ANALYSIS OF TURKISH STOCK MARKET WITH MARKOV REGIME SWITCHING VOLATILITY MODELS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET ALİ KARADAĞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF FINANCIAL MATHEMATICS

AUGUST 2008

Approval of the Graduate School of Applied Mathematics

Prof. Dr. Ersan AKYILDIZ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Ersan AKYILDIZ
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Dr. C. Coşkun KÜÇÜKÖZMEN
Co-advisor

Assist. Prof. Dr. Seza DANIŞOĞLU
Supervisor

Examining Committee Members

Assist. Prof. Dr. Seza DANIŞOĞLU

Dr. C. Coşkun KÜÇÜKÖZMEN

Prof. Dr. Gerhard Wilhelm WEBER

Assoc. Prof. Dr. Azize HAYFAVİ

Assist. Prof. Dr. Kasırga YILDIRAK

Assist. Prof. Dr. Ömür UĞUR

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Mehmet Ali KARADAĞ

Signature:

ABSTRACT

ANALYSIS OF TURKISH STOCK MARKET WITH MARKOV REGIME SWITCHING VOLATILITY MODELS

Karadağ, Mehmet Ali

M.Sc., Department of Financial Mathematics

Supervisor: Assist. Prof. Dr. Seza Danışoğlu

Co-advisor: Dr. C. Coşkun Küçüközmen

August 2008, 88 pages

In this study, both uni-regime GARCH and Markov Regime Switching GARCH (SW-GARCH) models are examined to analyze Turkish Stock Market volatility. We investigate various models to find out whether SW-GARCH models are an improvement on the uni-regime GARCH models in terms of modelling and forecasting Turkish Stock Market volatility. As well as using seven statistical loss functions, we apply Superior Predictive Ability (SPA) test of Hansen (2005) and Reality Check test (RC) of White (2000) to compare forecast performance of various models.

Keywords: Volatility, Markov Regime Switching GARCH models, Turkish Stock Market, Superior Predictive Ability test.

ÖZ

TÜRKİYE HİSSE SENETLERİ PİYASASININ MARKOV REGIME SWITCHING VOLATİLİTE MODELLERİ İLE ANALİZİ

Karadağ, Mehmet Ali

Yüksek Lisans, Finansal Matematik Bölümü

Tez Yöneticisi: Yar. Doç. Dr. Seza Danışoğlu

Tez Yönetici Yardımcısı: Dr. C. Coşkun Küçüközmen

Ağustos 2008, 88 sayfa

Bu çalışmada Türkiye Hisse Senetleri Piyasası volatilitesinin analiz edilmesi amacıyla *uni-regime GARCH* ve *Markov Regime Switching GARCH* (SW-GARCH) modelleri incelenmiştir. Türkiye Hisse Senetleri Piyasası volatilitesinin modellenmesi ve öngörürülmesi bakımından SW-GARCH modellerinin uni-regime GARCH modellerine göre daha iyi tahminler yapıp yapmadığı araştırılmıştır. Bir çok modelin öngörü performanslarını karşılaştırmak amacıyla yedi istatistiksel kayıp fonksiyonunun yanında *Superior Predictive Ability* (Hansen, 2005) ve *Reality Check* (White, 2000) testleri de kullanılmıştır.

Anahtar Kelimeler: Volatilite, Markov Regime Switching GARCH modelleri, Türkiye Hisse Senetleri Piyasası, Superior Predictive Ability testi.

To my family

Acknowledgements

I am grateful to my parents for their continuous support, encourage and motivation throughout my life.

I would like to express my sincere gratitude to my supervisor Assist. Prof. Dr. Seza DANIŞOĞLU who gave me the opportunity to complete my thesis.

I would like to thank my co-supervisor Dr. C. Coşkun KÜÇÜKÖZMEN for his ideas, suggestions and support during this study.

I want to thank the members of my committee, Prof. Dr. Gerhard Wilhelm WEBER, Assoc. Prof. Dr. Azize HAYFAVİ, Assist. Prof. Dr. Kasırga YILDIRAK and Assist. Prof. Dr. Ömür UĞUR for their helpful suggestions. Also, I want to thank all other members of the Institute of Applied Mathematics for their genialness and sincerity.

I have been fortunate to have the support of my great friends and colleagues throughout my study. I want to thank my friends at the Institute of Applied Mathematics, specifically Ethem GÜNEY, Hüseyin ŞENTÜRK, Özlem DURSUN, Dilek OĞUZ and Orçun KAYA.

Lastly but definitely not least, I am grateful to Emine TEKEREK for her valuable encouragement and support during this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Volatility can simply be represented as fluctuations in returns. This issue is extremely crucial for many financial activities such as risk management, derivative pricing, hedging, market making and portfolio management. Over the last decades, there has been enormous interest in modelling and forecasting volatility among both market professionals and academicians. Given the importance of volatility, many models have been developed. These models can be collected under three groups: econometric modelling such as Generalized Autoregressive Conditionally Heteroscedasticity (GARCH) type and stochastic volatility type models, implied volatility obtained from option prices and realized volatility obtained from high frequency data. GARCH type models are the most used ones for modelling time varying volatility in finance.

GARCH type models have been very popular since they are simple, easier to model and found quite successful in modelling time varying volatility. Also, they provide accurate volatility forecasts[1]. The first Autoregressive Conditionally Heteroscedasticity model (ARCH) was proposed by Engle (1982) and this study made him to win Nobel Prize in 2003 for his contributions to modelling volatility. Then, Bollerslev (1986) improved the ARCH models by introducing the Generalized ARCH models. The GARCH models mainly capture three characteristics of financial returns. First one is *volatility clustering* that large changes tend to be followed by large changes and small changes tend to be followed by small changes

[1]See Andersen and Bollerslev (1998)

(Mandelbrot (1963), Fama (1963)). Second is *fat tailedness* (excess kurtosis) that financial returns often display a fatter tail than a standard normal distribution and third one is *leverage effect* that negative returns result in higher volatility than positive returns of the same size.

Empirical studies suggest that parameter estimates of the GARCH models usually imply a high degree of persistence in conditional volatility of financial returns[2]. Hamilton and Susmel (1994) stated that the spurious high persistence problem in GARCH type models can be solved by combining Markov Regime Switching model with ARCH models and firstly introduced Markov Regime Switching ARCH models (SWARCH). Gray (1996) and Duaker (1997) extended this method to GARCH specification (SW-GARCH). The idea behind regime switching model is that as market condition changes, the factors that influence volatility change. For example, conditional volatility processes behave very differently in the period of crises (or recession) relative to the usual market conditions.

In the SW-GARCH setting, volatility level switches between two levels of volatility namely high and low volatility regimes[3]. All parameters of GARCH model take different values in each regime. In contrast to using dummy variable for pre-determined sub periods, the regimes are unobservable variables and estimated along with the other parameters of the model using maximum likelihood method.

In this thesis, uni-regime GARCH and Markov Regime Switching GARCH (SW-GARCH) models are examined to analyze Turkish Stock Market while early studies on Turkish Stock Market only consider the uni-regime GARCH models. We compare those models in order to see which ones are better in modelling the Turkish Stock Market volatility. We use four goodness of fit statistics and seven statistical loss functions to evaluate in-sample estimation performance of various models. Also, we attempt to detect if any structural breaks appear in volatility process of Turkish Stock Market. In order to proxy Turkish Stock Market, we

[2]See Bollerslev and Engle (1993), Ding and Granger (1996), Engle and Patton (2001)

[3]There can be more than two regimes. For simplicity, we assume presence of two regimes.

use Istanbul Stock Exchange 100 index (ISE-100)[4].

The main purpose of modelling volatility is forecasting future volatility[5]. Therefore, we also examine whether SW-GARCH models can contribute to the forecasting accuracy of Turkish Stock Market volatility. As well as using seven statistical loss functions, we apply Superior Predictive Ability (SPA) test of Hansen (2005) and Reality Check test (RC) of White (2000) to compare forecast performance of various models.

The thesis is organized as follows: uni-regime GARCH models including GARCH, EGARCH and GJR-GARCH models are presented in Chapter 2. In Chapter 3, Markov Regime Switching models are discussed in detail. Estimation and in-sample evaluation results are given in Chapter 4. In Chapter 5, Statistical loss functions, SPA test and RC test are described and out-of-sample forecasting performance of various models are discussed. Conclusion is given in Chapter 6.

1.2 Literature Review

As being basic risk measure in risk management, volatility have attracted enormous attention by researchers in recent years and vast literature has been accumulated on this subject. Since the introduction of ARCH model by Engle (1982) and its generalized version GARCH model by Bollerslev (1986), these type models have received considerable attention by research community. In ARCH models, current conditional volatility is determined by squared errors in previous p periods and a constant. The current conditional volatility in GARCH models is formulated as a linear function of squared errors in previous p periods and

[4]ISE is one of the fastest growing emerging stock markets. In recent years, market capitalization and foreign investment have noticeably increased. The ISE-100 index is an index consists of the largest and liquid 100 stocks, and regarded as a main indicator of Turkish Stock Market.

[5]A comprehensive literature review on volatility forecasting can be found in Poon and Granger (2003).

conditional variances in previous q periods.

The volatility of financial returns is usually affected asymmetrically from positive and negative shocks[6]. 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 normal distribution[7] 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, generalized error distribution (Nelson, 1991) and 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 proven success of GARCH type models 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 variance process. Following these ideas, Cai (1994) and Hamilton and Susmel (1994) have independently introduced Markov Regime Switching ARCH model (SWARCH) which combines Markov Switching model of Hamil-

[6]Black (1976), Engle and Ng (1993).

[7]Mandelbrot (1963), Fama (1963, 1965).

ton (1989, 1990) with ARCH specification. SWARCH model was designed to capture regime changes in volatility with unobservable state variable following 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 GARCH specification is better to fit financial data, Cai (1994) and Hamilton and Susmel (1994) used ARCH specification to overcome infinite path dependence problem arising in Markov Regime Switching GARCH model (SW-GARCH). On the other hand, Gray (1996) proposed a new approach that allows tractable estimation of the SW-GARCH model and eliminates the infinite path dependence problem. Also, Dueker (1997) took same approach as Gray (1996) to overcome infinite path dependence problem and introduced various alternative SW-GARCH models. Klaassen (1998) modified Gray's SW-GARCH model and argue that his specification improves forecasting performance of SW-GARCH models. Recently, Haas, Mittnik, and Paolella (2004) proposed a new method different from Gray's (1996) approach and claim that analytical tractability of their new model allows derivation of stationarity conditions and dynamic properties.

Hamilton and Susmel (1994) used weekly returns on New York Stock Exchange Index over the period 1962 to 1987 to test their SWARCH model with two to four regimes. They suggest that SWARCH specification is better to fit the data, to forecast volatility and to reduce volatility persistence than uni-regime GARCH type models.

Leon Li and William Lin (1994) and Wai Mun Fong (1996) applied the SWARCH model of Hamilton and Susmel (1994) to examine regime shifts and volatility persistence respectively in weekly Taiwan Stock Index (TAIEX) and weekly Japanese Stock Index (TOPIX). They conclude that SWARCH model provides a better description of the data and a much lower degree of volatility persistence than uni-regime GARCH type models. Moreover, SWARCH model have been applied to international stock markets by Fornari and Mele (1997), Schaller and Norden(1997), Susmel (1998a, 2000), Bautista (2003), Leon Li (2007) and 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 SWARCH models. Also, Kaufmann and Scheicher (2006) applied the SWARCH model performed within Bayesian framework to describe daily German Stock Index (DAX).

Gray (1996) extended SWARCH model to SW-GARCH case by developing a recombining method that merges conditional variances in different regimes into a single conditional variance. This makes SW-GARCH model path independent and allow for constructing a tractable likelihood function. Moreover, SW-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 concludes that the SW-GARCH model outperforms simple uni-regime models in forecasting performance and reduces persistence in volatility more than SWARCH model of Cai (1994) and Hamilton and Susmel (1994).

Dueker (1997) introduced a collapsing procedure based on Kim's (1994) algorithm for SW-GARCH and applied it to daily S&P500 index. A modification of Gray's model, which allows multi-step ahead volatility forecasting, was suggested by Klaassen (1998). In addition to normal distribution, he adopted student-t distribution for error terms and estimated his SW-GARCH specification with two regimes using daily U.S. dollar exchange rates. The results show that Klaassen's model improves volatility forecasts and volatility persistence is time-varying.

Recently, Marcucci (2005) compare a set of GARCH, EGARCH and GJR-GARCH models with a group of SW-GARCH in terms of their ability to forecast S&P100 volatility from one day to one month. Also, he assumed normal, student-t and generalized error distribution for the error terms. The main finding of Marcucci (2005) is that forecasting performance of SW-GARCH models are significantly better than uni-regime GARCH type models at shorter horizons while standard asymmetric GARCH is found better at longer horizon. Daouk and Guo (2004) extended SWARCH model to Markov Switching Regime Asymmetric GARCH (SW-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 SW-GARCH model contain Fong and See (2001, 2002), Yu (2001), Francq and Zakoian (2005), Elliott, Siu and Chan (2006), Liu (2006), Lee and Yoder (2007), Abramson and Cohen (2007a, 2007b), Brunetti, Mariano, Scotti and Tan (2007).

A number of empirical studies have investigated behavior of Istanbul Stock Exchange (ISE) by using various GARCH type models. Balaban (1995) examined the day of week effect on return and volatility for ISE with GARCH models. Yavan and Aybar (1998) and Okay (1998) focused on modeling volatility of ISE using GARCH type models. Muradoğlu, Berument and Metin (1999) argued that risk-return relationship and the factors determining risks in ISE change during a financial crisis. Harris and Küçüközmen (2001a, 2001b) stated that ISE returns are highly non-normal and display significant linear and nonlinear dependence. Güner and Önder (2002), Salman (2002), Yüksel (2002), Gündüz and Hatemi (2005) investigated relationship between volatility, return and trading volume of ISE. Kilic (2004) analyses long memory in ISE by using Fractionally Integrated GARCH (FIGARCH) model and claim that ISE volatility is a long memory process. Bildik and Elekdağ (2004) examined the effects of price limits on stock return volatility. Other studies on the ISE are Muradoğlu and Metin (1996), Yılmaz (1997), Muradoğlu (1999), Odabaş, Aksu and Akgiray (2004). In addition to these studies, Balaban (1999) firstly adopted seventeen models including random walk model, moving average models, regression models and GARCH type models to forecast monthly ISE volatility. Also, Mazıbaş (2004) compares forecasting performance of fifteen symmetrical and asymmetrical GARCH models for daily, weekly and monthly volatility in ISE.

CHAPTER 2

UNI-REGIME GARCH MODELS

The aim of this chapter is to present three main uni-regime GARCH type models used in this study: GARCH, EGARCH and GJR-GARCH. These models are derived from Autoregressive Conditional Variance (ARCH) model of Engle (1982). ARCH models are designed to capture volatility clustering and correlation. In ARCH model, conditional variance at time t depends on the past squared errors[1].

Let r_t be log-return at time t and assume conditional mean equation as

$$r_t = \mu + u_t,$$
$$u_t = \varepsilon_t \sqrt{h_t},$$

where μ is constant drift term, h_t is conditional variance of errors u_t, $\varepsilon_t|\psi_{t-1} \rightarrow iidD(0,1)$ and ψ_{t-1} refers all available information up to time $t-1$. The distribution D is generally assumed to be normal, student-t or GED.

Then, ARCH (q) process is specified as

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2. \tag{2.0.1}$$

To guarantee conditional variance is positive, it must be $\alpha_0 > 0$ and $\alpha_i > 0$. Unconditional variance of returns for ARCH (q) process can be computed as

$$\sigma^2 = \frac{\alpha_0}{1 - \sum\limits_{i=1}^{q} \alpha_i}.$$

The use of ARCH models is not practical since those models are highly outperformed by standard GARCH models (Alexander, 2001).

[1]See Hamilton (1994), Alexander (2001) and Tsay (2002) for more detailed information on ARCH/GARCH type models.

2.1 GARCH (1, 1) Model

Bollerslev (1986) introduced the generalized ARCH model by adding past conditional variances into the ARCH specification. The standard GARCH(p,q) model is specified as:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}. \tag{2.1.2}$$

Many papers that focus on GARCH modelling suggest that the use of $p = q = 1$ specification is quite successful in modelling most of the financial returns volatility[2]. Thus, we consider only GARCH (1, 1), EGARCH (1, 1), GJR-GARCH (1, 1) models. The GARCH (1, 1) model is as follows[3],

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}, \tag{2.1.3}$$

where $\alpha_0 > 0$, $\alpha_1 > 0$ and $\beta_1 > 0$ to ensure positive conditional variance. Bollerslev (1986) proposed that the inequality $\alpha_1 + \beta_1 < 1$ must be satisfied for stationary covariance process of returns[4]. In that case, unconditional variance of returns can be shown as follows:

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

The parameter estimates α_1 and β_1 reveal some information on volatility process. *The large β_1 indicate that shocks to the conditional variance take a long time to die out, so volatility is persistent. Large error coefficient α_1 means that volatility reacts quite intensely to market movements, and so if α_1 is relatively high and β_1 is relatively low then volatilities tend to be more spiky.* (Alexander,2001, p73).

[2]See Bollerslev, Chou and Kroner (1992), Hansen and Lunde (2001).

[3]In fact, GARCH (1, 1) model is equivalent to the ARCH model with infinite lag.

[4]If $\alpha_1 + \beta_1 = 1$, this process is known as "Integrated GARCH (IGARCH)" (Engle and Bollerslev, 1986).

For one-step-ahead, volatility forecasting from GARCH (1, 1) model is shown in equation (2.1.4),

$$\hat{h}_{t+1} = \alpha_0 + \alpha_1 u_t^2 + \beta_1 h_t. \tag{2.1.4}$$

In order to forecast volatility for 2-step-ahead, the fact $E[u_{t+1}^2|\psi_t] = \hat{h}_{t+1}$ is used[5]. Then,

$$\begin{aligned} \hat{h}_{t+2} &= \alpha_0 + \alpha_1 u_{t+1}^2 + \beta_1 \hat{h}_{t+1}, \\ \hat{h}_{t+2} &= \alpha_0 + (\alpha_1 + \beta_1)\hat{h}_{t+1}. \end{aligned} \tag{2.1.5}$$

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

$$\hat{h}_{t+k} = \alpha_0 \sum_{i=1}^{k-1} (\alpha_1 + \beta_1)^{i-1} + (\alpha_1 + \beta)^{k-1}\hat{h}_{t+1}. \tag{2.1.6}$$

2.2 EGARCH (1, 1) Model

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

$$\ln(h_t) = \alpha_0 + \alpha_1 \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta_1 \ln(h_{t-1}) + \xi \frac{u_{t-1}}{\sqrt{h_{t-1}}}, \tag{2.2.7}$$

where ξ is the asymmetry parameter to capture leverage effect. The EGARCH process is covariance stationary if the condition $\beta_1 < 1$ is satisfied[6].

[5]See Poon (2005, p.39).

[6]See Poon (2005, p.41).

One-step-ahead volatility forecast is computed as

$$\ln(\hat{h}_{t+1}) = \alpha_0 + \alpha_1 \left| \frac{u_t}{\sqrt{h_t}} \right| + \beta_1 \ln(h_t) + \xi \frac{u_t}{\sqrt{h_t}}. \tag{2.2.8}$$

Then, multi-step-ahead volatility forecast is computed[7] as

$$\ln(\hat{h}_{t+k}) = \alpha_0 + \beta_1 \ln(\hat{h}_{t+k-1}). \tag{2.2.9}$$

2.3 GJR-GARCH (1, 1) Model

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

$$\hat{h}_t = \alpha_0 + \alpha_1 u_{t-1}^2 (1 - I_{\{u_{t-1}>0\}}) + \beta_1 h_{t-1} + \xi u_{t-1}^2 I_{\{u_{t-1}>0\}}, \tag{2.3.10}$$

where $I_{\{u_{t-1}>0\}}$ is equal to one when u_{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[8]. Also, process is covariance-stationary if $(\alpha_1 + \xi)/2 + \beta_1 < 1$. Then, unconditional variance is defined as

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

One-step-ahead volatility forecast is computed as

$$\hat{h}_{t+1} = \alpha_0 + \alpha_1 u_t^2 (1 - I_{\{u_t>0\}}) + \beta_1 h_t + \xi u_t^2 I_{\{u_t>0\}}. \tag{2.3.11}$$

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

$$\hat{h}_{t+k} = \alpha_0 + \left(\frac{\alpha_1 + \xi}{2} + \beta_1 \right) \hat{h}_{t+k-1}. \tag{2.3.12}$$

[7] Alexander (2001, p.80) state that making volatility forecast with EGARCH models is extremely difficult since this model does not have analytic form for the volatility term structure. Therefore, we follow Marccuci (2005) to forecast volatility with EGARCH model.

[8] See Franses and Van Dijk (2000) for more detailed information.

