3.3 In-Sample Evaluation of Gold Price

We use various goodness-of-fit statistics to compare volatility models. These statistics are Akaike Information Criteria (AIC) Schwarz Bayesian Information Criteria (SBIC) and Log-likelihood (LOGL) values. In Table 3.5, the results of goodness-of- fit statistics and loss functions for all volatility models are presented. According to AIC, MRS-GARCH-GED is the best. GARCH-t is the best in SBIC, MRS-GARCH-2t is the best in LOGL, EGARCH -N is the best in MSE1 and MSE2. MRS-GARCH-t is the best in QLIKE. EGARCH-GED is the best in MAD2 and EGARCH-t is the best in MAD1 and HMSE. We found that different models were suitable for various loss functions.

3.3.4 Forecasting Volatility in Out-of-Sample of Gold Price

In this section, we investigate the ability of MRS-GARCH and GARCH type models to forecast volatility of Gold prices from out-of-sample. In Table 3.6, we present the result of loss function of out-of-sample with forecasting volatility for one day ahead, and we found the EGARCH and MRS-GARCH models perform best.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

3.3.5 Trading Future Contract with Forecast Volatility and Forecast Price

The aim of this study is to evaluate the profitability of applying different models to the volatility of gold prices. We assume that the market is a perfect market, and the positions in our strategy are not longer than one day as described below. We applied the Bollinger band indicator and we used samples of 21 days from 1 to 30 August 2011 (We trade one contract in GF10Q11 series is future contract in gold price with maturity date at August 2011) to trade in one contract with day by day and we did not include settlement, return do not include cost price i.e. margins, fees charged. The net daily rate of return for long position is computed as follows:

$$R_{t+1} = C_{t+1} - (O_{t+1} - m\sqrt{h_{t+1}}),$$

where $R_{t+1}, C_{t+1}, O_{t+1}$ are the return, close, open price, $h_{t+1}$is forecasting volatility at next day $(t+1)$ and $m \in \mathbb{Z}^+$ is constants.

The net daily rate of return on close position is computed as follows:

$$R_{t+1} = (O_{t+1} + m\sqrt{h_{t+1}}) - C_{t+1}.$$

Table 3.7 shows that the cumulative return with Markov Regime Switching in the GARCH-N model and the GJR-N model give higher cumulative return than the other models when we use $m = 30$.

3.4 Forecasting Volatility and Price of the SET50 Index using the Markov Regime Switching

In this section, we use GARCH, EGARCH, GJR-GARCH and MRS-GARCH models to forecast the volatility and price of the SET50 Index (The

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

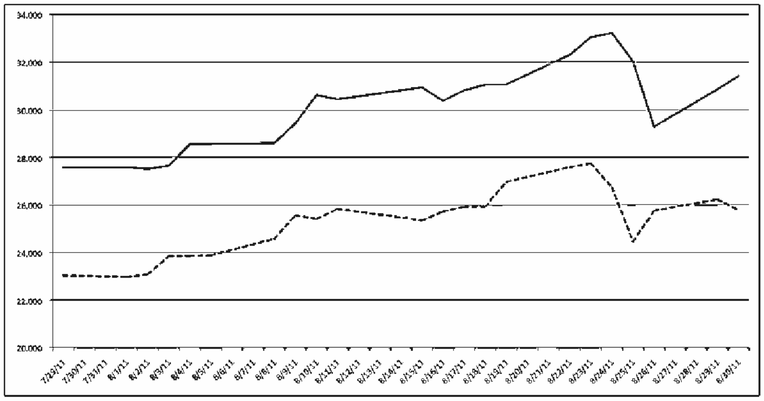


Figure 3.2 Forecasting closed gold price with MRS-GARCH-N (Dot line is closed price and line is forecast price).

stock prices of the top 50 listed companies on the SET (Stock Exchange of Thailand) in terms of large market capitalization) and to compare their performance with loss function.

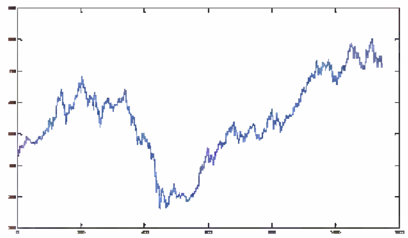
3.4.1 Empirical Methodology and Model Estimation Results of the SET50 Index

The data set used in this study is the daily closing prices of the SET50 Index (P_t) over the period 3/01/2007 through 30/03/2011 ($t = 1, ..., 1,038$ observations). The data set is obtained from the Stock Exchange of Thailand. The data set was divided into in-sample ($R = 1,016$ observations) and out-of-sample ($n = 22$ observations). The plot of P_t and its log returns series r_t are given in Fig. 3.3. P_t and r_t displays the usual properties of financial data series.

As expected, volatility is not constant over that period of the time and exhibits volatility clustered with large changes in the index often followed by large changes, and small changes often followed by small changes.

Descriptive statistics of r_t are presented in Table 3.8. As Table 3.8 shows, r_t has a quite small positive average return (about 0.04 percent). Standard deviation

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor



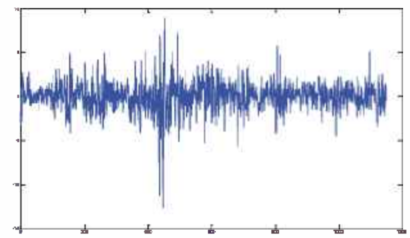


Figure 3.3 The left: Price of the SET50 Index , The right: Log returns series for the period 3/01/2007 through 31/03/2011.

of r_t is 2.09 percent. Moreover, we tested for the normality of r_t by using the Jarque-Bera test under null hypothesis r_t is normal distribution and we find that the test statistic value is 1,142.19 which confirms of rejecting null hypothesis. So r_t is not normally distribution.

Also, the skewness and kurtosis of r_t are 0.29 (not equal zero) and 8.11 (greater than 3) respectively. These values indicate that the returns are not normally distributed, namely, have fatter tails. Moreover, we test for stationarity of r_t by the Augmented Dickey-Fuller test. The test statistic value is -34.06 which indicates the stationarity of r_t .

The autocorrelation functions (ACF) are presented in Table 3.9, when we apply Ljung and Box (1978) to test serial correlation in P_t and r_t. We use the specified lag from the first to the tenth lags and the twenty-second lag. Serial correlation in P_t confirmed as non-stationary but r_t is stationary because of the ACF values decrease very fast when lag increase and is confirmed by Augmented Dickey-Fuller test in Table 3.9. We analyze the significance of autocorrelation in the squared mean adjusted return $(r_t - \mu_t)^2$ series by the Ljung-Box Q-test and apply the Engles ARCH (Auto Regressive Conditional Heteroskedasticity) test (1982) to test the ARCH effects. Therefore, the squared mean adjusted return is non-stationary which suggests the conditional heteroskedasticity.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

3.4.2 Estimator Parameters of the SET50 Index

This empirical part adopts the GARCH type and the MRS-GARCH (1,1) models to estimate the volatility of the P_t . The GARCH type models considered are GARCH (1,1), EGARCH (1,1) and GJR-GARCH (1,1). In order to account for the fat tails feature of financial returns, we consider three different distributions for the innovations: Normal (N), Student-t (t) and Generalized Error Distributions (GED).

- **GARCH type Models in SET50 Index**

Table 3.10 presents an estimation of the results for the GARCH type models. It is clear from the table that almost all parameter estimates, including δ, in the GARCH type models are highly significant at 1 percent. Only the leverage effect of the EGARCH and the GJR-GARCH models with the normal distribution, the δ are significant at 5 percent. However, the asymmetry effect term ξ in the EGARCH models is significantly different from zero, which indicates unexpected negative returns implying higher conditional variance as compared to same size positive returns. All models display strong persistence in volatility ranging from 0.9319 to 0.9655, that is, volatility is likely to remain high over several price periods once it increases.

- **Markov Regime Switching GARCH Models of the SET50 Index**

Estimation results and summary statistics of the MRS-GARCH models are presented in Table 3.11. Most parameter estimates in the MRS-GARCH are significant different from zero at least at 95 percent confidence level. But δ and α_1 are insignificantly different in some state. All models display strong persistence in volatility ranging from 0.6972 to 0.9646, that is, volatility is likely to remain high over several price periods once it increases.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

3.4.3 In-Sample Evaluation of the SET50 Index

We use various goodness-of-fit statistics to compare volatility models. These statistics are Akaike Information Criteria (AIC) Schwarz Bayesian Information Criteria (SBIC) and Log-likelihood (LOGL) values. In Table 3.12, the results of goodness-of- fit statistics and loss functions for all volatility models are presented. According to the AIC and the SBIC, the EGARCH model with the GED distribution performs best in modeling the SET50 Index Price volatility. However, in contrast the AIC and SBIC, the results suggest that the EGARCH with normal performs best in the QLIKE, and HMSE with t-distribution performs best in the MSE1 and MAD2, the GJR performs best in the MSE2. The MRS-GARCH with normal distribution performs best in the MAD1, with t-distribution (two degree of freedoms) in the LOGL.

3.4.4 Forecasting Volatility in Out-of-sample of the SET50 Index

In this section, we investigate the ability of the MRS-GARCH and GARCH type models to forecast volatility of the SET50 Index in out-of-sample. In Table 3.13, we present the results of loss function of out-ofsample with forecasting volatility for one day ahead, and we found that the MRS-GARCH-GED model performs best.

3.4.5 Forecasting Price in Out-of-sample of the SET50 Index

In this section, we investigate the ability of the MRS-GARCH, the GARCH type and random walk models to forecast the price of the SET50 In-

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

dex in out-of-sample with one day ahead. In Table 3.14, we present the result of the loss function of out-of-sample with forecasting price for one day ahead, and we found that the MRS-GARCH-2t model performs best.

3.5 Conclusion

In this chapter, we forecast the volatility of gold prices using Markov Regime Switching GARCH (MRS-GARCH) models. These models allow volatility to have different dynamics according to unobserved regime variables. The main purpose was to find out whether MRS-GARCH models are an improvement on the GARCH type models which include GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) in terms of modeling and forecasting gold price closed price volatility. We compared MRS-GARCH(1,1) models with GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1) models. All models were estimated under three distributional assumptions which are Normal, Student-t and GED. Moreover, Student-t distribution which takes different degrees of freedom in each regime was considered for MRS-GARCH models.

We first analyzed in-sample performance of various volatility models to determine the best form of the volatility model over the period 4/01/2007 through 30/08/2011. As expected, volatility is not constant over time and exhibits volatility clustering showing large changes in the price of an asset often followed by large changes, and small changes often followed by small changes.

Descriptive statistics of return series are represented by returns with fatter tails. The Augmented Dickey-Fuller test indicates gold price log returns are stationary. Serial correlation in the gold price confirms it is non-stationary but serial log returns of gold price are stationary. Serial correlation in the squared returns suggests conditional heteroskedasticity. This empirical part adopts GARCH type and MRS-GARCH models to estimate the volatility of the gold price.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In order to account for fat tailed features of financial returns, we considerd three different distributions for the innovations. Almost all parameter estimates in GARCH type models are highly significant at 1 percent. Most parameter estimates in MRS-GARCH are significantly different from zero at least at 95percents confidence level. However, the results of goodness-of- fit statistics and loss functions for all volatility models show different results.

The trading details we have used describe forecasts of the closing gold prices between 1/08/2011-30/08/2011 and trading in gold future contract (GF10Q11). We found the cumulative returns with the Markov Regime Switching GARCH-N model and the GJR-N model give us higher cumulative returns than the other models when we use $m = 30$.

For further study, three or four volatility regimes settings can be considered rather than two-volatility regimes. One way also combine Markov Regime Switching with other volatility models e.g. EGARCH, GJR. In addition, the performance of MRS-GARCH models can be hedged in future for long and short positions.

Finally, we forecast volatility of the SET50 Index using the Markov Regime Switching GARCH (MRS-GARCH) models. These models allow volatility to have different dynamics according to unobserved regime variables. The main purpose was to find out whether the MRS-GARCH models are an improvement on the GARCH type models in terms of modeling and forecasting the SET50 Index closing price volatility. We compared the MRS-GARCH (1,1) models with GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) models. All models are estimated under three distributional assumptions that are Normal, Student-t and GED. Moreover, Student-t distribution which takes different degrees of freedom in each regime was considered for the MRS-GARCH models.

The loss function of out-of-sample with forecasting volatility for one day

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

ahead, we found the MRS-GARCH with GED distribution model performs best. The loss function of out-of-sample with forecasting price for one day ahead, we found the MRS-GARCH with t-distribution (two degrees of freedoms) model performs best.

For further study, three or four volatility regimes setting can be considered rather than two-volatility regimes. One way also combine the Markov Regime Switching with other volatility models e.g. the EGARCH, the GJR. In addition, the performance of MRS-GARCH models can be compared in terms of their ability to forecast Value at Risk (VaR) for long and short positions.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.3 Summary results for estimation parameters of GARCH type models of gold price.

Parameter	GARCH			EGARCH			GJR		
	N	t	GED	N	t	GED	N	t	GED
δ	0.101^a	0.123^a	0.116^a	0.134^a	0.137^a	0.131^a	0.108^a	0.126^a	0.119^a
Std.	0.035	0.034	0.032	0.036	0.034	0.032	0.037	0.034	0.032
α_0	0.057^a	0.069^a	0.063^a	-0.097^a	-0.073^a	-0.084^a	0.055^a	0.064^a	0.060^a
Std.	0.010	0.017	0.017	0.010	0.018	0.018	0.010	0.016	0.016
α_1	0.082^a	0.076^a	0.078^a	0.143^a	0.109^a	0.124^a	0.061^a	0.056^a	0.058^a
Std.	0.009	0.018	0.017	0.014	0.026	0.025	0.012	0.020	0.021
β_1	0.891^a	0.890^a	0.889^a	0.046^a	0.049^a	0.046^a	0.894^a	0.895^a	0.893^a
Std.	0.009	0.018	0.018	0.011	0.017	0.017	0.009	0.018	0.017
ξ	na	na	na	0.718^a	0.432^a	0.624^a	0.099^a	0.091^a	0.091^a
Std.	na	na	na	0.004	0.005	0.006	0.013	0.025	0.024
ν	na	5.188^a	1.235^a	na	5.414^a	1.269^a	na	5.287^a	1.241^a
Std.	na	0.761	0.053	na	0.808	0.057	na	0.777	0.053
L.	−2087	−2033	−2038	−2370	−2160	−2185	−2086	−2033	−2037
P.	0.972	0.965	0.967	0.046	0.049	0.046	0.974	0.968	0.970
LB	32.64	32.64	32.64	32.64	32.64	32.64	32.64	32.64	32.64
LB^2	189.92	190.07	189.83	189.68	189.66	189.72	189.88	189.76	189.81

^{a,b} refer to significance at 99 percent, 95 percent confidence respectively

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22.