2.4 Distributions for Standardized Errors

Standard normal distribution sometimes may not be enough to describe fat-tail feature of the financial returns. In order to capture 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 such as normal distribution, they have fatter tails than normal distribution captures. In this study, we assume that standardized errors follow student-t and GED distributions as well as normal distribution.

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

$$f\left(u_t|u_{t-1},u_{t-2},..\right)=\frac{1}{\sqrt{2\pi h_t}}\exp\left(-\frac{1}{2}\cdot\frac{u_t^2}{h_t}\right). \tag{2.4.13}$$

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

$$f\left(u_t|u_{t-1},u_{t-2},..\right)=\frac{\Gamma\left[(\nu+1)/2\right]}{\sqrt{\pi(v-2)}\Gamma(v/2)}\frac{1}{\sqrt{h_t}}\left[1+\frac{u_t^2}{h_t(v-2)}\right]^{\frac{-(v+1)}{2}}. \tag{2.4.14}$$

where $\Gamma(.)$ is Gamma function. v is degree of freedom and must be greater than 2. When $v\rightarrow\infty$ student-t distribution becomes normal distribution. So, lower v implies fatter tails.

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

$$f\left(u_t|u_{t-1},u_{t-2},..\right)=\frac{v\exp\left[\left(\frac{-1}{2}\right)\left|\frac{u_t}{\delta\sqrt{h_t}}\right|^v\right]}{\delta 2^{\left(\frac{v+1}{v}\right)}\Gamma\left(1/v\right)\sqrt{h_t}}, \tag{2.4.15}$$

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

The parameters in GARCH type models are generally estimated by Maximum Likelihood Estimation (MLE) method[9]. 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 maximum of likelihood function can not be obtained analytically for GARCH type models, numerical optimization techniques are used to find set of parameters that maximize likelihood function[10].

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

For normal distribution,

$$L_{Normal} = -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(2\pi) + \ln(h_t) + \frac{u_t^2}{h_t}\right]. \tag{2.4.16}$$

For student-t distribution,

$$\begin{array}{c} L_{Student-t} = T\left\{\ln\left[\Gamma((v+1)/2)\right] - \ln\left[\Gamma(v/2)\right] - \frac{1}{2}\cdot\ln\left[\pi\left(\nu-2\right)\right]\right\} \\ -\frac{1}{2}\sum\limits_{t=1}^{T}\left[\ln\left(h_t^2\right) + (\nu+1)\ln\left(1 + \frac{u_t^2}{h_t^2(\nu-2)}\right)\right] \end{array}. \tag{2.4.17}$$

For GED distribution,

$$L_{GED} = \sum_{t=1}^{T}\left[\ln(v/\delta) - \frac{1}{2}\left|\frac{u_t}{\delta\sqrt{h_t}}\right|^{v} - \left(\frac{v+1}{v}\right)\ln(2) - \ln[\Gamma(1/v)] - \frac{1}{2}\ln(h_t)\right]. \tag{2.4.18}$$

[9]Other methods to estimate parameters of GARCH type models are Generalized Method of Moment (GMM) and Bayesian estimation technique.

[10]Bollerslev (1986) recommend Berndt-Hall-Hall-Hausmann (BHHH) algorithm.

CHAPTER 3

MARKOV REGIME SWITCHING MODELS

An important technique for analyzing structural breaks in financial return is Markov Switching model of Hamilton (1989, 1990). In his study, Hamilton extended Markov switching regression model of Goldfeld and Quandt (1973) to 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 first order Markov Chain process. In our study, we assume that there are two unobservable regimes.

3.1 Markov Regime Switching Models for Returns

3.1.1 Serially Uncorrelated Data

Let r_t is a financial return series and follows the model with structural breaks

$$r_t = \begin{cases} c_1 + \alpha_1 x_t + u_t & \text{if } s_t = 1 \\ c_2 + \alpha_2 x_t + u_t & \text{if } s_t = 2 \end{cases}$$

or, for shorthand notation,

$$r_t = c_{s_t} + \alpha_{s_t} x_t + u_t,$$

where $u_t \sim N(0, \sigma^2_{s_t})$, x is exogenous variable(s) and $s_t = 1$ if process is in regime 1 and $s_t = 2$ if process is in regime 2

The important point here is how the regime process is determined. If date of regimes 1 and 2 are known previously or determined by researcher, the model becomes a dummy variable model. Then, log-likelihood function of dummy variable model can expressed as

$$L = \sum_{t=1}^{T} \ln(f(r_t|s_t, \psi_{t-1})), \tag{3.1.1}$$

where

$$f(r_t|s_t, \psi_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2_{s_t}}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(r_t - c_{s_t} - \alpha_{s_t} x_t)^2}{\sigma^2_{s_t}}\right). \tag{3.1.2}$$

The parameters in the dummy variable models can be estimated by maximum likelihood estimation method using log-likelihood function in equation (3.1.1). However, if date of regimes are not known previously and regimes are determined by an unobservable variable s_t, the log-likelihood function can be constructed in two steps. Firstly, joint density of returns (r_t) and unobserved regime variable (s_t) can be written as follows,

$$f(r_t, s_t|\psi_{t-1}) = f(r_t|s_t, \psi_{t-1}) f(s_t|\psi_{t-1}), \tag{3.1.3}$$

where ψ_{t-1} refers the all available information up to $t-1$ and $f(r_t|s_t, \psi_{t-1})$ is given by equation (3.1.2).

Secondly, marginal density function of r_t can be constructed as follows;

$$\begin{aligned} f(r_t|\psi_{t-1}) &= \sum_{s_t=1}^{2} f(r_t, s_t|\psi_{t-1}) \\ &= \sum_{s_t=1}^{2} f(r_t|s_t, \psi_{t-1}) f(s_t|\psi_{t-1}) \\ &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(r_t - c_1 - \alpha_1 x_t)^2}{\sigma_1^2}\right) Pr(s_t = 1|\psi_{t-1}) \\ &+ \frac{1}{\sqrt{2\pi\sigma_2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(r_t - c_2 - \alpha_2 x_t)^2}{\sigma_2^2}\right) Pr(s_t = 2|\psi_{t-1}). \end{aligned} \tag{3.1.4}$$

Then, the log-likelihood function can be written as

$$L = \sum_{t=1}^{T} \ln \left(\sum_{s_t=1}^{2} f(r_t|s_t, \psi_{t-1}) Pr(s_t|\psi_{t-1}) \right). \tag{3.1.5}$$

The $Pr(s_t = \text{i}|\psi_{t-1})$ for $i = 1, 2$ in equation (3.1.5), called regime probability, is the probability that the process is in regime i at time t based on the all information up to time $t-1$. As seen from log-likelihood function in equation (3.1.5), regime probabilities must be computed to complete log-likelihood function. However, it is impossible to make inference about regime probabilities without any assumption on unobserved regime variable s_t (Kim and Nelson, 1999). There are mainly two assumptions on behavior of regime variable s_t: Independent Switching and Markov Chain Switching. In the case of independent switching, evolution of regime variable s_t is assumed independent from its own previous values[1]. In the case of Markov Chain Switching, it is assumed that regime switching is directed by first order Markov Chain process with constant transition probabilities, that is the current regime s_t only depends on the regime one period ago s_{t-1}. Then,

$$Pr(s_t|s_{t-1}, s_{t-2}...s_1, \psi_{t-1}) = Pr(s_t|s_{t-1}). \tag{3.1.6}$$

Considering only two regimes, the constant transition probabilities which are probability of switching from one regime to other regime can be defined as follows,

$$\begin{aligned}
&Pr(s_t = 1|s_{t-1} = 1) = p = \frac{\exp(p_o)}{1 + \exp(p_o)}, \\
&Pr(s_t = 2|s_{t-1} = 1) = 1 - p, \\
&Pr(s_t = 2|s_{t-1} = 2) = q = \frac{\exp(q_o)}{1 + \exp(q_o)}, \\
&Pr(s_t = 1|s_{t-1} = 2) = 1 - q.
\end{aligned} \tag{3.1.7}$$

In order to compute regime probabilities $Pr(s_t = \text{i}|\psi_{t-1})$ for $i = 1, 2$ in equation (3.1.5), following filter adopted by Kim and Nelson (1999) can be applied.

[1]More detailed information for the case of independent switching can be found at Kim and Nelson (1999).

For simplicity, we denote

$$\begin{aligned}
p_{1t} &= Pr(s_t = 1|\psi_{t-1}), \\
p_{2t} &= Pr(s_t = 2|\psi_{t-1}), \\
f_{1t} &= f(r_t|s_t = 1, \psi_{t-1}), \\
f_{2t} &= f(r_t|s_t = 2, \psi_{t-1}).
\end{aligned}$$

Step 1: Given $Pr(s_{t-1} = j|\psi_{t-1})$ for $j = 1, 2$ at the end of the time $t-1$ the regime probabilities $p_{it} = Pr(s_t = i|\psi_{t-1})$ for $i = 1, 2$ are computed as follows

$$Pr(s_t = i|\psi_{t-1}) = \sum_{j=1}^{2} \Pr(s_t = i, s_{t-1} = j|\psi_{t-1}).$$

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

$$\begin{aligned}
Pr(s_t = i|\psi_{t-1}) &= \sum_{j=1}^{2} \Pr(s_t = i|s_{t-1} = j) \Pr(s_{t-1} = j|\psi_{t-1}) \\
&= \sum_{j=1}^{2} p_{ji} \Pr(s_{t-1} = j|\psi_{t-1}).
\end{aligned}$$

Step 2: At the end of the time t, the $Pr(s_t = i|\psi_t)$ for $i = 1, 2$ is calculated as follows

$$Pr(s_t = i|\psi_t) = Pr(s_t = i|r_t, \psi_{t-1}),$$

where $\psi_t = \{\psi_{t-1}, r_t\}$, and by using bayesian arguments

$$\begin{aligned}
Pr(s_t = i|\psi_t) &= \frac{f(s_t = i, r_t|\psi_{t-1})}{f(r_t|\psi_{t-1})} \\
&= \frac{f(r_t|s_t = i, \psi_{t-1}) \Pr(s_t = i|\psi_{t-1})}{\sum_{i=1}^{2} f(r_t|s_t = i, \psi_{t-1}) \Pr(s_t = i|\psi_{t-1})} \\
&= \frac{f_{it} p_{it}}{\sum_{i=1}^{2} f_{it} p_{it}}.
\end{aligned}$$

Then, all regime probabilities p_{it} for $t = 1, 2, .., T$ can be computed by iterating these two steps. However, at the beginning of iteration, the $Pr(s_0 = \mathrm{i}|\psi_0)$ for

$i = 1, 2$ are necessary to start iteration. Hamilton (1989, 1990) suggest to use unconditional regime probabilities instead of $Pr(s_0 = \text{i}|\psi_0)$. These are given by

$$\begin{aligned} \pi_1 = Pr(s_0 = 1|\psi_o) = \frac{1-q}{2-p-q}, \\ \pi_2 = Pr(s_0 = 2|\psi_o) = \frac{1-p}{2-p-q}. \end{aligned} \tag{3.1.8}$$

3.1.2 Serially Correlated Data

In this section, Markov regime switching Autoregressive (AR) models are presented. We focus on a simple case: An AR model with first order autoregression $AR(1)$. For the case of general $AR(q)$ models, the procedure would be same as $AR(1)$. Let r_t is a financial return series, Hamilton's (1989) two state Markov regime switching $AR(1)$ model is as follows

$$r_t - \mu_{s_t} = \alpha_1(r_{t-1} - \mu_{s_{t-1}}) + u_t,$$

where $u_t \sim N(0, \sigma^2_{s_t})$ and $s_t = 1$ if process is in regime 1 and $s_t = 2$ if process is in regime 2.

As seen in Section 3.1.1, if date of regimes are known previously, model becomes a dummy variable model. Then, log-likelihood function can be constructed easily as follows

$$L = \sum_{t=1}^{T} \ln(f(r_t|s_t, s_{t-1}, \psi_{t-1})), \tag{3.1.9}$$

where

$$f(r_t|s_t, s_{t-1}, \psi_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2_{s_t}}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{[r_t - \mu_{s_t} - \alpha_1(r_{t-1} - \mu_{s_{t-1}})]^2}{\sigma^2_{s_t}}\right). \tag{3.1.10}$$

As before, if regimes are determined by an unobserved variable, log-likelihood function is constructed in two steps. Differently from the case of uncorrelated data, we must consider both of the regimes that is observed at time t and $t-1$ while constructing log-likelihood function.

Firstly, joint density of returns (r_t) and unobserved regime variables s_t and s_{t-1} can be written as follows;

$$f(r_t, s_t, s_{t-1}|\psi_{t-1}) = f(r_t|s_t, s_{t-1}, \psi_{t-1})f(s_t, s_{t-1}|\psi_{t-1}), \tag{3.1.11}$$

where ψ_{t-1} refers the all available information up to time $t-1$ and $f(r_t|s_t, s_{t-1}, \psi_{t-1})$ is given by equation (3.1.10).

Secondly, marginal density function of r_t can be expressed as follows;

$$\begin{aligned} f(r_t|\psi_{t-1}) &= \sum_{s_t=1}^{2}\sum_{s_{t-1}=1}^{2} f(r_t, s_t, s_{t-1}|\psi_{t-1}) \\ &= \sum_{s_t=1}^{2}\sum_{s_{t-1}=1}^{2} f(r_t|s_t, s_{t-1}, \psi_{t-1})f(s_t, s_{t-1}|\psi_{t-1}) \end{aligned} \tag{3.1.12}$$

Then, the log-likelihood function can be written as

$$L = \sum_{t=1}^{T} \ln[\sum_{s_t=1}^{2}\sum_{s_{t-1}=1}^{2} f(r_t|s_t, s_{t-1}, \psi_{t-1})Pr(s_t, s_{t-1}|\psi_{t-1})]. \tag{3.1.13}$$

In order to be able to compute the log-likelihood function in equation (3.1.13), obviously we have to calculate regime probabilities $Pr(s_t, s_{t-1}|\psi_{t-1})$ that is the probability of being in either regime at time t and $t-1$ given the information up to time $t-1$. Assuming the unobserved regime variable s_t is governed by first order Markov Chain, $Pr(s_t, s_{t-1}|\psi_{t-1})$ are computed by following two steps below.

Step 1 : Given $Pr(s_{t-1} = j|\psi_{t-1})$ for $j = 1, 2$ at the end of the time $t-1$, the regime probabilities $Pr(s_t = i, s_{t-1} = j|\psi_{t-1})$ for $i = 1, 2,\ j = 1, 2$ are computed as follows

$$Pr(s_t = i, s_{t-1} = j|\psi_{t-1}) = Pr(s_t = i|s_{t-1} = j, \psi_{t-1})Pr(s_{t-1} = j|\psi_{t-1}).$$

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

$$Pr(s_t = i, s_{t-1} = j|\psi_{t-1}) = Pr(s_t = i|s_{t-1} = j)Pr(s_{t-1} = j|\psi_{t-1}).$$

Step 2 : At the end of the time t, the $Pr(s_t = i, s_{t-1} = j|\psi_t)$ for $i = 1, 2,\ j = 1, 2$ are calculated as follows

$$Pr(s_t = i, s_{t-1} = j|\psi_t) = Pr(s_t = i, s_{t-1} = j|r_t, \psi_{t-1}),$$

where $\psi_t = \{\psi_{t-1}, r_t\}$, and by using bayesian arguments

$$\begin{aligned} Pr(s_t = i, s_{t-1} = j|\psi_t) &= \frac{f(s_t = i, s_{t-1} = j, r_t|\psi_{t-1})}{f(r_t|\psi_{t-1})} \\ &= \frac{f(r_t|s_t = i, s_{t-1} = j, \psi_{t-1})Pr(s_t = i, s_{t-1} = j|\psi_{t-1})}{\sum_{i=1}^{2}\sum_{j=1}^{2} f(r_t|s_t = i, s_{t-1} = j, \psi_{t-1})Pr(s_t = i, s_{t-1} = j|\psi_{t-1})}. \end{aligned}$$

Then,

$$Pr(s_t = i|\psi_t) = \sum_{j=1}^{2} Pr(s_t = i, s_{t-1} = j|\psi_t).$$

All regime probabilities $Pr(s_t = i, s_{t-1} = j|\psi_{t-1})$ for $t = 1, 2, .., T$ can be computed by iterating these two steps. In order to start iteration, as shown in Section 3.1.1, the unconditional regime probabilities in equation (3.1.8) can be used instead of starting probabilities $Pr(s_0 = \mathrm{i}|\psi_0)$ for $i = 1, 2$.

If the regime probabilities at time t are computed by using information up to time $t - 1$, they are called as *ex ante probability* ($\Pr(s_t = i|\psi_{t-1})$). These probabilities are useful for estimation and forecasting future regimes. When the regime probabilities are based on information up to time t, they are called as *filtered probability* ($\Pr(s_t = i|\psi_t)$). The estimation procedures of *ex ante* and *filtered probabilities* are explained in Section 3.1.1 and 3.1.2. On the other hand, Kim (1994) proposed *smoothed probability* ($Pr(s_t = i|\psi_T)$) which use all sample data to estimate regime probabilities at time t.

3.1.3 Kim's Smoothing Algorithm

After estimating transition probabilities, *filtered probabilities* and parameters in the model, the *smoothed probabilities* can be computed for each date t. These probabilities are generally used to make inference about which regime the process was in at a given time. Following Kim and Nelson (1999), the *smoothed probabilities* for Markov switching $AR(1)$ model are estimated as below

$$Pr(s_t = i|\psi_T) = \sum_{j=1}^{2} Pr(s_t = i, s_{t+1} = j|\psi_T).$$

Then, the joint density of $s_t = i$ and $s_{t+1} = j$ conditional on information up to time T are computed as

$$\begin{aligned} Pr(s_t = i, s_{t+1} = j|\psi_T) &= Pr(s_t = i|s_{t+1} = j, \psi_T)Pr(s_{t+1} = j|\psi_T) \\ &= Pr(s_t = i|s_{t+1} = j, \psi_t)Pr(s_{t+1} = j|\psi_T) \\ &= \frac{Pr(s_t = i, s_{t+1} = j|\psi_t)Pr(s_{t+1} = j|\psi_T)}{Pr(s_{t+1} = j|\psi_t)} \\ &= \frac{Pr(s_{t+1} = j|s_t = i, \psi_t)Pr(s_t = i|\psi_t)Pr(s_{t+1} = j|\psi_T)}{Pr(s_{t+1} = j|\psi_t)} \\ &= \frac{Pr(s_{t+1} = j|s_t = i)Pr(s_t = i|\psi_t)Pr(s_{t+1} = j|\psi_T)}{Pr(s_{t+1} = j|\psi_t)}. \end{aligned} \tag{3.1.14}$$

Given $\Pr(s_T = j|\psi_T)$ which is computed at last iteration of *filtered probability* algorithm, the *smoothed probabilities* for date $T-1, T-2, ..., 1$ are computed iteratively by following algorithm above.

The setting $Pr(s_t = i|s_{t+1} = j, \psi_T) = Pr(s_t = i|s_{t+1} = j, \psi_t)$ used in equation (3.1.14) can be shown as

$$\begin{aligned} Pr(s_t = i|s_{t+1} &= j, \psi_T) = Pr(s_t = i|s_{t+1} = j, \psi_t, r_{t+1}, r_{t+2}, .., r_T) \\ &= \frac{f(s_t = i, r_{t+1}, r_{t+2}, .., r_T|s_{t+1} = j, \psi_t)}{f(r_{t+1}, r_{t+2}, .., r_T|s_{t+1} = j, \psi_t)} \\ &= \frac{Pr(s_t = i|s_{t+1} = j, \psi_t)f(r_{t+1}, r_{t+2}, .., r_T|s_{t+1} = j, s_t = i, \psi_t)}{f(r_{t+1}, r_{t+2}, .., r_T|s_{t+1} = j, \psi_t)} \\ &= Pr(s_t = i|s_{t+1} = j, \psi_t). \end{aligned}$$

3.2 Markov Regime Switching Models for Volatility

3.2.1 Markov Regime Switching ARCH Model

Hamilton and Susmel (1994) and Cai (1994) proposed Markov Regime Switching ARCH (SWARCH) model independently by combining Markov Regime Switching model with ARCH models. In this model, each regime is characterized by a different ARCH (q) process and parameters of conditional variance take different values for each regime. Basically, two regime SWARCH (q) model can be written as follows[2]

$$
\begin{aligned}
r_t &= \mu_{s_t} + u_t, \\
u_t &= \tilde{u}_t \sqrt{g_{s_t}}, && (3.2.15) \\
\tilde{u}_t &= \varepsilon_t \sqrt{h_t}, \quad \varepsilon_t \sim iid(0,1), && (3.2.16) \\
h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \tilde{u}_{t-i}^2, && (3.2.17)
\end{aligned}
$$

where g_{s_t} is a regime factor with value of g_1 when the process is in the regime 1 and g_2 when the process is in the regime 2. The unobserved regime variable s_t follows a first order Markov Chain process and takes values of 1 and 2. ε_t is a zero mean, unit variance process.

As seen, $\tilde{u}_t$ is a standard ARCH process. The underlying ARCH process is multiplied by constant g_1 when the process is in the regime 1 and multiplied by constant g_2 when the process is in the regime 2. Generally, the constant g_1 is normalized to unity with $g_2 > 1$, and then g_2 is interpreted as the ratio of the average conditional variance when the process is in the regime 2 compared to that observed when the process is in the regime 1. So, the thought behind SWARCH is *to model changes in regime as changes in the scale of the process*[3].

[2]Extension to more than two regimes is straightforward.

[3]Hamilton and Susmel (1994).

The parameters of the SWARCH model are generally estimated by maximum likelihood method. The method explained in Section 3.1.2 can be used to construct likelihood function derived by Hamilton and Susmel (1994, appendix) and make inference about unobserved regime variable s_t . Based on information current and past regimes, the conditional variance implied for errors is

$$\begin{aligned} h_t(s_t, .., s_{t-q}, u_{t-1}, .., u_{t-q}) &= E[u_t^2 | s_t, ..., s_{t-q}, u_{t-1}, .., u_{t-q}] \\ &= g_{s_t}\left[\alpha_0 + \alpha_1\left(\frac{u_{t-1}^2}{g_{s_{t-1}}}\right) + \alpha_2\left(\frac{u_{t-2}^2}{g_{s_{t-2}}}\right) + .. + \alpha_q\left(\frac{u_{t-q}^2}{g_{s_{t-q}}}\right)\right]. \end{aligned} \tag{3.2.18}$$

3.2.2 Markov Regime Switching GARCH Model of Gray

The SW-GARCH model proposed by Gray (1996). The SW-GARCH model with two regimes simply can be represented as follows

$$r_t = \mu_{s_t} + u_t,$$

$$u_t \sim \varepsilon_t \sqrt{h_{t,s_t}}, \tag{3.2.19}$$

$$h_{t,s_t} = \alpha_{0,s_t} + \alpha_{1,s_t} u_{t-1}^2 + \beta_{1,s_t} h_{t-1}, \tag{3.2.20}$$

where $s_t = 1$ *or* 2. μ_{s_t} and h_{t,s_t} are the conditional mean and conditional variances respectively. Both are allowed to switch between two regimes. To ensure positivity of conditional variance in each regime, necessary conditions are similar to the necessary conditions in uni-regime GARCH (1, 1) model.

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

$$\begin{aligned} Pr(s_t = 1 | s_{t-1} = 1) &= p, \\ Pr(s_t = 2 | s_{t-1} = 1) &= 1 - p, \\ Pr(s_t = 2 | s_{t-1} = 2) &= q, \\ Pr(s_t = 1 | s_{t-1} = 2) &= 1 - q. \end{aligned} \tag{3.2.21}$$

In matrix notation,

$$P = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}. \tag{3.2.22}$$

Then, conditional distribution of return series r_t becomes a mixture-of-distribution model in which mixing variable is *ex ante probability* $(Pr(s_t = i|\psi_{t-1}))$ denoted by p_{it},

$$r_t|\psi_{t-1} = \begin{cases} f\left(r_t|s_t = 1, \psi_{t-1}\right) & \text{with probability} \quad p_{1t} \\ f\left(r_t|s_t = 2, \psi_{t-1}\right) & \text{with probability} \quad p_{2t} = 1 - p_{1t} \end{cases}$$

where $f\left(r_t|s_t, \psi_{t-1}\right)$ denotes one of the assumed conditional distributions for errors: Normal, Student-t or GED. ψ_{t-1} denotes the information at time $t-1$. p_{1t} is the *ex ante probability* of being in regime 1. The estimation procedure for computing p_{1t} is explained in details in Section 3.1.1.

The log-likelihood function for SW-GARCH model can be written as

$$L = \sum_{t=1}^{T} \ln[f(r_t|s_t = 1, \psi_{t-1})p_{1t} + f(r_t|s_t = 2, \psi_{t-1})(1 - p_{1t})]. \tag{3.2.23}$$

Both Hamilton and Susmel (1994) and Cai (1994) limited their estimation to the Markov Regime Switching ARCH model. The reason of this limitation is that there is an infinite path dependence problem inherent in SW-GARCH models. In SWARCH models, the conditional variance at time t depends on past q squared residuals and past q regime variables $(s_t, .., s_{t-q})$. However, in SW-GARCH model, the conditional variance at time t depends on the conditional variance at time $t-1$ and regime variable at time t (s_t) while the conditional variance at time $t-1$ depends on the conditional variance at time $t-2$ and regime variable at time $t-1$ (s_{t-1}), and so on. Therefore, the conditional variance at time t depends on the entire history of regimes up to time t. Both Hamilton and Susmel (1994) and Cai (1994) stated that path dependence nature of SW-GARCH model makes estimation infeasible and impossible for large sample size.

For example, in a SW-GARCH with M-regimes model, the number of paths enlarges by a factor of M in each period and integrating all possible paths is required to construct the likelihood function. For the t^{th} observation, there are M^t components of likelihood function and this makes estimation intractable for large sample sizes.

In order to solve problem of path dependence in SW-GARCH model, Gray (1996) proposed to use conditional expectation of the lagged conditional variance $E_{t-2}(h_{t-1})$ instead of lagged conditional variance h_{t-1}. *This approach preserves the natural essential of the GARCH process and allows tractable estimation of model*[4]. Gray's approach recombines $h_{t-1,s_{t-1}=1}$ and $h_{t-1,s_{t-1}=2}$ into h_{t-1} , and recombines $u_{t-1,s_{t-1}=1}$ and $u_{t-1,s_{t-1}=2}$ into u_{t-1} by taking conditional expectations of h_{t-1} and u_{t-1} based on the *ex ante probabilities*. That is,

$$\begin{aligned} h_{t-1} &= E_{t-2}(h_{t-1}) \\ &= E(r_{t-1}^2|\psi_{t-2}) - [E(r_{t-1}|\psi_{t-2})]^2 \\ &= p_{1t-1}\left[\mu_{s_{t-1}=1}^2 + h_{t-1,s_{t-1}=1}\right] + (1-p_{1t-1})\left[\mu_{s_{t-1}=2}^2 + h_{t-1,s_{t-1}=2}\right] \\ &- \left[p_{1t-1}\mu_{s_{t-1}=1} + (1-p_{1t-1})\,\mu_{s_{t-1}=2}\right]^2. \end{aligned} \tag{3.2.24}$$

Similarly, error terms u_{t-1} is given by

$$u_{t-1} = r_{t-1} - E(r_{t-1}|\psi_{t-2}) = r_{t-1} - p_{1t-1}\mu_{s_{t-1}=1} + (1-p_{1t-1})\,\mu_{s_{t-1}=2}. \tag{3.2.25}$$

Given equations (3.2.24) and (3.2.25), the conditional variance h_{t,s_t} in Gray's model can be written as

$$h_{t,s_t} = \alpha_{0,s_t} + \alpha_{1,s_t}u_{t-1}^2 + \beta_{1,s_t}h_{t-1}. \tag{3.2.26}$$

The use of conditional expectation of the lagged conditional variance $E_{t-2}(h_{t-1})$ instead of lagged conditional variance h_{t-1} makes conditional variance at time t depends on only current regime s_t and inference about s_{t-1}. Therefore, the Gray's *collapsing procedure* simplifies and makes tractable the estimation of SW-GARCH models.

[4] Gray (1996).

Given initial values for regime probabilities, conditional mean and conditional variance in each regime, the parameters of SW-GARCH model can be obtained by maximizing numerically the log-likelihood function given in equation (3.2.23). The log-likelihood function is constructed recursively similar to that in a uni-regime GARCH models

3.2.3 Markov Regime Switching GARCH Model of Klaassen

Klaassen (2002) introduced modification of Gray's SW-GARCH model. Differently from Gray (1996), he suggests using the conditional expected value $E_{t-1}(h_{t-1}|s_t)$ instead of $E_{t-2}(h_{t-1})$ to substitute for h_{t-1} . The setting of Klaassen for conditional expectation of the lagged conditional variance contains broader information than that of Gray. The specification of Klaassen for conditional variance can be given as[5]

$$\begin{aligned} h_{t,s_t=i} &= \alpha_{0,s_t=i} + \alpha_{1,s_t=i} u_{t-1}^2 + \beta_{1,s_t=i} h_{t-1} \\ &= \alpha_{0,s_t=i} + \alpha_{1,s_t=i} u_{t-1}^2 + \beta_{1,s_t=i} E_{t-1}(h_{t-1}|s_t = i), \end{aligned} \tag{3.2.27}$$

where

$$E_{t-1}(h_{t-1}|s_t = i) = \sum_{j=1}^{2} \tilde{p}_{ji,t-1} \left[\mu_{s_{t-1}=j}^2 + h_{t-1,s_{t-1}=j}\right] - \left[\sum_{j=1}^{2} \tilde{p}_{ji,t-1} \mu_{s_{t-1}=j}^2\right]^2, \tag{3.2.28}$$

and the probabilities $\tilde{p}_{ji,t-1}$ in 3.2.28 are computed as follows

$$\tilde{p}_{ji,t-1} = Pr(s_{t-1} = j|s_t = i, \psi_{t-1}) = \frac{p_{ji} Pr(s_{t-1} = j|\psi_{t-1})}{Pr(s_t = i|\psi_{t-1})} \tag{3.2.29}$$

with $p_{ji} = Pr(s_t = i|s_{t-1} = j)$ for $i, \ j = 1, \ 2.$

The specification of Klaassen has mainly two advantages over that of Gray. First one is that, when integrating out the previous regime (s_{t-1}), Klaassen's specification uses the information up to time $t-1$ and regime variable s_t while

[5] Marcucci (2005).

Gray's specification uses the information up to time $t-2$ and regime variable s_{t-1}. This makes Klaassen's specification more efficient (observable) use of conditional information. The second and most important one is that Klaassen's specification allows a convenient recursive formulation for multi-step ahead volatility forecast.

In SW-GARCH model with two regimes, volatility forecast for k-step-ahead conditional on information available at time $T-1$ is as follows[6]

$$\hat{h}_{T,T+k} = \sum_{i=1}^{2} Pr(s_{T+k} = i|\psi_{T-1})\hat{h}_{T,T+k,s_{T+k}=i}, \tag{3.2.30}$$

where $\hat{h}_{T,T+k,s_{T+k}=i}$ is k-step-ahead volatility forecast in regime i made at time T and computed as

$$\hat{h}_{T,T+k,s_{T+k}=i} = \alpha_{0,s_{T+k}=i} + (\alpha_{1,s_{T+k}=i} + \beta_{1,s_{T+k}=i})E_{T-1}(\hat{h}_{T,T+k-1}|s_{T+k} = i). \tag{3.2.31}$$

Also, the $\sum_{i=1}^{2} Pr(s_{T+k} = i|\psi_{T-1})$ in equation (3.2.30) is computed as

$$\begin{bmatrix} \Pr(s_{T+k} = 1|\psi_{T-1}) \\ \Pr(s_{T+k} = 2|\psi_{T-1}) \end{bmatrix} = P^{k+1} \begin{bmatrix} \Pr(s_{T-1} = 1|\psi_{T-1}) \\ \Pr(s_{T-1} = 2|\psi_{T-1}) \end{bmatrix}, \tag{3.2.32}$$

where P is given in equation (3.2.22).

Lastly, in order to compute expectation part $E_{T-1}(\hat{h}_{T,T+k-1}|s_{T+k} = i)$ in equation (3.2.31), the probability $\Pr(s_{T+k-1} = j|s_{T+k} = i, \psi_{T-1})$ is required, then

$$Pr(s_{T+k-1} = j|s_{T+k} = i, \psi_{T-1}) = \frac{p_{ji} \Pr(s_{T+k-1} = j|\psi_{T-1})}{Pr(s_{T+k} = i|\psi_{T-1})} \tag{3.2.33}$$

with $p_{ji} = Pr(s_t = i|s_{t-1} = j)$ for $i,\ j = 1,\ 2$.

[6] Marcucci (2005).

CHAPTER 4

EMPIRICAL METHODOLOGY AND GARCH ESTIMATION RESULTS

4.1 Data

The data set used in this study is the daily closing prices of value-weighted ISE-100 index over the period 03/01/1997 through 27/12/2007. The data set is obtained from the web site of the Central Bank of Republic of Turkey[1]. The data is divided into a ten year in-sample estimation period (2480 observations) and a subsequent one year out-of-sample forecasting period (248 observations):

$$t = \underbrace{-T+1, -T+2, ..., 0}_{estimation\ period} \qquad \underbrace{1, 2, ..., n}_{evaluation\ period} \quad .$$

Daily observations are converted into continuously compounded returns in a standard method as log differences:

$$r_t = 100 * \ln\left(\frac{P_t}{P_{t-1}}\right),$$

where P_t and P_{t-1} are closing values of ISE-100 index at time t and $t-1$.

The plot of return and price series are given in Figure 4.1 and Figure 4.2. ISE-100 return index displays usual properties of financial data series. As expected,

[1]http://evds.tcmb.gov.tr/

Figure 4.1: Graph of ISE-100 index for the period 1997 to 2006

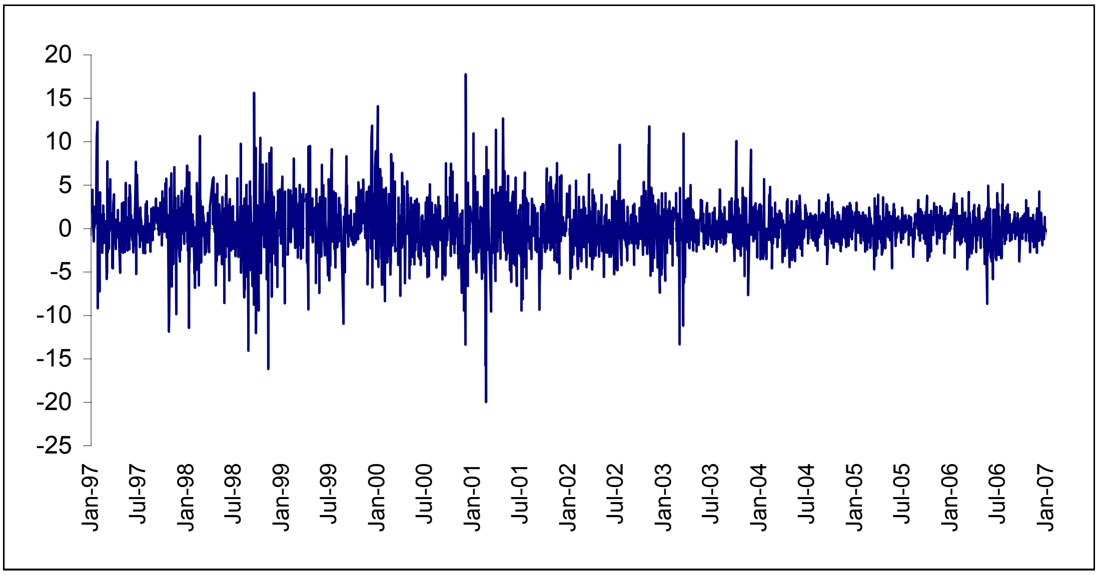

Figure 4.2: Graph of ISE-100 index returns for the period 1997 to 2006

volatility is not constant over the time and exhibits volatility clustering that is, as noted by Mandelbrot (1965), large changes in the price of an asset often followed by large changes, and small changes often followed by small changes. Moreover, the plot of returns show persistence in the return of ISE-100 index. There are huge spikes in 1998, late of 2000 and beginning of 2001 and March of 2003. The huge spikes in 1998 were mainly due to the crisis in Southeast Asian countries and transmission this crisis into Russia. Russia announced moratorium on debts and devaluated their currency in 1998 August. So, these economic crises has affected Turkey negatively and resulted in a capital outflows from the country. Moreover, spikes in the late 2000 and beginning of 2001 are caused by two crises in November and February respectively. These crises take root from domestic factors such as unsuccessful economic policies especially about curbing inflation, high current account deficits, short term foreign borrowing of companies and banks. External factors such as collapsing NASDAQ and other stocks in 2000, economic crises in Argentina and other Latin American countries, September 11 attacks upon the USA in 2001 have contributed to crises eruption. In addition, in the March of 2003, some developments can be said stock movements in Istanbul Stock Exchange. Possibility of an attack to Iraq by the USA and realizing in 20th of March, expectations about whether Turkey would involve in this process, the uncertainty about continuation of IMF economic program have been main reasons for fluctuation in financial markets.

Descriptive statistics of return series are represented in Table 4.1. As table shows, the index has a positive average return 0.156%. Daily standard deviation is 2.988%. The series also displays a negative skewness of -0.098 and an excess kurtosis of 4.255. These values indicate that the returns are not normally distributed, namely it has fatter tails. Also, Jarque-Bera test[2] statistic of 1864,827 confirms the non-normality of ISE-100 returns. These findings are consistent with other financial returns' properties.

As Alexander (2001) states, volatility is a concept that only applies to stationary processes. Therefore, we need to check whether return series is stationary or

[2]Jarque-Bera Normality test follows a χ^2 distribution with 2 degrees of freedom under the null hypothesis.

Table 4.1: Summary Statistics for ISE-100 returns (%)

Mean	0.156
Standard Error	0.060
Standard Deviation	2.988
Sample Variance	8.925
Excess Kurtosis	4.255
Skewness	-0.098
Range	37.752
Minimum	-19.979
Maximum	17.774
Jarque-Bera Normality Test	1864.827 (p=0.000)

Table 4.2: Augmented Dickey-Fuller Test Results

Inclusion in ADF test	ADF test Statistic		p-values	
	AIC	SBIC	AIC	SBIC
with Constant term	-11.229	-49.519	0.000	0.000
with Constant and Trend term	-11.231	-49.501	0.000	0.000

not before modelling volatility. In order to test stationarity, we apply Augmented Dickey-Fuller (Dickey and Fuller, 1981) test (ADF). The optimal lag length of ADF test is determined by both the Schwarz Bayesian Information Criterion (SBIC) and Akaike Information Criterion (AIC). We applied two versions of this test: with constant and with constant and trend terms. The null hypothesis of ADF test is that the series is non-stationary. Table 4.2 shows the results obtained and all tests indicate the stationarity of ISE-100 returns.

The autocorrelation functions (ACF) and partial autocorrelation functions (PACF) of the ISE-100 returns and squared ISE-100 returns are presented in Table 4.3. All ACF and PACF of return series are very small and less than

0.06. The very low autocorrelation indicate that the return series are almost uncorrelated. In order to test the significance level of autocorrelation, we apply Ljung and Box (1978) Q test[3]. According to results of the Ljung-Box Q test (LBQ) statistics, the null hypothesis of no serial correlation cannot be rejected up to fifth order lag for return series with 99% confidence level. However, there is significant serial dependence after fifth order lag.

Serial correlation in the squared returns suggests the conditional heteroskedasticity. Therefore, we analyze the significance of autocorrelation in the squared mean adjusted return $(r_t - \mu)^2$ series by Ljung-Box Q test[4]. The results reported in Table 4.3 show that there is significant correlation in squared returns up to fortieth order lag and this proves presence of ARCH effects in the returns. Moreover, we apply Engle's (1982) ARCH test and it confirms strong evidence of heteroskedasticity. Thus the use of GARCH type models for the conditional variance is justified.

4.2 Empirical Methodology

This empirical part adopts standard uni-regime GARCH and Markov Regime Switching GARCH models to estimate the volatility of the daily Turkish Stock Market. Standard uni-regime GARCH models contain GARCH, EGARCH and GJR-GARCH. In order to account fat tails feature of financial returns, we consider three different distributions for the innovations: Normal, Student-t and GED distributions. In the literature, GARCH (1, 1) is usually found good enough to describe a large number of financial returns[5]. So, we particularly focus on $p = q = 1$ specification for all GARCH and SW-GARCH models.

Since it allows a convenient recursive formulation for multi-step ahead volatil-

[3]The null hypothesis of the test is that there is no serial correlation in the series up to the specified lag.

[4]It refers the Breusch-Godfrey Lagrange Multiplier test.

[5]Bollerslev, Chou and Kroner (1992).