LB^2 is Ljung-Box test of squared innovation at lag 22 and Std. is standard error.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.4 Summary results for estimation parameters of MRS-GARCH models of gold price.

Parameter	MRS-GARCH							
	N		t		2t		GED	
State	$i=1$	$i=2$	$i=1$	$i=2$	$i=1$	$i=2$	$i=1$	$i=2$
$\delta^{(i)}$	0.083^b	0.180^b	0.114^a	0.170^b	0.114^a	0.170^b	0.171^b	0.109^a
Std.	0.040	0.093	0.039	0.077	0.039	0.077	0.078	0.037
$\alpha_0^{(i)}$	0.014^c	2.179^a	0.011	1.616^a	0.011	1.615^a	1.842^a	0.013
Std.	0.008	0.335	0.009	0.513	0.009	0.531	0.487	0.010
$\alpha_1^{(i)}$	0.046^a	0.365^a	0.038^b	0.317^a	0.038^b	0.317^a	0.324^a	0.042^b
Std.	0.013	0.103	0.016	0.115	0.016	0.117	0.126	0.018
$\beta_1^{(i)}$	0.944^a	0.000	0.954^a	0.184	0.954^a	0.186	0.102	0.949^a
Std.	0.015	0.112	0.018	0.177	0.018	0.180	0.140	0.020
p	0.9975^a		0.9981^a		0.9983^a		0.9981^a	
Std.	0.002		0.002		0.002		0.002	
q	0.9976^a		0.9983^a		0.9981^a		0.9983^a	
Std.	0.002		0.002		0.002		0.002	
$\nu^{(i)}$	na	na	6.058^a		6.079^a	6.013^a	1.323^a	
Std.	na	na	0.954		1.673	1.412	0.060	
L.	-2050		-2013		-2017		-2013	
σ^2	1.356	3.433	1.306	3.242	1.306	3.249	3.209	1.300
π	0.510	0.490	0.472	0.528	0.472	0.528	0.528	0.472
P.	0.990	0.365	0.992	0.501	0.992	0.503	0.426	0.990
LB.	34.996		34.996		34.996		34.996	
LB^2.	178.73^a		178.69^a		178.77^a		178.71^a	

^{a,b} refer to significance at 99 percent, 95 percent confidence respectively

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22.

LB^2 is Ljung-Box test of squared innovation at lag 22 and Std. is standard error.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.5 In-sample evaluation results for loss function.

Models	AIC	SBIC	MSE	QLIKE	MAD	HMSE
GARCH-N	3.509	3.526	1.381	1.665	2.738	0.870
GARCH-t	3.421	3.442	1.330	1.666	2.661	0.861
GARCH-GED	3.428	3.450	1.334	1.665	2.665	0.859
EGARCH-N	3.985	4.006	1.156	2.136	2.195	0.739
EGARCH-t	3.635	3.661	1.158	2.132	2.194	0.738
EGARCH-GED	3.676	3.702	1.161	2.156	2.195	0.739
GJR-N	3.510	3.531	1.390	1.664	2.752	0.871
GJR-t	3.422	3.448	1.333	1.665	2.666	0.860
GJR-GED	3.429	3.455	1.338	1.664	2.673	0.859
MRS-GARCH-N	3.457	3.500	1.300	1.615	2.655	0.843
MRS-GARCH-t	3.404	3.451	1.305	1.615	2.660	0.841
MRS-GARCH-2t	3.398	3.449	1.325	1.615	2.691	0.847
MRS-GARCH-GED	3.396	3.443	1.327	1.616	2.692	0.846

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.6 Result for loss function of out-of-sample with forecasting volatility for one day ahead.

Models	MSE	QLIKE	MAD	HMSE
GARCH-N	0.063	1.554	0.179	0.185
GARCH-t	0.055	1.538	0.17	0.181
GARCH-GED	0.056	1.539	0.167	0.182
EGARCH-N	0.057	1.537	0.217	0.240
EGARCH-t	0.047	1.525	0.183	0.218
EGARCH-GED	0.049	1.529	0.201	0.220
GJR-N	0.124	1.532	0.298	0.129
GJR-t	0.105	1.523	0.275	0.117
GJR-GED	0.109	1.525	0.279	0.12
MRS-GARCH-N	0.156	1.491	0.326	0.08
MRS-GARCH-t	0.133	1.487	0.25	0.071
MRS-GARCH-2t	0.132	1.487	0.25	0.079
MRS-GARCH-GED	0.086	1.492	0.213	0.073

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.7 Cumulative return of trading future contract of gold price with m =30 between 1 to 30 August 2011.

Date of treding	GARCH-GED	EGARCH-t	MRS-GARCH-t	MRS-GARCH-2t
1/8/2011	−40	−40	−50	−50
2/8/2011	−90	−90	−110	−110
3/8/2011	630	630	600	600
4/8/2011	710	710	670	670
5/8/2011	750	750	700	700
8/8/2011	1480	1490	1420	1420
9/8/2011	2540	2530	2460	2460
10/8/2011	2520	2500	2430	2430
11/8/2011	2200	2170	2120	2120
15/8/2011	1750	1700	1670	1670
16/8/2011	1440	1360	1350	1350
17/8/2011	1660	1560	1570	1570
18/8/2011	1760	1640	1670	1670
19/8/2011	2660	2520	2570	2570
22/8/2011	3500	3360	3410	3410
23/8/2011	3820	3670	3740	3740
24/8/2011	2850	2680	2800	2800
25/8/2011	5200	5000	5220	5220
26/8/2011	4370	4080	4450	4450
29/8/2011	5210	4870	5270	5270
30/8/2011	4890	4470	4960	4960

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.8 Descriptive statistics of the SET50 Index log returns series (r_t).

Statistic	Return
Min	-14.19
Max	11.46
Mean	0.04
Standard deviation	2.09
Skewness	-0.29
Kurtosis	8.11
Jarque-Bera Normality test	1,142.19 (P-value= 0.000)
Augmented Dickey-Fuller test	-34.06 (P-value= 0.000)

Table 3.9 ACF of the SET50 Index closing price, log returns series, square return and results for Engle's ARCH test.

Lags		1	5	10	22
	ACF	0.99	0.98	0.96	0.90
ACF of SET50	LBQ Test	1031.85	5073.77	9956.85	20882.11
	P-value	0.00	0.00	0.00	0.00
	ACF	0.05	-0.03	0.08	-0.02
ACF of return SET50	LBQ Test	2.26	8.05	31.28	66.24
	P-value	0.13	0.15	0.00	0.00
	ACF	0.31	0.23	0.27	0.03
ACF of square return SET50	LBQ Test	100.23	350.28	520.07	822.00
	P-value	0.00	0.00	0.00	0.00
	ARCH test	22.93	55.68	205.40	251.42
Result for Engle test	P-value	0.00	0.00	0.00	0.00

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.10 Summary results for estimation parameters of GARCH type models for forecasting volatility of the SET50 Index.

Parameter	GARCH			EGARCH			GJR		
	N	t	GED	N	t	GED	N	t	GED
δ	0.136^a	0.141^a	0.134^a	0.096^a	0.112^a	0.105^a	0.102^a	0.116^a	0.109^a
Std.	0.044	0.042	0.041	0.044	0.042	0.042	0.045	0.043	0.042
α_0	0.126^a	0.149^a	0.139^a	-0.151^a	-0.152^a	-0.152^a	0.148^a	0.172^a	0.162^a
Std.	0.027	0.044	0.041	0.022	0.031	0.030	0.030	0.046	0.045
α_1	0.142^a	0.152^a	0.147^a	0.234^a	0.241^a	0.237^a	0.196^a	0.215^a	0.206^a
Std.	0.020	0.032	0.030	0.028	0.042	0.040	0.030	0.046	0.044
β_1	0.81^a	0.792^a	0.80^a	0.976^a	0.959^a	0.962^a	0.805^a	0.786^a	0.794^a
Std.	0.018	0.034	0.031	0.008	0.013	0.012	0.020	0.035	0.033
ξ	na	na	na	-0.061^a	-0.072^a	-0.067^a	0.076^a	0.077^a	0.076^a
Std.	na	na	na	0.016	0.024	0.022	0.021	0.033	0.031
ν	na	8.298^a	1.445^a	na	8.981^a	1.491^a	na	8.614^a	1.463^a
Std.	na	1.972	0.083	na	2.414	0.091	na	2.075	0.084
L.	−1886	−1873	−1872	−1876	−1865	−1865	−1880	−1867	−1867
P.	0.952	0.943	0.946	0.966	0.959	0.962	0.941	0.932	0.935
LB	32.64	32.64	32.64	32.64	32.64	32.64	32.64	32.64	32.64
LB^2	189.92	190.07	189.83	189.68	189.66	189.72	189.88	189.76	189.81

^{a,b} refer to significance at 99 percent, 95 percent confidence respectively

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22.

LB^2 is Ljung-Box test of squared innovation at lag 22 and Std. is standard error.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.11 Summary results for estimation parameters of MRS-GARCH models for forecasting volatility of the SET50 Index.

Parameter	MRS-GARCH							
	N		t		2t		GED	
State	$i=1$	$i=2$	$i=1$	$i=2$	$i=1$	$i=2$	$i=1$	$i=2$
$\delta^{(i)}$	0.240^a	-0.027	0.242^a	$0-.023$	0.238	-0.024^a	0.236^a	-0.028
Std.	0.054	0.082	0.056	0.082	0.081	0.057	0.059	0.083
$\alpha_0^{(i)}$	0.214^a	0.209^a	0.206^b	0.222^b	0.214^b	0.234^b	0.209^b	0.209^b
Std.	0.076	0.081	0.068	0.108	0.112	0.066	0.094	0.105
$\alpha_1^{(i)}$	0.000	0.070^a	0.000	0.071^a	0.000	0.072	0.000	0.072^a
Std.	0.051	0.018	0.000	0.025	0.027	0.000	0.058	0.023
$\beta_1^{(i)}$	0.694^a	0.895^a	0.713^a	0.887^a	0.699^a	0.884^a	0.708^a	0.889^a
Std.	0.085	0.030	0.079	0.045	0.047	0.079	0.105	0.040
p	0.965^a		0.978^a		0.967^a		0.978^a	
Std.	0.016		0.012		0.012		0.011	
q	0.976^a		0.966^a		0.979^a		0.967^a	
Std.	0.010		0.018		0.017		0.018	
$\nu^{(i)}$	na	na	14.211^a		161.39^a	11.079^a	1.647^a	
Std.	na	na	0.121		4.948	4.932	0.122	
L.	-1865		-1861		-1861		-1861	
σ^2	0.699	5.912	0.717	5.224	0.708	5.205	0.715	5.367
π	0.404	0.596	0.398	0.602	0.396	0.604	0.396	0.604
P.	0.694	0.965	0.713	0.958	0.699	0.955	0.708	0.961

[a,b] refer to significance at 99 percent, 95 percent confidence respectively

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.12 Results for evaluation in-sample with forecasting volatility of the SET50 Index.

Models	AIC	SBIC	MSE	QLIKE	MAD	HMSE
GARCH-N	3.721	3.74	1.513	1.876	0.947	3.277
GARCH-t	3.696	3.72	1.498	1.876	0.944	3.326
GARCH-GED	3.695	3.72	1.502	1.876	0.944	3.322
EGARCH-N	3.702	3.726	1.443	1.856	0.932	2.958
EGARCH-t	3.683	3.712	1.433	1.857	0.931	3.023
EGARCH-GED	3.682	3.711	1.436	1.856	0.93	3.007
GJR-GARCH-N	3.711	3.735	1.47	1.864	0.938	3.212
GJR-GARCH-t	3.688	3.717	1.459	1.865	0.935	3.269
GJR-GARCH-GED	3.688	3.717	1.462	1.865	0.935	3.258
MRS-GARCH-N	3.691	3.739	1.477	1.861	0.929	3.333
MRS-GARCH-2t	3.687	3.745	1.473	1.859	0.932	3.166
MRS-GARCH-t	3.685	3.739	1.477	1.861	0.931	3.273
MRS-GARCH-GED	3.686	3.739	1.472	1.861	0.931	3.224

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.13 Result for loss function of out-of-sample with forecasting volatility of the SET50 Index for one day ahead.

Models	MSE	QLIKE	MAD	HMSE
GARCH-N	38.3879	14.1567	28.9676	21.9001
GARCH-t	38.4522	14.1545	28.9786	21.9
GARCH-GED	38.4204	14.1536	28.9724	21.9
EGARCH-N	43.5463	16.6204	30.7887	21.9105
EGARCH-t	44.1814	16.851	30.9848	21.9112
EGARCH-GED	43.6013	16.5907	30.7871	21.9102
GJR-GARCH-N	41.5379	15.6615	30.0785	21.9068
GJR-GARCH-t	41.8758	15.7649	30.1754	21.9071
GJR-GARCH-GED	41.6276	15.6497	30.0894	21.9066
MRS-GARCH-N	42.7165	15.6962	30.3001	21.9053
MRS-GARCH-2t	46.8062	18.1511	31.9257	21.9168
MRS-GARCH-t	47.8918	18.5713	32.2734	21.918
MRS-GARCH-GED	37.5699	13.7146	28.6605	21.8982

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 3.14 Result for loss function of out-of-sample with forecasting price of SET50 Index for one day ahead.