Table 4.3: ACF, PACF and Ljung and Box Q Test Results

	ISE-100 returns				ISE-100 Squared returns			
Lag	ACF	PACF	LBQ stat	p-value	ACF	PACF	LBQ stat	p-value
1	0.006	0.006	0.081	0.776	0.237	0.237	139.33	0.000
2	0.052	0.052	6.708	0.035	0.206	0.159	244.78	0.000
3	-0.017	-0.017	7.405	0.060	0.118	0.043	279.35	0.000
4	0.019	0.017	8.331	0.080	0.089	0.027	298.89	0.000
5	-0.041	-0.040	12.603	0.027	0.146	0.106	351.66	0.000
10	0.034	0.031	28.312	0.002	0.109	0.063	449.94	0.000
15	0.036	0.045	41.290	0.000	0.044	-0.001	511.860	0.000
20	-0.035	-0.031	54.504	0.000	0.088	0.053	585.99	0.000
25	0.021	0.016	56.466	0.000	0.075	0.039	647.99	0.000

Table 4.4: Results for Engle's ARCH test

LAG	ARCH Test stat	p-value
1	137.73	0.0000
2	199.0836	0.0000
3	202.6983	0.0000
4	204.3614	0.0000
5	230.4518	0.0000
10	250.3987	0.0000
15	259.0636	0.0000
20	264.2695	0.0000
25	269.9312	0.0000

ity forecasting, we apply SW-GARCH model of Klaassen (2002) in our analysis. We consider two volatility regimes: Low volatility (regime 1) and high volatility (regime 2). All parameters in the GARCH processes and mean equations are allowed to switch between different two regimes. For simplicity, first order Markov chain and constant transition probabilities are considered. In SW-GARCH model, standardized errors are assumed to follow normal, student-t, GED and student-t2 distribution in which each regime takes different degrees of freedom of a student-t distribution. All degrees of freedom and transition probabilities in models are determined by data.

All parameter estimates of the GARCH and SW-GARCH models are computed using Quasi Maximum Likelihood (QML) method with Broyden, Fletcher, Goldfarb, and Shanno (BFGS) algorithm. The negative log-likelihoods are minimized numerically by using MATLAB optimization routines[6]. Parameters in mean equation and variance equation are estimated jointly. For all models, 100 different sets of starting values are generated randomly and parameter estimates of those giving the highest likelihood value are used as initial values to get final parameter estimates. We find a single local maximum for all models.

To develop any GARCH type model, we have to provide two distinct specifications: one for the conditional mean equation and the other for the conditional variance equation. Since the major focus of this study is modelling and forecasting volatility of Turkish Stock Market, following Klaassen (2002) and Marcucci (2005), we specify the conditional mean as

$$\begin{aligned}
&r_t = \mu + u_t, \\
&u_t = \sqrt{h_t}\,\varepsilon_t, \\
&\varepsilon_t \sim Normal,\ \ Student - t\ \ or\ \ GED,
\end{aligned}$$

where $\mu = \mu_1 S_t + \mu_2(1 - S_t)$ for regime switching models and $S_t = 1$ when process is in the low volatility regime 1, $S_t = 0$ when process is in the high volatility regime 2.

[6] We thank Marcucci for providing his Matlab source codes which estimate SW-GARCH models's parameters and forecast volatility.

According to Ljung-Box Q test, there is no significant serial correlation up to 5 lags in returns. Also, ACF and PACF of returns seem very low. So, these results support our specification for conditional mean equation.

4.3 Empirical Results

4.3.1 Uni-regime GARCH Models

Table 4.5 present estimation results for uni-regime GARCH models. It is clear from the table that almost all parameter estimates including μ in uni-regime GARCH models are highly significant at 1%. Only the leverage effect ξ of EGARCH model with normal and GED errors are insignificant. However, the asymmetry effect term ξ in GJR-GARCH models is significantly different from zero, which indicates unexpected negative returns imply higher conditional variance as compared to same size positive returns.

The degree of volatility persistence for GARCH models can be obtained by summing ARCH and GARCH parameters estimates $(\alpha_1 + \beta_1)$ [7]. All models display strong persistence in volatility ranging from 0.980 to 0.987, that is, volatility is likely to remain high over several future periods once it increases.

If distribution assumptions for standardized errors are compared, it reveals that normality assumption is highly outperformed by other two fat-tailed distributions in terms of log-likelihood values. It is an anticipated result because of fat tails property of Turkish Stock Market. Overall, student-t distribution yields an improvement in fitting the data over the others and the GJR-GARCH model with student-t has the largest log-likelihood among uni-regime GARCH models.

If a GARCH model is successful at capturing volatility clustering, squared standardized residuals should have no autocorrelation. Applying Ljung-Box Q test to GARCH residuals, results presented in Table 4.5 show that p-values are

[7] For EGARCH (1, 1) and GJR-GARCH(1,1), persistence is equal to β_1 and $\frac{\alpha_1+\xi}{2} + \beta_1$ respectively.

Table 4.5: Summary results of uni-regime GARCH models

	GARCH			EGARCH			GJR-GARCH		
	Normal	***Student's t***	***GED***	***Normal***	***Student's t***	***GED***	***Normal***	***Student's t***	***GED***
μ	0.156***	0.178***	0.174***	0.131***	0.163***	0.160***	0.144***	0.165***	0.164***
std. err.	0.05	0.04	0.04	0.05	0.04	0.04	0.05	0.04	0.04
t stat	3.44	4.10	4.10	2.95	3.79	3.78	3.06	3.78	3.83
α_0	0.114***	0.136***	0.127***	-0.125***	-0.119***	-0.123***	0.116***	0.142***	0.130***
std. err.	0.02	0.03	0.03	0.01	0.02	0.02	0.02	0.03	0.03
t stat	5.98	4.48	4.32	-12.04	-6.86	-7.38	6.14	4.61	4.45
α_1	0.091***	0.094***	0.093***	0.209***	0.209***	0.210***	0.101***	0.115***	0.109***
std. err.	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.02	0.02
t stat	12.48	7.30	7.66	16.14	8.97	9.70	10.44	6.50	6.67
β_1	0.896***	0.889***	0.891***	0.983***	0.980***	0.981***	0.896***	0.887***	0.890***
std. err.	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01
t stat	126.58	71.71	75.00	271.62	170.24	173.14	127.37	70.38	74.58
ξ	-	-	-	-0.011	-0.026**	-0.018	0.081***	0.077***	0.076***
std. err.	-	-	-	0.01	0.01	0.01	0.01	0.01	0.01
t stat	-	-	-	-1.57	-2.12	-1.63	10.17	5.36	5.91
df	-	6.52***	1.37***	-	6.12***	1.36***	-	6.46***	1.37***
std. err.	-	0.77	0.04	-	0.72	0.05	-	0.76	0.04
t stat	-	8.45	31.37	-	8.49	30.47	-	8.52	31.39
Log(L)	-5927.99	-5861.45	-5871.68	-5930.30	-5859.81	-5872.05	-5926.78	-5859.04	-5870.09
Persistence	0.987	0.983	0.984	0.983	0.980	0.981	0.987	0.983	0.984
LBQ(25)	32.15 (0.15)	32.37 (0.15)	32.23 (0.15)	33.65 (0.12)	34.55 (0.10)	33.98 (0.11)	32.71 (0.14)	33.32 (0.12)	32.97 (0.13)
LBQ2(25)	20.11 (0.74)	19.87 (0.75)	19.83 (0.76)	33.60 (0.12)	33.15 (0.13)	32.95 (0.13)	18.29 (0.83)	17.03 (0.88)	17.49 (0.86)

Note: *, ** and *** refer the significance at 90%, 95%, 99% confidence level respectively. p-values for LBQ test are in parentheses.

high and indicate squared standardized residual appear to be serially uncorrelated. Thus, uni-regime GARCH models are well specified at modeling conditional variance. It is noteworthy that p-values of Ljung-Box Q test are too small in the case of EGARCH models compared with those from remaining models, which may imply EGARCH specification is not as good as other two models to eliminate ARCH effects in Turkish Stock Market. Also, Ljung-Box Q statistics for standardized residuals display no serial dependence.

4.3.2 Markov Regime Switching GARCH Models

Estimation results and summary statistics of SW-GARCH models are presented in Table 4.6. Almost all parameter estimates are significantly different from zero at least 95% confidence level. The conditional mean estimates in high volatility regime of SW-GARCH with normal and GED distributions are barely significant at 90% confidence level. However, ARCH parameters α_1 in both volatility regimes of SW-GARCH with normal distribution are insignificant.

To check accuracy of SW-GARCH models, we examine the autocorrelations in standardized residuals and squared standardized residuals from the models. According to results of Ljung-Box Q test given in Table 4.6, all SW-GARCH models appear to fit the data very well except SW-GARCH with normal distribution. Similar to the raw data, squared standardized residuals from SW-GARCH under normal distribution display significant autocorrelation. So, this result indicates that normality assumption for standardized errors in SW-GARCH model fails to capture heteroscedasticty in Turkish Stock Market. Therefore, although we report the results of model SW-GARCH with normal distribution, we will not take into consideration it while analyzing results and comparing estimation and prediction performance of volatility models. Hereafter, we imply SW-GARCH models with student-t, student-t2 and GED while discussing the SW-GARCH models.

In order to see existence of different volatility regimes, we compute the un-

Table 4.6: Summary results of Markov Regime Switching GARCH models

	Markov Regime Switching-GARCH							
	Normal		Student's t2		Student's t		GED	
	low volatility regime	high volatility regime	low volatility regime	high volatility regime	low volatility regime	high volatility regime	low volatility regime	high volatility regime
μ	0.174***	0.476*	0.195***	0.153**	0.198***	0.153**	0.205***	0.133*
std. err.	0.05	0.27	0.05	0.07	0.05	0.08	0.05	0.08
t stat	3.73	-1.77	3.72	2.09	3.83	1.99	4.10	1.78
α_0	0.199**	9.315***	0.150**	1.132***	0.149*	1.104***	0.150**	1.196***
std. err.	0.10	2.10	0.07	0.33	0.08	0.32	0.07	0.32
t stat	1.99	4.45	2.13	3.43	1.89	3.45	2.06	3.72
α_1	0.012	0.003	0.084***	0.149***	0.082***	0.148***	0.088***	0.143***
std. err.	0.03	0.06	0.03	0.03	0.03	0.03	0.03	0.03
t stat	0.48	0.06	2.82	5.11	2.58	5.31	2.90	5.50
β_1	0.787***	0.985***	0.870***	0.766***	0.880***	0.764***	0.868***	0.759***
std. err.	0.02	0.08	0.04	0.04	0.04	0.04	0.04	0.04
t stat	33.33	12.56	21.89	17.63	21.95	17.69	21.75	17.58
p	0.934***		0.999***		0.999***		0.999***	
std. err.	0.01		0.00		0.00		0.00	
t stat	70.79		1083.56		778.90		1057.39	
q	0.528***		0.999***		0.999***		0.999***	
std. err.	0.08		0.00		0.00		0.00	
t stat	6.40		1903.45		1815.68		1757.19	
df	-		10.32***	5.85***	6.88***		1.42***	
std. err.	-		3.25	0.93	0.92		0.05	
t stat	-		3.17	6.29	7.46		28.13	
Log(L)	-5878.97		-5845.22		-5846.56		-5854.76	
σ^2	0.990	765.486	3.188	13.290	3.722	12.526	3.432	12.231
π	0.878	0.122	0.438	0.562	0.451	0.549	0.440	0.560
Persistence	0.799	0.988	0.953	0.915	0.962	0.912	0.956	0.902
LBQ(25)	35.24 (0.084)		33.64 (0.116)		33.58 (0.117)		34.04 (0.107)	
$LBQ^2(25)$	46.89 (0.005)		20.67 (0.711)		19.56 (0.769)		19.95 (0.749)	

Note: *, ** and *** refer the significance at 90%, 95%, 99% confidence level respectively. p-values for LBQ test are in parentheses. σ^2 refers the unconditional variance in each regime. π is unconditional probability of being in associated regime.

conditional variances in each volatility regime. The results showed in Table 4.6 reveal that unconditional variance of high volatility regime is about four times higher than that of low volatility regime for all SW-GARCH models. These findings confirm that the volatility process of Turkish Stock Market is characterized by two different regimes. Also, the big difference between variance of each regime shows need of volatility models that allow regime switching.

The long term volatility level depends on the estimates of constant parameter α_0[8]. Results in Table 4.6 are consistent with this argument and display that there are huge differences between α_0 estimates of each volatility regime. The parameter estimates α_0 in high volatility regimes are nearly eight times greater than parameter estimates α_0 in low volatility regimes. Moreover, short run dynamics of volatility is determined by the ARCH parameter α_1 and GARCH parameter β_1. Large estimates of β_1 suggest that effect of shocks to future volatility die out in a long time, so volatility is persistent. Large values of α_1 display reaction of volatility to the recent price changes[9]. Comparing the low and high volatility regimes in all SW-GARCH models, the former volatility regimes have lower α_1 estimates but higher β_1 estimates than latter volatility regimes have. So, the GARCH processes in the low volatility regimes are more reactive but less persistent than that in the high volatility regime. In addition, it is interesting to notice that degree of volatility persistence $(\alpha_1 + \beta_1)$ within low volatility regime is higher compared to the high volatility regime[10]

Parameter estimates of transition probabilities p and q are statistically significant at 99% confidence level and close to unity, indicating that the volatility regimes are highly persistent. Moreover, we compute unconditional probability of being in low and high volatility regimes[11] and those are about 0.45 and 0.55 respectively.

[8]Alexander (2001), p.85.

[9]Alexander (2001), p.73.

[10]Persistence within each regime is calculated as $\alpha_1^i + \beta_1^i$ where i=1, 2.

[11]$p(s_t = 1) = \frac{1-q}{2-p-q}, \qquad p(s_t = 2) = \frac{1-p}{2-p-q}$

As expected conditional mean returns in low volatility regime are higher than that of high volatility regimes for all SW-GARCH models. So, the lower uncertainty in ISE-100 index gives chance of higher profit to the practitioners. This shows the importance of regimes switching models to model volatility.

In markov regime switching and uni-regime models, the degrees of freedom of GED distribution do not indicate any big difference. Also same inference can be attributed to Student-t distributions. However, SW-GARCH with student-t2 in which each regime takes different degrees of freedom displays that degrees of freedom are 10.32 and 5.85 respectively for low and high volatility regimes. That means high volatility regime has fatter tails than the low ones.

In order to display how volatility regimes have evolved over the estimation period, we report both *ex ante* and *smoothed probabilities* of being in high volatility regime in Figures 4.3, 4.4 and 4.5. To make inference about which volatility regime the process was in at a given time, the *smoothed probabilities* based on full sample data $Prob(s_t = i|y_{T,}y_{T-1}, ..., y_1)$ are generally used[12]. Following Hamilton (1989), we assume that process is in the high volatility regime at a given time if the associated *smoothed probability* of being in high volatility regime is greater than 0.5. The *ex ante probabilities* $Prob(s_t = i|y_{t-1,}y_{t-2}, ..., y_{1,})$ enables us to make inference about which volatility regime the process was in at date t based on observation obtained date $t-1$.

The graphs of the *smoothed probabilities* of being in high volatility regime clearly show up existence of two different volatility regimes. They also confirm that each regime is highly persistent. All SW-GARCH models identify that Turkish Stock Market starts in high volatility regime and then, switches permanently to the low volatility regime in the middle of 2003. So, the high volatility regime describes the long period from beginning 1997 to middle of 2003 while the remaining period from middle of 2003 to the end of 2006 is characterized by low

[12]Since *smoothed probabilities* use complete data, they give the most accurate answer to the question which regime the procees was likely in at a given time (Klaassen, 2002).

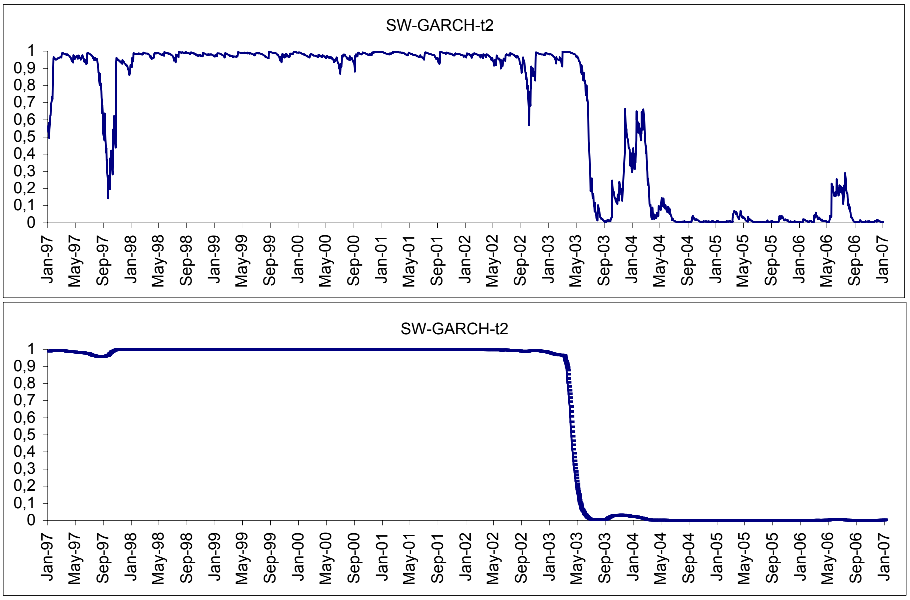

Figure 4.3: *Ex-ante* and *Smoothed probabilities* of being in high volatility regime for SW-GARCH-t2 respectively during the period 1997 to 2006

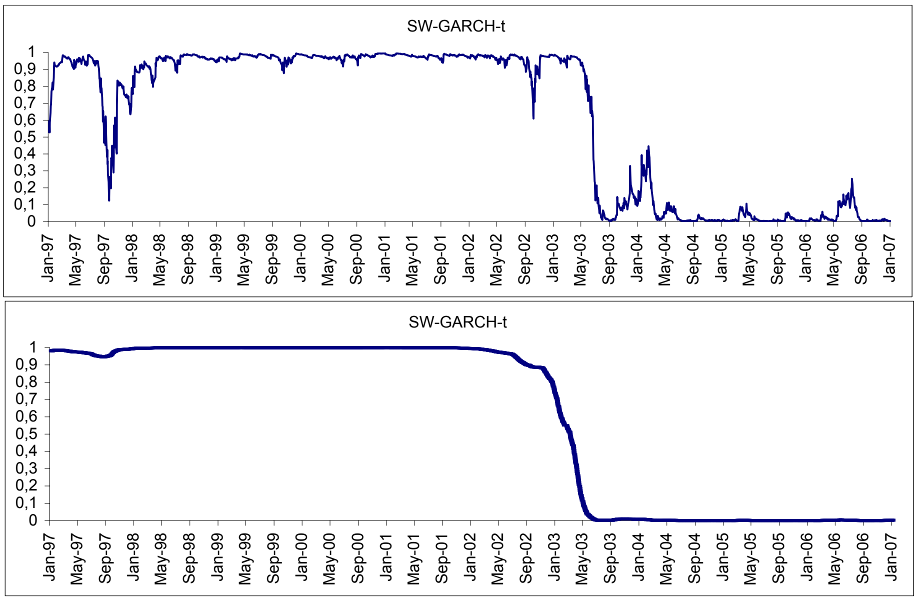

Figure 4.4: *Ex-ante* and *Smoothed probabilities* of being in high volatility regime for SW-GARCH-t respectively during the period 1997 to 2006

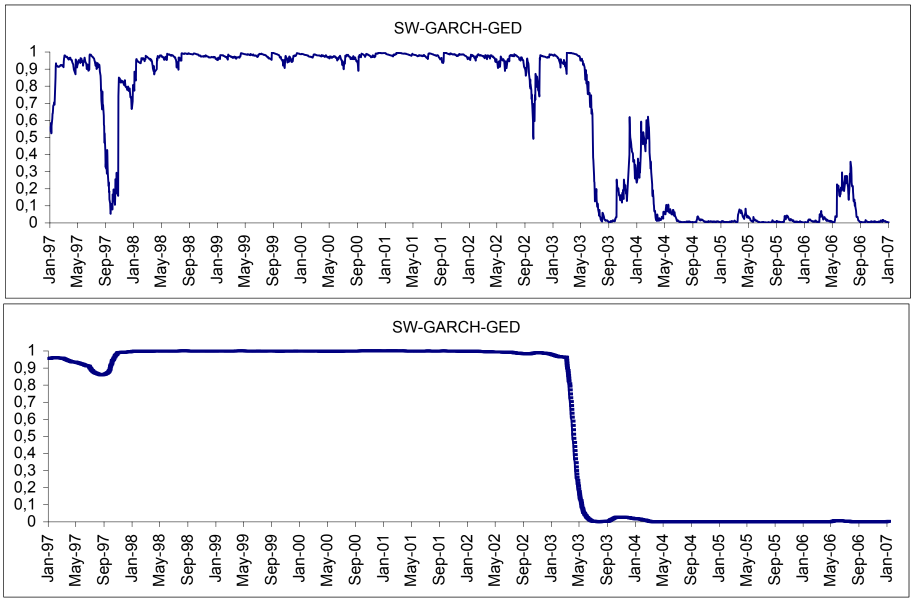

Figure 4.5: *Ex-ante* and *Smoothed probabilities* of being in high volatility regime for SW-GARCH-GED respectively during the period 1997 to 2006

volatility regime.

More importantly, Figures 4.3, 4.4 and 4.5 reveal that Turkish Stock Market exhibit a structural break around middle of 2003. There is a sharp decline in conditional variances after the break. Reasons of structural break and volatility reduction in Turkish Stock Market volatility can be quite complicated. Since our concern is estimating and forecasting Turkish Stock Market volatility, we do not examine the reasons behind the structural break.