Models	MSE	MAE	MAPE
GARCH-N	66.7153	6.608	0.9335
GARCH-t	66.8143	6.6195	0.9351
GARCH-GED	66.9819	6.6376	0.9377
EGARCH-N	67.6716	6.6999	0.9463
EGARCH-t	67.3822	6.6756	0.943
EGARCH-GED	67.5379	6.689	0.9448
GJR-GARCH-N	67.1641	6.6557	0.9402
GJR-GARCH-t	67.1693	6.6562	0.9402
GJR-GARCH-GED	67.2598	6.6646	0.9414
MRS-GARCH-N	66.023	6.5078	0.9196
MRS-GARCH-2t	65.9811	6.5046	0.9192
MRS-GARCH-t	66.0597	6.5109	0.92
MRS-GARCH-GED	66.3461	6.5557	0.9262
Random Walk	69.6096	6.8264	0.964

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER IV

FORECASTING THE STOCK EXCHANGE OF THAILAND USING DAY OF THE WEEK EFFECT AND MARKOV REGIME SWITCHING GARCH.

In the time series, the stock price is transformed to return series for stationary process which looked like white noise and forecasting was possible using the mean equation. The forecasting of daily returns has led to additional research in financial literature, specifically extending the analysis of the seasonal behavior to include the day of the week effect. This seasonality has been the subject of different studies which detected empirical evidence of abnormal yield distributions based upon the day of the week. The pioneering work was carried out as used in the analysis of seasonality and can be specifically seen in Miralles and Quiros (2000), they included five dummy variables, one for each day of the week.

Nevertheless two serious problems arise with this approach. The first problem is that the residuals obtained from the regression model can be autocorrelated, thus creating errors in the inference. The second problem is that the variances of the residuals are not constant and possibly time-dependent.

A solution to the first type of problem can be solved by introducing the returns with a one week delay into the regression model, as used in the works by Easton and Faff (1994) and Kyimaz and Berument (2001).

Moreover, Apolinario et al. (2006) and Ulussever et al. (2011) try to solve

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the second problem by modeling the residuals with the ARCH model in order to correct the variability in the variance of the residuals.

In this study, we reconsidered the two problems again. For the first problem, we modelled the SET Index returns by the mean equation with the day of the week effect and the autoregressive moving-average order p and q (ARMA (p, q)). For the second problem, we model the residuals by the GARCH, EGARCH, GJR-GARCH and MRS-GARCH models. Finally, we compare their performance by one day, one week, two weeks and one month.

In next section, we shall model the mean equation of the financial returns. In section 4.2, we describe the data. We estimate the parameters of the model and forecast volatility of returns and estimate the parameters with in-sample evaluation results in section 4.3. Moreover, statistical loss functions are described and out-of-sample forecasting performance of various models is discussed in section 4.4. The conclusion is presented in section 4.5.

4.1 Modelling Mean Equation of Financial Returns

In time series, let P_t denote the series of the financial price at time t on a probability space (Ω, F, P) and r_t be the log return of an asset at time index t (in percent) in (2.46), i.e.

$$\begin{aligned} r_t &= 100 \cdot ln(\frac{P_t}{P_{t-1}}) \\ &= \mu_t + \varepsilon_t. \end{aligned}$$

To put the volatility models in proper perspective, it is informative to

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

consider the conditional mean and variance of r_t given F_{t-1} ; that is,

$$\begin{aligned}\mu_t &= E(r_t|F_{t-1}),\\ h_t &= Var(r_t|F_{t-1})\\ &= E\left[(r_t-\mu_t)^2|F_{t-1}\right],\end{aligned}\tag{4.1}$$

where F_{t-1} refers to information up to time $t-1$. Therefore, the equation for μ_t in (4.1) should be simple, and we assume that r_t follows a simple time series model such as a stationary ARMA(p,q) model which includes five dummy variables, one for each day of the week, such that

$$\begin{aligned}r_t &= \mu_t+\varepsilon_t,\\ \mu_t &= \beta_1 D_{1t}+\beta_2 D_{2t}+\beta_3 D_{3t}+\beta_4 D_{4t}+\beta_5 D_{5t}+\sum_{j=1}^{p}\phi_j r_{t-j}-\sum_{l=1}^{q}\theta_l\varepsilon_{t-l},\end{aligned}\tag{4.2}$$

where D_{it} : $i=1,\ldots,5$: are dummy variables which take on the value of 1 if the corresponding return for day t is a Monday, Tuesday, Wednesday, Thursday or Friday, respectively and 0 otherwise.

$\beta_i : i=1,\ldots,5$: are coefficients which represent the average return for each day of the week.

$\phi_j : j=1,\ldots,p$, $\theta_l : l=1,\ldots,q$: are coefficients which represent the ARMA(p,q).

4.2 Empirical Methodology and Model Estimation Results of the SET Index

4.2.1 Data

The data set that was used in this study is the daily closing prices of the SET Index P_t over the period 3/01/2007 through 30/03/2011 ($t=1,\ldots,1,038$ observations). The data set was obtained from the Stock Exchange of Thailand.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

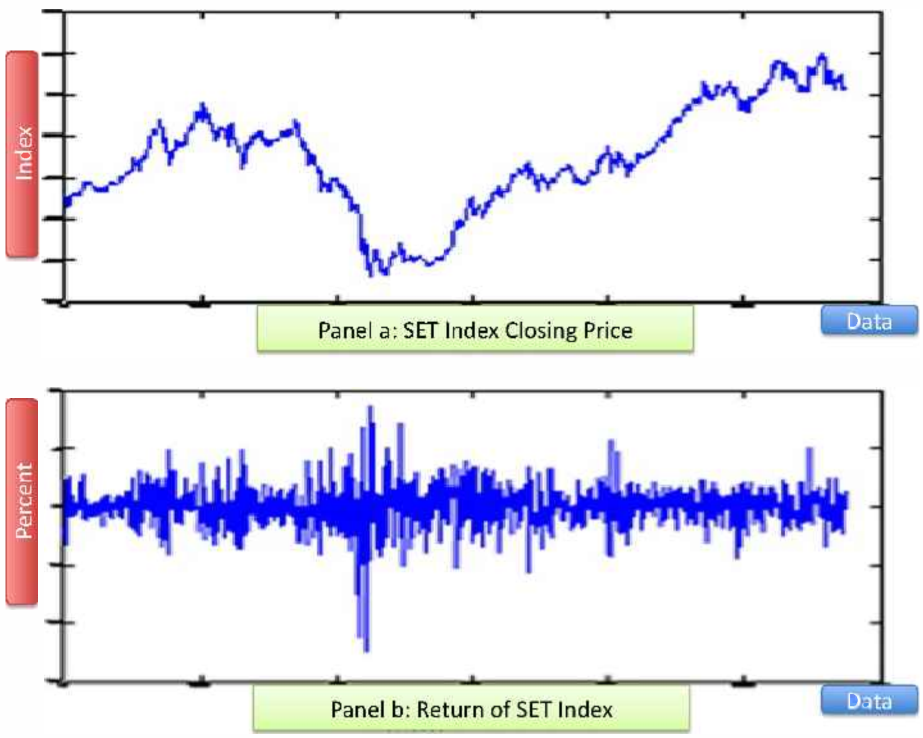


Figure 4.1 Graph of (a) SET Index closing prices and (b) log returns series for the period 3/01/2007 through 31/03/2011.

The data set is divided into in-sample (R =977 observations) and out-of-sample (n = 61 observations). The plot of P_t and its log returns series r_t (see equation (4.1)) are given in Figure 4.1. Plot P_t and r_t display the usual properties of financial data series. As expected, volatility is not constant over that period of time and exhibits volatility clustered with large changes in the index often followed by large changes, and small changes often followed by small changes.

Descriptive statistics of r_t are presented in Table 4.1. As Table 4.1 shows, overall, r_t has a quite small positive average return (about 0.0436 percent). Standard deviation of r_t is 1.5525 percent. The lowest average return is observed on Monday and the highest average return occurs on Friday.

Moreover, we tested for the normality of r_t by using the Jarque-Bera test

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(The Jarque-Bera Normality test is a goodness-of-fit measure of departure from normality and can be used to test which the Jarque-Bera test has a χ^2 distribution with 2 degrees of freedom under the null hypothesis that the data is from a normal distribution. The 5 percent critical value is, therefore, 5.99) under the null hypothesis r_t is normally distributed and we find that the test statistic value is 1,758.1080 which leads us to reject the null hypothesis. So r_t is not normally distributed. Also, the skewness and kurtosis of r_t are -0.7189 (not equal zero) and 6.2605 (greater than 3) respectively. These values confirm that the returns are not normally distributed, namely, have fatter tails.

Moreover, we test for the stationarity of r_t by using the Augmented Dickey-Fuller test (The Augmented Dickey-Fuller test is a test for a unit root in a time series sample, the null hypothesis of ADF test is that the series is non-stationary. The 1, 5 and 10 percent critical value are -3.44, -2.86 and -2.57 respectively). The test statistic value is -30.0801 which indicates the stationarity of r_t.

Table 4.1 Descriptive statistics of the SET Index log returns series (r_t).

Statistic	All days	Monday	Tuesday	Wednesday	Thursday	Friday
Mean	0.04	-0.04	-0.02	0.03	0.03	0.22
SD.	1.55	1.98	1.46	1.37	1.43	1.48
Min	-11.09	-11.09	-4.28	-7.13	-5.44	-10.10
Max	7.55	7.55	5.29	3.28	6.10	4.19
Skewness	-0.7189	-0.5511	0.2214	-1.0215	-0.2429	-1.9876
Kurtosis	6.2605	5.8511	1.7362	3.2096	2.7026	13.75
JB-test	1,758.11 (P-value= 0.000)					
ADF-test	-30.08 (P-value= 0.000)					

Table 4.2 reports the day of the week effects and ARMA(p,q) for returns

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

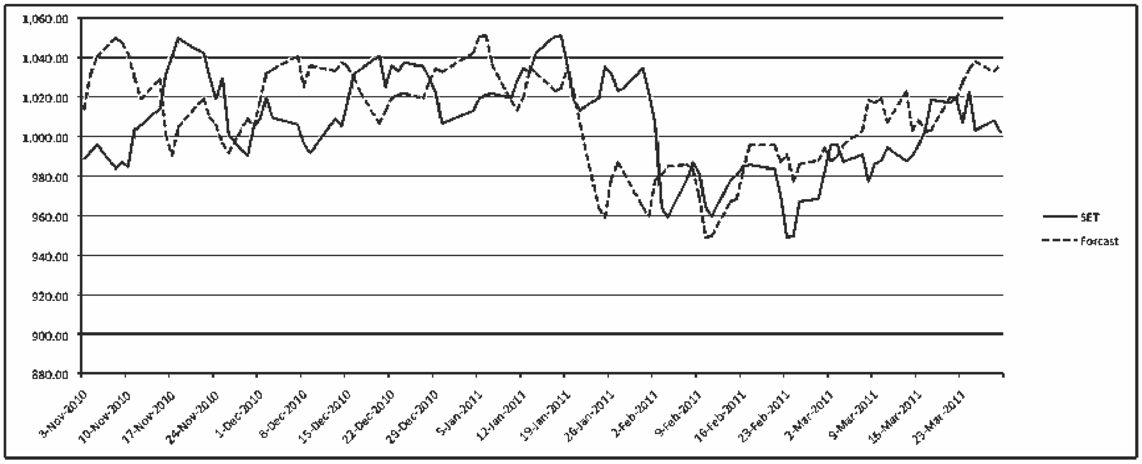


Figure 4.2 Graph of the SET Index closing prices and forecasting with day of the week effect and ARMA(3,3).

Panel A of Table 4.2 displays the first estimated coefficients of the day of the week effect ($\beta_i : i = 1, \ldots, 5$). From Table 4.2 (Panel A), we found the estimated coefficients are almost zero. Then we test under the null hypothesis that each coefficient $\beta_i : i = 1, \ldots, 5$ is zero. We find that the coefficient of Fridays' dummy variable is not zero significant at the 95 percent level and other days are insignificant. These observations suggest that only Friday is the day of effect of the SET Index. Panel B displays the estimated coefficients of the ARMA process and P-values. By using t-test under the null hypothesis that each coefficient AR(p) and MA(q) is zero, we found that the P-values are all zero then each coefficient is not zero significant at the 99 percent level. Hence the SET Index return can be modeled by the ARMA (3,3) process. Figure 4.2 showns the SET Index and forecasting with day of the week effect and ARMA(3,3).

The autocorrelation functions (ACF) are presented in Table 4.3, when we apply Ljung and Box (1978) to test serial correlation in P_t and r_t. We use the specified lag from the first to the tenth lags and the twenty-second lag. Serial correlation in is P_t confirmed as non-stationary but r_t is stationary because of

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.2 Day of the week effect and ARMA(p,q) in mean equation of the SET Index.

Panel A: Day of the week effect in mean equation of return

	Monday	Tuesday	Wednesday	Thursday	Friday
β	-0.064	-0.038	0.02	0.014	0.221^b
Std.	0.109	0.106	0.105	0.106	0.107

Panel B: ARMA models parametric estimates in mean equation of return

Variable	Coefficient	Std. Error	t-Statistic	P-value
AR(1)	2.5855	0.0579	44.6244	0.0000^a
AR(2)	-2.4248	0.1121	-21.6289	0.0000^a
AR(3)	0.8289	0.0617	13.4318	0.0000^a
MA(1)	2.5059	0.0732	34.2502	0.0000^a
MA(2)	-2.2667	0.1436	-15.7891	0.0000^a
MA(3)	0.7459	0.0799	9.3413	0.0000^a

^{a,b} refer to significance at 99 percent, 95 percent confidence respectively.

ACF values decrease very fast when the lag increases and is confirmed by the Augmented Dickey-Fuller test in Table 4.1. We analyse the significance of autocorrelation in the squared mean adjusted return $(r_t - \mu_t)^2$ series by using the Ljung-Box Q-test. Since the p-value is equal to zero then the squared mean adjusted return is non-stationary. Next, we apply Engle's ARCH test (1982) to test ARCH effects of the squared mean adjusted return. The p-value suggests conditional heteroskedasticity.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.3 ACF of the SET Index closing price, log returns series, squared mean adjusted return and results for Engle's ARCH test.

Lags		1	5	10	22
	ACF	0.9962	0.9798	0.9621	0.9113
ACF of P_t	LBQ Test	0.1033	0.5091	1.1011	2.1086
	P-value	0.0000	0.0000	0.0000	0.0000
	ACF	0.0672	-0.0261	0.0870	-0.0038
ACF of r_t	LBQ Test	4.6938	10.0980	31.8949	64.5758
	P-value	0.0303	0.0725	0.0004	0.0000
	ACF	0.2872	0.2031	0.2701	0.0168
ACF of $(r_t - \mu_t)^2$	LBQ Test	85.7126	306.58	456.47	726.43
	P-value	0.0000	0.0000	0.0000	0.0000
	ARCH test	85.4839	162.18	217.02	256.93
Result for Engle test	P-value	0.0000	0.0000	0.0000	0.0000

4.3 Empirical Methodology of the SET Index

This empirical part adopts the GARCH type and MRS-GARCH(1,1) models to estimate the volatility of the P_t . The GARCH type models considered are GARCH (1,1), EGARCH (1,1) and GJR-GARCH (1,1). In order to account for the fat tails feature of financial returns, we consider three different distributions for the innovations: Normal (N), Student-t (t) and Generalized Error Distributions (GED).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

4.3.1 GARCH Type Models of the SET Index

Table 4.4 presents an estimation of the results for GARCH type models. It is clear from the table that almost all parameter estimates are highly significant at 1 percent. However, the asymmetry effect term ξ in EGARCH models is significantly different from zero, which indicates unexpected negative returns implying higher conditional variance as compared to the same size positive returns. All models display strong persistence in volatility ranging from 0.8950 to 0.9521, that is, volatility is likely to remain high over several price periods once it increases.

Table 4.4 Summary results of GARCH type models of the SET Index.

Parameter	GARCH			EGARCH			GJR		
	N	t	GED	N	t	GED	N	t	GED
α_0	0.132^a	0.159^a	0.148^a	-0.161^a	-0.161^a	-0.163^a	0.158^a	0.183^a	0.172^a
Std.	0.027	0.044	0.043	0.022	0.033	0.032	0.031	0.047	0.047
α_1	0.153^a	0.166^a	0.161^a	0.248^a	0.254^a	0.251^a	0.217^a	0.242^a	0.232^a
Std.	0.021	0.036	0.034	0.030	0.046	0.044	0.032	0.053	0.051
β_1	0.785^a	0.761^a	0.770^a	0.952^a	0.945^a	0.949^a	0.776^a	0.751^a	0.760^a
Std.	0.021	0.039	0.036	0.010	0.015	0.015	0.024	0.041	0.039
ξ	na	na	na	-0.076^a	-0.089^a	-0.082^a	-0.075^a	0.076^a	0.076^a
Std.	na	na	na	0.016	0.026	0.024	0.022	0.036	0.034
ν	na	7.338^a	1.381^a	na	7.852^a	1.435^a	na	7.696^a	1.407^a
Std.	na	1.634	0.081	na	1.948	0.087	na	1.745	0.081
L.	-1683	-1668	-1667	-1673	-1660	-1660	-1677	-1663	-1663
P.	0.938	0.926	0.931	0.952	0.945	0.949	0.895	0.910	0.914
LB	63.325	63.325	63.325	63.325	63.325	63.325	63.325	63.325	63.325
LB^2	673.30	673.62	673.66	671.90	672.94	672.87	672.08	672.94	672.91

^{a,b} refer to the significance at 99 percent, 95 percent confidence respective.

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22 LB^2 is Ljung-Box test of squared innovation at lag 22 and Std. is standard error.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

4.3.2 Markov Regime Switching GARCH Models of the SET Index

Estimation results and summary statistics of MRS-GARCH models are presented in Table 4.5. Most parameter estimates in MRS-GARCH are significantly different from zero at least at the 95 percent confidence level. But α_0, α_1 and β_1 are insignificantly different in some states. All models display strong persistence in volatility ranging from 0.5842 to 0.9619, that is, volatility is likely to remain high over several price periods once it increases.

4.3.3 In-Sample Evaluation of the SET Index

We use various goodness-of-fit statistics to compare volatility models. These statistics are Akaike Information Criteria (AIC), Schwarz Bayesian Information Criteria (SBIC) and Log-likelihood (LOGL) values. In Table 4.6, the results of goodness-of-fit statistics and loss functions for all volatility models are presented. According to SBIC, the EGARCH model with GED-distribution performs best in modeling the SET Index volatility. However, the MSE1 and MSE2 suggest that the EGARCH with t-distribution performs best in the SET Index volatility. Also AIC and LOGL suggest that the MRS-GARCH with 2t-distribution performs best in the SET Index volatility. MAD1, MAD2 and HMSE suggest that the MRS-GARCH with t-distribution performs best in the SET Index volatility and in QLIKE the MRS-GARCH with GED-distribution performs best in the SET Index volatility.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

4.4 Forecasting Volatility in Out-of-Sample of the SET Index

In this section, we investigate the ability of MRS-GARCH and GARCH type models to forecast volatility of the SET Index in out-of-sample. In Table 4.7, we present the results of loss function of out-of-sample with forecasting volatility for one day ahead, five days ahead (a week), ten days ahead (two weeks) and twenty-two days ahead (a month). We found the GARCH-type models perform best in the short term (one day and a week) for forecasting volatility of the SET Index. Additionally, we have reported a particular sign-test, the Success Ratio (SR). The SR test is simply the fraction of volatility forecasts that have the same sign as volatility realizations. From the table we can see that the GARCH-type models do a great job in correctly predicting the sign of the future volatility in the short term. On the other hand, we found that the MRS-GARCH models perform best in the long term (two weeks and a month) for forecasting the volatility of the SET Index. Also, the SR test MRS-GARCH models do a great job in correctly predicting the future volatility in the long term.

4.5 Conclusion

In this study, we modelled the returns of the SET Index by mean equation with the day of the week effect and the autoregressive moving-average order p and q (ARMA(p, q)) and forecasted the volatility of the SET Index by the GARCH, EGARCH, GJR-GARCH and MRS-GARCH models. Moreover we compared their volatility forecast performance with one day, one week, two weeks and one month returns.

Friday is the day with the largest effect on the SET Index. The estimate of

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

return equation perform with ARMA (3, 3). The GARCH-type models perform best in the short term (one day and a week). On the other hand, the MRS-GARCH models perform best in the long term (two weeks and a month) for forecasting volatility of the SET Index.

For further study, three or four volatility regime settings can be considered rather than two-volatility regimes or using Markov Regime Switching with other volatility models e.g. EGARCH, GJR. In addition, the performance of the MRS-GARCH models can be compared in terms of their ability to forecast Value at Risk (VaR) for long and short positions.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.5 Summary results of MRS-GARCH models of the SET Index.

Parameter	MRS-GARCH							
	N		t		2t		GED	
State	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 1$	$i = 2$
$\alpha_0^{(i)}$	0.216^a	0.185^a	0.235^b	0.190^b	0.000	0.183^b	0.000	0.176^a
Std.	0.084	0.065	0.089	0.114	0.093	0.109	0.109	0.047
$\alpha_1^{(i)}$	0.000	0.075^a	0.000	0.076	0.903^a	0.075	0.962^a	0.068^b
Std.	0.056	0.018	0.026	69.205	0.027	50.228	0.355	0.035
$\beta_1^{(i)}$	0.605^a	0.885^a	0.584^a	0.875^a	0.000	0.761^a	0.000	0.776^a
Std.	0.121	0.030	0.046	0.178	0.047	0.177	0.021	0.039
p	0.9582^a		0.9603^a		0.9785^a		0.9822^a	
Std.	0.018		0.011		0.020		0.007	
q	0.9737^a		0.9776^a		0.4409^a		0.5696^a	
Std.	0.010		0.020		0.011		0.109	
$\nu^{(i)}$	na	na	11.252^a		9.141^a	8.375	1.4692^a	
Std.	na	na	0.445		4.297	29.679	0.095	
L.	-1658.07		-1652.69		-1651.18		-1654.06	
σ^2	0.548	1.713	0.565	3.897	0.000	1.123	0.000	0.996
π	0.386	0.614	0.361	0.639	0.037	0.963	0.040	0.960
Pers.	0.605	0.892	0.584	0.951	0.903	0.836	0.962	0.823
LB.	62.669		57.659		62.297		55.898	
LB^2.	678.936^a		725.076^a		677.794^a		720.355^a	

^{a,b} refer to significance at 99 percent, 95 percent confidence respectively.

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22.

LB^2 is Ljung-Box test of squared innovation at lag 22 and Std. is standard error.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.6 In-sample evaluation results of the SET Index.

Models	AIC	SBIC	MSE	QLIKE	MAD	HMSE
GARCH-N	3.453	3.473	1.194	1.608	7.414	0.836
GARCH-t	3.424	3.449	1.179	1.608	7.420	0.833
GARCH-GED	3.423	3.448	1.183	1.608	7.405	0.833
EGARCH-N	3.435	3.460	1.129	1.589	7.383	0.824
EGARCH-t	3.411	3.441	1.123	1.590	7.385	0.824
EGARCH-GED	3.410	3.440	1.124	1.589	7.371	0.823
GJR-N	3.444	3.469	1.164	1.597	7.369	0.830
GJR-t	3.416	3.446	1.154	1.598	7.373	0.829
GJR-GED	3.416	3.446	1.156	1.598	7.359	0.829
MRS-GARCH-N	3.415	3.465	1.152	1.588	7.251	0.817
MRS-GARCH-t	3.406	3.461	1.126	1.607	7.102	0.804
MRS-GARCH-2t	3.405	3.465	1.137	1.604	7.159	0.810
MRS-GARCH-GED	3.409	3.464	1.147	1.588	7.296	0.819

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.7 Results for loss function of out-of-sample with forecasting volatility for one day ahead of the SET Index.

Models	MSE	QLIKE	MAD	HMSE	SR
GARCH-N	0.931	2.332	1.026	14.994	0.533
GARCH-t	0.939	2.329	1.027	14.714	0.533
GARCH-GED	0.931	2.332	1.025	14.892	0.533
EGARCH-N	1.284	2.048	1.104	3.808	0.567
EGARCH-t	1.339	2.039	1.112	3.413	0.583
EGARCH-GED	1.307	2.042	1.107	3.608	0.583
GJR-N	1.340	2.045	1.113	3.928	0.567
GJR-t	1.384	2.033	1.120	3.570	0.550
GJR-GED	1.363	2.036	1.116	3.700	0.550
MRS-GARCH-N	1.238	2.102	1.105	4.432	0.500
MRS-GARCH-t	1.280	2.099	1.118	4.207	0.483
MRS-GARCH-2t	1.283	2.162	1.075	8.219	0.567
MRS-GARCH-GED	1.244	2.164	1.076	8.350	0.567

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.8 Results for loss function of out-of-sample with forecasting volatility for five days ahead of the SET Index (A week).

Models	MSE	QLIKE	MAD	HMSE	SR
GARCH-N	1.361	4.325	1.110	1.517	0.767
GARCH-t	1.365	4.327	1.123	1.444	0.767
GARCH-GED	1.361	4.324	1.114	1.476	0.767
EGARCH-N	0.983	4.134	0.864	0.580	0.883
EGARCH-t	0.977	4.124	0.849	0.529	0.917
EGARCH-GED	0.975	4.127	0.853	0.555	0.900
GJR-N	1.041	4.231	1.110	0.429	0.850
GJR-t	1.037	4.239	1.138	0.423	0.850
GJR-GED	1.034	4.233	1.120	0.423	0.850
MRS-GARCH-N	1.221	4.218	1.206	0.488	0.833
MRS-GARCH-t	1.264	4.251	1.277	0.513	0.850
MRS-GARCH-2t	1.277	4.218	0.994	0.658	0.833
MRS-GARCH-GED	1.278	4.214	0.982	0.642	0.800

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.9 Results for loss function of out-of-sample with forecasting volatility for ten days ahead (two weeks) of the SET Index.

Models	MSE	QLIKE	MAD	HMSE	SR
GARCH-N	1.910	5.285	1.628	0.560	0.767
GARCH-t	2.001	5.306	1.703	0.552	0.750
GARCH-GED	1.951	5.296	1.665	0.555	0.783
EGARCH-N	1.410	5.212	1.535	0.887	0.833
EGARCH-t	1.396	5.211	1.524	0.856	0.833
EGARCH-GED	1.400	5.214	1.537	0.886	0.833
GJR-N	2.083	5.316	1.803	0.404	0.783
GJR-t	2.217	5.338	1.875	0.413	0.800
GJR-GED	2.143	5.326	1.834	0.408	0.800
MRS-GARCH-N	1.906	5.264	1.694	0.378	0.850
MRS-GARCH-t	2.065	5.289	1.808	0.407	0.833
MRS-GARCH-2t	1.656	5.239	1.396	0.435	0.783
MRS-GARCH-GED	1.643	5.234	1.401	0.421	0.783

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 4.10 Results for loss function of out-of-sample with forecasting volatility for twenty-two days ahead (a month) of the SET Index.