Finally, we plot the estimated conditional volatility of all GARCH specifications in Figures 4.6, 4.7, and 4.8. All of the volatility models display similar patterns. Also, the significance decline about middle of 2003 confirms the structural break in Turkish Stock Market.

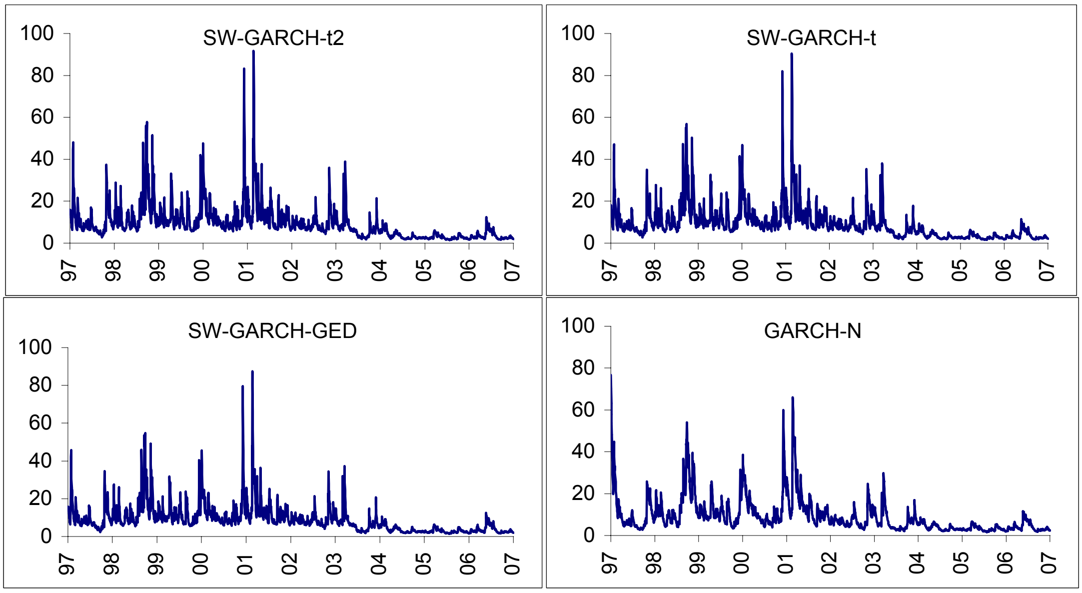

Figure 4.6: Conditional volatility of daily ISE-100 index returns over in-sample period 1997 to 2006.

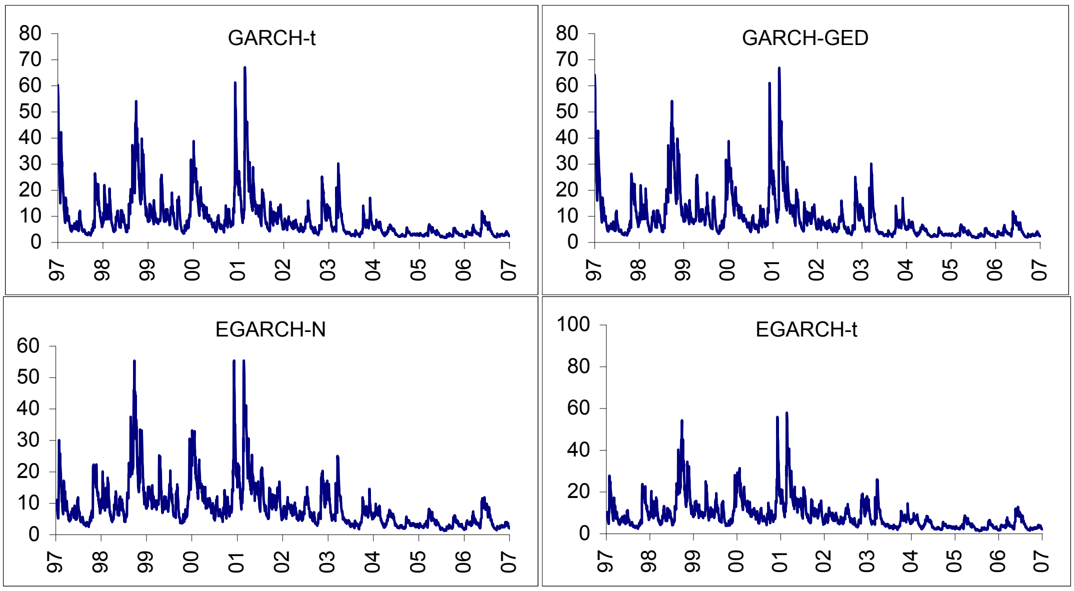

Figure 4.7: Conditional volatility of daily ISE-100 index returns over in-sample period 1997 to 2006.

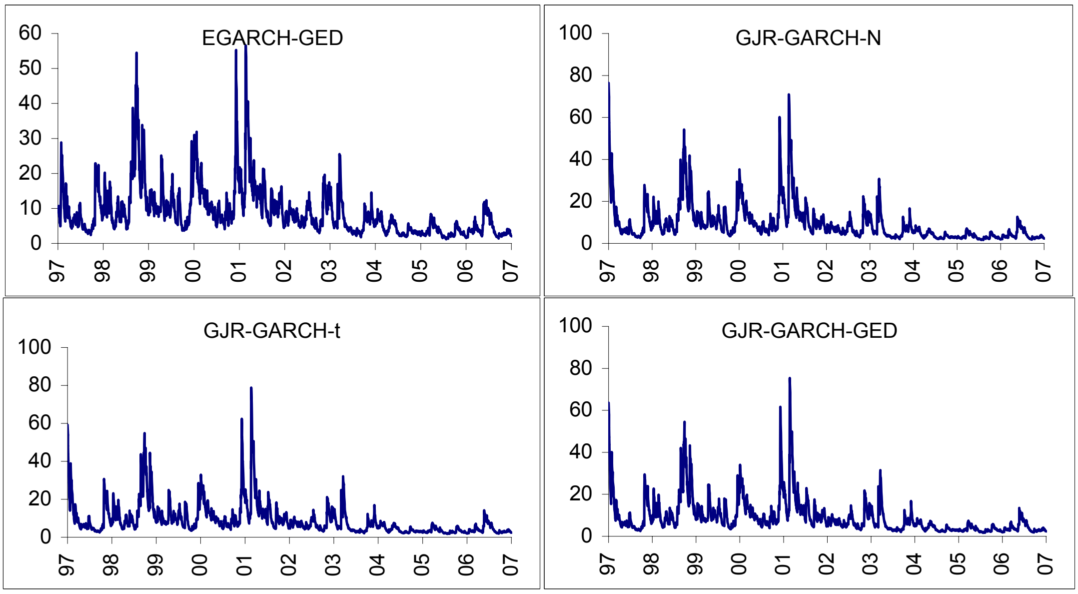

Figure 4.8: Conditional volatility of daily ISE-100 index returns over in-sample period 1997 to 2006.

4.3.3 In-Sample Evaluation

While making comparison between SW-GARCH and uni-regime GARCH models, standard Likelihood Ratio (LR) test is not applicable. Since the Markov transition probabilities are not identified under null hypothesis, the LR test statistics no longer follow χ^2 distribution (Hamilton and Susmel, 1994). Therefore, we use various goodness-of- fit statistics to compare volatility models. These statistics are Akaike information Criteria (AIC, Akaike, 1974) Schwarz Bayesian information criteria (SBIC, Schwarz, 1978), Hannan Quinn information criteria (HQIC, Hannan and Quinn, 1979) and Log-likelihood values.

The information criteria vary according to how they penalize number of estimated parameters. The SBIC penalize additional parameters larger than does the AIC. The SBIC and HQIC are consistent whereas AIC is not consistent. Also, SBIC is inefficient. The model with smaller value of information criteria is preferable[13]. The Akaike's, Schwarz's Bayesian and Hannan Quinn information criteria are computed as follows:

$$AIC = \frac{-2L}{T} + \frac{2k}{T}, \tag{4.3.1}$$

$$SBIC = \frac{-2L}{T} + \frac{k\ln(T)}{T}, \tag{4.3.2}$$

$$\mathrm{HQIC} = \frac{-2\mathrm{L}}{T} + \frac{2k\ln(\ln(T))}{T}, \tag{4.3.3}$$

where L is the value of the likelihood function, T is the number of observations, and k is the number of estimated parameters.

In Table 4.7, the results of goodness-of- fit statistics for all volatility models are presented. According to AIC and HQIC, SW-GARCH model with student-t and student-t2 perform best in modelling Turkish Stock Market volatility. However, in contrast the AIC and HQIC, the SBIC suggests that the uni-regime GARCH model with student-t provide the most accurate description of the data. In addition, three information criteria suggest that choice of student-t assumption for standardized errors is improves the fitting performance of all GARCH models. If log-likelihood function values are compared, SW-GARCH models highly outperform the uni-regime GARCH models and have considerably higher log-likelihood function values.

In addition to the goodness-of-fit statistics, we consider various statistical loss functions to analyze in-sample estimation performance of the volatility models[14]. We assume daily squared return as actual volatility. As seen Table 4.7, one of the SW-GARCH models obtains the highest ranking according to all statistical loss functions. Also, first three ranks are generally shared by SW-GARCH models.

[13] More detailed information on GARCH model selection criteria can be found in McKenzie and Mitchell (2003).

[14] The detailed information on statistical loss functions and actual volatility are introduced in Chapter 5.

Thus, evaluating in sample estimation results according to loss functions, as well as goodness-of-fit statstics, we conclude that SW-GARCH models perform better than uni-regime GARCH models in describing Turkish Stock Market volatility.

Lastly, as seen in third column of Table 4.7, comparing persistence of uni-regime GARCH models and SW-GARCH models, it is observed that the high persistence in the former specification is reduced by latter models[15]. This result is consistent with Lamoureux and Lastrapes's (1990) finding that is high persistent in volatility of GARCH is caused by regime shifts in the volatility process. Among all SW-GARCH models, SW-GARCH with student-t2 shows the largest decline in volatility persistence.

[15]For SW-GARCH models, following Marcucci (2005) we report the higher persistence value within each regime as model persistence value.

Table 4.7: In-Sample Evaluation Results

	# of Par.	Pers.	AIC	R	BIC	R	HQIC	R	LOGL	R	MSE	R	MAPE	R	QLIKE	R	R2LOG	R	MAE1	R	MAE2	R	HMSE	R
GARCH N	4	0,987	4,784	11	4,793	10	4,787	12	-5927,99	11	4,534	12	7,420	8	2,943	6	7,841	9	1,613	12	9,735	12	4,353	5
GARCH t	5	0,983	4,731	6	4,743	1	4,735	4	-5861,45	6	4,469	9	7,366	5	2,943	8	7,820	6	1,604	8	9,632	8	4,421	9
GARCH GED	5	0,984	4,739	8	4,751	5	4,744	7	-5871,68	8	4,479	10	7,362	4	2,943	7	7,811	4	1,604	9	9,647	10	4,446	10
EGARCH N	5	0,983	4,787	12	4,798	12	4,791	11	-5930,30	12	4,443	6	7,454	11	2,945	10	7,890	12	1,607	10	9,564	5	4,244	4
EGARCH t	6	0,980	4,730	5	4,745	3	4,736	5	-5859,81	5	4,412	4	7,474	12	2,946	12	7,886	11	1,601	4	9,507	3	4,356	6
EGARCH GED	6	0,981	4,740	9	4,754	8	4,745	9	-5872,05	9	4,405	3	7,429	9	2,945	11	7,859	10	1,598	3	9,494	2	4,385	8
GJR-GARCH N	5	0,987	4,784	10	4,795	11	4,788	10	-5926,78	10	4,524	11	7,430	10	2,942	4	7,835	8	1,611	11	9,726	11	4,383	7
GJR-GARCH t	6	0,983	4,730	3	4,744	2	4,735	3	-5859,04	4	4,459	7	7,389	7	2,943	9	7,814	5	1,602	6	9,628	7	4,511	12
GJR-GARCH GED	6	0,984	4,739	7	4,753	7	4,744	8	-5870,09	7	4,468	8	7,378	6	2,942	5	7,804	2	1,602	5	9,639	9	4,507	11
SW-GARCH t2	12	0,953	4,724	1	4,752	6	4,734	2	-5845,22	1	4,426	5	7,326	3	2,922	2	7,826	7	1,603	7	9,593	6	3,663	1
SW-GARCH t	11	0,962	4,724	2	4,750	4	4,733	1	-5846,56	2	4,385	2	7,306	2	2,922	1	7,810	3	1,594	2	9,514	4	3,714	2
SW-GARCH GED	11	0,956	4,730	4	4,756	9	4,740	6	-5854,76	3	4,362	1	7,272	1	2,923	3	7,781	1	1,589	1	9,467	1	3,837	3

Note: Pers. and R refers the persistence and Rank respectively.

CHAPTER 5

FORECASTING VOLATILITY

Volatility plays a key role in empirical finance and good forecasts of volatility is crucial for 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. In this chapter, we evaluate forecasting performance of SW-GARCH models with those of uni-regime GARCH models to determine which models make more accurate volatility prediction. To perform out-of-sample forecast, we divide sample into two parts. The first 2480 observations from 03/1/1997 to 31/12/2006 are used to estimate model parameters and remaining 248 observations from 04/01/2007 to 27/12/2007 are used for out-of-sample evaluation. Forecasts are based on the rolling window procedure and parameters of all models are re-estimated on each forecast date by adding next day observation and deleting first observation in the previous sample. So, all volatility models are estimated 227 times based on the samples of 2480 observations. Although many volatility forecasting papers compare accuracies at daily horizons, we evaluate forecasting performance of volatility models over daily ($k = 1$), weekly ($k = 5$), bi weekly ($k = 10$) and monthly ($k = 22$) horizons.

5.1 Realized Volatility

In order to assess forecasting performance of various models, firstly we have to define a proxy for actual volatility. Since volatility is not directly observable from market, unlike financial returns, it must be estimated. In the literature, general approach is to use squared daily (mean adjusted) returns as the measure of actual volatility, that is,

$$\sigma_t^2 = (r_t - \bar{r})^2, \tag{5.1.1}$$

where $\bar{r}$ is 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 estimate of unobserved volatility. Andersen and Bollerslev (1998) introduced an alternative volatility measure called Realized Volatility[1]. This measure has recently attracted 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 squared intraday returns and the higher frequency intraday data, the more noise reduction in the volatility estimate. Realized Volatility at day t can be formulated as follows

$$\sigma_t^2 = \sum\nolimits_{d=1}^{D} r_{t,d}^2, \tag{5.1.2}$$

where D is number of intraday return, such as D=24 for hourly data. However several assets are not traded whole day and changes during the out of trading hours must be considered. Then, 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,

$$\sigma_t^2 = R_{t,0}^2 + \sum\nolimits_{d=1}^{D} r_{t,d}^2, \tag{5.1.3}$$

where $R_{t,0}$ is the overnight return at day t. Hansen and Lunde (2005) stated that using overnight return, relatively large compared to the intraday return, leads to a noisy measure and suggest scaling estimator to obtain a measure for whole day

[1] Another alternative for actual volatility is *Implied Volatility* approach which is derived from matching trading prices of options.

volatility. So,

$$\sigma_t^2 = \hat{c} \sum\nolimits_{d=1}^{D} r_{t,d}^2 \quad where \quad \hat{c} = \frac{T^{-1} \sum\limits_{t=1}^{T} r_t^2}{T^{-1} \sum\limits_{t=1}^{T} \sum\nolimits_{d=1}^{D} r_{t,d}^2}. \tag{5.1.4}$$

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

Since intraday data of ISE-100 index is not available to us, we used squared daily returns as actual volatility for forecast horizon one. In order to calculate volatility over the k days, following Klaassen (2002), we sum squared daily returns over 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 actual volatility over the k days $t, ..., t+k-1$ as

$$\sigma_{t,K}^2 = \sum\nolimits_{i=t}^{t+k-1} (r_i - \bar{r})^2, \tag{5.1.5}$$

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

In practice, an investor, who has an investment horizon one month, generally concern with volatility forecast over the next 22 days rather than volatility forecast for day $t+22$ made on day t. So, we focus on volatility forecast over the next k days instead of k-step-ahead forecasts. In order to compute volatility forecast over the next k days, we aggregate k-step-ahead forecasts. Let $h_{t,K}$ denotes the volatility forecast over next k days, and then it can be formulated as follows,

$$h_{t,K} = \sum_{k=1}^{k} h_{t+k}, \tag{5.1.6}$$

where h_{t+k} denotes the k-step-ahead forecast made at time t [2].

[2]This calculation is permissible because of additive property of variance over the time (Brooks and Persand, 2003).

5.2 Statistical Loss Functions

After making forecasts and choosing a proxy for actual volatility, 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. 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 Error (MAE) and Heteroscedasticity-adjusted Mean Square Error (HMSE);

$$\begin{aligned}
MSE &= \frac{1}{n}\sum_{t=1}^{n}\left(\sigma_{t+K} - \sqrt{h_{t,K}}\right)^2, \\
MAPE &= 100\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\sigma_{t+K} - \sqrt{h_{t,K}}}{\sigma_{t+K}}\right|, \\
QLIKE &= \frac{1}{n}\sum_{t=1}^{n}\left(\ln\left(h_{t,K}\right) - \frac{\sigma^2_{t+K}}{h_{t,K}}\right), \\
R2LOG &= \frac{1}{n}\sum_{t=1}^{n}\left(\ln\left(\frac{\sigma^2_{t+K}}{h_{t,K}}\right)\right)^2, \\
MAE_1 &= \frac{1}{n}\sum_{t=1}^{n}\left(\sigma_{t+K} - \sqrt{h_{t,K}}\right), \\
MAE_2 &= \frac{1}{n}\sum_{t=1}^{n}\left(\sigma^2_{t+K} - h_{t,K}\right), \\
HMSE &= \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\sigma^2_{t+K}}{h_{t,K}} - 1\right)^2.
\end{aligned} \tag{5.2.7}$$

Since there is not any unique criterion that shows the best forecasting model[3], following Marcucci (2005), we used all of them rather than choosing a single loss function. That provides more comprehensive forecast evolution.

[3]See Bollerslev et al. (1994), Diebold and Lopez (1996) and Lopez (2001).

The MSE is the most widely used measure in forecast accuracy. The MAE_1 and MAE_2 are very similar to the MSE but they are more robust because of being less sensitive to large forecast errors. The main drawback of these loss functions is that they penalize both over forecast and under forecast equally. Bollerslev and Ghysels (1996) argued that MSE may be unreliable in the presence of heteroscedasticity and proposed HMSE[4]. 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 provided by Patton and Sheppard (2007). The lower loss functions values, the better forecasting performance.

As well as forecasting volatility accurately, predicting direction of volatility may be helpful for practitioners while constructing their investment strategies. For that purpose, we also evaluate out-of-sample forecasts by comparing fraction of the volatility forecast that have same sign of change as the actual volatility. We consider so-called Success Ratio (SR),

$$SR = \frac{1}{n}\sum_{j=1}^{n} I(\widehat{\sigma}^2_{t+j,K}.\widehat{h}_{t+j,K} > 0), \tag{5.2.8}$$

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

We apply Directional Accuracy (DA) test of Pesaran and Timmermann (1992) to test whether SR is significantly different from success ratio obtained in the case of independence of $\widehat{\sigma}^2_{t+j,K}$ and $\widehat{h}_{t+j,K}$ (SRI). Let $P = \frac{1}{n}\sum_{j=1}^{n} I(\widehat{\sigma}^2_{t+j,K} > 0)$

[4] *This type of performance measure is not appropriate if the absolute magnitude of the forecast error is a major concern. It is not clear why it is the predicted and not the actual volatility that is used in the denominator. The squaring of the error again will give greater weight to large errors*, Poon (2005).

and $\hat{P} = \frac{1}{n}\sum\limits_{j=1}^{n} I(\widehat{h}_{t+j,K} > 0)$, then SRI is computed as

$$SRI = P\hat{P} + (1-P)(1-\hat{P}). \tag{5.2.9}$$

DA test statistics is given as

$$DA = \frac{(SR - SRI)}{\sqrt{\text{var}(SR) - \text{var}(SRI)}} \xrightarrow{A} N(0,1), \tag{5.2.10}$$

where

$$\text{var}(SR) = \frac{1}{n} SRI(1 - SRI),$$

$$\text{var}(SRI) = \frac{1}{n}\left[(2\hat{P}-1)^2 P(1-P) + (2P-1)^2 \hat{P}(1-\hat{P}) + \frac{4}{n} P\hat{P}(1-P)(1-\hat{P})\right].$$

5.3 Tests for Forecasting Performance

The accuracy of volatility forecasting from different models can be measured by forecast error statistics (loss functions). However, when a forecast error statistics of a benchmark model is smaller than that of an alternative model we can not undoubtedly conclude that forecasting performance of benchmark model is superior to that of alternative model. To make a fair comparison among forecasting performance of different models, statistical significance of the observed diffrence in forecast error statistics should be investigated. For that purpose, Diebold and Mariano (1995) developed a test for equal predictive ability of two challenging models. Diebold and Mariano (DM) test is designed as follows:

H_0 : No difference between predictive performance of two models,

H_A : Predictive performance of two models are not equally accurate.

Let $(\mathrm{e}_{\mathrm{i,t}})_{t=1}^{n}$ and $(\mathrm{e}_{\mathrm{j,t}})_{t=1}^{n}$ denote forecast errors of two models i and j, then loss differential between two challenging models is defined as

$$d_t = g(\mathrm{e}_{\mathrm{i,t}}) - g(\mathrm{e}_{\mathrm{j,t}}) \qquad t = 1, 2, ..., n,$$

where $g(.)$ is the loss function. Assuming d_t is covariance stationary and has short memory, Diebold and Mariano (1995) show that mean difference between loss functions $\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$ is asymptotically distributed as

$$\sqrt{n}(\bar{d} - \mu) \xrightarrow{a} N(0, \hat{V}(\bar{d})),$$

where $\hat{V}(\bar{d}) = n^{-1}(\hat{\gamma}_0 + 2\sum_{i=1}^{k-1}\hat{\gamma}_i)$, $\hat{\gamma}$ is an estimate of the $i-th$ order autocovariance of the series d_t and k is the forecast horizon. So, DM test statistics under null hypothesis of equal predictive accuracy is given as

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \rightarrow N(0,1).$$

Harvey, Leybourne, and Newbold (1997) state that DM test can be over-sized when sample size is small and forecast horizon is long. To overcome the oversize problem, they propose Modified DM statistics (MDM) which is obtained by multiplying DM statistics with correction factor below

$$\sqrt{n^{-1}\left[n + 1 - 2k + n^{-1}\,k(k-1)\right]},$$

where n is evaluation period and k is the forecast horizon. Also, they suggest student-t distribution instead of normal distribution to compare test statistics.

DM test gives an opportunity to verify whether forecasting performance of two models is statistically same or not, but in practice a benchmark model is generally compared with multiple models. To make multiple comparison, White (2000) proposed Reality Check (RC) test for superior predictive ability. There are two main contributions of White's RC test that make it crucial in empirical finance. First, it enables researchers to implement joint hypothesis testing by using bootstrap methods . Second, it overcomes potential data snooping problem which arise when a given set of data is used more than once for purpose of model selection[5]. Null hypothesis of RC test is that a benchmark model is not outperformed by an alternative set of models according to pre-specified loss

[5]See White (2000) for more details on data snooping problem.

function. Let $\mathrm{L}(\sigma_{\mathrm{t}}^2, \mathrm{h_t})$ denotes loss function for prediction with model m, then relative forecasting performance of model m compared to the benchmark model 0 at time t is defined as

$$\mathrm{X_{m,t}}= \mathrm{L}(\sigma_{0,\,t}^2, \mathrm{h}_{0,\,t}) - \mathrm{L}(\sigma_{m,\,t}^2, \mathrm{h}_{\mathrm{m},\,t}),$$

where $m : 1, 2, ..., M; \quad t : 1, 2, ..., n.$

If $X_{m,t}$ is stationary, expected relative forecasting performance of model m can be defined as $\mu_m = E(X_{m,t})$. When the benchmark model 0 is outperformed by model m, μ_m takes positive values. So, the null hypothesis that benchmark model is not outperformed by any alternative model is formulated as follows

$$H_0 : max \quad \mu_m \leq 0 \qquad m : 1, 2, .., M.$$

The RC test statistics of White (2000) is given as

$$T_n^{RC} = \max_{m=1,..M} \sqrt{n}\bar{X}_m, \tag{5.3.11}$$

where $\bar{X}_m = n^{-1} \sum\limits_{t=1}^{n} X_{m,t}$.

However, it is very difficult to derive the theoretical distribution of test statistics T_n^{RC} under null since null distribution is not unique. For this reason, White (2000) suggest that empirical distribution of test statistics under null can be obtained by stationary bootstrap of Politis and Romano (1994). The appropriate p-values for testing null hypothesis can be computed by this method (White, 2000). After getting bootstrap samples X_m^b for $b : 1, 2, ..., B$ where B is number of bootstraps, the empirical distribution of test statistic T_n^{RC} under null can be identified as

$$T_b^{RC} = \max_{m=1,..M} \sqrt{n}(\bar{X}_m^b - \bar{X}_m), \tag{5.3.12}$$

where $\bar{X}_m^b = n^{-1} \sum\limits_{t=1}^{n} X_{m,t}^b, \quad b : 1, 2, ...B.$

Hansen (2005) indicates that the RC test is conservative since it is too sensitive to inclusion of poor models. When there is a poor model in the set of alternative models, RC p-values may remain large even after inclusion of better models. To tackle this problem, Hansen (2005) introduced the Superior Predictive Ability (SPA) test which is an additional development of RC. He suggests a method to obtain consistent estimate and lower bound of p-values of SPA test; those are more powerful and less sensitive to the inclusion of poor models. Hansen (2005) proposed to use following standardized test statistics to test null hypothesis

$$T_n^{SPA} = \max_{m=1,..M} \frac{\sqrt{n}\bar{X}_m}{\hat{w}_{mm}}, \tag{5.3.13}$$

where $\hat{w}_{mm}$ is consistent estimate of $w_{mm} = \lim_{n\to\infty} \operatorname{var}(\sqrt{n}\bar{X}_m)$.

Since the empirical distribution of test statistics under null is unknown, the consistent p-value of SPA test (SPA_C), lower bound for p-value of SPA test (SPA_L) and consistent estimator of w_{mm} can be computed with a stationary bootstrap of Politis and Romano (1994). After having bootstrap samples X_m^b, following modifications are considered by Hansen (2005) to obtain empirical distribution of test statistic T_n^{SPA}

$$T_b^{SPA_C} = \max_{m=1,..M} \frac{\sqrt{n}(\bar{X}_m^b - \bar{X}_m I_{(\bar{X}_m > -A_m)})}{\hat{w}_{mm}}, \tag{5.3.14}$$

$$T_b^{SPA_L} = \max_{m=1,..M} \frac{\sqrt{n}(\bar{X}_m^b - \max(\bar{X}_m, 0))}{\hat{w}_{mm}}, \tag{5.3.15}$$

where $\hat{w}_{mm} = \frac{n}{B}\sum_{b=1}^{B}\left(\bar{X}_m^b - \bar{X}_m\right)^2, \quad A_m = \frac{1}{4}n^{0,25}\hat{w}_{mm}, \quad b : 1, 2, ..., B.$

Hence, consistent estimate and lower bound of the p-values of SPA test are directly computed as below

$$p_C - value = \frac{1}{\mathrm{B}}\sum\nolimits_{b=1}^{B} I_{(T_b^{SPA_C} > T_n^{SPA})}, \tag{5.3.16}$$

$$p_L - value = \frac{1}{\mathrm{B}}\sum\nolimits_{b=1}^{B} I_{(T_b^{SPA_L} > T_n^{SPA})}, \tag{5.3.17}$$

where $I(.)$ is the indicator function, with value 1 when its argument is true and 0 otherwise.

5.3.1 Stationary Bootstrap

The bootstrap is an alternative technique for estimating distribution of an estimator or test statistic without making precise distributional assumption about data (Efron, 1979, 1982). This method is applied when conventional techniques are not valid or original sample size is small. The idea of the bootstrap is to generate many resamples by repeatedly sampling with replacement from the data at hand. Sample size of each resample is same with that of original data.

The methods for implementing the bootstrap depend on whether the data is a random sample from a distribution or a time series. Since our interest is in times series data, we do not focus on bootstrap methods for random sample. Detailed information about bootstrap methods for random sample can be found in Beran and Ducharme (1991), Hall (1992) and Efron and Tibshirani (1993).

There are mainly two bootstrap approaches that capture dependence structure in time series data: sieve bootstrap and block bootstrap. Each of them has many variants. Particularly, in block bootstrap methods, originally proposed by Kunsch (1989) , observed sample is divided into fixed length blocks of consecutive observations to generate resample. Then, blocks of consecutive observations are drawn with replacement from the set of blocks instead of only one observation. The blocks may be either overlapping or non-overlapping. In order to describe overlapping and non-overlapping blocks, let X_i for $i = 1, 2, ..., n$ be the original data set and block length l be 4. Then, non-overlapping and overlapping blocks can be represented respectively as below,

$X_1\ X_2\ X_3\ X_4\ X_5\ X_6\ X_7\ X_8 .... X_{n-3} X_{n-2}\ X_{n-1}\ X_n$ non-overlapping blocks

$X_1\ X_2\ X_3\ X_4\ X_5\ X_6\ X_7\ X_8 .... X_{n-3} X_{n-2}\ X_{n-1}\ X_n$ overlapping blocks

.
.
.

The choice of block length is the most important part of the block bootstrap.

However, selecting an appropriate block length is not an easy task always and may effects significantly performance of bootstrap (Lahiri, 1999). For instance, if block length is small, resamples may not imitate the pattern of dependence in original data. Another drawback of the block bootstrap is that resamples are not stationary even if original data is stationary (Hardle, Horowitz and Kreiss, 2003).

Politis and Romano (1994) introduced the stationary bootstrap which is a modification of overlapping block bootstrap method. They argue that resamples generated by this method become stationary and less sensitive to choice of expected block length. The stationary bootstrap uses random block length instead of fixed blocks length and the length of blocks are drawn independently from a geometric distribution with mean block length q. Then, random lengths become ideally small but sufficiently large to reflect serial dependence in the original data (Koopman, Jungbacker and Hol, 2000). Recent surveys of bootstrap methods for time series data include Horowitz (2003), Politis (2003), and Hardle, Horowitz, and Kreiss (2003).

To construct empirical distribution of SPA test statistics, following algorithm proposed by Politis and Romano (1994) can be implemented. Let X be $M \times n$ matrix consist of $X_{m,t}$ which is the relative forecasting performance of model m compared to the benchmark model at time t for $m = 1, 2, ..., M$ and $t = 1, 2, ...n$. Also, let q denotes expected block length which is an integer satisfying $0 < q < n$. Then $p = 1/q$.

1. Select a column $(X_{1,t}, X_{2,t}, ..., X_{M,t})$ randomly from the original matrix X.
2. Set first observation of resample data $\left(X^b_{1,1}, X^b_{2,1}, ..., X^b_{M,1}\right) = (X_{1,t}, X_{2,t}, ..., X_{M,t})$.
3. Draw an independent standard uniform variable U (between 0 and 1).
4. If U $\geq p$, then set $\left(X^b_{1,2}, X^b_{2,2}, ..., X^b_{M,2}\right) = (X_{1,t+1}, X_{2,t+1}, ..., X_{M,t+1})$ as the second observation of resample data; else select a new column $(X_{1,t}, X_{2,t}, ..., X_{M,t})$ randomly from original matrix X and

set $\left(X^b_{1,2}, X^b_{2,2}, ..., X^b_{M,2}\right) = (X_{1,t}, X_{2,t}, ..., X_{M,t})$ as the second observation of resample data.

5. Repeat the steps 3 and 4 to construct $\left(X^b_{1,3}, X^b_{2,3}, ..., X^b_{M,3}\right) = (X_{1,t}, X_{2,t}, ..., X_{M,t})$, and continue until n columns are drawn.

Repeating this procedure B times yields an empirical distribution for $\bar{X}^b_m$ with B realizations.

5.4 Out-of-Sample Evaluation

5.4.1 Results of Statistical Loss Functions

One of the main purpose of specifying a volatility model, as well as describing its some features, is forecasting future volatilities. Since volatility forecasting is crucial for option pricing, risk management and portfolio management etc., it has attracted much attention of investors over the recent decays. So far, we compared various GARCH models in terms of fitting data, capturing persistence and in-sample estimation performance. However, good fitting the data or superior in-sample estimation performance do not insure superior performance at volatility forecasting. In this section, we investigate ability of markov regime switching and uni-regime GARCH models to forecast Turkish Stock Market volatility at different forecast horizons. The forecast horizons 1, 5, 10 and 22 days are considered in this thesis.

In Table 5.1, we present the forecast error statistics for one day ahead. The six of seven forecast error statistics suggest that SW-GARCH models provide the most accurate volatility forecasts. In terms of MSE, MAPE, R2LOG, MAE1 and MAE2, the best forecasting performance belongs to the SW-GARCH model with GED; the second and third best models are SW-GARCH with student-t

Table 5.1: Out-of-Sample forecasting results for one day forecast horizon

	MSE	R	MAPE	R	QLIKE	R	R2LOG	R	MAE1	R	MAE2	R	HMSE	R	SR	DA
GARCH N	1.690	5	5.296	5	2.168	10	7.468	5	3.637	5	1.056	5	2.593	7	0.530	-2.226
GARCH t	1.692	6	5.316	6	2.169	12	7.480	6	3.637	6	1.057	6	2.643	8	0.540	-1.908
GARCH GED	1.686	4	5.293	4	2.169	11	7.459	4	3.629	4	1.054	4	2.651	9	0.530	-2.226
EGARCH N	1.723	9	5.514	8	2.150	8	7.572	8	3.745	7	1.081	7	2.112	5	0.540	-1.802
EGARCH t	1.740	12	5.545	11	2.151	9	7.613	11	3.781	12	1.088	12	2.079	4	0.560	-1.100
EGARCH GED	1.724	10	5.513	7	2.149	7	7.571	7	3.750	10	1.081	8	2.126	6	0.540	-1.487
GJR-GARCH N	1.718	8	5.543	10	2.146	2	7.608	10	3.749	9	1.083	10	1.984	1	0.560	-1.377
GJR-GARCH t	1.726	11	5.548	12	2.149	5	7.622	12	3.763	11	1.086	11	2.033	3	0.570	-1.161
GJR-GARCH GED	1.717	7	5.533	9	2.146	3	7.600	9	3.747	8	1.082	9	2.026	2	0.560	-1.377
SW-GARCH t2	1.658	3	5.107	3	2.144	1	7.260	2	3.607	3	1.034	3	2.791	10	0.590	-1.347
SW-GARCH t	1.610	2	5.102	2	2.148	4	7.268	3	3.515	2	1.025	2	2.814	11	0.570	-1.233
SW-GARCH GED	1.590	1	5.028	1	2.149	6	7.181	1	3.476	1	1.014	1	3.068	12	0.570	-1.161

Note: R refers the Rank.

and student-t2. These models are followed by uni-regime GARCH models. Also, QLIKE loss function suggests that SW-GARCH model with student-t2 ranks top. In contrast, according to HMSE, SW-GARCH models are highly outperformed by uni-regime models and GJR-GARCH models performs the best among all models. Overall, it can be suggested that SW-GARCH models improve volatility forecast at one day ahead compared to the uni-regime GARCH models. Lastly, if we consider the distribution assumptions for errors, models with student-t clearly perform the worst out-of-sample forecasts within each GARCH specifications on all statistical loss functions, although they are the best at fitting the data and in-sample estimation.

For the 5 days (one week) horizon, results of forecast error statistics are given in Table 5.2. According to all forecast error statistics except QLIKE and HMSE, EGARCH model with normal distribution shows the best forecasting performance and followed by models EGARCH with GED and student-t. SW-GARCH models with GED and student-t come only fourth and fifth. On the other hand, in terms of QLIKE, first three ranks are shared by SW-GARCH models. Although most of the loss functions choose GJR-GARCH model as the worst models, HMSE suggest that they have better forecasting performance than the all other models have. Overall, since EGARCH models outperform others with respect to five of seven loss function, it can be proposed that these models beat all other models in respect to forecasting performance.

Table 5.3 reports values and rankings of the statistical loss functions for forecasts horizon 10 days (two weeks). From the examination of Table 5.3, it is noted that none of the models clearly outperform the alternatives. Three of loss functions (MSE, R2LOG, and MAE2) indicate SWGARCH model with GED provide the most accurate forecasts, while two of them (MAPE and MAE1) favor EGARCH model with normal. Also, models SW-GARCH with student-t and GJR-GARCH with student-t rank first in terms of QLIKE and HMSE respectively. However, in general, it seems that GJR-GARCH models are dominated

Table 5.2: Out-of-Sample forecasting results for one week (5 days) forecast horizon

	MSE	R	MAPE	R	QLIKE	R	R2LOG	R	MAE1	R	MAE2	R	HMSE	R	SR	DA
GARCH N	3.296	7	0.557	7	3.787	7	1.104	8	1.433	7	11.946	7	0.657	5	0.590	0.606
GARCH t	3.313	8	0.562	9	3.788	11	1.112	9	1.442	9	12.009	8	0.652	4	0.580	0.393
GARCH GED	3.291	6	0.557	8	3.787	9	1.103	7	1.434	8	11.944	6	0.660	6	0.590	0.606
EGARCH N	2.845	1	0.469	1	3.787	8	0.928	1	1.308	1	10.718	1	0.983	11	0.590	0.973
EGARCH t	2.867	3	0.475	3	3.785	5	0.939	3	1.318	3	10.824	3	0.946	10	0.580	0.642
EGARCH GED	2.849	2	0.469	2	3.787	10	0.930	2	1.310	2	10.745	2	0.983	12	0.600	1.147
GJR-GARCH N	3.433	10	0.584	10	3.783	4	1.147	10	1.486	10	12.526	10	0.540	2	0.620	1.356
GJR-GARCH t	3.526	12	0.596	12	3.788	12	1.171	12	1.515	12	12.810	12	0.535	1	0.610	1.143
GJR-GARCH GED	3.468	11	0.588	11	3.785	6	1.155	11	1.496	11	12.637	11	0.540	3	0.610	1.264
SW-GARCH t2	3.324	9	0.527	6	3.777	1	1.038	6	1.411	6	12.084	9	0.684	7	0.610	0.244
SW-GARCH t	3.067	5	0.520	5	3.778	2	1.025	5	1.368	5	11.341	5	0.734	8	0.580	0.139
SW-GARCH GED	3.033	4	0.508	4	3.779	3	1.003	4	1.354	4	11.216	4	0.788	9	0.590	0.227

Note: R refers the Rank.

Table 5.3: Out-of-Sample forecasting results for two weeks (10 days) forecast horizon

	MSE	R	MAPE	R	QLIKE	R	R2LOG	R	MAE1	R	MAE2	R	HMSE	R	SR	DA
GARCH N	4.808	6	0.419	7	4.489	4	0.714	7	1.745	7	20.623	6	0.382	6	0.580	0.504
GARCH t	4.864	8	0.425	9	4.491	6	0.724	9	1.766	9	20.856	8	0.377	4	0.580	0.396
GARCH GED	4.810	7	0.420	8	4.489	5	0.715	8	1.748	8	20.643	7	0.381	5	0.580	0.504
EGARCH N	4.466	4	0.318	1	4.607	11	0.632	3	1.626	2	18.186	1	1.493	11	0.530	-0.786
EGARCH t	4.437	3	0.320	3	4.597	10	0.630	2	1.626	3	18.247	2	1.410	10	0.520	-1.503
EGARCH GED	4.483	5	0.319	2	4.608	12	0.636	4	1.634	4	18.267	3	1.494	12	0.520	-1.099
GJR-GARCH N	5.264	10	0.449	10	4.497	7	0.767	10	1.854	10	22.221	10	0.349	2	0.580	0.176
GJR-GARCH t	5.503	12	0.462	12	4.504	9	0.795	12	1.902	12	22.876	12	0.348	1	0.560	-0.327
GJR-GARCH GED	5.361	11	0.453	11	4.500	8	0.777	11	1.871	11	22.466	11	0.349	3	0.580	0.064
SW-GARCH t2	4.994	9	0.391	6	4.482	3	0.664	6	1.708	6	20.890	9	0.426	7	0.580	-0.460
SW-GARCH t	4.311	2	0.380	5	4.480	1	0.639	5	1.635	5	19.131	5	0.455	8	0.570	-0.121
SW-GARCH GED	4.264	1	0.371	4	4.481	2	0.626	1	1.619	1	18.955	4	0.490	9	0.570	-0.029

Note: R refers the Rank.

by others.

As well as short forecast horizons, we consider forecasting performance of various GARCH models at longer horizon 22 days (one month). Results are presented in Table 5.4. The rankings for one month horizon are quietly similar to that of the one day horizon. According to all loss functions except HMSE, SW-GARCH model with GED is the best model in forecasting volatility while SW- GARCH model with student-t ranks second and SW- GARCH model with student-t2 ranks third. Following markov regime switching models, standard uni-regime GARH models are ranked as fourth, fifth and sixth. On the other hand, HMSE suggests that top three volatility forecasters are standard uni-regime GARCH models.

It is important to note that there is substantial difference between results of HMSE and other statistical loss functions if model comparisons are considered. Most of time, result of HMSE are completely opposite to that of others. Marcucci (2005) has confronted with similar results and stated that *HMSE loss is not particularly suitable for evaluating different volatility forecasts and it should be expected to give weird results*[6].

Finally, we examine ability of volatility models to forecast sign of Turkish Stock Market volatility relative to its average volatility. We apply DA test and results are shown at last column of Tables 5.1-5.4. At all horizons, all models have the success ratios changing between 0.58 and 0.64. But none of the models are successful to predict sign of Turkish Stock Market volatility accurately relative to its average over the all forecasts horizons even 90% significance level.

5.4.2 Results of SPA test

So far, we evaluate forecasting performance of volatility models in terms of seven statistical loss functions. However, these loss functions give us only an idea to evaluate forecasting performance of models. Without any formal statistical

[6] For more discussion, see Patton (2005).

Table 5.4: Out-of-Sample forecasting results for one month (22 days) forecast horizon

	MSE	R	MAPE	R	QLIKE	R	R2LOG	R	MAE1	R	MAE2	R	HMSE	R	SR	DA
GARCH N	7.435	5	0.324	4	5.294	4	0.433	4	2.256	4	40.204	5	0.209	2	0.580	0.688
GARCH t	7.620	6	0.331	6	5.298	6	0.446	6	2.296	6	40.864	6	0.210	3	0.580	0.784
GARCH GED	7.429	4	0.325	5	5.294	5	0.435	5	2.259	5	40.192	4	0.208	1	0.580	0.784
EGARCH N	13.201	11	0.340	8	5.997	12	1.034	12	2.998	11	43.840	9	5.199	12	0.520	-0.806
EGARCH t	12.849	10	0.335	7	5.944	10	0.987	10	2.959	10	43.525	8	4.741	10	0.530	-0.738
EGARCH GED	13.225	12	0.341	9	5.993	11	1.033	11	3.009	12	44.015	10	5.139	11	0.530	-0.727
GJR-GARCH N	8.797	7	0.351	10	5.311	7	0.494	7	2.411	7	43.482	7	0.214	4	0.600	1.082
GJR-GARCH t	9.450	9	0.365	12	5.321	9	0.525	9	2.497	9	45.223	12	0.221	6	0.590	0.691
GJR-GARCH GED	9.030	8	0.356	11	5.315	8	0.505	8	2.438	8	44.036	11	0.217	5	0.590	0.791
SW-GARCH t2	7.294	3	0.273	3	5.277	3	0.358	3	2.023	3	37.440	3	0.240	7	0.640	1.505
SW-GARCH t	5.658	2	0.268	2	5.273	2	0.335	2	1.972	2	34.522	2	0.251	8	0.580	0.494
SW-GARCH GED	5.490	1	0.259	1	5.271	1	0.323	1	1.925	1	33.637	1	0.264	9	0.570	0.199

Note: R refers the Rank.

test, we can not answer the question that relative differences between forecasting performance of volatility models are fairly large or not. So, just by ranking the forecast error statistics, we can not undoubtedly suggest any model is superior or outperformed significantly.

If a benchmark model is compared with a single alternative model, DM test can be applied. However, if we compare a benchmark model with several models, DM test may give wrong result because of data snooping problem[7]. Therefore, we apply the RC of White (2000) and SPA test of Hansen (2005) since they overcome the data snooping problem and allow simultaneous comparison of many models. The null hypothesis of these two tests is that none of the alternative models outperform the benchmark model in terms of pre-specified loss function.

In Tables 5.5-5.8, results from RC and SPA test are presented in the form of p-values. One by one, each model is defined as a benchmark and tested against other remaining models in terms of seven statistical loss functions. Benchmark models are shown in the rows. While the RC in tables denotes the p-values of RC[8] test, SPA_L and SPA_C refer to the lower and consistent p-values of SPA test.

We use stationary bootstrap algorithm proposed by Politis and Romano (1994) with 10.000 re-samples to compute p-values. Following Marcucci (2005), three different expected block length 3, 5 and 10 are considered. Since there is no considerable difference between p-values, we report only the result of expected block length 10.

Hansen (2005) indicates that inclusion of poor models in the set of alternative models may artificially increase p-values of the RC test, and then make it

[7]Data snooping problem occurs since the DM test statistics become mutually dependent because of using same data set more than once. White (2000) stated that *When data reuse occurs, there is always the possibility that any satisfactory results obtained may simply be due to chance rather than to any merit inherent in the method yielding the results.*

[8]SPA test includes RC as special case and upper bound for p-values of SPA test is p-values of RC test (Marcucci, 2005).

conservative. Moreover, he shows that the SPA test is a more powerful test than the RC[9]. Therefore, although we present results of RC test, following Hansen and Lunde (2005), we consider only the p-values based on SPA_C test to assess relative performance of volatility models. As seen in Tables 5.5-5.8, p-values of RC test based on the all loss functions except HMSE display too conservative results[10]. All of the models can not be beaten by alternative models in terms of any loss functions. These results are consistent with Hansen (2005) findings.

Table 5.5 reports p-values of SPA_C tests for forecast horizon one day. When uni-regime models are benchmark, in terms of all loss functions except QLIKE, the p-values are very small and indicate clearly that all uni-regime GARCH models are outperformed by other models at least 94% confidence level. However, when SW-GARCH models are defined as benchmark, all SPA_C tests based on all loss function except HMSE have high p-values (at least 0.14) and indicating SW-GARCH models under each distribution can not be beaten. Thus, there is strong evidence that SW-GARCH models superior uni-regime GARCH models. Moreover, the higher p-values indicate the more accurate forecasting ability. Since p-values of SPA_C obtained when SW-GARCH model with GED is benchmark are dramatically higher than that of remaining models, we can suggest this model is best to forecast volatility of Turkish Stock market at one day forecast horizon.

In Table 5.6, we present p-values of SPA_C test for five days (a week) forecast horizon. All GARCH and GJR-GARCH models are outperformed significantly by other models at least 93% confidence level in terms of all loss functions except QLIKE. It is interesting to notice that although EGARCH models have the smallest loss functions values[11], their p-values are smaller than that of SW-GARCH models. Moreover, EGARCH models are outperformed significantly by other

[9]For more discussion, see Hansen (2005), Hansen and Lunde (2005) and Hsu and Kuan (2005).

[10]The p-value of less than α indicate the rejection of Null hypothesis with $(1-\alpha)\%$ confidence level.

[11]Results are given in Table 5.2.

models in terms of MAPE, R2LOG and HMSE, while any model can not superior the SW-GARCH models in terms of all loss functions except HMSE. Therefore, overall, it can be concluded that SW-GARCH models are the best specifications to forecast volatility for five day forecast horizon. On the other hand, when the SW-GARCH models with GED is the benchmark, p-values display substantial increase. Thus, it can be concluded that forecast from SW-GARCH model with GED is the most preferable among all models for 5 days forecast horizons.

If we consider the p-values of SPA_C test for ten days forecast horizon (given in Table 5.7), GARCH and GJR-GARCH models again have very small p-values and display significantly worse forecasting performance. On the other hand, when SW-GARCH or EGARCH model is benchmark, p-values are very high and indicating that these models can not be outperformed by other models. Moreover, since p-values are too close to each others in both cases, we can not claim that SW-GARCH models are superior to the EGARCH models or vice-versa. It can be concluded that there is not a huge difference in ability of these two models to forecast volatility for two weeks horizons.

Lastly, performances of volatility models are tested by SPA_C for long forecast horizon 22 days. Results are given in Table 5.8. GARCH and GJR-GARCH models are outperformed by other models in terms of all loss functions except QLIKE. However, SW-GARCH and EGARCH models have high p-values according to all loss function except HMSE when they are benchmark. As can be seen in Table 5.4, loss function values of SW-GARCH models are smaller than that of EGARCH models, but latter ones have dramatically higher p-values than former ones in terms of all loss functions when they are benchmark. Thus, it is implied that EGARCH models are the best forecaster for one month forecast horizons.

Besides, it can be suggested that the choice of model specification effects quality of forecasting more importantly than that of distribution assumption. On the other hand, the loss functions HMSE and QLIKE do not reveal any information on forecasting performance of models. Independent of benchmark model and forecast horizon, the SPA_C test is always rejected in terms of HMSE while not rejected in terms of QLIKE.

Table 5.5: SPA and RC test results for one day forecast horizon

Benchmark	Test	MSE	MAPE	QLIKE	R2LOG	MAE1	MAE2	HMSE
GARCH N	SPA_L	0.014	0.038	0.284	0.000	0.001	0.013	0.002
GARCH N	SPA_C	0.033	0.038	0.284	0.000	0.001	0.013	0.002
GARCH N	RC	0.508	0.437	0.505	0.256	0.317	0.500	0.002
GARCH t	SPA_L	0.020	0.040	0.246	0.000	0.000	0.014	0.002
GARCH t	SPA_C	0.020	0.040	0.246	0.000	0.000	0.014	0.002
GARCH t	RC	0.510	0.423	0.472	0.245	0.310	0.502	0.002
GARCH GED	SPA_L	0.021	0.048	0.268	0.000	0.001	0.016	0.001
GARCH GED	SPA_C	0.021	0.048	0.268	0.000	0.001	0.016	0.001
GARCH GED	RC	0.523	0.451	0.488	0.257	0.331	0.499	0.001
EGARCH N	SPA_L	0.011	0.046	0.664	0.000	0.000	0.000	0.001
EGARCH N	SPA_C	0.011	0.046	0.675	0.000	0.000	0.000	0.001
EGARCH N	RC	0.478	0.289	0.845	0.200	0.227	0.457	0.001
EGARCH t	SPA_L	0.006	0.030	0.603	0.000	0.000	0.000	0.004
EGARCH t	SPA_C	0.006	0.030	0.603	0.000	0.000	0.000	0.004
EGARCH t	RC	0.472	0.257	0.827	0.174	0.205	0.429	0.004
EGARCH GED	SPA_L	0.011	0.043	0.667	0.000	0.000	0.000	0.002
EGARCH GED	SPA_C	0.011	0.043	0.670	0.000	0.000	0.000	0.002
EGARCH GED	RC	0.476	0.287	0.842	0.201	0.221	0.451	0.002
GJR-GARCH N	SPA_L	0.003	0.054	0.777	0.000	0.000	0.000	0.002
GJR-GARCH N	SPA_C	0.003	0.054	0.790	0.000	0.000	0.000	0.002
GJR-GARCH N	RC	0.484	0.275	0.985	0.147	0.215	0.441	0.002
GJR-GARCH t	SPA_L	0.003	0.040	0.716	0.000	0.000	0.000	0.003
GJR-GARCH t	SPA_C	0.003	0.040	0.749	0.000	0.000	0.000	0.003
GJR-GARCH t	RC	0.470	0.259	0.949	0.144	0.208	0.440	0.003
GJR-GARCH GED	SPA_L	0.003	0.055	0.784	0.000	0.000	0.000	0.001
GJR-GARCH GED	SPA_C	0.003	0.055	0.861	0.000	0.000	0.000	0.001
GJR-GARCH GED	RC	0.490	0.270	0.981	0.159	0.221	0.452	0.001
SW-GARCH t2	SPA_L	0.144	0.206	0.786	0.178	0.117	0.098	0.004
SW-GARCH t2	SPA_C	0.144	0.206	0.861	0.186	0.117	0.105	0.004
SW-GARCH t2	RC	0.607	0.616	0.900	0.593	0.538	0.570	0.004
SW-GARCH t	SPA_L	0.127	0.194	0.707	0.139	0.083	0.046	0.003
SW-GARCH t	SPA_C	0.426	0.214	0.760	0.139	0.251	0.371	0.003
SW-GARCH t	RC	0.803	0.705	0.954	0.607	0.700	0.801	0.003
SW-GARCH GED	SPA_L	0.576	0.765	0.532	0.671	0.649	0.568	0.003
SW-GARCH GED	SPA_C	0.765	0.896	0.646	0.860	0.864	0.568	0.003
SW-GARCH GED	RC	0.992	1.000	0.717	1.000	1.000	0.999	0.003

Table 5.6: SPA and RC test results for one week (5 days) forecast horizon

Benchmark	Test	MSE	MAPE	QLIKE	R2LOG	MAE1	MAE2	HMSE
GARCH N	SPA_L	0.049	0.000	0.207	0.001	0.005	0.018	0.001
GARCH N	SPA_C	0.049	0.000	0.207	0.001	0.005	0.018	0.001
GARCH N	RC	0.535	0.345	0.469	0.337	0.364	0.497	0.001
GARCH t	SPA_L	0.054	0.000	0.180	0.001	0.004	0.014	0.001
GARCH t	SPA_C	0.054	0.000	0.180	0.001	0.004	0.014	0.001
GARCH t	RC	0.554	0.337	0.460	0.325	0.360	0.514	0.001
GARCH GED	SPA_L	0.065	0.001	0.191	0.002	0.005	0.022	0.002
GARCH GED	SPA_C	0.065	0.001	0.191	0.002	0.005	0.022	0.002
GARCH GED	RC	0.545	0.346	0.457	0.346	0.371	0.513	0.002
EGARCH N	SPA_L	0.190	0.010	0.829	0.030	0.109	0.117	0.000
EGARCH N	SPA_C	0.190	0.010	0.867	0.030	0.125	0.132	0.000
EGARCH N	RC	0.676	0.409	0.964	0.435	0.476	0.598	0.000
EGARCH t	SPA_L	0.165	0.002	0.722	0.010	0.059	0.065	0.000
EGARCH t	SPA_C	0.165	0.002	0.792	0.010	0.067	0.069	0.000
EGARCH t	RC	0.619	0.385	0.925	0.399	0.427	0.533	0.000
EGARCH GED	SPA_L	0.195	0.009	0.907	0.033	0.103	0.112	0.000
EGARCH GED	SPA_C	0.195	0.009	0.965	0.033	0.119	0.123	0.000
EGARCH GED	RC	0.676	0.414	0.988	0.431	0.463	0.580	0.000
GJR-GARCH N	SPA_L	0.016	0.000	0.575	0.000	0.001	0.000	0.000
GJR-GARCH N	SPA_C	0.016	0.000	0.612	0.000	0.001	0.000	0.000
GJR-GARCH N	RC	0.515	0.260	0.952	0.264	0.304	0.469	0.000
GJR-GARCH t	SPA_L	0.010	0.000	0.519	0.000	0.000	0.000	0.000
GJR-GARCH t	SPA_C	0.010	0.000	0.521	0.000	0.000	0.000	0.000
GJR-GARCH t	RC	0.500	0.241	0.890	0.245	0.292	0.470	0.000
GJR-GARCH GED	SPA_L	0.011	0.000	0.560	0.000	0.000	0.000	0.001
GJR-GARCH GED	SPA_C	0.011	0.000	0.586	0.000	0.000	0.000	0.001
GJR-GARCH GED	RC	0.496	0.257	0.937	0.270	0.306	0.474	0.001
SW-GARCH t2	SPA_L	0.181	0.162	0.586	0.197	0.164	0.146	0.001
SW-GARCH t2	SPA_C	0.181	0.162	0.614	0.197	0.217	0.157	0.001
SW-GARCH t2	RC	0.607	0.574	0.746	0.589	0.566	0.584	0.001
SW-GARCH t	SPA_L	0.243	0.116	0.512	0.165	0.119	0.089	0.002
SW-GARCH t	SPA_C	0.243	0.155	0.532	0.220	0.421	0.089	0.002
SW-GARCH t	RC	0.851	0.643	0.708	0.674	0.752	0.832	0.002
SW-GARCH GED	SPA_L	0.587	0.724	0.455	0.728	0.658	0.591	0.001
SW-GARCH GED	SPA_C	0.944	0.724	0.455	0.728	0.995	0.991	0.001
SW-GARCH GED	RC	0.975	1.000	0.622	0.999	0.998	0.996	0.001

Table 5.7: SPA and RC test results for two weeks (10 days) forecast horizon

Benchmark	Test	MSE	MAPE	QLIKE	R2LOG	MAE1	MAE2	HMSE
GARCH N	SPA_L	0.106	0.009	0.412	0.012	0.029	0.048	0.001
GARCH N	SPA_C	0.106	0.009	0.412	0.012	0.029	0.048	0.001
GARCH N	RC	0.574	0.398	0.673	0.398	0.406	0.540	0.001
GARCH t	SPA_L	0.106	0.007	0.359	0.013	0.027	0.053	0.001
GARCH t	SPA_C	0.106	0.007	0.359	0.013	0.027	0.053	0.001
GARCH t	RC	0.571	0.386	0.647	0.386	0.397	0.552	0.001
GARCH GED	SPA_L	0.117	0.008	0.382	0.014	0.034	0.048	0.001
GARCH GED	SPA_C	0.117	0.008	0.382	0.014	0.034	0.048	0.001
GARCH GED	RC	0.598	0.396	0.654	0.406	0.409	0.556	0.001
EGARCH N	SPA_L	0.771	0.488	0.811	0.742	0.263	0.253	0.000
EGARCH N	SPA_C	0.874	0.496	0.906	0.790	0.342	0.253	0.000
EGARCH N	RC	0.983	0.785	0.933	0.910	0.681	0.787	0.000
EGARCH t	SPA_L	0.470	0.286	0.834	0.404	0.187	0.145	0.000
EGARCH t	SPA_C	0.491	0.316	0.866	0.450	0.237	0.194	0.000
EGARCH t	RC	0.874	0.644	0.957	0.736	0.588	0.689	0.000
EGARCH GED	SPA_L	0.677	0.475	0.895	0.698	0.261	0.243	0.000
EGARCH GED	SPA_C	0.725	0.475	0.922	0.724	0.332	0.243	0.000
EGARCH GED	RC	0.965	0.753	0.959	0.922	0.670	0.771	0.000
GJR-GARCH N	SPA_L	0.022	0.000	0.541	0.001	0.001	0.001	0.000
GJR-GARCH N	SPA_C	0.022	0.000	0.572	0.001	0.001	0.001	0.000
GJR-GARCH N	RC	0.522	0.301	0.890	0.321	0.312	0.465	0.000
GJR-GARCH t	SPA_L	0.012	0.000	0.440	0.002	0.000	0.001	0.000
GJR-GARCH t	SPA_C	0.012	0.000	0.452	0.002	0.000	0.001	0.000
GJR-GARCH t	RC	0.512	0.290	0.815	0.306	0.291	0.453	0.000
GJR-GARCH GED	SPA_L	0.016	0.000	0.500	0.002	0.001	0.001	0.001
GJR-GARCH GED	SPA_C	0.016	0.000	0.527	0.002	0.001	0.001	0.001
GJR-GARCH GED	RC	0.502	0.297	0.868	0.318	0.289	0.452	0.001
SW-GARCH t2	SPA_L	0.189	0.232	0.597	0.266	0.175	0.131	0.000
SW-GARCH t2	SPA_C	0.191	0.232	0.639	0.266	0.187	0.131	0.000
SW-GARCH t2	RC	0.627	0.604	0.776	0.617	0.524	0.566	0.000
SW-GARCH t	SPA_L	0.378	0.250	0.610	0.283	0.265	0.241	0.001
SW-GARCH t	SPA_C	0.378	0.286	0.657	0.310	0.548	0.319	0.001
SW-GARCH t	RC	0.840	0.724	0.801	0.717	0.806	0.843	0.001
SW-GARCH GED	SPA_L	0.473	0.567	0.529	0.511	0.717	0.661	0.000
SW-GARCH GED	SPA_C	0.547	0.796	0.563	0.662	0.992	0.981	0.000
SW-GARCH GED	RC	0.867	0.971	0.668	0.891	0.993	0.988	0.000

Table 5.8: SPA and RC test results for one month (22 days) forecast horizon

Benchmark	Test	MSE	MAPE	QLIKE	R2LOG	MAE1	MAE2	HMSE
GARCH N	SPA_L	0.047	0.001	0.544	0.007	0.010	0.013	0.000
GARCH N	SPA_C	0.047	0.001	0.544	0.007	0.010	0.013	0.000
GARCH N	*RC*	0.540	0.322	0.771	0.354	0.355	0.475	0.000
GARCH t	SPA_L	0.048	0.002	0.510	0.006	0.009	0.012	0.000
GARCH t	SPA_C	0.048	0.002	0.510	0.006	0.009	0.012	0.000
GARCH t	*RC*	0.535	0.313	0.761	0.347	0.348	0.465	0.000
GARCH GED	SPA_L	0.052	0.001	0.524	0.007	0.009	0.016	0.001
GARCH GED	SPA_C	0.052	0.001	0.524	0.007	0.009	0.016	0.001
GARCH GED	*RC*	0.538	0.326	0.762	0.361	0.356	0.469	0.001
EGARCH N	SPA_L	0.603	0.510	0.738	0.543	0.621	0.592	0.000
EGARCH N	SPA_C	0.842	0.630	0.762	0.656	0.815	0.854	0.000
EGARCH N	*RC*	0.959	0.882	0.789	0.881	0.925	0.953	0.000
EGARCH t	SPA_L	0.346	0.083	0.861	0.168	0.352	0.245	0.000
EGARCH t	SPA_C	0.571	0.368	0.891	0.168	0.531	0.461	0.000
EGARCH t	*RC*	0.886	0.709	0.948	0.751	0.822	0.845	0.000
EGARCH GED	SPA_L	0.590	0.451	0.758	0.471	0.585	0.458	0.000
EGARCH GED	SPA_C	0.764	0.451	0.765	0.471	0.867	0.827	0.000
EGARCH GED	*RC*	0.948	0.907	0.804	0.905	0.918	0.920	0.000
GJR-GARCH N	SPA_L	0.015	0.000	0.597	0.004	0.005	0.004	0.000
GJR-GARCH N	SPA_C	0.015	0.000	0.632	0.004	0.005	0.004	0.000
GJR-GARCH N	*RC*	0.509	0.267	0.930	0.305	0.307	0.448	0.000
GJR-GARCH t	SPA_L	0.008	0.000	0.507	0.003	0.002	0.002	0.000
GJR-GARCH t	SPA_C	0.008	0.000	0.514	0.003	0.002	0.002	0.000
GJR-GARCH t	*RC*	0.495	0.238	0.858	0.284	0.284	0.430	0.000
GJR-GARCH GED	SPA_L	0.012	0.001	0.565	0.002	0.004	0.003	0.000
GJR-GARCH GED	SPA_C	0.012	0.001	0.589	0.002	0.004	0.003	0.000
GJR-GARCH GED	*RC*	0.505	0.257	0.907	0.296	0.291	0.435	0.000
SW-GARCH t2	SPA_L	0.140	0.046	0.758	0.079	0.151	0.118	0.000
SW-GARCH t2	SPA_C	0.142	0.046	0.776	0.079	0.151	0.119	0.000
SW-GARCH t2	*RC*	0.586	0.455	0.904	0.492	0.509	0.536	0.000
SW-GARCH t	SPA_L	0.142	0.022	0.869	0.061	0.162	0.094	0.000
SW-GARCH t	SPA_C	0.142	0.022	0.899	0.069	0.303	0.094	0.000
SW-GARCH t	*RC*	0.722	0.457	0.991	0.506	0.634	0.685	0.000
SW-GARCH GED	SPA_L	0.233	0.063	0.742	0.104	0.277	0.197	0.000
SW-GARCH GED	SPA_C	0.233	0.114	0.742	0.104	0.496	0.197	0.000
SW-GARCH GED	*RC*	0.770	0.564	0.839	0.599	0.745	0.758	0.000

CHAPTER 6

CONCLUSION

In this thesis, volatility of Turkish Stock Market for the period of 1997 to 2007 is examined. Daily ISE-100 index returns are used to proxy Turkish Stock Market. Differently from previous works, we adopt Markov Regime Switching GARCH models. These models allow volatility to have different dynamics according to unobserved regime variables.

The main purpose of this thesis is to find out whether SW-GARCH models are an improvement on the uni-regime GARCH models in terms of modelling and forecasting Turkish Stock Market volatility. We compare SW-GARCH models with standard 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 is considered for SW-GARCH models.

We first analyze in-sample performance of various volatility models to determine the best form of the volatility model over the period 1997 to 2006. SW-GARCH models under fat tailed distributions offer a better statistical fit to the data in terms of Log-likelihood, AIC and HQIC[1]. However, BIC suggest that uni-regime GARCH models with Student-t distribution provide more accurate description of the data. It is noteworthy that choice of Student-t assumption for standardized errors increase in-sample performance of volatility models. In addition, we compare in-sample volatility estimation performance of all models

[1]The squared standardized residuals from SW-GARCH model under normal distribution display highly significant autocorrelation. Therefore, we never account for this model while evaluating the results and comparing the competing models.

in terms of seven different statistical loss functions. The volatility models which show the lowest value for a loss function are considered to be most accurate, and then all loss functions favor the SW-GARCH models.

The use of SW-GARCH models also reveals the presence of two different volatility regimes in Turkish Stock Market. The unconditional variance in high volatility regime is found nearly four times greater than that in low volatility regimes. Moreover, the graphs of smoothed probabilities being in high volatility regime confirm the existence of two volatility regimes and clearly show that there is a structural break in Turkish Stock Market around middle of 2003. Before middle of 2003, volatility process is in high volatility regime, and then it switches the low volatility regime permanently.

Another improvement of the SW-GARCH models is that they reduce the high persistence in uni-regime GARCH models. The sum of ARCH and GARCH parameter estimates in the SW-GARCH models implies lower degree of volatility persistence than uni-regime GARCH models. These results are consistent with findings of Lamoureux and Lastrapes (1990) that regime shifts in volatility can lead to spuriously high levels of volatility persistence. On the other hand, under all distribution assumptions, estimated transition probability of each regime is extremely close to one and indicating that each regime is highly persistent.

Finally, we evaluate out-of-sample forecasting performance of SW-GARCH models compared to the uni-regime GARCH models for one day, one week, two weeks and one month forecast horizons over the period 2006 to 2007. Superior Predictive Ability (SPA) test of Hansen (2005) is applied by using seven statistical loss functions. This test is a powerful tool and widely used to compare jointly forecast performance of several models. The main results are: Firstly, for short horizons one day and one week, overall, uni-regime GARCH models are highly outperformed by SW-GARCH models. Also, among all models, the most accurate forecasts are obtained with SW-GARCH model under GED. Secondly, if we turn to two weeks of forecast horizon, we notice that standard GARCH and GJR-

GARCH models were significantly outperformed by SW-GARCH and EGARCH models. No superior models are identified between EGARCH and SW-GARCH models. Lastly, when forecasting performance of models at one month horizon are considered, EGARCH and SW-GARCH models perform better than other models. Moreover, we have some evidence that EGARCH models are better than SW-GARCH models in forecasting at one month horizon.

For further study, three or four volatility regimes setting can be considered rather than two-volatility regimes. Also, time varying transition probabilities can be adopted instead of constant ones. In addition, the performance of SW-GARCH models can be compared in terms of their ability to forecast Value at Risk (VaR) for long and short positions.

REFERENCES

[1] Abramson A. and I. Cohen. State smoothing in markov-switching time-frequency GARCH models. *IEEE Signal Processing Letters*, 13(6):377–380, 2006.

[2] Abramson A. and I. Cohen. On the stationarity of markov-switching GARCH processes. *Econometric Theory*, 23:485–500, 2007.

[3] Akaike H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions*, 19(6):716–723, 1974.

[4] Akgiray V., C. Aksu, and A. Odabası. The statistical evolution of prices on the Istanbul Stock Exchange. *The European Journal of Finance*, 10(6):510–525, 2004.

[5] Alexander C. *Market Models, A Guide to Financial Data Analysis.* John Wiley and Son, UK, 2001.

[6] Andersen T. G. and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905, 1998.

[7] Andersen T. G., T. Bollerslev, F. X. Diebold, and P. Labys. Modelling and forecasting realized volatility. *Econometrica*, 71:579–625, 2003.

[8] Ane T. and L. U. Rangau. Stock market dynamics in a regime-switching asymmetric power GARCH model. *International Review Of Financial Analysis*, 15:109–129, 2006.

[9] Areal N. M. P. C. and S. J. Taylor. The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets*, 22(7):627–648, 2002.

[10] Aybar C. B. and Z. A. Yavan. İMKB'de oynaklık. *İMKB Dergisi*, 2(6):35–47, 1998.

[11] Balaban E. Day of the week effects: Evidence from an emerging stock market. *Applied Economics Letters*, 2:139–142, 1995.

[12] Balaban E. Forecasting stock market volatility : Evidence from Turkey. *The ISE Finance Award Series*, 1, 1999.

[13] Barucci E and Reno R. On measuring volatility of diffusion processes with high frequency data. *Economics Letters*, 74:371–378, 2002.

[14] Bautista C. C. Stock market volatility in the Philippines. *Applied Economics Letters*, 10(5):315–318, 2003.

[15] Bera A. K. and M. L. Higgins. ARCH models: Properties, estimation and testing. *Journal of Economic Surveys*, 7(4):305–365, 1993.

[16] Beran R. and G. R. Ducharme. *Asymptotic Theory for Bootstrap Methods in Statistics.* Les Publications CRM, 1991.

[17] Berument H., K. Metin, and G. Muradoglu. An empirical investigation stock returns and determinants of risk in an emerging market: Istanbul Stock Exchange. 1998. Computations in Economics, Finance and Engineering:Economics Systems Preprints, IFAC:Society for Computational Economics Meeting, UK.

[18] Bildik R. and S. Elekdağ. Effects of price limits on volatility: Evidence from the Istanbul Stock Exchange. *Emerging Markets Finance and Trade*, 40(1):5–34, 2004.

[19] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[20] Bollerslev T. A Conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69(3):542–547, 1987.

[21] Bollerslev T., R. Y. Chou, and K. P. Kroner. ARCH modelling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52:5–59, 1992.

[22] Bollerslev T. and R. F. Engle. Common persistence in conditional variances. *Econometrica*, 61(1):167–186, 1993.

[23] Bollerslev T., R. F. Engle, and D. B. Nelson. ARCH models. *Handbook of Econometrics*, 4:2959–3038, 1994. in: Engle, R.F. and D.L. McFadden.

[24] Bollerslev T. and E. Ghysels. Periodic autoregressive conditional heteroscedasticity. *Journal of Business and Economic Statistics*, 14(2):139–151, 1996.

[25] Brooks C. and G. Persand. Volatility forecasting for risk management. *Journal of Forecasting*, 22:1–22, 2003.

[26] Brunetti C., R. S. Mariano, C. Scotti, and A. H. H. Tan. Markov switching GARCH models of currency turmoil in Southeast Asia. *Board Of Governors Of The Federal Reserve System International Finance Discussion Papers*, (889), 2007.

[27] Buhlmann P. Bootstraps for time series. *Statistical Science*, 17(1):52–72, 2002.

[28] Cai J. A markov model of unconditional variance in ARCH. *Journal of Business and Economic Statistics*, 12(3):309–316, 1994.

[29] Chan L., R. J. Elliott, and T. K. Siu. Option pricing for GARCH models with markov switching. *International Journal Of Theoretical And Applied Finance*, 9(6):825–841, 2006.

[30] Daouk H. and J. Q. Guo. Switching asymmetric GARCH and options on a volatility index. *Journal of Futures Markets*, 24(3):251–282, 2004.

[31] Dickey D. A. and W. A. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica,*, 49(4):1057–1072, 1981.

[32] Diebold F. X. and J. A. Lopez. Forecast evaluation and combination. *Statistical Methods in Finance*, 14:241–268, 1996. In Handbook of Statistics.

[33] Diebold F. X and Mariano R.S. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–263, 1995.

[34] Ding Z., R. F. Engle, and C. W. J. Granger. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

[35] Ding Z. and C. W. J. Granger. Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73:185–215, 1996.

[36] Dueker M. J. Markov switching in GARCH processes and mean-reverting stock market volatility. *Journal of Business and Economic Statistics*, 15(1):26–34, 1997.

[37] Efron B. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[38] Efron B. *The Jackknife, the Bootstrap and Other Resampling Plans.* Society for Industrial and Applied Mathematics, Philadelphia, 1982.

[39] Efron B. and R. Tibshirani. *An Introduction to the Bootstrap.* Chapman and Hall, New York, 1993.

[40] Engle R. F. Autoregressive conditional heteroscedasticity with estimates of U.K. inflation. *Econometrica*, 50(4):987–1008, 1982.

[41] Engle R. F. and T. Bollerslev. Modelling the persistence of conditional variances. *Econometric Reviews*, 5:1–50, 1986.

[42] Engle R. F. and V. K. Ng. Measuring and testing the impact of news on volatility. *Journal of Finance*, 48:1749–78, 1993.

[43] Engle R. F. and A. J. Patton. What good is a volatility model? *Quantitative Finance*, 1(2):237–245, 2001.

[44] Fama E. F. Mandlebrot and stable paretian hypothesis. *American Economic Review*, (65):269–282, 1963.

[45] Fama E. F. The behavior of stock market prices. *Journal of Business*, 38:34–105, 1965.

[46] Fong W. M. Volatility persistence and switching ARCH in Japanese stock returns. *Financial Engineering and the Japanese Markets*, 4:37–57, 1997.

[47] Fong W. M. The dynamics of DM/£ exchange rate volatility: A SWARCH analysis. *International Journal of Finance and Economics*, 3:59–71, 1998.

[48] Fong W. M. and K. H See. Modelling the conditional volatility of commodity index futures as a regime switching process. *Journal Of Applied Econometrics*, 16:133–163, 2001.

[49] Fong W. M. and K. H. See. A markov switching model of the conditional volatility of crude oil futures prices. *Energy Economics*, 24:71–95, 2002.

[50] Fornari F. and A. Mele. Sign and volatility-switching ARCH models: Theory and application to international stock markets. *Journal of Applied Econometrics*, 12:49–67, 1997.

[51] Francq C. and J. M. Zakoian. The L2-structures of standard and switching-regime GARCH models. *Stochastic Processes And Their Applications*, 115:1557–1582, 2005.

[52] Franses P. H. and D. Van Dijk. *Non-Linear Time Series Models in Empirical Finance.* Cambridge University Press, Cambridge, 2000.

[53] Glosten L. R., R. Jagannathan, and D. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5):1779–1801, 1993.

[54] Granger C. W. J. and S. H. Poon. Forecasting financial market volatility: A review. *Journal of Economic Literature*, 41(2):478–539, 2003.

[55] Gray S. Modelling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1):27–62, 1994.

[56] Gündüz L. and A. Hatemí. Stock price and volume relation in emerging markets. *Emerging Markets Finance and Trade*, 41:29–44, 2005.

[57] Güner N. and Z. Önder. Information and volatility: Evidence from an emerging market. *Emerging Markets Finance and Trade*, 38:26–46, 2002.

[58] Haas M., S. Mittnik, and M. S. Paollela. A new approach in markov-switching GARCH models. *Journal Of Financial Econometrics*, 2(4):493–530, 2004.

[59] Hall P. *The Bootstrap and Edgeworth Expansion.* Springer Verlag, New York, 1992.

[60] Hamilton J. D. A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica*, 57:357–84, 1989.

[61] Hamilton J. D. Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45:39–70, 1990.

[62] Hamilton J. D. *Time Series Analysis.* Princeton University Press, New Jersey, US, 1994.

[63] Hamilton J. D. and R. Susmel. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1-2):307–33, 1994.

[64] Hannan E. J. and B. G. Quinn. The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, 41(2):190–195, 1979.

[65] Hansen P. R. A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23:365–380, 2005.

[66] Hansen P. R. and A. Lunde. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20:873–889, 2005.

[67] Hardle W., J. Horowitz, and J. P. Kreiss. Bootstrap methods for time series. *International Statistical Review*, 71(2):435–459, 2003.

[68] Harris R. D. F and Küçüközmen C. The empirical distribution of stock returns: Evidence from an emerging European market. *Applied Economics Letters*, 8:367–371, 2001.

[69] Harris R. D. F and Küçüközmen C. Linear and nonlinear dependence in Turkish equity returns and its consequences for financial risk management. *European Journal of Operational Research*, 134:481–92, 2001.

[70] Harvey D., S. Leybourne, and P. Newbold. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291, 1997.

[71] Hentschel L. F. All in the family: Nesting linear and nonlinear GARCH models. *Journal of Financial Economics*, 39:139–64, 1995.

[72] Hol E., B. Jungbacker, and S. J. Koopman. Forecasting daily variability of the S&P100 stock index using historical, realised and implied volatility measurements. *Tinbergen Institute Discussion Paper*, 2004.

[73] Horowitz J. The bootstrap in econometrics. *Statistical Science*, 18(2):211–218, 2003.

[74] Hsu P. H. and C. M. Kuan. Reexamining the profitability of technical analysis with data snooping checks. *Journal of Financial Econometrics*, 3(4):606–628, 2005.

[75] Kaufmann S. and M. Scheicher. A switching ARCH model for the German DAX index. *Studies in Nonlinear Dynamics and Econometrics*, 10(4), 2006.

[76] Kaufmann S. and S. F. Schnatter. Bayesian analysis of switching ARCH models. *Journal of Time Series Analysis*, 23(4):425–458, 2002.

[77] Kılıc R. On the long memory properties of emerging capital markets: Evidence from Istanbul Stock Exchange. *Applied Financial Economics*, 14:915–22, 2004.

[78] Kim C. J. Dynamics linear models with markov switching. *Journal of Econometrics*, 60:1–22, 1994.

[79] Kim C. J. and C. R. Nelson. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* MIT Press, Cambridge, MA, 1999.

[80] Klaassen F. Improving GARCH volatility forecasts. *Empirical Economics*, 27(2):363–394, 2002.

[81] Kunsch H. R. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3):1217–1241, 1989.

[82] Lahiri S. N. Second order optimality of stationary bootstrap. *Statistics Probability Letters*, 11:335–341, 1991.

[83] Lamoureux C. G. and W. D. Lastrapes. Persistence in variance, structural change and the GARCH model. *Journal of Business Economics Statistics*, 8:225–234, 1990.

[84] Lee H. T. and J. K. Yoder. A Bivariate markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. *Applied Economics*, 39(10):1253–1265, 2007.

[85] Leon Li M. Y. Volatility states and international diversification of international stock markets. *Applied Economics*, 39(14):1867–1876, 2007.

[86] Leon Li M. Y and H. W William Lin. 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, 2003.

[87] Ljung G. M. and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrica*, 65:297–303, 1978.

[88] Lopez J. A. Evaluation of predictive accuracy of volatility models. *Journal of Forecasting*, 20(1):87–109, 2001.

[89] Mandelbrot B. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.

[90] Marcucci J. Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, 9(4), 2005.

[91] Martens M. Measuring and forecasting S&P500 index futures volatility using high frequency data. *Journal of Futures Markets*, 22(6):497–518, 2002.

[92] Mazıbaş M. İMKB piyasalarındaki volatilitenin modellenmesi ve öngörülmesi: Asimetrik GARCH modelleri ile bir uygulama. 2005. VII. Ulusal Ekonometri ve İstatistik Sempozyumu, 26-27 Mayıs 2005.

[93] McKenzie M. and H. Mitchell. GARCH model selection criteria. *Quantitative Finace*, 3:262–284, 2003.

[94] Metin K. and G. Muradoglu. Efficiency of the Turkish stock market with respect to monetary variables: A Cointegration analysis. *European Journal of Operational Research*, 90:566–576, 1996.

[95] Mincer J. and V. Zarnowitz. *The Evaluation of Economic Forecasts.* NBER, New York, in J.Mincer edition, 1969.

[96] Nelson D. B. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991.

[97] Norden S. V. and H. Schaller. Regime switching in stock market returns. *Applied Financial Economics*, 7:177–191, 1997.

[98] Okay N. Asymmetric volatility dynamics: Evidence from Istanbul Stock Exchange. 1998. Business and Economics for the 21st Century-Volume 2, Selected Papers: 1998 Business and Economics Society International Conference.

[99] Pagan A. R. The econometrics of financial markets. *Journal of Empirical Finance*, 3:15–102, 1996.

[100] Patton A. J. Volatility forecast comparison using imperfect volatility proxies. *University of Technology Research Paper Series*, (175), 2005.

[101] Patton A. J. and K. Sheppard. Evaluating volatility and correlation forecasts. *Oxford Financial Research Centre Working Papers Series*, 2007.

[102] Perez G. Q. and A. Timmermann. Firm size and cyclical variations in stock returns. *The Journal of Finance*, 55(3):1229–1262, 2000.

[103] Pesaran M. H. and A. Timmermann. A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics*, 10(4):461–465, 1992.

[104] Politis D. N. The impact of bootstrap methods on time series analysis. *Statistical Science*, 18(2):219–230, 2003.

[105] Politis D. N. and J. P. Romano. The stationary bootstrap. *Journal of The American Statistical Association*, 89(428):1303–1313, 1994.

[106] Poon S. H. *A Practical Guide to Forecasting Financial Market Volatility.* John Wiley and Son, UK, 2005.

[107] Salman F. Risk-Return-Volume relationship in an emerging stock market. *Applied Economics Letters*, 9:549–552, 2002.

[108] Schwarz G. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.

[109] Sentana E. Quadratic arch models. *Review of Economic Studies*, 62:639–61, 1995.

[110] Susmel R. Switching volatility in Latin American emerging markets. *Emerging Markets Quarterly*, 2, 1998a.

[111] Susmel R. Switching volatility in private international equity markets. *International Journal of Finance and Economics*, 5(4):265–283, 2000.

[112] Tsay R. S. *Analysis of Financial Time Series.* John Wiley and Son, Canada, 2002.

[113] White H. A reality check for data snooping. *Econometrica*, 68(5):1097–1126, 2000.

[114] Yılmaz M. K. Hisse senedi oynaklığı ve fiyat oynaklığının vade yapısı: Türkiye için genel bir değerlendirme. *İMKB Dergisi*, 1(3):25–45, 1997.

[115] Yu S. W. Index futures trading and spot price volatility. *Applied Economics Letters*, 8:183–186, 2001.

[116] Yüksel A. The performance of the Istanbul Stock Exchange during the Russian crisis emerging markets. *Emerging Markets Finance and Trade*, 38:78–99, 2002.

[117] Zakoian J. M. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955, 1994.