Models	MSE	QLIKE	MAD	HMSE	SR
GARCH-N	14.177	6.314	4.117	0.651	0.550
GARCH-t	14.890	6.346	4.246	0.660	0.550
GARCH-GED	14.509	6.329	4.182	0.657	0.550
EGARCH-N	9.560	6.553	4.087	4.965	0.550
EGARCH-t	9.697	6.554	4.096	4.901	0.550
EGARCH-GED	9.658	6.573	4.104	5.082	0.550
GJR-N	17.960	6.429	4.509	0.651	0.517
GJR-t	18.656	6.453	4.611	0.659	0.500
GJR-GED	18.237	6.438	4.552	0.655	0.500
MRS-GARCH-N	16.716	6.401	4.300	0.628	0.550
MRS-GARCH-t	16.687	6.403	4.363	0.640	0.633
MRS-GARCH-2t	12.290	6.223	3.689	0.737	0.500
MRS-GARCH-GED	12.586	6.230	3.713	0.723	0.500

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER V

FORECASTING FINANCIAL MARKET WITH PRINCIPAL COMPONENT ANALYSIS AND MARKOV REGIME SWITHCING

The aim of this study was to forecast the returns for the Stock Exchange of Thailand (SET) Index by adding some explanatory variables and stationary Autoregressive Moving-average order p and q (ARMA (p,q)) in the mean equation of returns. In addition, we used the Principal Component Analysis (PCA) to remove possible complications caused by multicollinearity.

In order to forecast the return r_t for their specifid purposes, many researchers have made different assumptions for μ_t appearing in equation (2.48). For example, Kyimaz et al. (2001) assume μ_t to be a regression model with a one week delay, Supoj (2003) assumes μ_t to be an autoregressive process, Mehmet (2008) assumes to be a constant, and Sattayatham et al. (2012) assume μ_t to be an ARMA process with a one week delay.

The financial returns r_t ($r_t := 100 \cdot ln(P_t/P_{t-1})$) for $t = 1, 2, ..., T-1, P_t$ denoting the financial price at time t depend concurrently and dynamically on many economic and financial variables. Since the returns have a statistically significant autocorrelation themselves, lagged returns might be useful in predicting future returns. In order to model these financial returns, Tsay (2005) assume that r_t follows a simple time series model such as a stationary ARMA (p,q) model with some explanatory variables X_{it} . In other words, r_t satisfies the following equation

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(2.48) as:

$$\begin{aligned} r_t &= \mu_t + \varepsilon_t, \\ \mu_t &= \mu_0 + \sum_{i=1}^{n} \beta_i X_{it} + \sum_{s=1}^{p} \phi_s r_{t-s} - \sum_{m=1}^{q} \theta_m \varepsilon_{t-m}, \end{aligned} \qquad (5.1)$$

where

$$X_{it} = 100 \cdot ln(\frac{P_{it}}{P_{i(t-1)}}).$$

Here P_{it} denotes the financial price asset i for $i = 1, 2, ..., n$ at time t, r_{t-s}, $s = 1, 2, ..., p$ is the returns at the sth lag, ε_t represents errors assumed to be a white noise series with an $i.i.d.$ mean of zero and a constant variance σ_ε^2, $\mu_0, \beta_i, \phi_s, \theta_m$ are constants and n, p, q are positive integers.

Note that the variance of errors ε_t in the model 5.1 is assumed to be a constant; some authors use this assumption in the modeling of ground-level ozone (Agirre-Basurko et al., 2006; Pires et al., 2007). In this chapter, we consider the case where the variance of ε_t is not constant. That is, we introduce the heteroskedasticity model to forecast the volatility of returns using GARCH, EGARCH, GJR-GARCH and Markov Regime Switching GARCH (MRS-GARCH) with distribution normal, student-t and general error distribution (GED).

The objective of this chapter is to forecast returns for the Stock Exchange of Thailand (SET) Index by using model 5.1. We vary the process μ_t using four different types and compare the performance of the different types. Moreover, we forecast the volatility of returns with heteroskedasticity models.

In the next section, we present the basics of principal component analysis to remove possible complications caused by the multicollinearity of explanatory variables. The empirical methodology and formulae for model estimation are given in section 5.2. Forecasting of the returns is discussed in section 5.3. Forecasting the volatility of returns is described in section 5.4 and the conclusions are presented

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

in section 5.5.

5.1 Principal Component Analysis : PCA

Given an n-dimensional random variable $X_t = (X_{1t}, X_{2t}, ..., X_{nt})'$ with covariance matrix, Σ_X, $(.)'$ denotes the transposed matrix. Principal component analysis (PCA) is concerned with using fewer linear combinations of X_t to explain the structure of Σ_X . If X_{it} denotes returns as appears in equation (2.48) for $i = 1, 2, ..., n$, then PCA can be used to study the source of variations of these n returns.

Let $(\lambda_1, e_1), ..., (\lambda_n, e_n)$ be the eigenvalue-eigenvector pairs of Σ_X, with the eigenvalues λ_i set up in decreasing order $\lambda_1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$. Then the $i-th$ principal component of X_t is given by $Z_{it} = e_i' X_t = \Sigma_{j=1}^{n} e_{ij} X_{jt}$ for $i = 1, ..., n$.

We note that

$$\begin{aligned} Var(Z_{it}) &= e_i' \Sigma_X e_i = \lambda_i, \\ Cov(Z_{it}, Z_{jt}) &= e_i' \Sigma_X e_j = 0, \end{aligned} \qquad (5.2)$$

for $i \neq j, i, j = 1, 2, ..., n$ and $e_i' = (e_{i1}, ..., e_{in})$ are orthonormal vectors.

In order to cope with the problem of multicollinearity, we transform the explanatory variables in model in equation (2.48) into the principal components. Then the new model for forecasting r_t is

$$r_t = \mu_0 + \sum_{i=1}^{n} \beta_i Z_{it} + \sum_{s=1}^{p} \phi_s r_{t-s} - \sum_{m=1}^{q} \theta_m \varepsilon_{t-m} + \varepsilon_t, \qquad (5.3)$$

where Z_{it} are $i-th$ principal component of explanatory variables at time t. We follows Tsay (2005) by assuming that the asset return series r_t is a weakly stationary process.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.2 Empirical Studies and Methodology.

Naturally, the Thai stock market has a unique characteristic, so the factors influencing the prices of stocks traded in this market are different from the factors influencing other stock markets (Chaigusin et al., 2008). Examples of factors that influence the Thai stock market and the statistics used by researchers who have studied these factors in forecasting the SET Index are shown in Table 5.2.

5.2.1 Data

The data sets used in this study are the daily return closing prices for the SET Index at time t (dependent variables) and the daily return closing prices for twelve factors (explanatory independent variables). These twelve factors are the following:

1. The Dow Jones Index at time $t-1$ ($DJIA$).
2. The Financial Time 100 Index at time $t-1$ ($FSTE$).
3. The SP 500 Index at time $t-1$ (SP).
4. The Nikkei225 Index at time t (NIX).
5. The Hang Seng Index at time t ($HSKI$).
6. The Singapore Straits Time Industrial Index at time t ($SES009$).
7. The Taiwan Stock Weighted Indes at time t ($TWII$).
8. The South Korea Stock Exchange Index at time t ($KOSPI$).
9. The Oil Price in NYMEX (New York Mercantile Exchange) at time t (OIL).
10. The Gold Price in NYMEX at time t ($GOLD$).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

11. The Currency Exchange Rate in Thai Baht to one US dollar at time t (THB/USD).

12. The Currency Exchange Rate in Thai Baht to one Hong Kong dollar at time t (THB/HKD).

The actual closing prices for these twelve factors were obtained from http://www.efinancethai.com. We used data sets from April 5, 2000 to July 5, 2012. We divided these data into two disjoint sets. The first set, from April 5, 2000 to December 30, 2011 was used as a sample (2,873 observations). The second set, from January 3,2012 to July 5, 2012 was used as out-of-sample (125 observations). The plot for the SET Index closing prices and returns is given in Figure 5.1.

Descriptive statistics and the correlations matrix are given in Tables 5.2 and 5.3. As can be seen from Table 5.3, there are highly significant correlations [we mean when test significant $(p < 0.01)$] between the dependent variables and the explanatory variables. Therefore, these explanatory variables were used to predict the SET Index. Also, there are highly significant correlations $(p < 0.01)$ among the explanatory variables. These correlations provide a measure for the linear relations between two variables and also indicate the existence of multicollinearity between the explanatory variables. However, multiple regression analysis based on this dataset also shows that there was a multicollinearity problem with the variance inflation factor $(VIF >= 5.0)$ as shown in Table 5.2. One approaches to avoid this problem is PCA. Hence, we used twelve explanatory variables to find the principal components, and overall descriptive statistics for selected principal components (PCs) as shown in Tables 5.4 and 5.5, respectively.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

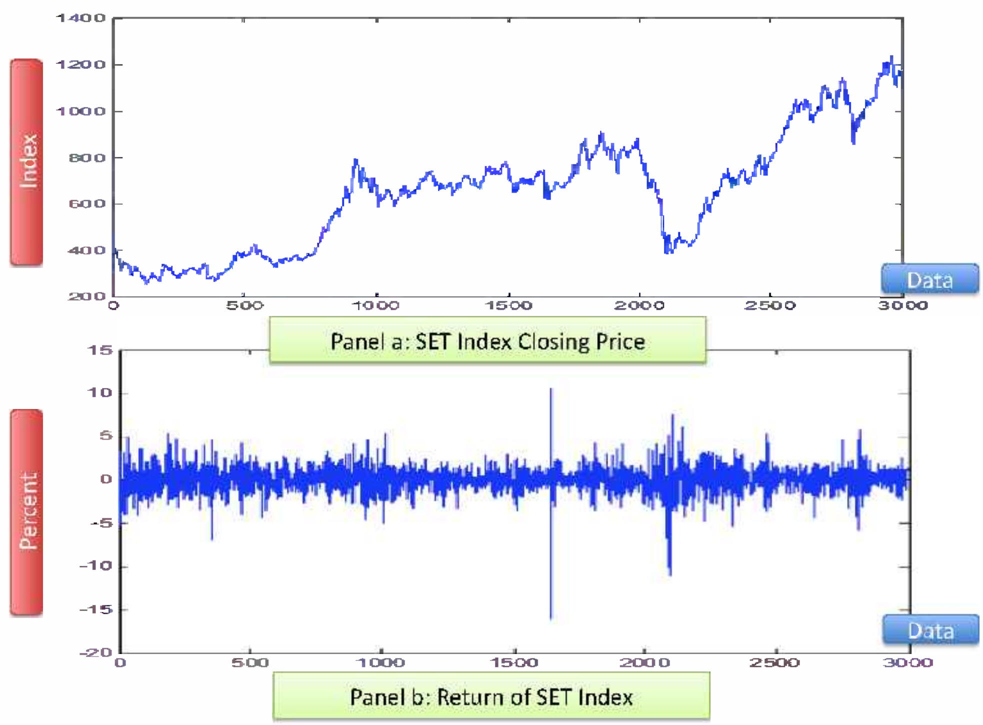


Figure 5.1 Graph of (a) SET Index closing prices and (b) log returns series for the period 5/04/2000 through 5/07/2012.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.2.2 Results of Principal Component Analysis

The Bartletts sphericity test for testing the null hypothesis where the correlation matrix is an identity matrix was used to verify the applicability of PCA. The value of the Bartletts sphericity test for the SET Index was 18,167.07, which implies that the PCA is applicable to our datasets (Table 5.2). Moreover, Kaiser's measure of sampling adequacy was also computed as 0.788, which indicates that the sample sizes were sufficient for us to apply the PCA. The results for PCA (Table 5.4) indicate that there are twelve principal components (PCs) for multiple regression analysis.

5.3 Forecasting the Returns the SET Index by Mean Equations

In this section, we forecast the returns for the SET Index ($r_t = \mu_t + \varepsilon_t$) using four mean equations (μ_t): constant, ARMA (1,1), multiple regression based on PCA, and ARMA (1,1), which includes multiple regression based on PCA. Afterwards, we compare the errors using two loss functions, i.e. mean square error (MSE) and mean absolute error (MAE). The parameters for mean equations for forecasting the SET Index and the value of loss functions are shown in Table 5.5. We found that the mean equation ARMA (1,1) that includes multiple regression based on PCAs (Table 5.5, No. 4) has the best performance (MSE=0.5393, MAE=0.5947). So, we use this mean equation for forecasting the returns for the SET Index.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

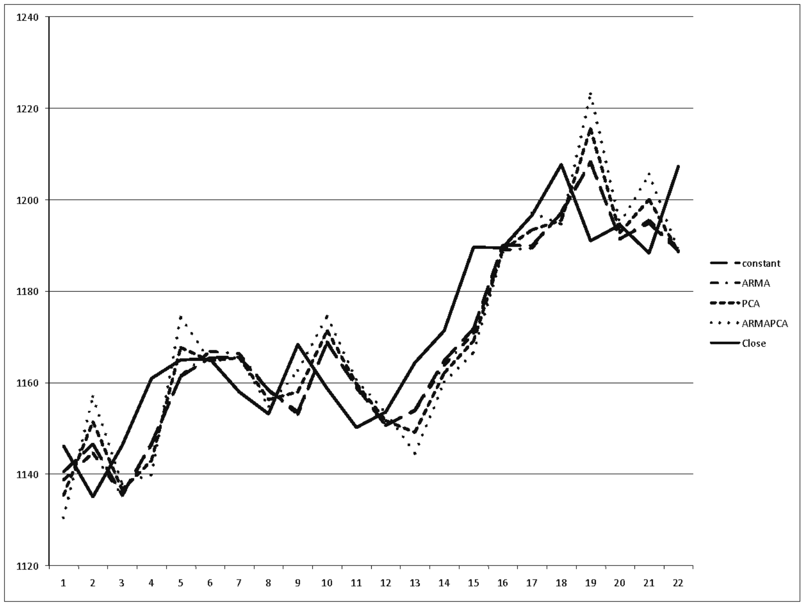


Figure 5.2 Forecasting SET Index with various models.

5.4 Forecasting the Volatility of Returns for the SET Index

We applied Ljung and Box (1978) to test serial correlation for returns (r_t) and squared mean returns adjusted ($(r_t - \mu_t)^2$) where is the mean equation in Table 5.5 (No. 4). We used a specified lag from the first to the tenth lags and we used the twenty-second lag. Serial correlation for returns is confirmed as stationary because the autocorrelation function (ACF) values decrease very fast when lags increase and this is confirmed by the Augmented Dickey-Fuller Test ($-52.76**$) as in Table 5.2. We analyzed the significance of autocorrelation in the squared mean adjusted returns series with the Ljung-Box Q-test and used Engles ARCH test to test the ARCH effects. Therefore, the squared mean for the adjusted return is non-stationary, which suggests conditional heteroskedasticity.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.4.1 Empirical Methodology

This empirical section adopts the GARCH type and the MRS-GARCH (1,1) models to estimate the volatility of the returns on the SET Index. The GARCH type models considered are the GARCH (1,1), EGARCH (1,1) and GJR-GARCH (1,1). In order to account for the fat-tailed feature of financial returns, we considered three different distributions for the innovations: Normal (N), Student-t (t) and Generalized Error Distributions (GED).

GARCH Type Models

Panel A of Table 5.7 presents an estimation of the results for the GARCH type models. It is clear from the table that almost all parameter estimates in the GARCH type models are highly significant at 1 percent. However, the asymmetry effect term ξ in the EGARCH models is significantly different from zero which indicates unexpected negative returns, implying higher conditional variance as compared to the same-sized positive returns. All models display strong persistence in volatility ranging from 0.8895 to 0.9572, that is, volatility is likely to remain high over several price periods once it increases.

Markov Regime Switching GARCH Models.

The estimated results and summary statistics for the MRS-GARCH models are presented in Panel B of Table 5.7. Most parameter estimates in the MRS-GARCH are significantly different from zero at least at the 95 percent confidence level. But α_1, β_1 are not significantly different in some areas. All models display strong persistence in volatility ranging from 0.6650 to 0.9892, that is, volatility is likely to remain high over several price periods once it increases.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.4.2 In-Sample Evaluation

We used various goodness-of-fit statistics to compare volatility models. These statistics are the Akaike Information Criteria (AIC), the Schwarz Bayesian Information Criteria (BIC), and the Log-likelihood (LOGL) values. Table 5.8 presents the results for the goodness-of-fit statistics and loss functions for all volatility models. According to the BIC, the MSE2, and the QLIKE, the GJR model performs best in modelling SET Index volatility. However, the contrast in the AIC, the LOGL, the MSE1, the R2LOG, the MAD2 and the MAD1 suggests that the MRS-GARCH performs best.

5.4.3 Forecasting Volatility for the Out-of-sample in SET Index

In this section, we investigate the ability of the GARCH, EGARCH, GJR-GARCH, and MRS-GARCH models to forecast volatility for the SET Index out-of-sample set. In Table 5.10, we present the results for loss function for out-of-samples in forecasting volatility for one day ahead, five days ahead (short term), ten days ahead, and twenty-two days ahead (long term). We found the GARCH model performs best for one day ahead; the EGARCH model performs best for five days, ten days, and twenty-two days ahead.

5.5 Conclusions

We considered the problem of forecasting returns for the SET Index by using a stationary Autoregressive Moving-average order p and q (ARMA (p,q)) with some explanatory variables. After considering four types of mean equations, we found that ARMA (1,1), which includes multiple regressions based on PCA, has

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the best performance (MSE=0.5393, MAE=0.5947). In forecasting the volatility of the returns for the SET Index, GARCH type models such as GARCH (1,1), EGARCH (1,1), GJR-GARCH (1,1) and MRS-GARCH (1,1) were considered. We found that the GARCH (1,1) model performs best for one day ahead, and the EGARCH (1,1) model performs best for five days, ten days and twenty-two days ahead respectively.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.1 Impact factors on the Stock Exchange of Thailand Index (SET Index).

Variables	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
The Nasdaq Index	na	na	na	X	na	na	na	na
The Down Jones Index	X	X	X	X	X	X	X	X
The S and P 500 Index	na	na	na	X	na	na	na	na
The Nikkei Index	X	X	X	na	X	na	X	X
The Hang Seng Index	X	X	X	na	X	na	X	X
The Straits Times industrial Index	X	X	X	na	na	na	na	na
The Currency Exchange Rate in Thai Baht to one US dollar	na	X	X	na	na	X	X	na
The Currency Exchange Rate in Thai Baht to 100 Japan Yen	na	X	X	na	na	na	na	na
The Currency Exchange Rate in Thai Baht to one Hong Kong dollar	na	na	X	na	na	na	na	na
The Currency Exchange Rate in Thai Baht to one Singapore dollar	na	na	X	na	na	na	na	na
Gold prices	na	X	na	na	X	na	X	na
Oil Prices	na	X	X	na	na	X	na	na
Minimum Loan Rates	na	X	na	na	X	X	X	X

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.2 Descriptive statistics for the SET Index and return explanatory variables.

Explanatory	Mean	S.D.	Skewness	Kurtosis	Correlation	VIF
SET	0.0373	1.4644	-0.69	9.194	1	na
DJIA	0.0047	1.2792	-0.017	7.626	0.219^a	14.581
FSTE	-0.0043	1.328	-0.169	5.718	0.166^a	1.527
SP	-0.0031	1.3647	-0.128	7.764	0.239^a	15.197
NIX	-0.0273	1.5986	-0.499	7.609	0.369^a	2.01
HSKI	0.0053	1.6593	-0.067	8.96	0.495^a	2.405
SES900	0.0122	1.3011	-0.337	7.674	0.507^a	2.15
TWII	-0.0096	1.5716	-0.202	3.348	0.351^a	1.618
KOSPI	0.0272	1.7733	-0.867	9.737	0.410^a	2.152
OIL	0.0413	2.5662	0.087	7.578	0.119^a	1.057
GOLD	0.0581	1.1831	0.137	6.383	0.077^a	1.068
THB/USD	-0.0063	0.4258	0.511	20.223	-0.152^a	2.197
THB/HKD	-0.0059	0.5304	0.57	32.596	-0.107^a	2.175
JB-test			10,741.72 (P-value= 0.000)			
ADF-test			-52.76 (P-value= 0.000)			
KMO			0.79			
Bartlett's Test			18,167.07,df=66, (P-value= 0.000)			

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.3 Correlation matrix for the SET Index and explanatory variables.

r	SET	DJIA	FSTE	SP	NIX	HSKI	SES900	TWII	KOSPI	OIL	GOLD	THB/USD	THB/HKD
SET	1.00												
DJIA	0.22^{a}	1.00											
FSTE	0.17^{a}	0.55^{a}	1.00										
SP	0.24^{a}	0.96^{a}	0.56^{a}	1.00									
NIX	0.37^{a}	0.45^{a}	0.39^{a}	0.48^{a}	1.00								
HSKI	0.50^{a}	0.37^{a}	0.29^{a}	0.40^{a}	0.59^{a}	1.00							
SES900	0.51^{a}	0.33^{a}	0.20^{a}	0.35^{a}	0.53^{a}	0.69^{a}	1.00						
TWII	0.35^{a}	0.30^{a}	0.23^{a}	0.32^{a}	0.45^{a}	0.49^{a}	0.47^{a}	1.00					
KOSPI	0.41^{a}	0.31^{a}	0.26^{a}	0.34^{a}	0.59^{a}	0.61^{a}	0.57^{a}	0.57^{a}	1.00				
OIL	0.12^{a}	0.01	-0.01	0.01	0.06^{a}	0.10^{a}	0.11^{a}	0.06^{a}	0.06^{a}	1.00			
GOLD	0.08^{a}	0.04^{b}	0.03	0.05^{a}	0.07^{a}	0.09^{a}	0.07^{a}	0.02	0.07^{a}	0.20^{a}	1.00		
THB/USD	-0.15^{a}	-0.07^{a}	-0.05^{a}	-0.08^{a}	-0.08^{a}	-0.12^{a}	-0.12^{a}	-0.10^{a}	-0.13^{a}	-0.04^{b}	-0.13^{a}	1.00	
THB/HKD	-0.11^{a}	0.00	0.00	-0.02	0.00	-0.07^{a}	-0.10^{a}	-0.11^{a}	-0.08^{a}	-0.12	-0.02	-0.10^{a}	1.00

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.4 Eigenvalues for PCAs.

	Initial Eigenvalues			Weight for the PCs											
PC	Total	λ_i	$\sum_{i=1}^{j} \lambda_i$	DJIA	FSTE	SP	NIX	HSKI	SES900	TWII	KOSPI	OIL	GOLD	THB/USD	THB/HKD
1	4.266	35.548	35.548	0.732	0.572	0.753	0.775	0.777	0.727	0.656	0.742	0.104	0.119	-0.224	-0.212
2	1.734	14.453	50.001	0.257	0.224	0.249	0.071	-0.036	-0.068	-0.02	-0.065	-0.158	-0.279	0.854	0.842
3	1.482	12.347	62.348	-0.538	-0.451	-0.518	0.155	0.329	0.377	0.32	0.372	0.22	0.039	0.262	0.294
4	1.151	9.595	71.944	0.097	0.069	0.097	-0.029	-0.037	-0.057	-0.149	-0.126	0.731	0.714	0.126	0.153
5	0.789	6.575	78.519	-0.058	0.009	-0.047	0.071	0.042	-0.014	-0.05	0.059	-0.614	0.622	0.05	0.068
6	0.607	5.056	83.576	-0.145	0.367	-0.147	0.015	-0.231	-0.361	0.462	0.167	0.046	0.043	-0.011	0.007
7	0.570	4.749	88.325	0.232	-0.453	0.214	-0.255	-0.148	-0.03	0.41	-0.008	-0.037	0.079	0.002	0.025
8	0.448	3.736	92.061	-0.051	0.264	-0.055	-0.458	0.197	0.256	0.141	-0.193	-0.033	0.026	-0.001	0.021
9	0.355	2.960	95.020	0.04	0.014	0.038	-0.297	0.017	-0.08	-0.199	0.465	0.016	-0.009	0.023	0.000
10	0.299	2.494	97.514	-0.009	0.066	-0.016	0.017	-0.406	0.339	-0.023	0.072	-0.005	0.018	0.068	-0.06
11	0.264	2.198	99.712	-0.002	-0.012	-0.001	-0.005	0.066	-0.048	0.024	-0.021	0.003	0.012	0.359	-0.357
12	0.035	0.288	100.000	0.13	0.001	-0.133	0.003	0.003	0.000	-0.001	0.001	0.000	0.001	0.000	0.000

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.5 Mean equations for return for the SET Index and loss functions.

Model	Mean Equation	MSE	MAE	SR
1.Constant mean	$\mu_t = \bar{r}, \mu_t = 0.0373$	0.8914	0.7576	0.5169
2.ARMA(1,1)	$\mu_t = \phi_0 + \phi_1 r_{t-1} - \theta_1 \varepsilon_{t-1}$	0.8963	0.7627	0.5179
	$\mu_t = 0.5454 r_{t-1} + 0.4951 \varepsilon_{t-1}$			
3.Multiple regression	$\mu_t = \phi_0 + \sum_{i=1}^{n} \alpha_i Z_{it}$			
based on PCAs.	$\mu_t = 0.163 Z_{1t} - 0.055 Z_{2t} + 0.259 Z_{3t} + 0.499 Z_{4t} + 0.059 Z_{5t} + 0.124 Z_{6t}$	0.5444	0.5963	0.5225
	$+0.272 Z_{7t} + 0.215 Z_{8t} + 0.077 Z_{9t} - 0.146 Z_{10t} + 0.41 Z_{11t}$			
4.ARMA (1,1) and Multiple	$\mu_t = \phi_0 + \sum_{i=1}^{n} \alpha_i Z_{it} + \sum_{i=1}^{p} \phi_i r_{t-1} - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$			
regression based on PCAs.	$\mu_t = 0.162 Z_{1t} - 0.054 Z_{2t} + 0.258 Z_{3t} + 0.5 Z_{4t} + 0.059 Z_{5t} + 0.124 Z_{6t}$	0.5393	0.5947	0.5242
	$+0.271 Z_{7t} + 0.214 Z_{8t} + 0.080 Z_{9t} - 0.146 Z_{10t} + 0.405 Z_{11t}$			
	$-0.991 r_{t-1} - 0.996 \varepsilon_{t-1}$			

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.6 The ACF of the SET Index returns series, squared mean adjusted return and results for Engles ARCH test.

Lags	ACF of return			ACF of $(r_t - \mu_t)^2$			Engles ARCH test	
	ACF	LBQ	P	ACF	LBQ	P	ARCH Test	P
1	0.036	3.925	0.048	0.316	300.283	0.000	262.048	0.000
2	0.073	19.764	0.000	0.057	310.008	0.000	273.691	0.000
3	0.007	19.890	0.000	0.037	314.078	0.000	274.194	0.000
4	-0.018	20.862	0.000	0.017	314.974	0.000	274.183	0.000
5	-0.004	20.911	0.001	0.037	319.091	0.000	274.155	0.000
6	-0.048	27.860	0.000	0.023	320.727	0.000	274.116	0.000
7	0.006	27.970	0.000	0.006	320.841	0.000	274.074	0.000
8	-0.016	28.781	0.000	0.036	324.781	0.000	274.036	0.000
9	0.034	32.321	0.000	0.049	332.100	0.000	274.037	0.000
10	0.043	37.820	0.000	0.316	300.283	0.000	262.048	0.000
22	-0.005	62.147	0.000	0.010	358.620	0.000	273.467	0.000

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.7 Summary of results for estimation parameters of the GARCH type models.

Parameter	GARCH			EGARCH			GJR		
	N	t	GED	N	t	GED	N	t	GED
α_0	0.188^a	0.090^a	0.113^a	-0.110^a	-0.139^a	-0.131^a	0.226^a	0.105^a	0.135^a
Std.	0.017	0.018	0.022	0.013	0.018	0.019	0.019	0.019	0.024
α_1	0.113^a	0.118^a	0.118^a	0.227^a	0.214^a	0.219^a	0.198^a	0.168^a	0.180^a
Std.	0.012	0.016	0.019	0.021	0.025	0.029	0.023	0.024	0.026
β_1	0.792^a	0.835^a	0.825^a	-0.011^a	-0.056^a	-0.073^a	0.762^a	0.821^a	0.807^a
Std.	0.018	0.019	0.023	0.010	0.014	0.015	0.021	0.021	0.025
ξ	na	na	na	0.890^a	0.946^a	0.932^a	0.044^a	0.075^a	0.067^a
Std.	na	na	na	0.010	0.010	0.013	0.011	0.019	0.020
ν	na	7.194^a	1.307^a	na	7.515^a	1.331^a	na	7.485^a	1.328^a
Std.	na	0.599	0.025	na	0.642	0.026	na	0.635	0.025
L.	-4982	-4832	-4863	-4957	-4824	-4853	-4957	-4822	-4852
P.	0.908	0.957	0.946	0.890	0.946	0.931	0.891	0.946	0.932
LB	63.325	63.325	63.325	63.325	63.325	63.325	63.325	63.325	63.325
LB^2	673.30	673.62	673.66	671.90	672.94	672.87	672.08	672.94	672.91

[a,b] refer to significance at 99 percent, 95 percent confidence respectively.

L. refers to loglikelihood. P. refers to persistence. LB is Ljung-Box test of innovation at lag 22.

LB^2 is Ljung-Box test of squared innovation at lag 22 and Std. is standard error.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.8 Summary of results for estimation parameters of MRS-GARCH models.

Parameter	MRS-GARCH							
	N		t		2t		GED	
State	$i=1$	$i=2$	$i=1$	$i=2$	$i=1$	$i=2$	$i=1$	$i=2$
$\alpha_0^{(i)}$	6.061^a	0.077^a	0.046^a	1.197^a	0.046^a	1.448^a	0.073^a	1.259^a
Std.	1.057	0.021	0.018	0.468	0.020	0.808	0.021	1.017
$\alpha_1^{(i)}$	0.189	0.060^a	0.067	0.296	0.069^a	0.417^b	0.068^a	0.041
Std.	0.117	0.018	0.020	0.107	0.018	0.193	0.018	0.134
$\beta_1^{(i)}$	0.000	0.835^a	0.883^a	0.359^a	0.873^a	0.300	0.846^a	0.948^a
Std.	0.365	0.019	0.022	0.175	0.021	0.254	0.020	0.453
p	0.571^a		0.983^a		0.980^a		0.990^a	
Std.	0.144		0.008		0.009		0.005	
q	0.983^a		0.907^a		0.823^a		0.447^a	
Std.	0.004		0.040		0.069		0.219	
$\nu^{(i)}$	na	na	8.259^a		11.112^a	3.694^a	1.524^a	
Std.	na	na	0.968		2.692	0.818	0.061	
L.	-4847.87		-4808.71		-4815.98		-4812.81	
σ^2	7.472	0.742	0.924	3.469	0.775	5.121	0.854	6.556
π	0.491	0.509	0.157	0.843	0.103	0.898	0.017	0.983
Pers.	na	0.895	0.950	0.655	0.941	0.717	0.914	0.989

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.9 In-sample evaluation results.

	N	PERS	AIC	BIC	LOGL	MSE1	MSE2	QLIKE	R2LOG	MAD2	MAD1	HMSE
GARCH-N	4	0.908	3.471	3.480	-4982.540	1.085	49.027	1.630	8.916	2.323	0.802	20.064
GARCH-t	5	0.957	3.367	3.378	-4832.110	1.104	49.838	1.649	8.685	2.353	0.792	37.633
GARCH-GED	5	0.946	3.389	3.400	-4863.530	1.099	49.476	1.639	8.781	2.348	0.797	30.080
EGARCH-N	5	0.890	3.455	3.465	-4957.540	1.064	48.025	1.615	8.831	2.285	0.794	16.533
EGARCH-t	6	0.946	3.362	3.375	-4824.090	1.068	48.798	1.631	8.644	2.285	0.782	30.572
EGARCH-GED	6	0.931	3.383	3.395	-4853.300	1.069	48.574	1.622	8.723	2.288	0.788	24.424
GJR-N	5	0.891	3.454	3.465	-4957.140	1.066	47.478	1.615	8.841	2.308	0.796	17.621
GJR-t	6	0.946	3.361	3.374	-4822.700	1.087	48.756	1.631	8.638	2.336	0.787	31.358
GJR-GED	6	0.932	3.382	3.394	-4852.160	1.083	48.309	1.622	8.726	2.334	0.792	25.243
MRS-GARCH-N	10	0.900	3.382	3.403	-4847.870	1.053	48.120	1.641	8.636	2.282	0.782	32.025
MRS-GARCH-t2	12	0.936	3.356	3.381	-4808.710	1.101	50.600	1.639	8.665	2.343	0.791	34.437
MRS-GARCH-t	11	0.958	3.360	3.383	-4815.980	1.093	49.873	1.655	8.598	2.332	0.784	42.232
MRS-GARCH-GED	11	0.953	3.358	3.381	-4812.810	1.059	48.175	1.651	8.606	2.285	0.780	41.623

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.10 Result of loss function for out-of-sample with forecasting volatility.

Panel A: Result loss function of out-of-sample with forecasting volatility for one day ahead.

Model	MSE1	MSE2	QLIKE	MAD1	MAD2	HMSE	SR
GARCH-N	0.742	1.264	0.116	0.651	0.791	0.598	0.273
GARCH-t	0.745	1.277	0.119	0.653	0.794	0.598	0.273
GARCH-GED	0.741	1.263	0.115	0.651	0.790	0.598	0.273
EGARCH-N	0.930	1.958	0.245	0.728	0.984	0.598	0.240
EGARCH-t	0.957	2.104	0.258	0.737	1.011	0.598	0.248
EGARCH-GED	0.941	2.026	0.249	0.732	0.995	0.598	0.248
GJR-GARCH-N	0.955	2.212	0.250	0.734	1.008	0.598	0.273
GJR-GARCH-t	0.975	2.347	0.259	0.740	1.029	0.598	0.265
GJR-GARCH-GED	0.965	2.287	0.254	0.737	1.019	0.598	0.265
MRS-GARCH-N	0.908	1.853	0.221	0.718	0.961	0.598	0.240
MRS-GARCH-2t	0.882	1.987	0.198	0.703	0.934	0.598	0.273
MRS-GARCH-t	0.936	1.913	0.249	0.732	0.990	0.598	0.231
MRS-GARCH-GED	0.881	2.051	0.193	0.701	0.933	0.598	0.273

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.11 Result of loss function for out-of-sample with forecasting volatility (Cont.).

Panel B: Result loss function of out-of-sample with forecasting volatility for five days ahead.

Model	MSE1	MSE2	QLIKE	MAD1	MAD2	HMSE	SR
GARCH-N	3.884	33.761	1.129	1.499	4.184	0.598	0.248
GARCH-t	3.947	34.597	1.140	1.513	4.250	0.598	0.248
GARCH-GED	3.904	33.960	1.133	1.504	4.205	0.598	0.248
EGARCH-N	3.274	24.855	1.022	1.371	3.547	0.598	0.240
EGARCH-t	3.335	26.017	1.030	1.382	3.609	0.598	0.256
EGARCH-GED	3.288	25.268	1.023	1.373	3.561	0.598	0.248
GJR-GARCH-N	4.828	54.530	1.244	1.663	5.159	0.598	0.240
GJR-GARCH-t	4.938	57.343	1.257	1.681	5.273	0.598	0.240
GJR-GARCH-GED	4.877	55.893	1.249	1.671	5.209	0.598	0.240
MRS-GARCH-N	4.659	48.704	1.216	1.632	4.985	0.598	0.248
MRS-GARCH-2t	3.990	37.690	1.138	1.513	4.291	0.598	0.248
MRS-GARCH-t	4.799	50.266	1.243	1.663	5.131	0.598	0.240
MRS-GARCH-GED	4.030	39.615	1.138	1.517	4.331	0.598	0.231

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.12 Result of loss function for out-of-sample with forecasting volatility (Cont.).

Panel C: Result loss function of out-of-sample with forecasting volatility for ten days ahead.

Model	MSE1	MSE2	QLIKE	MAD1	MAD2	HMSE	SR
GARCH-N	8.219	145.378	1.589	2.192	8.856	0.598	0.198
GARCH-t	8.428	150.989	1.607	2.223	9.075	0.598	0.223
GARCH-GED	8.299	147.145	1.597	2.204	8.940	0.598	0.207
EGARCH-N	4.667	51.194	1.251	1.640	5.141	0.597	0.182
EGARCH-t	4.737	53.020	1.259	1.652	5.214	0.597	0.182
EGARCH-GED	4.669	51.557	1.250	1.640	5.143	0.597	0.182
GJR-GARCH-N	9.863	216.313	1.686	2.391	10.555	0.598	0.190
GJR-GARCH-t	10.091	225.828	1.701	2.419	10.790	0.598	0.182
GJR-GARCH-GED	9.949	219.959	1.692	2.402	10.643	0.598	0.182
MRS-GARCH-N	9.671	204.750	1.664	2.360	10.353	0.598	0.174
MRS-GARCH-2t	7.391	120.506	1.528	2.077	7.994	0.598	0.207
MRS-GARCH-t	9.930	210.715	1.688	2.400	10.625	0.598	0.182
MRS-GARCH-GED	7.514	127.357	1.534	2.090	8.120	0.598	0.190

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table 5.13 Result of loss function for out-of-sample with forecasting volatility (Cont.).

Panel D: Result loss function of out-of-sample with forecasting volatility for twenty two days ahead.

Model	MSE1	MSE2	QLIKE	MAD1	MAD2	HMSE	SR
GARCH-N	19.886	808.025	2.132	3.431	21.418	0.598	0.231
GARCH-t	20.559	852.974	2.154	3.493	22.119	0.598	0.248
GARCH-GED	20.152	823.431	2.141	3.456	21.695	0.598	0.248
EGARCH-N	5.801	82.676	1.412	1.834	6.616	0.594	0.198
EGARCH-t	5.928	86.247	1.424	1.854	6.751	0.594	0.198
EGARCH-GED	5.803	83.054	1.412	1.834	6.618	0.594	0.198
GJR-GARCH-N	22.451	1046.189	2.198	3.637	24.070	0.598	0.215
GJR-GARCH-t	22.920	1082.107	2.211	3.678	24.558	0.598	0.207
GJR-GARCH-GED	22.559	1051.835	2.201	3.647	24.183	0.598	0.207
MRS-GARCH-N	22.658	1072.927	2.193	3.640	24.278	0.598	0.182
MRS-GARCH-2t	14.815	453.499	1.967	2.966	16.135	0.598	0.223
MRS-GARCH-t	23.064	1093.953	2.208	3.681	24.703	0.598	0.174
MRS-GARCH-GED	15.129	477.250	1.977	2.994	16.461	0.598	0.207

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER VI

CONCLUSION

In the time series, the stock price was transformed to return series for stationary process. For the purpose of forecasting, one normally uses the mean equation. However, the constant mean equation cannot be used for forecasting due to inaccuracy of the financial data, since the financial returns depend concurrently and dynamically on many economic and financial variables. The return has a statistically significant autocorrelation which indicates that the lagged returns might be useful in predicting of returns. Thus, we considered to choose some explanatory variables and stationary Autoregressive Moving-average order p and q (ARMA (p,q)) by adding them into the mean equation for increased accuracy. During the process approach, we found two problems. The first problem is that some explanatory variables occur multicollinearity, and there is high correlation in a regression model. We use Principal Component Analysis (PCA) to remove possible complications caused by multicollinearity. The other problem is that the variances of the residuals are not constant and possibly time-dependent (Heteroskedasticity). We used the volatility models to forecast volatility of this approach, considering the models of volatility such as GARCH, EGARCH, GJR-GARCH and MRS-GARCH (Markov Regime Switching GARCH) models.

In Chapter III, we forecasted volatility of the SET50 Index by using the Markov Regime Switching GARCH (MRS-GARCH) models. These models allow volatility to have different dynamics according to unobserved regime variables. The main purpose of this chapter was to find out whether the MRS-GARCH models are

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

an improvement on the GARCH type models in terms of modelling and forecasting the SET50 Index closing price volatility. We compared the MRS-GARCH (1,1) models with GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) models. All models were estimated under three distributional assumptions that are Normal, Student-t and GED. Moreover, Student-t distribution which takes different degrees of freedom in each regime is considered for the MRS-GARCH models. We first analyzed in-sample performance of various volatility models to determine the best form of the volatility model over the period 4/01/2007 through 30/08/ 2011. As expected, volatility is not constant over time and exhibits volatility clustering showing large changes in the price of an asset often followed by large changes, and small changes often followed by small changes.

This empirical part adopts GARCH type and MRS-GARCH models to estimate the volatility of the gold price. In order to account for fat tailed features of financial returns, we considered three different distributions for the innovations. Almost all parameter estimates in GARCH type models are highly significant. Most parameter estimates in MRS-GARCH are significantly different from zero. However, the results of goodness-of-fit statistics and loss functions for all volatility models show different results. The trading details we have used describe forecasts of closed price of gold price between 1/08/2011-30/08/2011 and trading in gold future contract (GF10Q11). We found the cumulative returns with the Markov Regime Switching GARCH-N (MRS-GARCH-N) model and the GJR-N model give us higher cumulative returns than the other models.

In Chapter IV, we modelled the returns of the SET Index by mean equation with the day of the week effect and the autoregressive moving-average order p and q (ARMA(p, q)) and forecasted the volatility of the SET Index by the GARCH, EGARCH, GJR-GARCH and MRS-GARCH models. Moreover we compared their

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

volatility forecast performance with one day, one week, two weeks and one month returns. We found that Friday is the day which effect on the SET Index and ARMA (3, 3) the best process for forecasting in the mean equation. The GARCH-type models perform best in the short term (one day and a week). On the other hand, the MRS- GARCH models perform best in the long term (two weeks and a month) for forecasting volatility of the SET Index.

In Chapter V, we considerd the problem of forecasting returns of the SET Index using a stationary Autoregressive Moving-average order p and q (ARMA (p,q)) with some explanatory variables. By considering the four types of mean equations, we found that ARMA (1,1) which includes multiple regression based on PCA has best performance (MSE=0.5393, MAE=0.5947). For the forecasting volatility of the returns SET Index, and the GARCH type models such as GARCH (1,1), EGARCH (1,1), GJR-GARCH (1,1) and MRS-GARCH (1,1) have been considered. We found that the GARCH (1,1) model was the best performed for one day ahead, the EGARCH (1,1) model was the best performed for five days, ten days and twenty-two days ahead respectively.

For further study, three or four volatility regimes settings can be considered rather than two-volatility regimes. Also, one may use the Markov Regime Switching with other volatility models e.g. the EGARCH, the GJR. In addition, the performance of the MRS-GARCH model can be compared in terms of their ability to forecast Value at Risk (VaR) for long and short positions.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

REFERENCES

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

REFERENCES

Abramson A. and Cohen, I. (2007a). State smoothing in Markov-switching time frequency GARCH models. **IEEE Signal Processing Letters**. 13(6): 377-380.

Abramson A. and Cohen, I. (2007b). On the stationarity of Markov-switching GARCH processes. **Econometric Theory**. 23: 485-500.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. **International Economic Review**. 39: 885-905.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003). Modelling and forecasting realized volatility. **Econometrica**. 71: 579-625.

Ane T. and Rangau, L. U. (2006). Stock market dynamics in a regime-switching asymmetric power GARCH model. **International Review Of Financial Analysis**. 15: 109-129.

Apolinario, R. M. C. Santana, O. M. and Sales, L. J. (2006). Day of the week effect on European stock markets. **International Research Journal of Finance and Economics**. 2: 53-70.

Areal, N. M. P. C. and Taylor, S. J. (2002). The realized volatility of FTSE-100 futures prices. **Journal of Futures Markets**. 22(7): 627-648.

Barucci, E. and Reno, R. (2002). On measuring volatility of diffusion processes with high frequency data. **Economics Letters**. 74: 371-378.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Bautista, C. C. (2003). Stock market volatility in the Philippines. **Applied Economics Letters**. 10(5): 315-318.

Bera, A. K. and Higgins, M. L. (1993). ARCH models: Properties, estimation and testing. **Journal of Economic Surveys**. 7(4): 305-365.

Black, F. (1976). Studies in stock price volatility changes, Proceeding of the 1976 Business Meeting of the Business and Economics Statistics Section, **American Statistical Association**. 177-181.

Bollerslev, T., Chou, R. Y. and Kroner, K. P. (1992). ARCH modelling in finance: A review of the theory and empirical evidence. **Journal of Econometrics**. 52: 5-59.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, **Journal of Econometrics**. 31(3): 307-327.

Box, G. E. D. Jenkins, G. M. and Reinsel, G. C. (1994). **Time Series Analysis : Forecasting and Control**. 3rd edition. Prentice Hall, Englewood Cliffs. NJ.

Brunetti, C. Mariano, R. S. Scotti, C. and Tan, A. H. H. (2007). Markov switching GARCH models of currency turmoil in Southeast Asia. **Board of Governors of the Federal Reserve System** International Finance Discussion Papers.

Cai, J. (1994). A markov model of unconditional variance in ARCH. **Journal of Business and Economic Statistics**. 12(3): 309-316.

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). **The Econometrics of Financial Markets**. Princeton University Press, Princeton, NJ.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Campbell, J. Y. and Kyle, A. S. (1993). Smart money, noise trading and stock price behavior, **Review of Economic Studies**. 60: 1-34.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance foe speculative prices, **Econometrica**. 41: 135-156.

Daouk H. and Guo, J. Q. (2004). Switching asymmetric GARCH and options on a volatility index. **Journal of Futures Markets**. 24(3): 251-282.

Ding, Z. Engle, R. F. and Granger, C. W. J. (1993). A long memory property of stock market returns and a new model. **Journal of Empirical Finance**. 1: 83-106.

Dueker, M. J. (1997). Markov switching in GARCH processes and mean-reverting stock market volatility. **Journal of Business and Economic Statistics**. 15(1): 26-34.

Easley, D. and Ohara, M. (1992). Time and the process of security price adjustment, **Journal of Finance**. 47: 577-605.

Easton, S. and Faff, R. (1994). An examination of the robustness of the day of the week effect in Australia. **Applied Financial Economics**. 4: 99-110.

Edel, T. and Brian, M. (2005). An APGARCH Investigation of the Main Influences on the Gold Price. [Online]. Available : http://ssrn.com/abstract=792205 or http://dx.doi.org/10.2139/ssrn.792205.html.

Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. **Econometric Reviews**. 5: 1-50.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, **Econometrica**. 50(4): 987-1008. doi:10.2307/1912773 .

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, **Journal of Finance**. 48: 1749-1801.

Fama, E. F. (1963). Mandelbrot and the stable paretian distribution, **Journal of Business**. 36: 420-429.

Fama, E. F. (1965). The behavior of stock market prices, **Journal of Business**. 38: 34-105.

Fong, W. M. and See, K. H. (2001). Modelling the conditional volatility of commodity index futures as a regime switching process. **Journal of Applied Econometrics**.16: 133-163.

Fong, W. M. and See, K. H. (2002). A markov switching model of the conditional volatility of crude oil futures prices. **Energy Economics**. 24: 71-95.

Fong, W. M. (1998). The dynamics of DM per pounds exchange rate volatility: A SWARCH analysis. **International Journal of Finance and Economics**. 3: 59-71.

Fornari, F. and Mele, A. (1997). Sign and volatility-switching ARCH models: Theory and application to international stock markets. **Journal of Applied Econometrics**. 12: 49-67.

Francq, C. and Zakoian, J. M. (2005). The L2-structures of standard and switching regime GARCH models. **Stochastic Processes and their Applications**. 115: 1557-1582.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Franses, P. H. and Van Dijk, D. (2000). **Non-Linear Time Series Models in Empirical Finance**. Cambridge University Press, Cambridge.

Gallant, A. R., Rossi, P. E. and Tauchen, G. (1992). Stock prices and volume, **Review of Financial Studies**. 5: 199-242.

Gallant, A. R., Rossi, P. E. and Tauchen, G. (1993). Nonlinear dynamic structures, **Econometrica**. 61: 871-907.

Glosten, L. R., Jagannathan, R. and Runkle, D. (1993), On the relation between the expected value and the nominal excess return on financials, **Journal of Finance**. 48(5): 1779-1801. doi:10.2307/2329067.

Granger, C. W. J. and Poon, S. H. (2003). Forecasting financial market volatility: A review. **Journal of Economic Literature**. 41(2): 478-539.

Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process, **Journal of Financial Economics**. 42(1): 27-62. doi:10.1016/0304-405X(96)00875-6.

Guo, Z. F. and Cao, L. (2011). A smooth transition GARCH model with asymmetric transition phases, proceedings of international conference of financial engineering, London, 6-8 July 2011. [On-line]. Available : http://www.iaeng.org/publication/WCE2011/WCE2011.html.

Guo, Z. F. and Cao, L. (2011). An asymmetric smooth transition GARCH model, **IAENG Journals**. [On-line]. Available : http://www.iaeng.org/IJAM/issuesv41/issue4/IJAM41410.html.

Haas, M., Mittnik, S. and Paollela, M. S. (2004). A new approach in markov switching GARCH models. **Journal Of Financial Econometrics**. 2(4): 493-530.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. **Econometrica**. 57: 357-384. doi: 10.2307/1912559.

Hamilton, J. D. (1990). Analysis of time series subject to change in regime, **Journal of Econometrics**. 45(1-2): 39-70. doi:10.1016/0304-4076(90)90093-9.

Haminton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and change in regime, **Journal of Econometrics**. 64(1-2): 307-333. doi:10.1016/0304-4076(94)90067-1.

Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH (1,1) **Journal of Applied Econometrics**. 20: 873889.

Hentschel L. F. (1995). All in the family: Nesting linear and nonlinear GARCH models. **Journal of Financial Economics**. 39: 139-64.

Jarque, C. M. and Bera, A. K. (1987). A test of normality of observations and regression residuals. **International Statistical Review**. 55: 163-172.

Jolliffe, I. T. (2002). **Principal Component Analysis**. 2nd ed. Springer-Verlag New York Inc.

Kaufmann, S. and Scheicher, M. (2006). A switching ARCH model for the German DAX index. **Studies in Nonlinear Dynamics and Econometrics**. 10(4).

Kaufmann, S. and Schnatter, S. F. (2002). Bayesian analysis of switching ARCH models. **Journal of Time Series Analysis**. 23(4): 425-458.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Kim, C.J. and Nelson, C.R. (1999). **State-space models with regime switching: Classical and Gibbs-sampling approaches with applications.** 1st ed., MIT Press, Cambridge, MA., ISBN-10: 0262112388, 297.

Klaanssen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH, **Empirical Economics**. 27(2): 363-394. doi:10.1007/s00181010010.

Kyimaz, H. and Berument, H. (2001). The day of the week effect on stock market volatility. **Journal of Economics and Finance**. 25(2): 181-193.

Lamoureux, C. G. and Lastrapes, W. D. (1990). Persistence in variance, structural change and the GARCH model. **Journal of Business Economics Statistics**. 8: 225-234.

Lee, H. T. and Yoder, J. K. (2007). A Bivariate markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. **Applied Economics**, 39(10): 1253-1265.

Leon Li, M. Y. and William Lin, H. W. (2003). Examining the volatility of Taiwan stock index returns via a three-volatility-regime markov switching ARCH model. **Review of Quantitative Finance and Accounting**. 21(1): 123-139.

Leon Li, M. Y. (2007). Volatility states and international diversification of international stock markets. **Applied Economics**. 39(14): 1867-1876.

Mandelbrot, B. (1963). The variation of certain speculative prices, **Journal of Business**. 36: 394-416.

Mandelbrot, B. and Taylor, H. (1967). On the distribution of stock prices differences. **Operations Research**. 15: 1057-1062.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Marcucci, J. (2005). Forecasting financial market volatility with regime-switching GARCH Model, **Working Paper**. University of California, San Diego.

Martens, M. (2002). Measuring and forecasting S and P 500 index futures volatility using high frequency data. **Journal of Futures Markets**. 22(6): 497-518.

Mehmet, A. (2008). Analysis of Turkish financial market with markov regime switching volatility models, **Master degree Thesis**. The Middle East Technical University, Ankara.

Mendes, M. (2009). Multiple linear regression models based on principal component scores to predict slaughter weight of broiler. **European Poultry Science Archiv fr Geflgelkunde Revue de Science Avicole Europenne**. 73(2): 139-144.

Miralles, J. L. M. and Quiros, M. D. M. M. (2000). An empirical analysis of the weekday effect on the Lisbon stock market over trading and non-trading periods. **Portuguese Review of Financial Markets**. 3: 5-14.

Mukhopadhyay, P. (2009). **Multivariate Statistical Analysis**. World Scientific Publishing Co. Pte. Ltd., London.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A New Approach. **Econometrica**. 59(2): 347-370. doi:10.2307/2938260.

Pagan, A. R. (1996). The econometrics of financial markets. **Journal of Empirical Finance**. 3: 15-102.

Perez, G. Q. and Timmermann, A. (2000). Firm size and cyclical variations in stock returns. **The Journal of Finance**. 55(3): 1229-1262.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Pires, J. (2008). Selection and validation of parameters in multiple linear and principal component regressions. **Environmental Modelling and Software**. 1: 50-55.

Rama, C. (2001). Empirical properties of asset returns: stylised facts and statistical issues. **Quantitative Finance**. 1: 223-236.

Sattayatham, P., Sopipan, N. and Premanode B. (2012). Forecasting the stock exchange of Thailand using day of the week effect and Markov regime switching GARCH. **American Journal of Economics and Business Administration.** 4: 84-93. doi:10.3844/ajebasp.2012.84.93

Sentana, E. (1995). Quadratic ARCH models. **Review of Economic Studies**. 62: 639-661.

Snedecor, G.W. and Cochran, W.G. (1980). **Statistical Methods**. 7th edition, Iowa State University Press, Ames ,IA.

Sopipan, N., Sattayatham, P. and Premanode, B. (2012). Forecasting volatility of gold price using markov regime switching and trading strategy, **Journal of Mathematical Finance**. 2(1): 121-131. doi: 10.4236/jmf.2012.21014.

Supoj, C. (2003). Investigation on regime switching in stock market: Case study of Thailand and US. **Thesis in Thammasat University**, Bangkok, Thailand.

Susmel, R. (1998). Switching volatility in Latin American emerging markets. **Emerging Markets Quarterly**. 2.

Susmel, R. (2000). Switching volatility in private international equity markets. **International Journal of Finance and Economics**. 5(4): 265-283.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Tsay, S. (2005). **Analysis of Financial Time Series**. 2nd ed. Wiley-Interscience, USA.

Ulussever, T., Yumusak, I. G. and Kar, M. (2011). The day-of-the-week effect on the Saudi stock exchange: A non-linear GARCH analysis. **Journal of Economic and Social Studies**. 1: 9-23.

Yu, S. W. (2001). Index futures trading and spot price volatility. **Applied Economics Letters**. 8: 183-186.

Zakoian, J. M. (1994). Threshold heteroskedastic models. **Journal of Economic Dynamics and Control**. 18(5): 931-955.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CURRICULUM VITAE

Name: Mr. Nop Sopipan

Office Address: Faculty of Science and Technology, Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima, 30000 , Thailand.

Educational Background:

- 2005. M.Sc. in Applied Statistics, Khon Kaen University, Khon Kaen , Thailand.
- 2003. B.Ed. in Mathematics Education, Khon Kaen University, Khon Kaen, Thailand.

Work Experience:

- 2011. Special Lecturer: Investment and Forecasting in Business in Prince of Songkla University, Pattani Campus,Thailand.
- 2011. Practical Training Program conducted by Plantheon Trading Company Limited in the capacity of Future Trade Analyst.Bangkok , Thailand.
- 2006-2009. Lecturer in Statistics at Nakhon Ratchasima Rajabhat University., Nakhon Ratchasima, Thailand.
- 2005. Lecturer in Statistics at Udornthani Rarjabhat University,Thailand.

Scholarships:

- 2009. Lecturer Development of Nakhon Ratchasima Rajabhat University., Nakhon Ratchasima, Thailand.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor