

UNIVERSITY OF VAASA
FACULTY OF BUSINESS STUDIES
DEPARTMENT OF ACCOUNTING AND FINANCE

Mo Zhang

**VOLATILITY FORECASTING COMPARISON BETWEEN
IMPLIED VOLATILITY AND MODEL BASED FORECASTS**

Master's Thesis in
Accounting and Finance

Finance

VAASA 2010

TABLE OF CONTENTS	Page
1. INTRODUCTION	7
1.1. Purpose of the study	8
1.2. Previous studies	9
1.2.1. Stock market indices	10
1.2.2. Individual stocks	18
1.2.3. Other asset	20
1.3. Structure of the study	22
2. VOLATILITY	23
2.1. Definition of volatility	23
2.1.1. Volatility measurements	23
2.1.2. Misperceptions of volatility	27
2.2. Characteristics of financial market volatility	28
2.2.1. Fat tails and a high peak	28
2.2.2. Volatility clustering	29
2.2.3. Mean-reversion	30
2.2.4. Long memory effect	31
2.2.5. Volatility asymmetry	32
2.2.6. Cross-border spillovers	33
3. VOLATILITY FORECASTING	35
3.1. Model based forecast	35
3.1.1. Historical volatility models	36
3.1.2. ARCH family	39
3.1.3. Stochastic volatility models	47
3.2. Implied volatility forecasting	48
3.2.1. Volatility implied by Black-Scholes option pricing model	49
3.2.2. Features of implied volatility	51
3.2.3. Drawbacks of volatility implied by B-S model	54

4. EVALUATION CRITERIA	57
4.1. Loss functions	58
4.1.1. Symmetric error measures	58
4.1.2. Asymmetric error measures	61
4.2. Regression based evaluation	62
5. DATA AND METHODOLOGY	64
5.1. Data	64
5.2. Methodology	69
5.2.1. Computing actual volatility	69
5.2.2. Models in the competition	70
5.2.3. The evaluation of predictabilities	73
6. RESULTS	75
7. SUMMARY AND LIMITATIONS	79
REFERENCES	81
APPENDIX 1	97
APPENDIX 2	104

UNIVERSITY OF VAASA**Faculty of Business Studies**

Author:	Mo Zhang
Topic of the Thesis	Volatility Forecasting Comparison Between Implied Volatility and Model Based Forecasts
Name of the Supervisor:	Professor Sami Vähämaa
Degree:	Master of Science in Economics and Business Administration
Department:	Department of Accounting and Finance
Major Subject:	Accounting and Finance
Line:	Finance
Year of Entering the University:	2008
Year of Completing the Thesis:	2010

Pages: 106**ABSTRACT**

The purpose of this study is to compare the forecasting performance between implied volatility and model based forecasts (MBFs) in the U.S. stock market. During recent thirty years, volatility forecasting has always been a hot and important issue in both practical and academic areas, but there is no final conclusion on the best forecasting method. This study aims to use the long enough and updated data from Jan 1990 to Dec 2009 to reexamine this significant topic. Moreover, by reviewing ample literatures, the author found that the efficiency of option markets developed by leaps and bounds after severe financial crisis. Therefore, this study also throws a light on testing whether the efficiency of the U.S. option market has been improved since 2007 financial crisis burst.

The empirical study consists of monthly volatility forecasting and the predictive power comparison. Model based forecasts are given by several econometrical models including: random walk, *Riskmetrics*TM, GARCH (1, 1) and GJR (1, 1) by using the daily closing prices of S&P 500 index. VIX index implied by options on S&P 500 is used as the representative of the implied volatility forecast. Forecasting performance is compared by three error measures-mean square error, mean absolute percentage error, QLIKE, and regression based evaluation.

Two hypotheses are tested here: firstly, implied volatility performs better on the volatility forecasting than MBFs do; secondly, the efficiency of option market improved after 2007 financial crisis. The empirical evidence rejects the first hypothesis and finds that GJR (1, 1) model dominates other methods as the best forecast. Implied volatility is even inferior to GARCH (1, 1) model. Meanwhile, more sophisticated models are superior to simple historical models on monthly forecasting. The second hypothesis is strongly supported. The U.S. option market realized an obvious improvement after 2007 financial crisis.

KEYWORDS: volatility forecasting, implied volatility, model based forecasts

1. INTRODUCTION

Volatility refers to the uncertainty of a variable, which is closely related to risk. It is often expressed as the sample standard deviation or variance. As the most basic statistical risk measure, it is widely used in both practical and academic fields. In daily practice, volatility is generated for risk management on immense individual financial instruments as well as portfolios. Investors also take future volatility into account in decision-making and portfolio creation. Not only traders, investors and risk managers rely on future volatility estimate, but monetary policy makers also need volatility prediction as the important reference to achieve appropriate policy establishment. In research area, volatility forecasting is indispensable to derive option prices. And it is also needed as important input to get hedge ratios for derivative portfolios as well as for value-at-risk model.

Because of its wide and important applications, volatility forecasting has been a hot issue over the last thirty years. Most studies explore this topic in two methods: model based forecasts (hereafter referred to as MBFs) based on historical information and implied volatility derived from option prices. Theoretically implied volatility as the market expectation of volatility should be the best prediction of the future volatility and reflect all of available information in the markets, incorporating the historical information. Many studies support that implied volatility is better than MBFs (Poon & Granger 2003). Although it seems that implied volatility is the predominate method on volatility forecasting, it cannot be neglected that it works well only on specific time horizons for a limited set of assets (Ederington & Guan 2005: 466). Those basic assets trading in the tiny market, which don't have relative derivatives, cannot make benefits from this way.

In contrast, MBFs are preferred as the more flexible method, which can be applied on any asset to forecast volatility in any time-horizon; however, it still has some limitations. One is the trade-off between model complexity and forecasting error. More

sophisticated models capture the volatility structure more accurately in the in-sample estimation, but may also induce additional out-of-sample forecasting error due to additional parameters. Besides that, there is another trade-off which is the proper weighting of recent versus older observations. If just keep eyes on the recent observations, the results reflect up-to-date information, but omit some certain patterns exists in the past structure. On the other hand, if the historical information is more relied on, some extremes and noise may be averaged, while some updated changes may be overlooked. It seems that there is no absolutely superior model among these numerous methods. The closer examine is demanded on this significant issue.

1.1. Purpose of the study

Concerning on the predictabilities of different methods on volatility, the researchers have different opinions. As the market expectation of future volatility, if the market is efficient, implied volatility should definitely be the best forecast for realized volatility. There is the heated debate on the issue whether financial market is informational efficient. Implied volatility may not be an unbiased and efficient forecast. However, this study just focuses on the predictabilities of different methods. One biased forecasting method can still have powerful predicting ability. As many researchers (Lamoureux & Lastrapes 1993; Vasilellis & Meade 1996; Christensen & Prabhala 1998; Blair, Poon & Taylor 2001) state, although implied volatility is not a perfect forecast, it is still superior to MBFs. However, Becker et al. (2006, 2007, 2008, 2009) using more recent data and distinguishing evaluation criteria give a serial of sound controversial evidences. It is worthy of checking this issue by using up-to-date data. Furthermore, the author finds an interesting phenomenon from the previous studies, after 1987 stock crash and 1995 Japanese financial crisis, the efficiency of option markets realized the quality leap (Christensen & Prabhala 1998: 127., Corrado & Miller 2005:366), called “awakening” of the option markets (Poon & Granger 2003:500). It demonstrated that investors in the markets would improve their ability on risk management and forecasting after the severe financial crisis. It is well known that the U.S. subprime mortgage crisis from 2007 brought disasters throughout the global financial system and the worldwide

economy. Therefore, now it is an appropriate time to reexamine the performance of implied volatility on volatility prediction and whether the efficiency of option market develops after the 2007 financial crisis.

The main purpose of this study is to compare the forecasting performance between implied volatility and MBFs in the U.S. stock market. This study uses the CBOE Market Volatility Index (VIX) underlying on the S&P 500 options as the representative of implied volatility. Random walk, *RiskMetricsTM*, General Autoregressive Conditional Heteroskedasticity (GARCH) (1, 1) and GJR (1, 1) (asymmetric GARCH model proposed by Glosten, Jagannathan & Runkle 1993) models are used to compare with the implied volatility. The daily S&P 500 index returns are used to generate the volatility process. In this study, the sample period is from 2 January 1990 to 31 December 2009. The forecasting horizon is 30 calendar days.

Based on the discussion above, two main hypotheses are outlined below:

- (1) Implied volatility performs better on the volatility forecasting than MBFs do.
- (2) The efficiency of option market improved after 2007 financial crisis.

1.2. Previous studies

In recent decades, there are dozens of researches on volatility forecasting. The main competition is between volatility implied by option prices and MBFs by using historical information. The advocators for market efficiency believe that implied volatility is absolutely the most accurate method on volatility forecasting. The opponents declare that econometric models should be better. Compared with MBFs, the volatility predictions drawn from the implied volatility are more complicated. The test on the forecasting power of implied volatility is actually a joint test of option market efficiency and the correction of the option pricing model (Poon & Granger 2003:499). Due to different trading frictions across assets, some types of options are easier to trade and hedge than others. It is reasonable to anticipate different levels of efficiency and distinct

forecasting competences for options written on various assets. Therefore, the author will review the previous studies based on different asset classifications.

1.2.1. Stock market indices

To reduce the measurement errors and prevent low liquidity problems, many researchers focus on volatility implied by stock market index options. Day and Lewis (1992) examine the incremental information content of volatilities implied from call option relative to GARCH and Exponential GARCH (EGARCH) model. The dividend-adjusted version of the Black-Scholes model is used to estimate the implied volatilities on the S&P100 index. Different from the previous cross-sectional studies, they study this issue from the time-series setting. The sample period started from 11 March 1983 through 31 December 1989. Their in-the-sample results state that implied volatilities, GARCH and EGARCH models all reflect incremental information about the weekly future volatility. But neither of them can completely characterize within-sample conditional stock market volatility (1992:350). On one more step, they compare the relative predictive power of implied volatility forecasts to ex post volatility. The results indicate that the short-run market volatility is difficult to predict (1992:349). Since they predict on one week horizon and that does not necessarily correspond with the life of the option, it easily causes the measurement errors which attribute to some combinations of specification error, maturity mismatch, and random estimation error (1992:343).

To solve this problem, Canina and Figlewski (1993) test the predictability of implied volatility and historical returns on the volatility over the remaining time of option contract. They use the daily closing prices for all call options on the S&P 100 index (OEX) from March 15, 1983 through March 28, 1987, and then derive the implied volatility from a binomial model that adjusts for dividends and early exercise. They believe that there exist some systematic factors which drive investors to price particular options high or low relative to others (1993:667). Therefore, it is inappropriate to simply form a weighted-average implied standard deviation (WISD) using multiple options with different expirations or measured on different dates and look them as if

they are just multiple noisy observations on the same parameter (1993:667). Unlike early studies, they separate option data into 32 groups according to maturity (one-, two-, three-, and four-month maturities) and intrinsic value, then make the predicting horizons match option maturity. They report that implied volatility has no statistically significant correlation with real volatility at all. Even it does not contain the information which is indicated by the available historical volatility forecast. These findings can be interpreted as OEX almost has no predictive power of future volatility. It is somewhat unexpected since OEX options at that time were the most active trading options in the world. Canina and Figlewski attribute these results to the inability of option model to capture the net effect of many factors which influence option supply and demand on the market pricing process (1993:677).

Christensen and Prabhala (1998) criticize that the studies of both Day and Lewis (1992) and Lamoureux and Lastrapes (1993) suffer from the overlapping sample problem as well as maturity mismatch problem (1998:126). In addition, they also doubt the surprising conclusions generated by Canina and Figlewski (1993) probably due to no incorporation of the data after 1987 crash and the adoption of overlapping sample (1998:126-127). They reexamine the relation between implied volatility and future volatility for the OEX option market. In contrast with previous studies, their study differs in two ways. Firstly, they use the longer sample period (from November 1983 through May 1995) than previous studies in order to increase the statistical power and allow for evolution in the efficiency of the market for OEX options since their introduction in 1983 (1998:127). Secondly, they utilize the lower (monthly) frequency data and produce the non-overlapping sample. This can avoid overlapping and maturity mismatch problem and guarantee more reliable regression estimates. They find that implied volatility in at-the-money one-month OEX call options is an unbiased and efficient forecast in out-of-sample prediction after the 1987 stock market crash. Their study also throws a light on the effect of 1987 crash on the volatility forecasting. They document that “implied volatility is more biased before the crash than after” (1998:127). Since Christensen and Prabhala (1998) just focus on at-the-money call options, Christensen and Hansen (2002) extend the Christensen and Prabhala’s study (1998) by testing the robustness of the unbiasedness and the efficiency of implied volatility for

both call and put OEX options. It is the first time that the information content of volatility implied in put options is checked (2002:189). They choose at-the-money, in-the-money and out-of-the-money options in five day period between 1993 and 1997 to construct the implied volatility. Christensen and Prabhala's finding in 1998 is confirmed. Furthermore, they (2002:204) prove that although call implied volatility is a better volatility forecast than put implied volatility, put option prices also contain valuable volatility information.

In 1993 the CBOE Market Volatility Index (VIX) was firstly introduced as a new measure of market volatility by Professor Robert E. Whaley of Duke University. VIX was constructed from the implied volatilities of eight OEX options based on the Black-Scholes (1973)/Merton (1973) option valuation framework. The cash-dividend adjusted binomial method was used to calculate the component implied volatilities (Whaley 1993). Fleming, Osdiel and Whaley (1995) investigate the statistical properties of VIX and test its predictive power for one month interval from 1986 to 1992. They declare that VIX behaves well with little evidence of seasonality, which also has a strong negative and asymmetric association with contemporaneous stock market return. Regarding the predicting performance, VIX dominates the historical volatility as a high quality forecast of future stock market volatility. Its upward bias is constant or estimable. Relative to previous measurement of implied volatility, VIX does not incur the usual time-variation resulting from moneyness and time-to-expiration effects. The relatively constant VIX forecast bias can be sufficiently corrected by a naïve adjustment based on a rolling average of past forecast errors. Again, Fleming (1998) use a volatility measure similar to VIX to show that implied volatility outperforms historical information.

After doing many efforts on the implied volatility constructions, the researchers consider predictability comparison issue from another angle. Because the latent volatility cannot be observed, the common way is to use the volatility proxy to compare with volatility forecasts. The most popular proxy is daily squared return. Does this imperfect volatility proxy cause the confused results? And is there any more accurate proxy than those existed? Andersen and Bollerslev (1998) firstly detected that high-

frequency data contained more information but less noises. It can be used to measure latent volatility process and generate volatility estimates. Blair, Poon and Taylor (2001) compare the information content of implied volatility and ARCH models with both 5-min interval intraday return and low-frequency data, in the context of forecasting volatility over horizon from 1 to 20 days from 1987 to 1999. In agreement with previous evidences provided by Christensen and Prabhala (1998) and Fleming (1998), they report that in-sample forecasts from VIX index provide nearly all relevant information compared with that from ARCH models using low-frequency data. Moreover, to extent the historical information set, they include high-frequency (5-min) returns and show that high-frequency data is highly informative, whereas implied volatility is even more informative than 5-min return. The similar result is also observed in out-of-sample period. VIX generates more accurate forecasts than either low- or high-frequency historical index return through all of predicting horizon. Just a combination of VIX forecasts and index return forecasts illustrates that probably some incremental forecasting information in daily returns existed when forecasting 1-day ahead. However, for 20-day forecasting horizon, VIX estimates subsume all of relevant information.

So far, researchers have obtained fruitful achievement; however, they are not easily satisfied with what they have had. They are curious to know the correction of Black-Scholes model and its effect on implied volatility. Is Black-Scholes model sufficient to capture the option pricing process? If not, what does this misspecification impact on information content and predictability test on implied volatility? And what can be done to eliminate the influence? Researchers notice that although the information content and predictive power test on VIX index are free from the moneyness and time-to-expiration effects as well as dividend and early-exercise problem, these tests are still the joint tests on market efficiency and the correction of Black-Scholes model. Because VIX on S&P 100 is based on the Black-Scholes model, these studies are still subject to model misspecification errors. To address this question, Jiang and Tian (2005) conduct direct tests of the informational efficiency of the option market using an alternative implied volatility measure that is independent of option pricing models. This measure is derived by Britten-Jones and Neuberger (2000) under diffusion assumptions. They extend

Britten-Jones and Neuberger (2000) by taking random jumps into account. They demonstrate empirical tests using S&P 500 index (SPX) options traded on the CBOE and minimize measurement errors by using tick-by-tick data, commonly used data filters, non-overlapping samples as well as realized volatility estimated from high-frequency index return. They report the model-free implied volatility reflected all information subsumed in both the B-S implied volatility and historical data estimate and is a more efficient forecast for future realized volatility. Their results prove informational efficiency of the option market.

Since December 2003, CBOE replace an earlier version of implied volatility index based on the Black-Scholes model with the new version VIX underlying on the S&P 500 options by the model-free method. Becker, Clemants and White (2006) examine whether S&P 500 implied volatility index is in fact efficient with respect to common available conditioning information over the period 2 January 1990 and 17 October 2003. This study provides a supplementary analysis of forecasting efficiency to Jiang and Tian (2005), as a much wider set of conditioning information is utilized. Moreover, they also take the possible volatility risk premium into account as discussed firstly by Chernov (2001). Their results are in line with the previous studies reporting a significant positive correlation between the VIX index and future volatility. Unlike Jiang and Tian's finding (2005), they illustrate that VIX is not an efficient volatility forecast. In that sense, other available information can improve on the VIX forecasts.

Furthermore, Becker, Clemants and White (2007) look into the informational content of implied volatility beyond that available from MBFs with the same data series of their study in 2006. They adopt new approach which is different from the traditional forecast encompassing approach. They consider the chosen set of MBFs as a comprehensive set of forecasts, while the previous method compared the implied volatility to the individual MBF. They argue that the apparent superiority of implied volatility may be attributed to the shortcomings of individual MBF used in the comparisons. Therefore, they decompose the implied volatility into two parts: VIX^{MBF} , information in VIX that is captured by MBF, and VIX^* , information in VIX not captured by MBF. Then they conduct the orthogonality test between VIX^* and realized volatility to see whether VIX

contained additional information that could not be obtained from the totality of information reflected in MBF. Their empirical results indicate that VIX does not contain any incremental information beyond that captured in a wide array of MBFs (2007:2548). However, no forecast comparison is undertaken in Becker et al. (2007), and they merely speculate that the VIX may be viewed as a combination of MBFs.

To address this question, Becker and Clements (2008) examine the forecast performance of VIX, compared to a general set of MBFs and combination forecasts on the basis of both implied volatility and MBFs. To make the results comparable with Becker et al. (2007), the same data series is considered. The practical evidence shows that when the best MBFs are combined, they are superior to both individual MBF and VIX estimates. The most precise S&P 500 volatility forecast is generated from a combination of short and long memory models of realized volatility. This study claims that VIX not only contain no additional information, it cannot also efficiently reflect the information incorporated in MBFs. VIX cannot be treated as the best combination of all MBFs.

So far, implied volatility is being discussed as risk neutral forecast of spot volatility, whereas the time-series models are estimated by the risk-adjusted or real world data of the underlying assets. Since the forecasting target is the real world, it seems that implied volatility has an inherently disadvantage. Becker, Clements and Coleman-Fenn (2009) specifically investigate the effect of volatility risk premium on the predicting performance of implied volatility. They adopt the method proposed by Bollerslev, Gibson and Zhou (2008) to transform the unadjusted risk-neutral implied volatility into risk-adjusted implied volatility, and then test whether risk-adjusted forecasts are statistically superior to the unadjusted risk-neutral forecasts as well as a wide range of MBFs. Their research period is from 2 January 1990 to 31 December 2008. The empirical evidence says that risk-adjusted implied volatility provides the better prediction rather than the risk-neutral implied volatility. However, they also find implied volatility with adjusted risk premium has the equal prediction accuracy to the MBFs (2009:17).

In previous encompassing regressions for estimating the information content of implied volatility, the historical volatility used in the model is often a rather crude measure (lagged realized volatility). Some researchers wonder to know whether more sophisticated measures of historical volatility would improve the precision of regression and affect the conclusion. Corrado and Miller (2005:348) add the several instrument variables in the information content test of implied volatility in order to deal with the econometric error problem in historical volatility as well as implied volatility. They use lagged realized volatility, lagged VPA (referred to formula (3)) proposed by Parkinson (1980) to capture the information in the high-low price range and lagged VRS proposed by Rogers and Satchel (1991) to convey the information in open-close price differences. These three instrumental variables are applied together to represent the historical information set. Lagged VIX, lagged VXO (volatility index on S&P 100) and lagged VXN (volatility index on NASDAQ 100) are employed to reflect the whole information set of implied volatility. Finally, they find that the CBOE volatility indexes on S&P 100 and S&P500 options appear to contain significant forecast errors in the pre-1995 period, while from 1995-2003 there is no indication of significant forecast error variances for any of CBOE volatility indexes (2005:367). They conclude that volatility indexes corresponding to S&P 100 and S&P 500 are biased but more efficient in terms of mean squared forecast errors rather than historical volatility.

Different from Corrado and Miller (2005), Giot and Laurent (2006) take the price jump effect into account and decompose the historical volatility into continuous component and jump component. These components are also arranged to reflect a 'time-structure' (daily, weekly, monthly component) for each volatility component. They assess whether the continuous/jump components of historical volatility and its time structure affect the explanatory power and information content of implied volatility based on the S&P 100 and S&P 500 index options. The empirical evidence suggests that the weekly and monthly continuous decomposition express more information rather than implied volatility. However, although the coefficient of the monthly jump component is in some cases significantly negative and takes a rather large negative value, implied volatility still shows the very high information content with large R^2 close to 70% and even decomposed measure of realized volatility does not bring valuable additional

information. As far as forecasting is concerned, the jump decomposition does not contain incremental information. The similar study is also done by Becker, Clements and McClelland (2009). Compared with Giot and Laurent (2006), they involve more MBFs except for GARCH model and allowed for the time-varying risk premium by adding the current level of volatility to vector of explanatory variables. They are in line with Giot and Laurent (2006) reporting that VIX does reflect the past jump activity in the S&P 500 and its forecast errors are indeed uncorrelated to past available information relating to jump activity. In other words, VIX appears incremental information content, relative to MBFs, for explaining the future jump activity.

Out of U.S. market, some researchers switch to the smaller markets to investigate the forecasting power of implied volatility in tiny markets, including Australian, Danish, Germany, Hong Kong, Japanese and Spanish markets. Hansen (2001) analyzes whether volatility implied in the KFX (Denmark equity index) option prices is more informative than the historical volatility about the subsequently realized KFX volatility forecast, in spite of the option's illiquidity in Danish market. They declare that after the measurement errors are diminished, the implied volatility appears to be the better estimate rather than the historical data. Classen and Mittnik (2002) focus on the informational efficiency of German DAX-index option market and information content of volatility on DAX-index options (VDAX). In-sample fitting and out-of sample forecasting results show that VDAX is the superior estimate beyond the past return data. As most evidences of U.S. markets, they find the positive bias exists in the implied volatility forecast in Germany market. Nishina et al. (2006), using the similar model-free method with VIX to develop the implied volatility index for Japanese market, assess the forecasting ability of implied volatility index relative to alternative GARCH models. Relying on the better forecasting performance in out-of-sample, implied volatility index yields GARCH model as well as historical volatility to be the best estimate for future volatility. However, Dowling and Muthuswamy (2003) provide the contradictory evidence. They construct the volatility index for Australian stock market with the similar method of VIX and find that this volatility index underperformed the historical volatility with respect of predictive power. Likewise, Gonzalaez and Novales (2007) who proposed the implied volatility index VIBEX (non- model free) and

VIBEX-NEW (model free) for Spanish market, conclude that their volatility index is the inferior predictor, since high mean forecasting error suggests that forecasting ability of VIBEX-NEW is unreliable. Yu, Liu & Wang (2010) are interested on the efficiency of stock index options traded over-the-counter (OTC) and on the exchanges in Hong Kong and Japan. They compare the information content of implied volatility with historical volatility and GARCH (1, 1) forecasts. The predictive power of implied volatility traded OTC is investigated on the first time. They support that implied volatility is superior to historical volatility and GARCH (1, 1) forecasts. Implied volatility subsumes all the information in historical volatility as well as GARCH (1, 1) prediction. Furthermore, they take a close look at the efficiency of OTC markets in Hong Kong and Japan. They find that OTC market is more efficient than exchange-traded market in Japan, but that is not the case in Hong Kong.

Recently, Siriopoulos and Athanasios (2009) study the information content of all publicly available implied volatility indices across the world and investigate international market integration by examining equity co-movements in terms of implied volatility but not realized returns or variances. They report that all of the sampled volatility indices are biased estimates of future realized volatility, whereas contains more predictive power than past realized volatility. What's more, they confirm that there is a world-wide integration from the aspect of market expectation of future uncertainty. The change in implied volatility in U.S. equity market spread across other markets. Therefore, VIX is the leading source of uncertainty in the world. In addition, the volatility of Euro zone stock markets, as proxy by VSTOXX, is the leading source of uncertainty among European markets.

1.2.2. Individual stocks

Latane and Rendleman (1976) are the pioneers discovering the forecasting capability of implied standard deviation (ISD). They use actual closing option prices of 24 companies to generate weighted implied standard deviations (WISD). Individual ISD is derived from the Black-Scholes model. Then they compare forecasting ability of WISD with volatility predictors based on the historical stock data. They conclude that although

Black-Scholes model cannot fully capture the actual process in option pricing, WISD still outperforms historical standard deviation estimate on the future volatility prediction (1976:381).

Following Latane and Rendleman's step, Schmalensee and Trippi (1978) investigate the weekly data of six common stocks and corresponding American call option prices from 1974 to 1975. They want to find the determinants of changes in the market's expectations of common stock volatility. Firstly they corroborate Latane and Rendleman's findings. Additionally they conclude that increase in the stock price is accompanied by decrease on the volatility expectation associated with its options (1978:145). Moreover, they stress that the implied volatilities of different stocks have the positive correlation (1978:146). However, due to the limited observations, these studies suffer statistical significant problem in terms of forecasting power from the time-series perspective. Chiras and Manaster (1978) and Beckers (1981) also find forecasts from implied volatility can explain a large amount of the cross-sectional variations of individual stock volatilities. Lamoureux and Lastrapes (1993: 324) take a time-series perspective to examine the joint hypothesis of a class of stochastic volatility option pricing models and information efficiency in the option market. By using daily returns for 10 individual stocks in U.S. over the period April 19, 1982, to March 30, 1984, they conclude that, although the option market is not informational efficient and Black-Scholes group models are imperfect equilibriums of options pricing, implied volatility still contain the useful information to generate better equity volatility forecasts than time series models produced.

In line with the US studies, Gemmill (1986) report that the in-the-money ISD is the marginally best forecast of subsequent volatility by using call option prices and underlying stocks on thirteen companies in the U.K. from 1978 to 1983. Furthermore, out-of-the-money options contain no useful forecasting information. Although combinations of ISDs and historical based forecasts are examined, no combined forecast is found to be superior to the individual forecasts. Nevertheless, this study is not such solid evidence, because at this researching period London derivative market was actually the thin market which easily leads to low liquidity problem. After then, in the

90's, the London derivative markets boomed. The trading volume increased a lot. Vasilellis and Meade (1996) again examine twelve common stocks quoted on the London Stock Exchange and the corresponding options. On one side, they confirm that weighting scheme of implied volatility have the better performance on predicting future volatility than historical return time series in individual models for three-month investment horizon. On the other hand, they also get some contrary evidence compared with Gemmill. Combination of GARCH and implied volatility forecasts significantly outperforms its components. This finding implied that option markets do not embrace all the information and equity option market is not informational efficiency, which is consistent with Lamoureux and Lapstrapes' (1993) conclusion.

In summary, due to the low liquidity using estimates implied from individual stock option prices tends to suffer a lot from measurement errors and bid-ask spread. That is the reason why the conclusions exhibit inconsistent. (Poon & Granger 2003:500).

1.2.3. Other asset

The strongest supporting power for implied volatility is from currency markets (Poon & Granger 2003:501). Numerous studies state that implied volatility is the dominant method for volatility forecasting in currency markets rather than historical average forecasts (Wei & Frankel 1991) as well as ARCH family models (Jorion 1995, 1996; Pong et al. 2002; Xu & Taylor 1995). However, Li (2002) compares the forecasting power of option-implied volatility from at-the-money forward currency options on the deutschemark, the Japanese yen, and the British pound to the forecasting power of historical volatility-based predictions model over different forecasting horizons. And then he gives the counterevidence. Their results reveal that AR (FI) MA model is more suitable for forecasting future volatility than implied volatility in the long-memory situation.

Edey and Elliot (1992), Fung and Hsieh (1991), and Amin and Ng (1997) throw the light on the forecasting power of volatility implied from interest rate options. Edey and

Elliot (1992), and Fung and Hsieh (1991) employ the Black model (a modified version of Black-Scholes model) to derive the implied volatility, while the single factor Heath-Jarrow-Morton model is used by Amin and Ng (1997). All three studies report the significant forecasting power in implied volatility of interest rate options over a short horizon.

Most studies on the forecasting power of implied volatility mainly aim at the fundamental assets, such as individual stocks, stock index, interest rate or currency. Unlike the previous studies, Szakmary et al. (2003) studies the predictive power of volatility embedded in 35 futures options. The classes of studied futures options include equity index, interest rate, energy, industrial and agricultural futures options across eight exchanges. GARCH model and historical forecasts are used to compare with implied volatility. This study improves two shortcomings in previous work. Firstly, the futures and options contracts trade on the same exchange. Therefore, their closing prices are less likely to suffer the non-synchronous trading problem. Secondly, transaction costs on futures are lower relative to equity or currency trading, which has less trading frictions. 34 out of 35 futures options demonstrated the positive constant term, and all slope coefficients for implied volatility are positive and highly significant but less than unity. This can be interpreted as implied volatility is biased but contains useful information on future volatility prediction. The predictive power of implied volatilities is superior to historical volatilities in 34 out of 35 futures options. Historical volatilities provide additional information relative to implied volatilities in only 6 out of 35 futures options. As far as GARCH model forecast is concerned, although it demonstrates some incremental information that is not contained in implied volatility, it does not add much predictive power in the majority cases. In addition, the main predictive power is from the implied volatility.

In appendix 1, the author summarized the main empirical results of the previous literatures mentioned above. Although the information content and predictability of implied volatility have been examined on different horizons over various sample periods, even across several markets, the researchers cannot get the consistent

conclusion whether implied volatility is the efficient and superior estimate beyond MBFs.

1.3. Structure of the study

The thesis consists of the theoretical part and the empirical part. The theoretical part discusses the relative theories, models to lay the solid theoretical foundation for the empirical part. The empirical part presents data, studying methodology, the empirical results and conclusions in this study.

The previous researches relating to this issue has been reviewed. The rest part of this thesis is organized as follows. In chapter 2, the author will define the concept of volatility and state the characteristics of volatility. The third chapter demonstrates existed volatility forecasting techniques. The first part of this chapter presents the time-series models using to estimate the volatility and generate the forecasts. Then the features of implied volatility and model of implied volatility estimation are discussed. The fourth chapter concentrates on how to evaluate the volatility forecasting performance. Next chapter starts the empirical study, introducing the data, methodology. In this chapter, the statistical characteristics of the data set are firstly presented. In addition, methodology used in the thesis will be described in details. The sixth chapter analyses the empirical results. Finally, this study is summarized and some limitations are stated in the last chapter.

2. VOLATILITY

This chapter tries to answer two questions: the first one is what volatility is; the second one is what kinds of features the volatility does have. It starts with an explanation of the concept of volatility, mainly for the purpose of clarifying the scope of this thesis.

2.1. Definition of volatility

The precise definition of the *volatility* of an asset is “*an annualized measure of dispersion in the stochastic process that is used to model the log returns*” (Alexander 2008: 90). However, true process of volatility cannot be observed, because the pure volatility is not traded in the market. It can be only estimated and forecasted.

2.1.1. Volatility measurements

Statistically, volatility is often used to refer to the standard deviation of the returns in the sample period,

$$(1) \quad \hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_t - \bar{r})^2}$$

where r_t is the return on day t , and \bar{r} is the average return over the T-day period (Poon 2005: 1). This sample standard deviation $\hat{\sigma}$ is a distribution free parameter representing the second moment characteristic of the sample.

In practice, the predictions of price variations of financial assets are very hard. So the usual way is to assume that the distributions of successive returns are relatively

independent of each other. The one-period log returns are normally distributed with mean μ and standard deviation σ . Since the dispersion will increase as the holding period h increases, this means that standard deviation of n-days returns cannot compare with standard deviation of m-days. It is necessary to transform the sample standard deviation into annualized form in order to make it comparable. The annualized standard deviation is called the annual volatility, or simply the volatility, defined as follows

$$(2) \quad \text{Annual volatility} = (100\sigma\sqrt{A})\%$$

where A is an annualizing factor, the number of returns per year (Alexander 2001:5). However, the above transformation of standard deviation is only valid when returns are i.i.d (independent and identical distribution), which implies that volatility is constant. The constant volatility process exactly corresponds to the assumption for Black-Scholes-Merton type option pricing models and moving average statistical volatility, whereas this assumption is not realistic in the real world. Nonetheless, the annual volatility as quoting volatility has been the market convention and been widely employed to forecast standard deviation, not considering whether it is based on an i.i.d assumption for returns (Alexander 2008:92).

Even though annualized sample standard deviation is broadly utilized, some academicians query its applicability, especially when the small sample is taken into account. For example, when monthly volatility is needed and daily data is obtainable, it is simple to use formula (1) to calculate the standard deviation. However, when daily volatility is considered and only daily data is accessible, the problem appears. Figlewski (1997) points out that since sample mean has the inherent statistical properties making it very imprecise, when used as the estimate of the true mean. To address this issue, many researchers turn to use daily squared return to represent daily volatility, generated from the market closing prices. By employing this daily squared return method as the measurement of the latent volatility process, Cumby et al. (1993), Figlewski (1997), and Jorion (1995, 1996) find that despite they get highly significant in-sample parameter estimates, their standard volatility models do poor performance in out-the-sample forecasts. Andersen and Bollerslev (1998:886) give the explanation as follow. Set the

return innovation as $r_t = \sigma_t * z_t$, where z_t denotes an independent mean zero, unit variance stochastic process. While the latent volatility, σ_t , changes simultaneously with the specific model described. If the model for σ_t is correctly specified, then $E_{t-1}(\sigma_t^2 * z_t^2) = E_{t-1}(r_t^2) = \sigma_t^2$. It is apparently reasonable to adopt daily squared return innovation as a proxy for ex-post volatility. Whereas the error component, z_t^2 , varies observation-by-observation in a large degree. Hence, the squared innovation may become the quite noisy measurement. When r_t^2 is as the measure of ex-post volatility, the poor predictability of volatility models may well be due to the inherent noise in the return process, but not incompetent models.

As mentioned above, daily returns are generated by the close prices. Some researchers think that it is not sufficient. Different prices contain ample information. Just using close price is incapable of reflecting this information set. Parkinson (1980) proposed *extreme value* method, also called *high-low* measurement, to estimate volatility. The basic idea is to use the highest- and lowest price for a unit time interval to capture the relative information. The specific formula is as follows

$$(3) \quad \hat{\sigma}_t^2 = \frac{(\ln H_t - \ln L_t)^2}{4 \ln 2}$$

Denote H_t and L_t respectively as the highest and lowest price on t time interval. The extreme value method is very easy to apply in practice, because daily, weekly, and in some cases, monthly highs and lows are published for every stock by major newspaper. Although high-low method captures more information and is convenient to implement, it is still founded on the normal distribution assumption which is invalid for financial market returns. Financial market return exhibits a long tail. Therefore, this method is very sensitive to outliers. When it is applied on the volatile procedures, it will be inefficient.

Pitfalls above motivate the researchers to find a new and more accurate way to measure the latent volatility process. Fung and Hsieh (1991), Andersen and Bollerslev (1998) take the initiative in using the term *realized volatility*, which means “*the sum of intraday*

squared returns at the short intervals such as fifteen- or five-minutes.” (Poon & Granger 2003: 481). This total new method reduced the noise dramatically and makes a radical improvement in temporal stability relative to the method based on daily return (Andersen & Bollerslev 1998:887). According to Poon (2005), such a volatility estimator has been shown to be the accurate estimate of the latent volatility process. However, high-frequency data is not readily accessible, especially impossible in some illiquid markets. Table 1 summarized respective strengths and weaknesses of different methods to measure the latent volatility.

Table 1. Summary of possible measurements of the latent volatility process.

Measurements	Formula	Strengths	Drawbacks
Sample standard deviation	$\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_t - \bar{r})^2}$	No distribution assumption. Unconditional volatility.	Unavailable in short-interval sample. Inaccurate estimation of the true means.
Daily squared returns	$\hat{\sigma}^2 = r_t^2$	Daily estimation available. No mean estimation.	r_t contains innovation term which tends to be much noisy.
High-low measure	$\hat{\sigma}_t^2 = \frac{(\ln H_t - \ln L_t)^2}{4 \ln 2}$	Using high-low price to consider microstructure.	Very sensitive to outliers. Just useful on the trimming procedures.
High-frequency data	$\hat{\sigma}^2 = \sum_{j=1, \dots, m} r_{m, t+j/m}^2$	The most accurate method so far.	Inconvenient to get the data.

2.1.2. Misperceptions of volatility

In financial market, the investors usually tend to translate the volatility as *risk*. However, they are not exactly the same. Volatility only refers to the spread of a distribution but nothing with its shape. If and only if the distribution is normal or log-normal distribution, volatility is a sufficient statistic for the dispersion of the returns distribution. Otherwise, the shape or the entire function of the distribution needs to be known to determine the dispersion. Figure 1 (a) plots the return distributions of Dow index from 1st October 1928 to 8th March 2010. The normal distribution simulated with the same mean and standard deviation of the financial asset returns is drawn on Figure 1 (b) to facilitate comparison. Compared with normal distribution, the distribution of Dow returns has fat left tail and higher peak with skewness (-0.596960) and kurtosis (27.81276). The volatility is not the sole determinant input for the dispersion of the Dow return distribution. This thesis is only interested in volatility. Although volatility is not the sole determinant of the asset return distribution, it is also a key input to many significant financial applications.

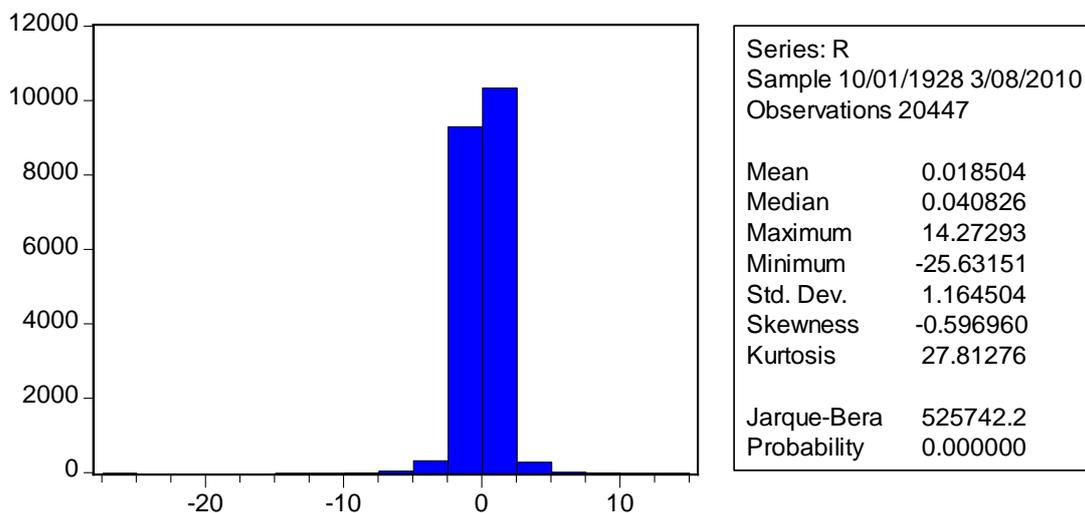


Figure1. (a) Return distribution of Dow index from 1928 to 2010.

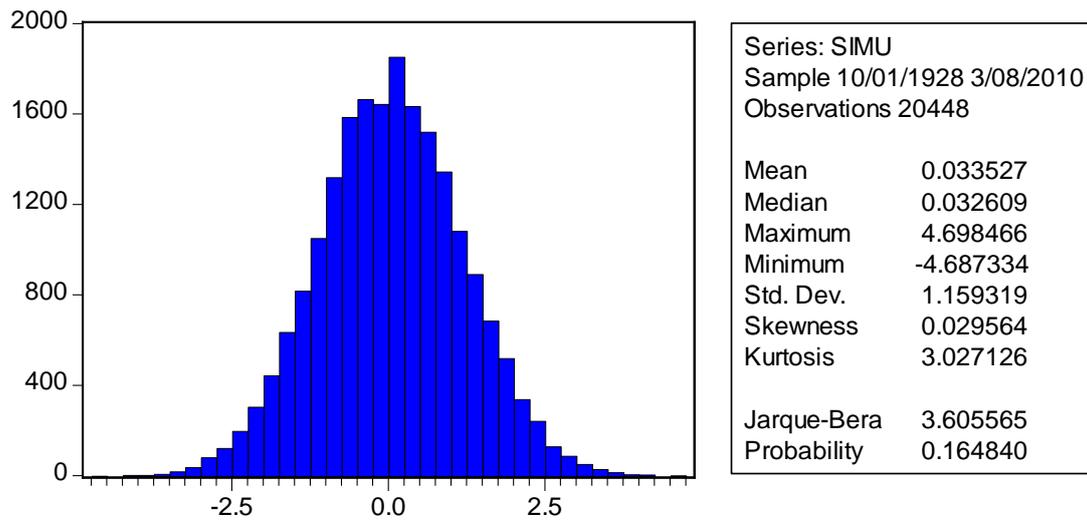


Figure1. (b) The simulated normal distribution with the same mean and standard deviation of Dow index.

2.2. Characteristics of financial market volatility

In the markets, financial time series such as asset return displays different behaviors unlike the theoretical assumption. These features widely exist in different assets. This section discusses the characteristics of volatility in the real world which may affect the volatility model selection, estimation and forecasting.

2.2.1. Fat tails and a high peak

In contrast with the assumption of financial theory, most asset returns are not normally distributed. However, Mandelbort (1963) firstly questioned the normal distribution assumption of asset returns. He cited other example (1963: 395, Fig 1) to document empirical leptokurtosis. He thought that high kurtosis contains certain information and should not be simply overlooked. Cootner (1964) found return distribution with the longer tail rather than normal distribution and developed the whole theory to explain it. After that, numerous literatures investigated the features of stock return distribution. Two most obvious features are fat tails and a high peak (Poon 2005:4). Moreover, these

two features are interdependent, because extreme values gain large weights in the variance of the distribution. It indicates that there are more observations around the mean of the distribution compared with the normal distribution with the same mean and variance (Taylor 2005: 69-71). In another word, stock return varies in the smaller range but extreme values occur more frequently than it is assumed in theory. These fat-tailed and leptokurtosis effects should be taken into account appropriately, when forecasting the future volatility.

2.2.2. Volatility clustering

Volatility clustering refers to the phenomenon that a turbulent trading day tends to be followed by another turbulent day; similarly, a stable period tends to be persistent by another stable period. It is obvious from Figure 2 (a) that fluctuations of financial asset returns are lumpier than the even variations of normally distributed variable in Figure 2 (b). This observation is firstly noted by Mandelbrot (1963) and Fama (1965). Then this autoregressive conditional heteroskedasticity is widely found across equity, commodity and foreign exchange markets at the daily, even the weekly frequency (Alexander 2001: 65). For instance, Chou (1988) investigates the volatility persistence in U.S. equity market with GARCH technique. According to his study, the volatility persistence of shocks is so high that even the test cannot decide whether the volatility process is stationary or not. After that, Schwert (1989) confirms Chou's conclusion with the longer sample data. Haan and Spear (1998) document that the volatility of monthly real interest rates has the persistent characteristic. They explain this phenomenon by the business cycle and the spread between the borrowing and the lending rate. Recently, Andersen, Bollerslev, Diebold & Labys (2003) employ the high-frequency data to generate realized volatility and also detect the volatility clustering pattern in the exchange rate market.

Volatility clustering implies that return successive distributions are not serially independent and identical; hence volatility is absolutely not constant over time. This implication is a negation of the constant volatility models that refers to the

unconditional volatility of a return process. To address this pitfall, Engle (1982) proposed ARCH (autoregressive conditional heteroskedasticity) model to firstly capture this type of volatility persistence and gained Nobel Prize. After Engle, Bollerslev (1986) introduced more appropriate model-GARCH that is fit for financial data better. These models will be discussed in details in the third chapter.

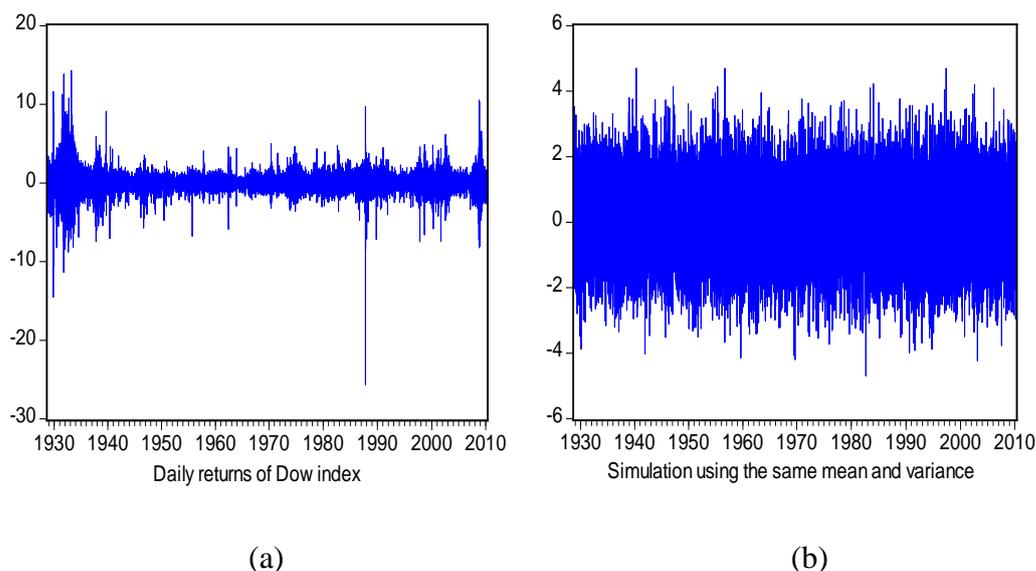


Figure 2. Time series of daily returns on Dow index and a simulated random variable.

2.2.3. Mean-reversion

Volatility clustering indicates that volatility moves up and down. Thus a period of high volatility will eventually fall. Likewise a period of low volatility is quite likely to rise in the following step. This *mean-reversion* behavior in volatility implies that there is a normal level of volatility to which volatility converges at length (Engle & Patton 2001:239). For very long-run prediction of volatility, it should converge to this normal level regardless of the time when they are made (Engle & Patton 2001:239). In another word, the current shock cannot affect the long-term volatility forecasts.

There are abundant evidences of volatility mean-reversion. According to Fouque, Papanicolaou and Sircar (2000:29), volatility of S&P 500 index reverts to mean value very fast. They found volatility can be modeled well by a fast mean-reverting stochastic process. In currency option pricing, Sørensen (1997) advocates mean reversion through the dynamics in the domestic and foreign term structures of interest rates. Similarly, Wong and Lau (2008) document that exchange rate has the mean-reverting feature. It has the substantial effect on option pricing. Recently, Bali and Demirtas (2008) hire continuous GARCH model to investigate the degree of mean reversion in financial market volatility. The empirical findings indicate that the conditional variance, log-variance, and standard deviation of futures writing on S&P 500 index approach to some long-run average level over time (Bali & Demirtas 2008:23).

2.2.4. Long memory effect

As mentioned above, volatility persistence is described by ARCH and GARCH group models. Autocorrelation of conditional variance in GARCH model decays at an exponential rate. However, the autocorrelations of $|r_t|$ and r_t^2 decay at the much slower rate than the exponential rate, just as Figure 3 demonstrates. The positive autocorrelations remain in very long lags. This is defined as *the long memory effect of volatility* (Granger & Joyeux 1980; Hosking 1981; Bailie 1996). That means the effect of volatility shocks lasts for the longer time than GARCH model describes and impacts on future volatility over a long horizon. The volatility shocks are much more powerful than the common sense.

The integrated GARCH (IGARCH) model developed by Engle and Bollerslev (1986) captures this effect. With a drawback, a shock in IGARCH model affects future volatility in the infinite horizon. At the same time, there is no unconditional variance for this model (Poon 2005:45). In addition, many nonlinear short memory volatility models, such as break model (Granger & Hyung, 2004), the volatility component model (Engle & Lee, 1999), and the regime-switching model (Hamilton & Susmel, 1994), can also mimic the long memory effect in volatility as well. Details of some models are provided

in the next chapter. Regarding *long memory effect*, one more interesting phenomenon is known as *Taylor effect*. Taylor (1986) noted firstly that the absolute return $|r_t|$ has a longer memory relative to the squared returns r_t^2 . For explaining this phenomenon, researchers are still working on process.

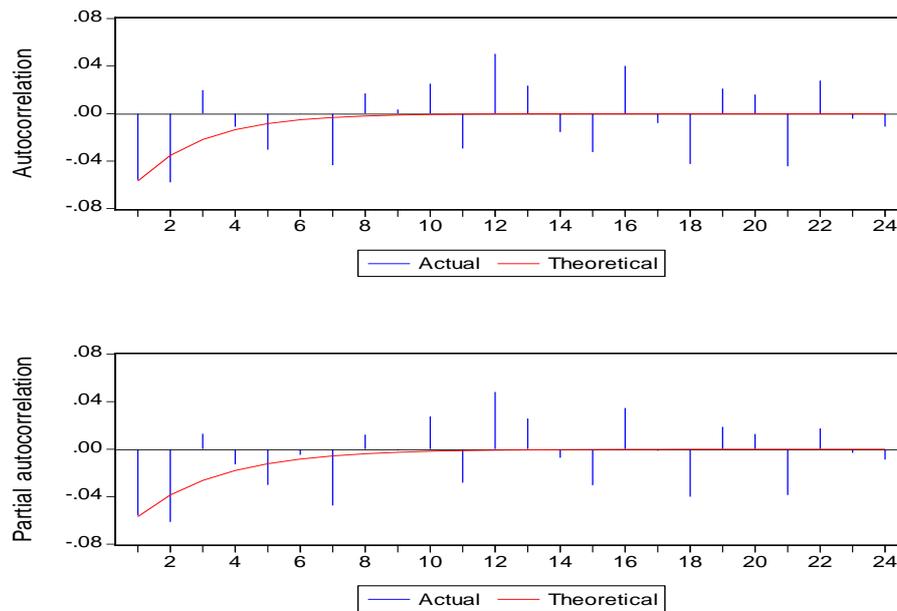


Figure 3. Autocorrelation and partial autocorrelation of daily squared returns on S&P 500 index.

2.2.5. Volatility asymmetry

A number of volatility models assume that the market responds symmetrically to the positive and negative shocks. One typical instance is GARCH (1, 1) model, in which the conditional volatility depends on the lagged shock, but there is no distinction between good or bad news. However, in equity markets, it is quite noticeable that a negative shock leads to higher conditional volatility in the following period than a positive shock does (Black 1976; Alexander 2001:68; Poon 2005:41; Alexander 2008:147). Markets tend to respond far greater to a large negative return than to the same amount of

positive return. This phenomenon is always pronounced during large falls (Poon 2005:8).

Black (1976) and Christie (1982) interpret this asymmetric response with *the leverage effect*. When the stock price declines, debt keeps constant in the short interval. Therefore the debt/equity ratio increases. Based on the capital structure theory, financial leverage of the company becomes higher. That implies that the risk of the equity rises so that the future of the company is uncertain. Hence the stock price behaves more turbulent and vice versa. However, there is also some questioning sound. Figlewski and Wang (2000) give the evidence that there is a strong "leverage effect" associated with falling stock prices, but for positive news a very weak or nonexistent leverage effect as the explanation. Furthermore, they found no apparent effect on volatility when leverage changes due to the new issue of debts or stocks, only when the share price changes. They attribute the reason of volatility asymmetry to "down market effect" (Figlewski & Wang 2000:23).

There are still some debates on its reason, but no one can deny that volatility asymmetry is the important feature of volatility process. After the early reference-Black (1976) of this phenomenon, it has been found repeatedly since then by authors such as Christie (1982), Schwert (1989), Glosten, Jagannathan and Runkle (1993), Braun, Nelson and Sunier (1995), and many others. It appears both in the volatility of realized stock returns and also in implied volatilities from stock options. That is why plenty of asymmetric GARCH models, such as exponential GARCH (EGRACH) model by Nelson (1991), GJR-GARCH model by Glosten et al. (1993) and so on, are created to capture this phenomenon. Some of these models will be clarified in 3.1.2.

2.2.6. Cross-border spillovers

The means and volatilities of different assets (e.g. individual stocks), even different markets (bond vs. equity markets in one or more nations), are inclined to move together (Poon 2005:8). This is called international financial integration (Hamao, Masulis, Ng

1990: 281). Literally dozens of researches shed the light on the correlation of asset prices and volatilities across international markets. Hillard (1979) examines the contemporaneous and lagged correlation in daily closing price changes across 10 major stock markets. They confirm that there exists, to some extent, the relation among the different markets; especially most intra-continental prices move simultaneously (Hillard 1979:113). Jaffe and Westerfield (1985) study daily stock market returns in the U.K., Japan, Canada, and Australia. Eun and Shim (1989) investigate daily stock returns across nine national stock markets and try to figure out the transmission mechanism of stock market movements via vector autoregression (VAR) analysis. The empirical evidence indicates that there is actually a substantial amount of interdependence among regional stock markets. And American market is the leading market. The innovation from American market affects other markets, but no one market can explain American market innovations. Barclay, Litzenberger, and Warner (1990) examined daily price volatility and volume for common stocks dually listed on the New York and Tokyo stock exchanges. They report the evidence of positive correlations in daily close-to-close returns across individual stock exchanges. More evidence on equity market integration are also detected by King, Sentana and Wadhwani (1994); Karolyi (1995); Koutmos and Booth (1995); Forbes and Chinn (2004). The similar phenomenon is also plotted in exchange rates (Hong, 2001) and interest rates (Tse and Booth, 1996).

3. VOLATILITY FORECASTING

In last chapter, the target of forecasting has been made clear. Then the specific paths to implement forecasts will be described in this part. According to Canina and Figlewski (1993:659), Poon and Granger (2003:493-499), Becker and Clements (2008:122), Becker, Clements and Fenn (2009:2), the ways of the volatility forecasting are generally divided into two groups. On one hand, predictions of future volatilities can be generated from econometrical models by hiring historical information (model based forecasts, MBFs). In contrast with MBFs, implied volatilities derived from options can also be utilized as volatility forecasts. The latter one is usually looked as the market expectation of the future volatility. If market efficiency hypothesis is accepted, it should be the best prediction of the future volatility. However, the main purpose of this study is not targeted to test the efficiency of the option market. So the author just raises the competition between implied volatility and model based forecasts. Even if the market is not efficient, implied volatility may be still the superior predictor rather than MBFs.

3.1. Model based forecast

There are some models, for example, Capital Asset Price Model (CAPM), which are a kind of “theory based” models. While the models discussed in this section concentrate on capturing the main features of volatility found in actual markets, but not the theoretical basis. According to Poon and Granger (2003), Poon (2005), the models of the volatility prediction can be typed into three groups: the first one is historical volatility models; secondly, ARCH family; the last is stochastic volatility (SV) models. These models become more and more sophisticated from group one to group three. Following this logic, this section will introduce the easiest one firstly- historical volatility model.

3.1.1. Historical volatility models

Historical volatility models (HIS) are the less artificial models, which are easier to be manipulated and constructed rather than the other types of volatility models. If volatility estimates are already available, HIS do not usually use the return information. It is quite different from ARCH and SV models in which volatility mainly relates to return inputs. They demand less computational work and restriction as well as requirements on input data so that it is pretty much attractive (Poon 2005: 31). However, because of the simple structure, HIS cannot capture some specific features of the realistic markets. That is why statisticians take some more complicated models into account. Nevertheless, HIS do provide a good benchmark when examining the more sophisticated models.

a) Random walk

The simplest historical model is *random walk* model. This model assumes that the difference between consecutive period volatilities is just a random noise. It indicates that today's volatility is the best available forecast for tomorrow's volatility.

$$(4) \quad \sigma_t = \sigma_{t-1} + v_t$$

$$(5) \quad \hat{\sigma}_{t+1} = \sigma_t$$

Where σ_t represents the prediction of σ_{t+1} . Random walk model is such a simple model which includes the most up-to-date information, say, just one lagged observed volatility. To extend this idea, *historical average* model is introduced as follows.

b) Historical average

Different from random model, historical average model generates a prediction on the basis of the entire history. It is a typical supporter of mean-reversion. The basic idea is that the best forecast of future volatility should be the long term volatility, which can be represented by historical mean of past volatilities. The underlying assumption of this

model is the existence of the constant conditional expectation of volatility (Yu 2002). Historical average model includes all of the sample observations to increase the amount of information and gives the old and new information in the same weights.

$$(6) \quad \hat{\sigma}_{t+1} = \frac{1}{t}(\sigma_t + \sigma_{t-1} + \dots + \sigma_1)$$

c) Moving average

Similar to the historical average model, moving average model needs to calculate the arithmetic historical mean. The only difference is that older information is discarded. The lag length (the value of t) can be decided subjectively or based on minimizing in-sample estimation error, $\varepsilon_{t+1} = \sigma_{t+1} - \hat{\sigma}_{t+1}$. Like historical average model, moving average model uses equally weighted average to calculate the historical mean. This way tends to overweight the extreme events. Because no matter extreme events occurred yesterday or at any other time in average period, they are just as the same importance for current estimates (Alexander 2001: 52).

$$(7) \quad \hat{\sigma}_{t+1} = \frac{1}{t}(\sigma_t + \sigma_{t-1} + \dots + \sigma_{t-\tau-1})$$

There is also another noticeable point in moving average volatility estimate. It is easy to find that considerable differences between volatility estimates obtained from equally weighted averages of different time lengths. In small sample, the effect of an extreme event is more pronounced than large one, because an extreme event is averaged over just a few observations. However, this effect lasts for a relatively short period of time. Volatility estimates made from the short term period is more volatile than the estimates obtained from the large sample. The problem is that they estimate the same thing-the unconditional volatility. Under the constant unconditional volatility assumption, there should be little difference between historical volatility estimates of different sample lengths. (Alexander 2001:53). Therefore, it is necessary to remove extreme events or make the sample period as long as possible, when such kind of models is applied (Alexander 2001:52).

d) Exponentially weighted moving average

As mentioned above, equally weighted average models do not include the dynamic properties of returns. They are typical static models. An exponentially weighted moving averaged (EWMA) model stresses on more recent observations. Additionally, it accounts for the dynamic ordering in returns. In another word, a time-varying framework is involved.

$$(8) \quad \hat{\sigma}_{t+1} = \sqrt{(1 - \lambda)r_t^2 + \lambda\hat{\sigma}_t^2}$$

Where λ is the smoothing parameter, $0 < \lambda < 1$. EWMA model reflects that volatility estimates will react immediately following an unusual return. Then the impact of the shock will decay along with time. High λ means that volatility estimate reacts little to actual market events, but gives the great persistence. While low λ indicates that volatility reacts rapidly but quickly diminishes away. One considerable restriction of EWMA is that the reaction and smoothing parameters are interdependent, since the sum of them is one. This assumption may be not the universal relevance for all the markets (Alexander 2001:59).

The most suitable value for the smoothing parameter is a worthy topic to be discussed. Poon (2005) asserts that it should be estimated by minimizing the in-sample forecast errors. Depending on a rule of thumb, the value of smoothing parameters should be in the range of 0.75-0.98 (Alexander 2001:59). JP Morgan also uses a simplification of EWMA model as the volatility forecasting tool in their *RiskMetricsTM* risk management system. Unlike the EWMA above, they utilize predetermined value of the smoothing parameter instead of continual estimation. They choose the same parameters for all assets. For daily forecast, they set $\lambda = 0.94$

$$(9) \quad \hat{\sigma}_{t+1} = \sqrt{0.94\hat{\sigma}_t^2 + 0.06r_t^2}$$

If the forecasting horizon, ΔT , exceeds one day, daily forecast is scaled by $\sqrt{\Delta T}$ (*RiskMetricsTM* 1996:80-85). It is a simple and effective volatility prediction approach. However, it works well just for short-horizon forecasts, for example forecasting horizon less than one month. However, for monthly forecasting, λ is set as 0.97. Regarding to long-term forecasting, the optimal decay factor must be adjusted to the desired horizon ΔT . This occurs because a long term forecast must apply more information from the distant past than a short term forecast. (Zumbach 2007:4).

e) Simple regression

Simple regression model also predicts volatility as weighted average of historical volatility models above; expect that weighting schemes of this type model are not pre-specified. It is estimated by OLS (ordinary least squares) regression of observed volatilities.

$$(10) \quad \hat{\sigma}_{t+1} = \gamma + \beta_1\sigma_t + \beta_2\sigma_{t-1} + \dots + \beta_n\sigma_{t-n+1}$$

Where γ is the constant mean of volatility, $\beta_1, \beta_2, \dots, \beta_n$ are the estimated coefficients of past observed volatilities $\sigma_t, \sigma_{t-1}, \dots, \sigma_{t-n+1}$. If $\beta_2 = \beta_3 = \dots = \beta_n = 0$, it is the simplest autoregression model, AR (1).

3.1.2. ARCH family

Concerning on HIS, one assumption is that the volatility is constant. So they don't take the time-varying conditional volatility into account, while just provide the unconditional volatility estimates. Although some changes are observed in the estimates, they are attributed to "noise" or sampling errors in these models. Except for that, there is nothing allowed for variation in volatility (Alexander 2001:63). However, as mentioned in 2.2.1,

financial market volatility is known to cluster. It cannot be ignored in both in-sample and out-of-sample forecasts. Some more sophisticated models are developed to capture this volatility persistence.

a) ARCH(q)

As a pioneer, Engle (1982) firstly introduced ARCH model, which allows for variations of conditional volatilities. It can be simply described as today's shock gives information about the tomorrow's volatility forecast (Engle 1982: 987). ARCH processes are “*mean zero, serially uncorrelated processes with non-constant variances conditional on the past, but constant unconditional variances.*” (Engle 1982: 987). ARCH models formulate conditional variance, $h_t = \sigma_t^2$, of asset returns by maximum likelihood estimation. Firstly demonstrate return, r_t , as

$$(11) \quad \begin{aligned} r_t &= \mu + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} z_t \end{aligned}$$

Where $z_t \sim D(0, 1)$ is a white noise process; μ is the constant mean of the returns. The conditional variance, h_t , is a function of past squared residuals of returns, which scales the process z_t . In the ARCH (q) process proposed by Engle (1982),

$$(12) \quad h_t = \omega + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2$$

Where $\omega > 0$ and $\alpha_j \geq 0$ guarantee that h_t is strictly positive; ε_t is the error term of return estimation at time t . With these non-negative restrictions (particularly $\omega > 0$), if a major shock happened one-lagged period, two-lagged period or up to j periods ago, the impact will increase recent conditional variance. However, no matter the market movement is positive or negative, due to the *squared* return shock on the right-hand side of formula (12), the variance responds symmetrically. This indicates that after one

extreme value, another extreme value follows to unpredictable direction. Because of the construction of conditional volatility, h_t is known at time $t-1$. Therefore, the one-step-ahead prediction is readily at hand.

In practical application of the standard ARCH models, short lags lead to too variable volatility estimates. Thus, the long lag length is usually demanded. (Alexander 2001:71). Whereas, it causes another problem that as the lag increase, the parameter estimation becomes more difficult, because the likelihood function becomes very flat (Alexander 2001:96). Alexander (2001:71) approves that “*the ARCH model with exponentially decaying lag coefficients is equivalent to a GARCH model*”. In addition, she also testifies that as lag length increases, ARCH (q) models converge to GARCH (1, 1) (Alexander 2001:74). GARCH (1, 1) equals to an infinite ARCH process and has less parameter. So GARCH models are more favorable in practice.

b) GARCH (p, q)

The standard GARCH model is proposed by Bollerslev (1986) to solve the problem caused by the long lag structure in the ARCH process. In GARCH (p, q) process, the conditional variance depends not only on q lagged error square, but also on p lagged historical conditional variances. The return process is defined as the same form in the ARCH process. The volatility process in GARCH is defined as follows,

$$(13) \quad h_t = \omega + \sum_{i=1}^p \beta_i h_{t-i} + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2$$

Where h_{t-i} represents i lagged conditional variances, $i=1, 2, \dots, p$. β_i is the coefficient of historical conditional variance at time $t-i$. $\omega > 0$. For higher orders GARCH, the constraints on α_j, β_i are much complicated and hardly expressed. The scales of the coefficients α and β determine the short-term dynamics of the resulting volatility time series (Alexander 2001:73). α indicates the speed of reaction of volatility to market shocks. The larger the value of α is, the quicker market reacts to the market movements.

β implies the persistence of volatility. The larger the value of GARCH lag coefficient β is, the longer the effects of shocks last for. (Alexander 2001: 73).

According to Alexander (2001: 72-73), Poon and Granger (2003: 484), many financial institutions go for GARCH (1, 1) in practice, which is fully competent to model most market data. For GARCH (1, 1), the constraints $\alpha_1 \geq 0, \beta_1 \geq 0$ are used to ensure that h_t is strictly positive. It is mentioned above that GARCH (1, 1) is equivalent to an infinite ARCH process with exponentially declining weights, but with a much more flexible lag structure (Bollerslev 1986:43, 55; Alexander 2001: 71-72).

$$(14) \quad \begin{aligned} h_t &= \omega + \beta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 \\ \omega &> 0, \alpha_1, \beta_1 \geq 0 \end{aligned}$$

Take a look at the unconditional variance of GARCH (1, 1). The sum $\alpha_1 + \beta_1$ should be less than 1 to make the return process is stationary. Only under these constraints, GARCH volatility term structure will converge to a long-term average level. Otherwise, it will go infinitely, which means that today's shock will affect the future volatility forever. However, in some special cases, the estimates of $\alpha_1 + \beta_1$ are very close to one. Such kind of processes would better being modeled by a different GARCH model, such as IGARCH (integrated GARCH) introduced in next part, in order to capture long memory.

$$(15) \quad \sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

c) IGARCH

As mentioned in 2.2.3, long memory is an obvious characteristic of volatility process, which is observed across different markets (Engle and Bollerslev 1986: 28; Alexander 2001:75; Mikosch et.al 2004:378). Especially in currency and commodity markets, they may not mean-revert at all (Alexander 2001:75). The volatility processes in these markets are purely random walk. While standard GARCH model forecasts tend to be

mean-reversion. It cannot capture this feature appropriately. Therefore, IGARCH is introduced to use for such kind of non-stationary volatility process.

When $\alpha_1 + \beta_1 = 1$ and put $\beta = \lambda$, the equation (14) can be rewritten as

$$(16) \quad h_t = \omega + (1 - \lambda)\varepsilon_{t-1}^2 + \lambda h_{t-1}$$

In this situation, the unconditional variance is infinite. Meanwhile, conditional forecasts do not converge to the long-run average level. In another word, today's information is significant for prediction forever. The impact of shocks is permanent. (Engle & Bollerslev 1986:27-28). If $\omega = 0$, the IGARCH model (16) is identical with EWMA model. When $\omega > 0$, this GARCH model is integrated in variance with trend. While intuitively unconvincing as a volatility process with infinite unconditional variance, EWMA and IGARCH are still powerful in volatility forecasting due to no limitation of a mean level of volatility. It can adjust promptly to changes in unconditional volatility (Poon 2005: 40).

d) Asymmetric GARCH

The standard GARCH (p, q) model responses to the past innovations symmetrically. GARCH models just account for the magnitude but not the positivity or negativity of unexpected excess returns when determining conditional variances. Thus it is incapable of capturing the asymmetric characteristic of volatility written in 2.2.4. The drawback of standard GARCH model motivates researchers to build some asymmetric models. Exponential GARCH (EGARCH) is one of the most well-known asymmetric models by Nelson (1991). With the same return process of models above, volatility process can be defined as

$$(17) \quad \begin{aligned} \ln h_t &= \alpha_0 + \sum_{j=1}^q \beta_j \ln h_{t-j} + g(\epsilon_{t-1}) \\ g(\epsilon_{t-1}) &= \sum_{k=1}^p \left[\theta_k \epsilon_{t-k} + \gamma_k \left(|\epsilon_{t-k}| - \sqrt{2/\pi} \right) \right] \\ \epsilon_t &= \varepsilon_t / \sqrt{h_t} \end{aligned}$$

Define the notations: α_0 is the constant mean of the conditional variance process; β_j is the autoregressive parameter; ϵ_t is the standardized residual. θ_k, γ_k are the parameters of the standardized residual and $|\epsilon_t| - \sqrt{2/\pi}$ separately. The terms $g(\epsilon_t)$ have zero mean, because $\sqrt{2/\pi}$ is the expectation of $|\epsilon_t|$ since ϵ_t is a standard normal variable. (Nelson 1991:91; Taylor 2005:236). For simplicity, consider the one lagged situation. Formula (17) be transformed as

$$(18) \quad \begin{aligned} \ln h_t &= \alpha_0 + \beta_1 \ln h_{t-1} + g(\epsilon_{t-1}) \\ g(\epsilon_{t-1}) &= \theta_1 \epsilon_{t-1} + \gamma_1 (|\epsilon_{t-1}| - \sqrt{2/\pi}) \end{aligned}$$

Over the range $0 < \epsilon_{t-1} < \infty$, $g(\epsilon_{t-1})$ is linear in ϵ_{t-1} with slope $\theta + \gamma$, and over the range $-\infty < \epsilon_{t-1} \leq 0$, $g(\epsilon_{t-1})$ is linear in ϵ_{t-1} with slope $\theta - \gamma$ (Nelson 1991:91). With this conditioning of parameters, $g(\epsilon_{t-1})$ allows the conditional variance process to react asymmetrically to market rises and falls.

EGARCH capture the asymmetric feature of volatility process appropriately, however, it is not a good candidate for volatility forecasting. Because of logarithmic transformation, there is no analytic form for the volatility term structure in EGARCH model (Alexander 2001:80). Hence, Engle and Ng (1993) proposed the asymmetric GARCH (A-GARCH) model with simple analytic volatility term structure. The conditional variance function of A-GARCH (1, 1) is

$$(19) \quad h_t = \omega + \alpha(\varepsilon_{t-1} - \lambda)^2 + \beta\sigma_{t-1}^2 \quad (\omega > 0, \alpha, \beta, \lambda \geq 0)$$

Where λ is a constant parameter. If the information shock is negative, the squared scale effect will be larger than the time when the shock is positive. This probably describes the asymmetric phenomenon happened in the actual volatility process.

Many other asymmetric GARCH models have been provided, such as the GJR model by Glosten, Jagannathan and Runkle (1993), the quadratic GARCH (QGARCH) model by Sentana (1995). In GJR model, it draws into dummy variable to measure different effects of positive and negative innovations.

$$(20) \quad h_t = \omega + \alpha \varepsilon_{t-1}^2 + \alpha^- S_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Set S_{t-1} as a dummy, when the return is below its conditional expectation, $S_{t-1} = 1$, the squared return residual is timed by $\alpha + \alpha^-$. Oppositely, if the return is above its conditional expectation, $S_{t-1} = 0$, the slope of the squared return residual is α . The constraints of parameters are $\omega \geq 0, \alpha + \alpha^- > 0, \beta \geq 0$.

As the most popular volatility forecasting methods, historical volatility forecasts and ARCH family have many members. In terms of easy comparison, Table 2 gives a brief summary of the common used volatility models in historical and ARCH groups.

Table 2. Comparison of major historical volatility forecasts and ARCH family models.

Models	Long- /short-run	Targeted volatility features	Unconditional variance	Forecasting difficulty
Random Walk	Long	None	Finite	Easy
Historical Average	Long	Mean reversion; Long memory	Finite	Easy
Moving Average	Long	Mean reversion	Finite	Easy
EWMA	Short	Long memory; volatility clustering	Not exist	Easy
Simple Regression	Both	Mean reversion	Finite	Easy
ARCH	Short	Volatility clustering	Finite	Easy
GARCH	Short	Volatility clustering	Finite	Easy
IGARCH	Short	Volatility clustering; Long memory	Not exist	Middle
EGARCH	Short	Volatility clustering; Asymmetric reaction	Finite	Hard
GJR	Short	Volatility clustering; Asymmetric reaction	Finite	Middle

3.1.3. Stochastic volatility models

Volatility changes so frequently that it is reasonable to view and model it as a random variable. Additionally, as Black and Scholes (1972: 416) stated “... *there is evidence of non-stationary in the variance. More work must be done to predict variances using the information available.*” Hence, the more flexible process is demanded to measure the latent volatility, not only in the discrete time framework, but also in terms of continuous time. Different from ARCH models that specify a process for the conditional variance of returns, stochastic volatility (SV) models lead in a stochastic process to allow for continuous-time and non-stationary variance. For clear explanation, here considers the discrete time SV model,

$$(21) \quad \begin{aligned} r_t &= \mu + \sigma_t u_t \\ \sigma_t^2 &= N_t \theta^2 \end{aligned}$$

Suppose the volatility on day t , denoted by σ_t , is partially determined by unpredictable news on the same day. The number of news on day t is denoted by a random variable, represented by N_t . θ^2 denotes the variance of price changes, when every news i arrives in the market. Then u_t is a standard normal random variable which is independent of the random variable σ_t . (Taylor 2005: 194, 268-269). There are two random shocks per unit time in SV model: one is u_t ; the other one is N_t , which is the partial determinant of σ_t . Therefore, it is impossible to obtain the true value of σ_t and u_t from return r_t directly. (Taylor 2005:269). So the SV parameter estimation poses a challenge in implementation field, because it cannot be obtained by maximum likelihood method (Poon 2005:59; Shephard 2005:2; Taylor 2005:267). More complicated estimation methods are called for, for example, general method of moments (GMM), and Monte Carlo Markov Chain (MCMC) approach. These methods are hard to be handled and manipulated. Except for hard estimation method, SV model has a high requirement on the data series. If low frequency data is used, the forecasts are more volatile and noisy. Because, unlike ARCH family models, the predictive distribution of returns is specified indirectly, via information set available (Shephard 2005:2). Low frequency data does not contain

enough information to obtain the accurate parameter estimation. With high-frequency data which can enrich the information set, the accuracy of SV model forecast can be improved a lot. However, high-frequency data is not easily reachable in many cases. Thus, Poon (2005: 59) argues that “*SV model is a theoretical model rather than a practical and direct toll for volatility forecast*”. Nevertheless, the important application of SV models in option pricing and foreign exchange markets cannot be overlooked.

Three different types of volatility models have been introduced above. They base on different assumptions and target different features. Hence, they have goodness as well as relevant downsides. Even some models have good performance on in-sample forecasts, it does not mean that it outperforms on out-of-sample prediction. It's hard to say which the absolute leading forecasting model on volatility is. In practice, the model should be chosen based both on the data features and forecasting horizon.

3.2. Implied volatility forecasting

Volatility implied by option prices is the alternative of MBFs on the volatility prediction. Traditionally, implied volatility is calculated from either the Black-Scholes (1973) option pricing model or the Cox-Ross-Rubinstein (1979) binomial model. It is usually perceived as a market's expectation of the future volatility without parameter estimations. Based on the market efficiency theory, implied volatility, a market-based prediction, should have involved all information available and be better than model based forecasts. However, it is definitely a premature conclusion. In order to produce useful volatility estimates, implied volatility forecast requires some assumptions to guarantee the validity of option theory. Some of these assumptions may be not realistic, which probably causes biases to some extent. Furthermore, implied volatility may be biased due to many market-driven pricing irregularities. (Poon 2005:115). The details are discussed as follows. 3.2.1. elaborates the detailed procedure of generating implied volatility on the basis of Black-Scholes (B-S) model. 3.2.2 discusses the features of the implied volatility process. 3.2.3 explores the drawbacks of implied volatility caused by unrealistic assumptions in the B-S model.

3.2.1. Volatility implied by the Black-Scholes option pricing model

Black and Scholes (1973) propose a precise model for determining the equilibrium value of an option. Merton (1973) made some contributions to extend the B-S model. The B-S model has a great influence on trading and hedging options. It is also fruitful for the development of financial engineering from 1980s to 1990s. (Hull 2006:281).

The B-S pricing model bases on some strict assumptions of markets. The assumptions are: (1) the stock price follows the continuous time stochastic process where the probability distribution of returns is normal and volatility is constant over time. (2) Short selling of securities is available. (3) No transactions costs or taxes. All securities are perfectly divisible. (4) Assume no dividend payments during the life of the derivative. (5) No riskless arbitrage opportunities exist. (6) Continuous trading. (7) Risk-free interest rate is constant and the same for all maturities. (Hull 2006: 290-291).

The B-S option pricing model for European call option on a stock without dividend payment is showed as follows:

$$(22) \quad c = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Where c is call option value, S_0 is the stock price at time zero, K represents the strike price, r is the continuously compounded risk-free rate, T is the time to expiration, σ represents stock price volatility, the function $N(d)$ is the cumulative probability distribution function for a standardized normal distribution. (Hull 2006: 295-296).

The market option price as well as underlying stock price S_0 , can be observed. Strike price, time to maturity and risk-free interest rate has already known. In B-S model, all the variables except volatility are at hand. If assume that the market option price is equal to the model price, then put five known variables into the B-S formula, the volatility can be obtained by solving the option pricing equation. This volatility is known as *implied volatility*. “*Implied volatility is the volatility of underlying asset price process that is implicit in the market price of an option*” (Alexander 2001:22). To a certain extent, it is a market prediction on volatility with the horizon equal to the maturity of the option. In this sense, implied volatility should be viewed differently from MBFs, even though both of them are predictions of future volatility of underlying assets. Implied volatility uses the current option price containing all the forward expectations of investors. These expectations should be rational and include all the historical information. Under the assumption (1), the model for implied volatility treats stock return process as the continuous time process. However, most MBFs rely on the historical data of the underlying asset returns and use discrete time models for the variance of time series. (Alexander 2001: 28). Additionally, statistical volatility models usually have the exact and clear formulas to generate the prediction. However, since implied volatility cannot be defined as a linear function of option price, strike price and other variables in B-S formula, there is no closed-form solution for implied volatility. It can just be given as an implicit function of the known quantities:

$$(23) \quad \textit{Implied volatility} = f(C, K, S_0, T, r)$$

using the same denotation with equation (22). (Alexander 2001:26). Table 3 summarized the differences between implied volatility forecasts and MBFs.

Table 3. The comparison between implied volatility forecast and MBFs.

Methods	Predicting view	Data	Continuous or discrete time	Defined function
Implied volatility	Market expectation	Current option price	Continuous time	Implicit, general
MBFs	Inference from history	Historical return process	Discrete time (most models)	Clear and exact

3.2.2. Features of implied volatility

In the real markets, when the B-S model is used, one generates different implied volatilities depending on different strike prices and maturities on the same underlying asset. This phenomenon can be attributed to the violation of the assumptions of the B-S model, especially the normal distribution of return process and constant volatility assumption. Because with the constant volatility assumption for the underlying assets, the same implied volatility should be obtained from the prices of the same type options on the same underlying asset. However the markets question the assumptions. The market prices reflect properties of the price process that are not assumed in the B-S model. Instead of adjusting pricing model, it is more convenient to change the only unknown factor-the volatility to match the model price with the market price. That leads to some typical patterns of implied volatility.

1) Volatility smile and skew

The *volatility smile* refers to the empirical fact that a plot of implied volatility of an option is a function of its strike price, which has a smile shape (Alexander 2001:30; Hull 2006:375). Figure 4 illustrates the shape of volatility smile. Volatility smile is usually used by traders in equity and foreign currency markets. As Figure 1a demonstrates and 2.2.1 discusses, the stock return distribution is not normal but fat-tailed with high kurtosis. Additionally, section 2.2.2 has argued that volatility is

definitely not constant. Both arguments violate assumption (1) in the B-S model. That means large price changes may happen more frequently than that the B-S model assumes. Consequently, there should be higher probabilities on which an out of money option becomes in-the-money than is assumed in the B-S model. Based on formula (22), the B-S model price will be less than the market price for an out-of-money option. To fix the difference, the only parameter, volatility, needs to be increased. Therefore, implied volatilities for out-of-money options are higher than at-the-money implied volatilities on the same underlying asset.

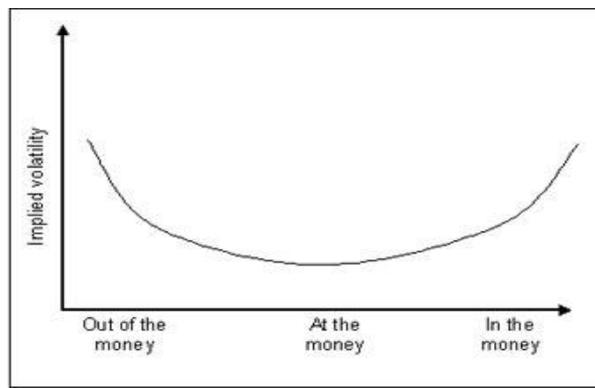


Figure 4. Implied volatility smile.

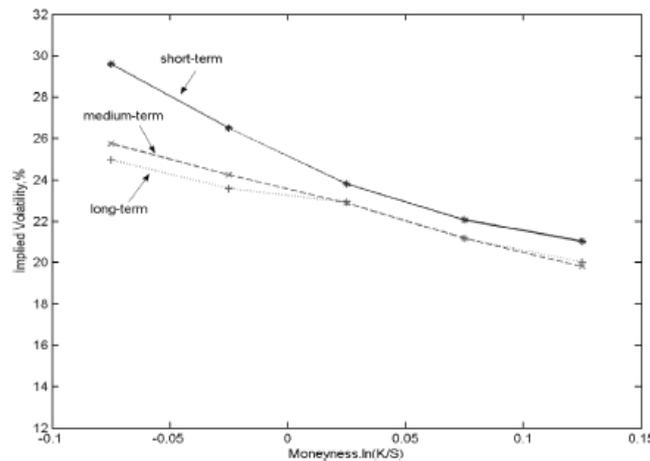


Figure 5. Equity Volatility skew of options on FTSE 100 index.

(Source: Lin, Bing-Huei., Ing-Jye Chang & Dean A. Paxson 2008.)

The volatility smile for equity options are studied by Rubinstein (1994), Jackwerth and Rubinstein (1996). Unlike the symmetric smile in implied volatility of foreign currency, volatility smile implied by equity options is asymmetric, called *volatility skew* (Figure 5). This negative skew appears the higher implied volatility as the strike price decreases. (Hull 2006:380). This is ascribed to asymmetric responses in equity markets, which has been discussed in details in 2.2.5. Another explanation given by Rubinstein (1994:487) is crashphobia. Investors are more sensitive to bad news than to good news in the same magnitude. When a very large price fall comes, investors tend to hold over-pessimistic attitude on the possibility of a similar market crash. Risk-averse investors seek for insurance and even pay more for out-of-the-money puts. (Alexander 2001:31).

2) Volatility term structure

In addition to the volatility smile, implied volatility also behaves another pattern depending on the varying time to maturities for a fixed strike price. As a consequence of a mean-reverting behavior of underlying asset volatility (illustrated in 2.2.3), implied volatility often appears the same characteristic in the term structure (Alexander 2001:31; Taylor 2005:381).

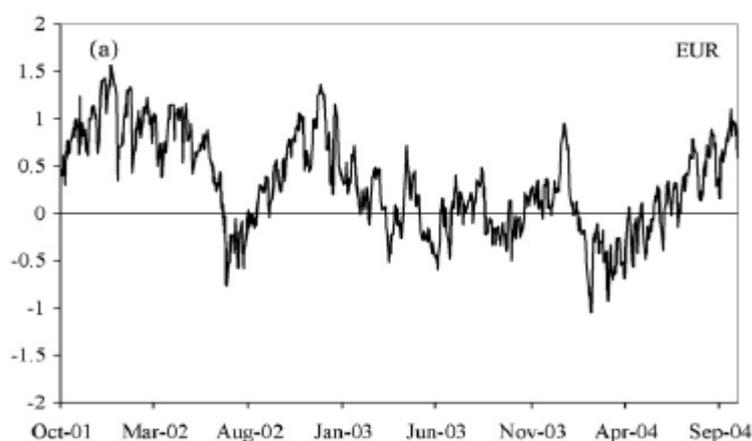


Figure 6. Implied volatility term structure.

(Source: Krylova, Elizaveta., Jussi Nikkinen & Sami Vähämaa 2009.)

Generally speaking, the volatility term structure converges to a long-term average volatility level. Thus implied volatilities tend to decline as maturity increases, when short-term implied volatilities are historically high. Although current period is turbulent, then short-dated volatility will persist above the long-term average level (as volatility clustering), for long-run, the market expectation of volatility will fall back to its normal level. Similarly, if the market is tranquil with the low short-term volatility, volatility will increase to the average level over a long period. (Alexander 2001:31; Taylor 2005:381; Hull 2006:381-382). Figure 6 shows the term structure of implied volatility.

3) Volatility surfaces

A three-dimensional plot of implied volatility, which combines volatility smile or skew with volatility term structure, is called a *volatility surface* (Alexander 2001:32-33; Hull 2006:382). According to the constant volatility assumption in the B-S model, volatility implied by options with different maturities and strike prices should be the same. That means that this surface plot ought to be flat. However, figure 7 demonstrates that it is not realistic. The surface shows dynamic and waving. The shape of volatility smile or skew relates to the option maturity. In figure 7, as time to maturity increases, the smile and skew effects become less noticeable. When a new option is valued, the financial engineer usually consults volatility surface generated by the previous market data to find the appropriate volatility. Then they put this volatility in the B-S pricing model or binominal tree to get the reasonable option prices. (Alexander 2001:34; Hull 2006:382).

3.2.3. Drawbacks of volatility implied by the B-S model

The B-S model definitely made a great breakthrough in the option pricing due to its easy operation and clear expression. However, it has some downsides caused by its unrealistic assumptions. In this section, consequences of these unrealistic assumptions will be studied.

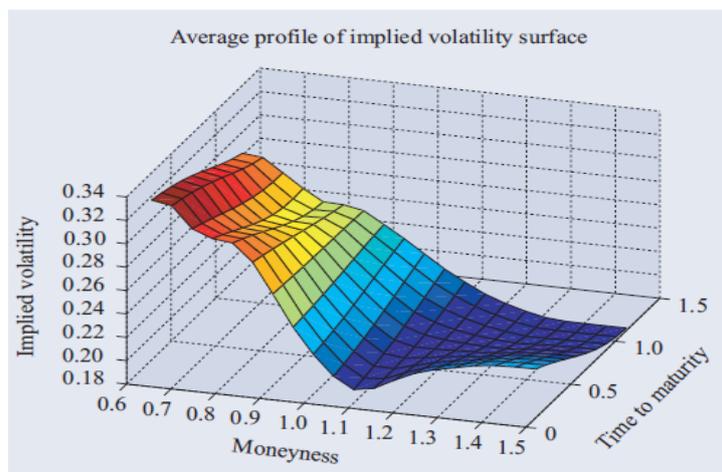


Figure 7. Volatility surface of options underlying on the S&P 500 index.

(Source: Cont, R. & J. Fonseca 2002.)

Both empirical evidences given in 3.2.2 and many researchers criticize strict restrictions in the B-S model, which not only affect the option pricing, but also bias the accuracy of implied volatility on the prediction. The most criticized assumptions are the normal probability distribution of underlying asset returns and the constant volatility. The normal distribution assumption leads the B-S model to overvalue a deep-out-of-the-money call but undervalue a deep in-the-money put (Hull: 2006:381). In another word, if looked volatility as the output, compared with the realized volatility, the B-S model underestimates the implied volatility on a deep-out-of-the-money call, while overestimates volatility implied by a deep in-the-money put. Thereby, implied volatility appears as the smile or skew shape (in 3.2.2), which swings constant volatility assumption as well. In term of volatility forecasts by using implied volatility, these out-of-the-money (OTM) and in-the-money (ITM) options will bring much noise.

Then let throw light on the constant volatility assumption. As mentioned above, it is unrealistic by the evidence of volatility term structure and volatility surface. How does it impact on the volatility prediction? In B-S model, it assumes that there is no free arbitrage opportunity due to continuous perfect hedge. Thus investors are risk neutral and just demand the risk free return. However, if volatility is a stochastic process, then

it will be impossible to hedge every type of asset perfectly. Additionally, investors are unable to adjust continuously to keep the risk free position. Beyond that, taxes and transaction cost, margin treatments of securities, constraints on margin purchases and short sales all keep the market far from perfect hedge. Thus the foundation of risk neutrality hypothesis is turbulent. It is reasonable to believe that rational investors demand higher returns for some more risky assets than for some relative safe securities. Whereas, the B-S model uses risk-free interest rate for all the assets to generate implied volatility. This risk neutral implied volatility is no longer appropriate. Risk premium should be taken into account to adjust the original risk neutral implied volatility.

Even if the B-S model has these pitfalls, it is still preferred by many traders for the following reasons. Firstly, implied volatilities on OTM and ITM options are biased, while the short-run at-the-money implied volatility can reflect the average level of volatility appropriately (Bodie and Merton 1995; Hull and White 1987). Apart from this, as implied volatility index appears, this volatility smile and term structure effects, transaction costs problem are minimized (Chicago Board Options Exchange 2009). On the other hand, it is handy to estimate the parameters. This reduces the estimation errors caused by using the stochastic volatility process with more realistic assumptions. This view is advocated by Jarrow and Wiggins (1989). In terms of the problem brought by risk-neutral hypothesis, Bollerslev, Gibson and Zhou (2008) provide the method to transform the risk-neutral implied volatility into risk-adjusted form.

At last, here has to say that with the exclusion of market efficiency and correct specification of option pricing model, there is still some limitations on implied volatility as a predicting tool. Neither every security has the exactly corresponding option, nor are all options highly liquid. If one security does not have particular and liquid options, investors tend to pick up the volatility implied by option of the similar type of assets as its volatility forecast. But this generalization is incapable of capturing some own features of this particular security. And besides, the forecasting horizon of implied volatility should match with time to maturity of options. These inherent downsides lead that implied volatility cannot be as flexible as MBFs.

4. EVALUATION CRITERIA

In terms of forecasting volatility, Chapter 3 provides a plenty of choices to implement this task. The predictability of the model itself is crucial. Apart from that, appropriate evaluation criteria chosen to measure the forecasting accuracy play another important role in the volatility predictability comparison. In some circumstances, the model itself is not a bad model, while the misjudgment is due to the inappropriate evaluation criterion, which brings some noise and assesses inadequately the predicting accuracy of the model. It is also normal that different evaluation criteria give contradictory results about the forecasting accuracy of the same model (Braisford & Faff 1996: 432). Thus, it is significant to choose the suitable evaluation criteria based on forecasting purposes and properties of the asset itself (Makridakis 1993:527).

Alexander (2001:118-125) divides the evaluation measurements into two groups: one is statistical criteria from the pure technical view to judge the predicting performance; the other is operational criteria, which is more subjective on the ground of a trading or a risk management performance. In this study, the author just focuses on the forecasting volatility process itself and compares out-of-sample forecasting powers of different methods, but nothing is concerning on the particular practice of option trading or risk management in the real markets. Therefore, here adopts the statistical criterion which is the best overall measure used in the great majority of situations and satisfies both theoretical and practical concerns.

After narrowing down the range of application, one more thing is also noticeable. In forecasting evaluations, it is vital to distinguish in-sample and out-of-sample forecasts. In-sample forecasts are usually related to parameter estimations with all data in the sample. It assumes that parameter is stable through time. However, actually parameters may change as time varies. So a good model should maintain the robustness of an out-of sample test. Out-of sample forecasts test how the model is close to the reality. (Poon & Granger 2003:492). According to Makridakis (1986), a model that best fits the historical data is not always the best candidate for the post-sample forecasts. So forth, this study

pays more attention to model performance in out-of-sample forecasts, not in in-sample fitness.

Generally there are usually two statistical ways used to judge the performance of volatility forecasting. The first common one is to employ loss functions. A model that provides a smaller average loss is more accurate and thus preferred. The second choice is to run a regression of the squared returns on the volatility predictions, then check R^2 . The highest R^2 model is the best candidate. However, these two ways are far from perfect. They get their inherent advances as well as downsides. The following part states them in detail. Loss functions (4.1) are categorized into two types: one is symmetric error measures (4.1.1); the other is asymmetric measures (4.1.2). Section 4.2 focuses on the regression based evaluation.

4.1. Loss functions

In statistics, a loss function measures the difference between the estimation and the true value. They are grouped depending on the different treatments on over- and under-predictions.

4.1.1. Symmetric error measures

Symmetric error measures give the equal penalty both on over-prediction and under-predictions of the same magnitude. Although they receive variety of doubts, such as Christoffersen and Diebold (1996), Christoffersen and Diebold (1997), Aretz, Bartram and Pope (2009), symmetric loss functions are also prevalent in practice due to their simple forms and easy computations. Following the previous studies, the most common symmetric measures are defined here.

$$(24) \quad ME = \frac{1}{T} \sum_{i=1}^T (\hat{\sigma}_i - \sigma_i)$$

Denote $\hat{\sigma}_i$ as predicting value of σ_i , σ_i is the actual value of volatility, T is the number of observations. The mean error (ME) measure is the simplest one, which is the mean of the differences between the forecasts and observations. This is a measure of the overall bias of the forecasts, which is usually used as a general guide as to the direction of over/under-prediction (Brailsford & Faff 1996: 432). However, the similar magnitude of positive and negative forecasting errors may cancel out each other so that it decreases the overall forecasting error. If one model produces large but symmetric individual forecasts errors, the other one gives small but asymmetric individual predicting errors, ME favors the former, even the latter one is more accurate. This is a kind of misleading.

To avoid this problem, the mean absolute error (MAE) measure is introduced next. MAE is the average magnitude of the errors in a set of forecasts, without considering their direction.

$$(25) \quad MAE = \frac{1}{T} \sum_{i=1}^T |\hat{\sigma}_i - \sigma_i|$$

On one more step further, scale the MAE by dividing the actual volatility to achieve unit-free aim. This measure is called mean absolute percentage error (MAPE), which is most welcomed unit-free measure (Armstrong & Collopy 1992:70). However, when the actual values approach to zero, some absolute percentage errors (APE) can become extremely large and distort the competitions in forecasting competitions or empirical studies (Armstrong & Collopy 1992: 70; Makridakis 1993:529). Fortunately, this problem can be fixed somehow by reporting the MAPEs with and without outliers. Furthermore, Makridakis (1993:529) argues that it does not matter in practice, because large errors and outliers must be known by decision-makers.

$$(26) \quad MAPE = \frac{1}{T} \sum_{i=1}^T \left| \frac{\hat{\sigma}_i - \sigma_i}{\sigma_i} \right|$$

Taking the square form of the forecasting error can also eliminate the problem in ME loss function. The usually used functions are MSE (mean square error) and RMSE (root mean square error). MSE stands for the arithmetical average of the square of the differences between the forecasts and observations. Similar with MSE, RMSE is another one of the most frequently used measures for judging forecasting methods (Armstrong & Collopy 1992:70; Makridakis & Hibon 1979:39-40). RMSE is the square root of MSE.

$$(27) \quad MSE = \frac{1}{T} \sum_{i=1}^T (\hat{\sigma}_i - \sigma_i)^2$$

$$(28) \quad RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (\hat{\sigma}_i - \sigma_i)^2}$$

Because volatility is the second moment of return series, mean square error of volatility forecasts is the fourth moments of the original series. Estimation of the fourth moment for non-normal distribution is very complicated. Apart from that, the latent volatility process cannot be observed directly from the markets. The common way to deal with this problem is to use squared returns instead of actual volatility. This adds more noise in the application of RMSE. That is why RMSE is unreliable as reported by Armstrong & Collopy 1992: 70, 72; Alexander 2001:122; Brailsford & Faff 1996:432-433. Generally speaking, RMSE is more suitable for mean estimation rather than volatility forecasting (Alexander 2001:122).

The error measures demonstrated above are all self-explanatory, comparing the forecasting errors with its actual value but not with another model. Due to large shocks occurring over the forecast horizon, they bring more noise and make the forecasting competition harder. One way is to discard those large changes. Another way is to use the relative error, which is generated by comparing the model with a benchmark model, usually random walk. This kind of measures is named as *relative error measures* (Armstrong & Collopy 1992:71). The well-known one is Theil's *U*-coefficient measure, which compares the MSE for a proposed model with the MSE for the random walk.

$$(29) \quad \text{Theil} - U = \frac{\sum_{i=1}^T (\hat{\sigma}_i - \sigma_i)^2}{\sum_{i=1}^T (\sigma_{i-1} - \sigma_i)^2}$$

When U -coefficient is less than 1, it means that the proposed model is better than naïve method. When U -coefficient is equal to 1, the proposed model has the same forecasting power with random walk model. If U -coefficient is larger than 1, it can be interpreted as the proposed model is even worse than naïve method. The closer the U -coefficient value approaches to zero, the better the model is. Theil's U -coefficient is not widely used due to its difficult interpretation. The same as MAPE, it doesn't exist an upper bound, so a few extremely large errors can readily distort the comparison results (Makridakis & Hibon 1979:40). Additionally, it's hard to interpret the practical meaning of the Theil's U -coefficient, since it is the forth power of the observed returns.

4.1.2. Asymmetric error measures

Unlike return forecasting, volatility prediction is somehow related to risk. Insufficient estimation of risk may cause a fatal loss, especially in turbulent time. Whereas over-prediction on volatility will not bring such large harms as under-prediction does. It is sensible to penalize heavier when negative forecasting errors come out and vice versa. This view is also advocated by some researchers (Granger & Newbold 1986:125-126; Christoffersen & Diebold 1996:561; Christoffersen & Diebold 1997:808) as well as practitioners (Aretz, Bartram & Pope 2009:19). Therefore, asymmetric loss functions have been developed. Pagan and Schwert (1990) propose logarithmic loss function (LL). With this natural logarithmic transformation, even the impact of outliers can be reduced (Poon & Granger 2003:493).

$$(30) \quad LL = \frac{1}{T} \sum_{i=1}^T [\ln(\sigma_i) - \ln(\hat{\sigma}_i)]^2$$

Bollerslev and Ghysels (1996) define a heteroskedasticity-adjusted version of MSE (HMSE). HMSE does a good job to capture the asymmetric reality, whereas Poon and Granger (2003:493) think that this type of performance measure is inappropriate when mainly concerning the absolute magnitude of the forecast errors.

$$(31) \quad HMSE = \frac{1}{T} \sum_{i=1}^T \left(\frac{\sigma_i}{\hat{\sigma}_i} - 1 \right)^2$$

The quasi-likelihood-based (QLIKE) loss is named for its close relation to the Gaussian likelihood estimation. QLIKE is the relative new loss function, but attracts lots of researchers' notice, such as Becker and Clements (2008), Brownlees, Engle & Kelly (2009), Patton (2006), Patton & Sheppard (2008).

$$(32) \quad QLIKE = \frac{\sigma_i}{\hat{\sigma}_i} + \log \hat{\sigma}_i$$

In summary, these loss functions have their own advantages and disadvantages in the different scopes of applications. As Granger and Pesaran (2000) point out, the applicability of any loss function hinges on the purpose of the forecast.

4.2. Regression based evaluation

Regression based evaluation is to run a regression of the volatility proxy, usually squared returns, on the volatility forecasts, then check the information content of volatility forecasts. The regression function is as follows

$$(33) \quad \sigma = \alpha + \beta \hat{\sigma} + \varepsilon$$

Where σ is actual volatility, $\hat{\sigma}$ is the volatility prediction, ε is the estimation error, α is the estimation bias, β is the coefficient. If the forecast is perfect, then the intercept

α should be zero, the coefficient β ought to be one. That means this prediction is unbiased and efficient. However, due to imperfect proxy of latent volatility and model inability or some other reasons, the coefficient β is unlikely to be one. In practice, it just needs to be positive. Notice that biasness and predictive power are not the same. A biased forecast can have predictive power, only if the bias can be corrected. However, an unbiased forecast with big forecasting error is worthless. (Poon & Granger 2003:491). When testing the predictability, one does not only depend on the coefficient β , while concerns more on R^2 . The higher the value of R^2 is, the more the forecast is preferred.

$$(34) \quad R^2 = 1 - \frac{SS_{err}}{SS_{tot}}$$

$$SS_{tot} = \sum_{i=1}^T (\sigma_i - \bar{\sigma}_i)^2$$

$$SS_{err} = \sum_{i=1}^T (\sigma_i - \hat{\sigma}_i)^2$$

$$SS_{reg} = \sum_{i=1}^T (\hat{\sigma}_i - \bar{\sigma}_i)^2$$

Where R^2 is the coefficient of determination, SS_{err} is the sum of squares of residuals, SS_{reg} is the regression sum of squares, SS_{tot} is the total sum of squares, $\bar{\sigma}_i$ is the mean of actual volatility, $\bar{\hat{\sigma}}_i$ is the mean of volatility forecasts. This method follows slightly different logic from the loss function. It tests in which degree the forecasts can explain the actual volatility, while loss functions focus on whether the forecasts are accurate compared to actual volatility. In practice, squared returns are usually as the volatility proxy. As mentioned in 2.1.1., squared returns are very noisy measurements of latent volatility. Excessive variations may bias down the value of R^2 (Alexander 2001:123). When the imperfect volatility proxy is used, to some extent, low R^2 does not mean that the model is inaccurate (Alexander 2001: 124). The use of realized volatility may reduce the noise and force R^2 to approach to its true value.

5. DATA AND METHODOLOGY

This chapter begins to look at the empirical evidence of the volatility forecasting issue. Firstly, the data selection and data properties will be introduced. Secondly, the volatility proxy and different volatility forecasting methods used in this thesis will be calculated. At last, evaluation measures chosen to compare the model accuracy will be presented.

5.1. Data

There are two data series examined here. One is the VIX index as the representative of implied volatility; the other is the underlying stock index price (S&P 500 index), which is used to compose the monthly volatility proxy. The sample period is from 1st January, 1990 to 31st December 2009.

(1) Data for stock index options

Section 2.2.6 mentioned that due to international market integration, asset price and volatility tends to spill over different geographic markets. And U.S. market is usually the leading market, while other markets have no influence on the U.S market (Eun and Shim 1989; Knif & Pynnönen 1998). This study targets U.S market in order to diminish the effect from other markets. Noticeably, S&P 500 options are the most active index option in the U.S. (Whaley 2008:3). The illiquid problem won't bother it. At the same time, S&P 500 index is accepted as the benchmark of the U.S. stock market return, which contains rich macroeconomic information and is widely related to many investment decisions. It is very convenient for portfolio hedging, even for individual stock hedging. This market is most likely to be efficient. Except for that, S&P 500 options are European-style, which can only be exercised at the expiration, making them easier to value (Whaley 2008:2). So VIX index provided by CBOE is chosen as the representative of implied volatility in this study. It is constructed by Chicago Board of Options Exchange from S&P500 index options and derived from near- and next-term

out-of-the-money put and call options, usually in the first and second S&P 500 contract months. As the market's expectation of future volatility, VIX measures 30-day expected volatility of the S&P 500 Index. (Chicago Board Options Exchange, 2009:3-5). VIX index is independent of any model and has no relation to B-S option pricing model (Costas & Athanasios 2009:5). This makes VIX avoid suffering the problems caused by B-S model misspecifications.

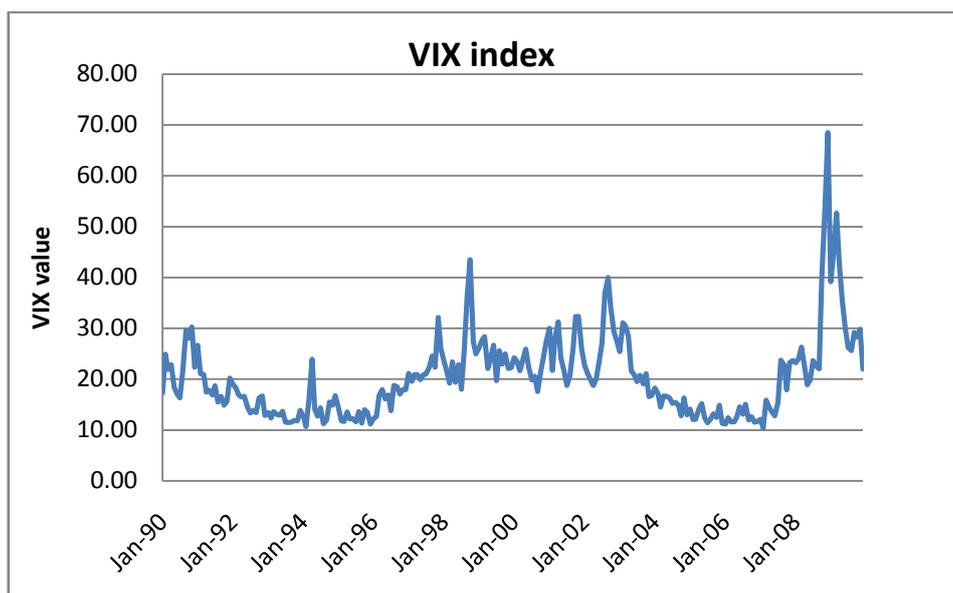


Figure 8. VIX index values from January 1990 to December 2009.

Figure 8 gives the general view of the implied volatility process. VIX moves up and down, but goes around the mean level about the value of 20%. From January 1991 to January 1997, it stayed at the relative low level, while, across the following five years, it stood at the higher stage. Then tranquil time came at the beginning of 2004 until September 2007.

(2) Data for S&P 500 index

Since the latent volatility process cannot be observed, S&P 500 index daily prices on the same sample period are chosen to constitute the volatility proxy and estimate

parameters of the econometric models. Daily returns are calculated from the S&P 500 index daily close price obtained from Yahoo finance website. Following the most common way, return is defined as the natural logarithm of price on successive days, $r_t = \ln(P_t/P_{t-1})$.

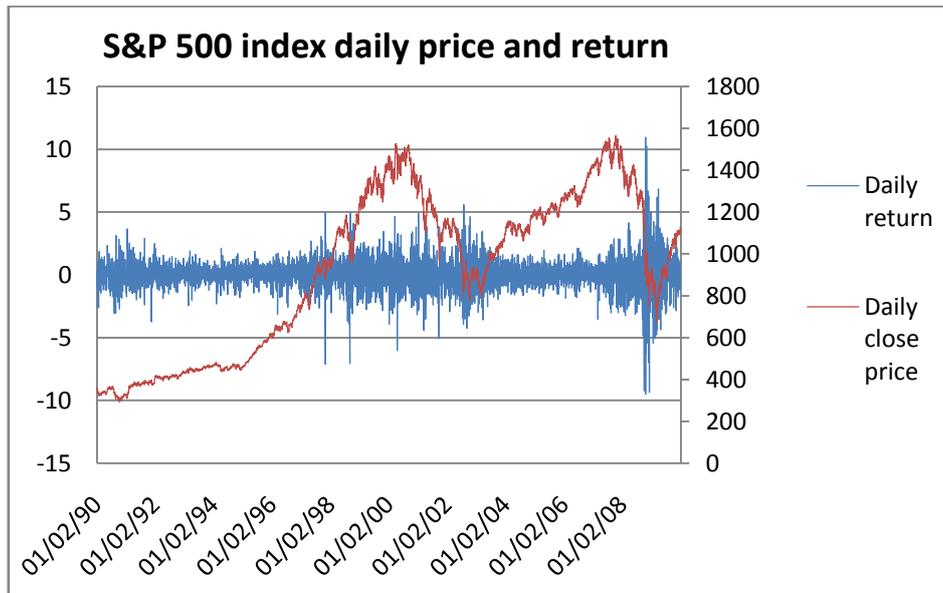


Figure 9. S&P 500 index daily prices and returns from January 1990 to December 2009.

Figure 9 demonstrates how S&P 500 index develops through these twenty years. It experienced a remarkable growth through January 1990 and September 2000 from the lowest point (295.46) to the highest point (1527.46), and then suffers a huge decline between October 2000 and September 2003. It is familiar as the IT bubble burst in 2000. After this recession, the market recovered and continued growing until September 2007. The shock happened in September 2007 is well known by almost everyone, called subprime mortgage crisis, later becomes global financial crisis. It is easy to find that huge price decline is accompanied by the great increase of volatility.

After knowing about evolution of S&P 500 index for past twenty years, now it is time to check the statistical properties of the return series. The mean of the daily returns is

0.022441%, which is close to 0. But the extreme values deviate far from the mean, respectively 10.9572% for maximum and -9.469514% for minimum. The standard deviation for daily data is 1.172553%. The skewness value is negative (-0.198609), which means that S&P 500 return distribution is a little bit asymmetric with a small left fat tail. The high positive kurtosis (12.17409) is worth highlighting. It coincides with the fat tails and a high peak characteristic mention in section 2.2.1. And Jarque-Bera value strongly rejects the normal distribution hypothesis.

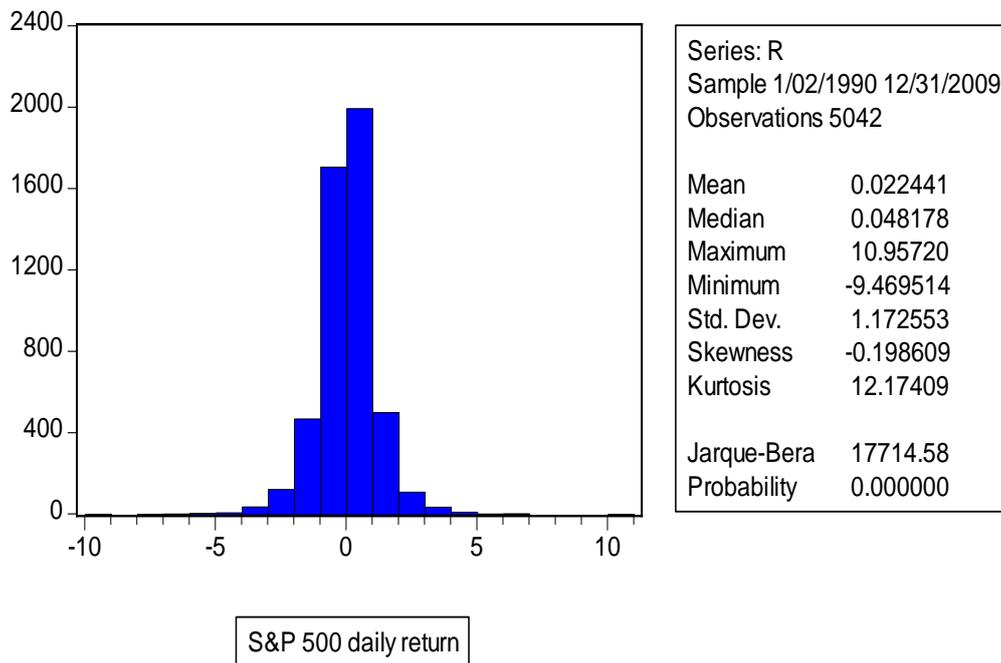


Figure 10. Distribution and descriptive statistics of S&P 500 index daily returns.

In random walk and *RiskmetricsTM* model, there is no stationary requirement on the data process. However, when GARCH group models are taken into account, the return process must be stationary. Thus, before generating the volatility forecasts, one more thing needed to be done is to run the Dickey-Fuller test for examining whether the return process is stationary. If not, GARCH (1, 1) and GJR (1, 1) models cannot be used here. The test is run by EViews 5. The result is reported in table 4. The unit-root hypothesis is strongly rejected (p-value=0.0001) at any confident level. That means that

S&P 500 daily return process is definitely stationary. Under this situation, GARCH (1, 1) and GJR (1, 1) model can be employed to forecast the future volatility.

Table 4. The result of Dickey-Fuller test for S&P 500 index daily returns.

Null Hypothesis: RETURN has a unit root		
Exogenous: Constant		
Lag Length: 1 (Automatic based on SIC, MAXLAG=31)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-54.81636	0.0001
Test critical values: 1% level	-3.431466	
5% level	-2.861918	
10%level	-2.567014	
*MacKinnon (1996) one-sided p-values.		

(3) Sampling Procedure

Christensen and Prabhala (1998) criticize that earlier studies conducted by Day and Lewis (1992), Lamoureux and Lastrapes (1993), Canina and Figlewski (1993) all suffer the overlapping sampling problem, when they construct their daily data sample. This overlapping sampling may cause a serial problem of serial correlation. It tends to favor the historical volatility forecasts as an efficient prediction of future volatility, because the correlations easily bias the standard error downwards in the coefficient estimation. (Yu, Liu & Wang 2010:3). This study adopts a non-overlapping sample on a monthly basis to conquer the problem. Since VIX index predict the fixed horizon-30days, using monthly data is also a good choice to overcome the mismatch problem. Therefore, VIX index on the first trading day of each month is used as the prediction for every month. For instance, 2nd January 1990 is the first trading day in January 1990. VIX index on this day is chosen as the prediction for the whole January and then VIX value on 1st February 1990 for the whole February and so on.

In addition, the whole sample is divided into two parts. The data from January 1990 to December 1999 is for the in-sample estimation. The sample through January 2000 and December 2009 is for out-of-sample comparison. The parameters of the econometric models are not constant, which varies through the time. As regarding to MBFs, the parameters should be estimated by rolling window method. In this study, the estimation window is kept for ten years. When generating a new forecast, the oldest information is discarded, at the same time, including one up-to-date monthly data. The model parameters are estimated renewedly for each forecast.

5.2. Methodology

This part will introduce the calculation of actual volatility and the set of forecasting models as well as the evaluation criteria used in this study.

5.2.1. Computing actual volatility

Since the latent volatility process is invisible, the actual volatility can be only represented by a proxy. Section 2.1.1. introduces and compares three popular methods of measuring the actual volatility. Without a doubt, realize volatility is a highly efficient and unbiased proxy. Although high-frequency data is not available in this study, the author adopts the method mimicking realized volatility method. Andersen et al. (2003) propose the realized volatility as an empirical measure of daily return variability by summing five- or fifteen-minutes intraday squared returns. Monthly volatility is needed in this study. It can be obtained by summing daily squared returns on that month.

$$(35) \quad \sigma_T = \sqrt{\sum_{t=1}^{N_T} r_t^2}$$

$$(36) \quad \sigma_A = 100 * \sqrt{12} \sigma_T$$

Define r_t is the daily return on day t and N_T is the number of trading days on month T .

In this way, monthly realized volatilities are obtained. The formula (36) is to annualize the monthly data. In figure 11, circles mark the relative high volatility periods, which is typical volatility-clustering phenomenon. Volatility deviates from the mean level, nevertheless, it always goes around the mean. This mean-reversion is also existed in this study sample. This convinces the author to include some sophisticated models for capturing these features.

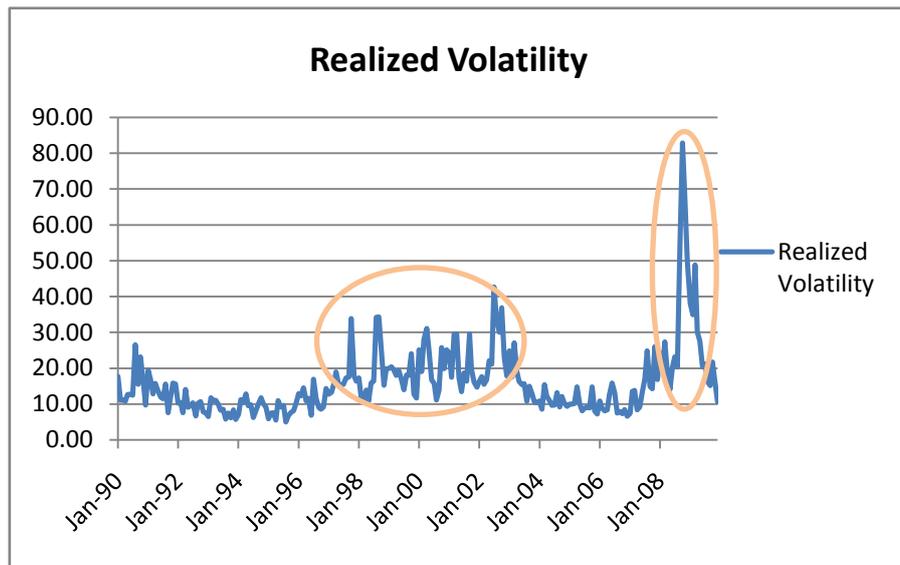


Figure 11. Annualized monthly volatility of S&P 500 index through January 1990 and December 2009.

5.2.2. Models in the competition

Chapter 3 introduces enormous econometric models for capturing different characteristics of volatility. Based on previous literature review and features of data process, this study chooses the models-Random walk, *Riskmetrics*TM model, GARCH (1, 1), GJR (1, 1) as the representatives of MBFs. Random walk is the simplest model chosen as the benchmark. *Riskmetrics*TM model is an excellent practical model used by professional financial stuffs. It can be employed to model the long-memory feature. It is well known that GARCH (1, 1) model performs well in the in-sample estimation

and is widely used in academic researches, which can capture clustering as well as mean-reversion in volatility process. GJR (1, 1) model is introduced in order to reflect asymmetric characteristic in the volatility procedure.

The first model considered in this study is the easiest one-random walk. In this case, random walk model assumes that the best volatility forecast for next month is the volatility of this month. With this way, the parameter estimation is not needed. The one-lagged observed volatility, σ_{t-1} , is naturally used as the forecast at the time t . For the monthly horizon, the forecast needs to be scaled by the square root of 30 (assuming 30 calendar days per month).

$$(37) \quad \hat{\sigma}_t(RW) = \sigma_{t-1}$$

Similarly with random walk model, *Riskmetrics*TM forecasting model has the predetermined coefficients without the in-sample parameter estimation. *Riskmetrics*TM prediction at time t is composed by one-lagged period forecast, $\hat{\sigma}_{t-1}$, and squared return on time $t-1$. As reported in the *Riskmetrics*TM Technical Document (1996), for monthly forecasts, 0.97 is the appropriate value for the coefficient of $\hat{\sigma}_{t-1}$, and then 0.03 is the suitable value for the coefficient of previous day's squared return. By using the daily data, firstly the one day *Riskmetrics*TM forecasts are calculated, and then multiplied by the square root of 30. *Riskmetrics*TM model is the same as IGARCH (1, 1) without the constant, which has no constant unconditional variance and can be used to reflect the long memory existed in volatility process.

$$(38) \quad \hat{\sigma}_t(RM) = \sqrt{0.97\hat{\sigma}_{t-1}^2 + 0.03r_{t-1}^2}$$

As the most popular model in GARCH group, GARCH (1, 1) is relatively easy to estimate and generally has robust coefficients that are interpreted naturally in terms of long-term volatilities and short-run dynamics (Alexander 2001:75). It is a competitive forecasting model among MBFs. In this study, daily returns in the first 10 years are used

to generate the parameter estimation. By this way, the first day of the out-of-sample prediction (3rd January 2000) is obtained as formula (40) presents. Then the second step is to use the first-day forecast and then produce the second-day prediction as formula (41) describes. After that, repeat the second step for getting the third-day prediction and so on until the last day of a month. Finally, sum all daily forecasts in one month as the monthly prediction. The sample is rolled over month-by-month and always maintains 10-year length window. The model parameters are re-estimated once a month.

$$(39) \quad r_t = \mu + \varepsilon_t$$

$$\varepsilon_t \sim (0, \sigma_t^2)$$

$$(40) \quad \hat{\sigma}_t(GARCH) = \sqrt{\omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2}$$

$$(41) \quad \hat{\sigma}_{t+k}(GARCH) = \sqrt{\omega + (\alpha_1 + \beta_1) \hat{\sigma}_{t+k-1}^2} \quad (k > 1)$$

Standard GARCH model is capable of modeling volatility clustering and mean-reversion characteristics. However, it fails to capture the volatility asymmetry. Thus, asymmetric GARCH model is called for. There are different forms of asymmetric GARCH models as mentioned on page 43-45. Engle and Ng. (1993) run a competition among EGARCH, GJR and some other symmetric ARCH models to test that which one is the best candidate for modeling the impact of news. They find that GJR model outperforms all other models. (Engle & Ng. 1993:165). Then Hagerud (1997) compare the performance of seven asymmetric ARCH and GARCH models by 45 Nordic stocks. These models are EGARCH, GJR, TGARCH, A-PARCH, GQARCH, VS-ARCH and LS-TGARCH. They conclude that GJR, TGARCH and GQARCH are superior in estimating the asymmetric dynamic of conditional variance (Hagerud 1997:10). Based on the previous studies, here chooses GJR (1, 1) model as the representative of asymmetric models. GJR (1, 1) forecast is composed by three parts. The first procedure is the in-sample parameter estimation in which the forecast for time $t+1$ has been already contained as formula (42) presents. Then following the formula (43) can generate the forecasts for time $t+2, t+3 \dots t+k$, where k is the last day in one month. The last step is to sum these daily forecasts from $t+1$ until $t+k$ as a monthly forecast. The sampling procedure and parameter re-estimation follows the same way as GARCH (1,

1) model does. The return process is identified with formula (39) shows. Now GJR (1, 1) forecasts are given below:

$$(42) \quad \hat{\sigma}_t = \sqrt{\omega + \alpha \varepsilon_{t-1}^2 + \alpha^- S_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1}}$$

$$(43) \quad \hat{\sigma}_{t+k} = \sqrt{\omega + \left(\frac{1}{2}(\alpha + \alpha^-) + \beta\right) \hat{\sigma}_{t+k-1}^2} \quad (k > 1)$$

All MBFs should be annualized in percentage forms by multiplying by $100 * \sqrt{12}$.

Implied volatility is the market expectation of future volatility. It is the inherent volatility prediction without parameter estimation. As far as VIX index is concerned, it always forecasts 30 days forward, while the value reports on every trading day. Here uses the index value on the first day of one month as the prediction for the whole month. The reason has been discussed in page 68.

$$(44) \quad \hat{\sigma}_t(IV) = VIX_{1st\ day, t}$$

5.2.3. The evaluation of predictabilities

This studies use both loss functions and regression based evaluation as evaluation criteria. There are numerous loss functions for predictability comparisons. It is hard to say which one is the most suitable for volatility forecasting evaluation. So the most common and widely used loss functions are employed in this study. As far as accuracy measures are concerned, Makridakis (1993:527) recommends mean absolute percentage error (MAPE) from the theoretical and practical point of view. Regarding to comparisons of univariate and multivariate volatility forecasts, the most robust loss functions are MSE and QLIKE, even when the imperfect volatility proxy is used (Brownlees, Engle & Kelly 2009:9; Patton 2006: 14; Patton & Sheppard 2008: 22-23). The loss functions chosen here are (1) mean squared error (MSE) (referred to formula (27)), (2) mean absolute percentage error (MAPE) (referred to formula (26)), (3) quasi-

likelihood-based loss (QLIKE) (referred to formula (32)). The best forecast should be the one with the lowest forecasting error, in another word, with the minimal values of the loss functions.

Regression based evaluation is also the popular method used in previous studies. Following this way, this study runs the regression of realized volatility on the forecast volatility and estimates the coefficient by OLS. The regression equation is exactly the same as the formula (33). The R -squares of models are then compared. The model with the highest value of R -square is the most accurate forecast. In this study, the author also wonders to know whether the efficiency in U.S. option market has been improved after 2007 financial crisis. This aim is achieved by comparing the change of R -squares and coefficients in pre-crisis sample and after-crisis sample. The available data after 2007 financial crisis is from Jan 2008 to Dec 2009. For giving the same power to two testing samples, the pre-crisis sample is chosen from January 2005 to December 2006 with the same amount of observations.

6. RESULTS

The main interest of this study is to compare forecasting performances of implied volatility and econometric models mentioned above. The main results are reported in table 5 and 6. Table 5 presents the comparative results for ten-year's (2000-2009) monthly forecasts by using MSE, MAPE and QLIKE loss functions. When conducting the comparison by loss functions, the model with the minimal forecasting error is favorable. Relative to the best forecast, there is no absolute conclusion. MSE and QLIKE criteria recommend GJR (1, 1) model as the most accurate model, while MAPE points that implied volatility is actually superior. MSE and QLIKE even rank implied volatility lower than GARCH (1, 1) model, as the third accurate forecast. It is noticeable that Brownlees, Engle & Kelly (2009), Patton (2006), Patton & Sheppard (2008) conduct a series of researches to prove that MSE and QLIKE are the most efficient evaluation criteria for volatility forecasting comparison, even when imperfect volatility proxy is used. In their researches, they have demonstrated that for this particular issue-volatility forecasting, MSE and QLIKE are better criteria relative to MAPE. Therefore, the author relies more on MSE and QLIKE measurements.

However, among different MBFs, three loss functions all rank GJR (1, 1) model as the first one. The rest ranking order is GARCH (1, 1), *RiskmetricsTM*, and at last random walk. According to the superior performance of GJR (1, 1) model, it can confirm that volatility asymmetry exists in the U.S. stock market. That means that investors are more sensitive to price decline than to price increase. As the most common used practical model, *RiskmetricsTM* model seems less competently to forecast monthly volatility in the U.S. market, relative to GARCH group models. This may be due to the constant coefficient, which is more suitable for short horizon forecasts but not for middle- and long-term. In addition, the extremely large QLIKE value of random walk model is very impressive. It is caused by some outliers existed in the forecasting data set, which are close to zero. Because random walk forecast uses squared return on the last day of month $t-1$ and then multiply with 30 as the forecast of month t . The returns of some days are close to zero. Even after multiplied by 30 and then annualized, it still

approaches to zero. If replaced these outliers by sample average value, the QLIKE value of random walk model decrease to 1960.520592. It is much less than the original one, but it is still quite high compared with other forecasts. Therefore, it is reasonable to conclude that for the monthly horizon, the sophisticated models performance better than simple historical models.

Table 5. Results of the evaluation by loss functions for monthly forecasts (Jan 2000-Dec 2009).

<i>Forecasting methods</i>	<i>MSE</i>	<i>MAPE</i>	<i>QLIKE</i>
Random walk	92.13375547	43.19519	1287203.294
<i>RiskmetricsTM</i>	58.7758969	27.17587	- 284.5610494
GARCH (1, 1)	48.35899221	25.68331	- 290.6233808
GJR (1, 1)	45.14803835	25.0862	- 291.4190116
Implied volatility	54.1583545	23.04208	- 288.4122312

Table 6. Results of regression based evaluation for monthly forecasts (Jan 2000-Dec 2009).

<i>Forecasting methods</i>	<i>R²</i>	<i>α</i>	<i>β</i>
Random walk	0.335578	12.06077 (0.0000)	0.475991 (0.0000)
<i>RiskmetricsTM</i>	0.576138	2.916866 (0.0452)	0.690904 (0.0000)
GARCH (1, 1)	0.651259	1.771482 (0.1806)	0.858281 (0.0000)
GJR (1, 1)	0.674415	1.75225 (0.1653)	0.879591 (0.0000)
Implied volatility	0.609437	- 2.342135 (0.1711)	0.960649 (0.0000)

Table 6 presents the results of regression based evaluation. When R^2 values are checked, the author gets the similar results given by MSE and QLIKE loss functions. Regression based evaluation again presents GJR (1, 1) model as the best volatility

forecast with the highest R^2 equal to 0.674415. Apart from that, GARCH (1, 1) model also has higher R^2 value (0.651259) than implied volatility does (0.609437). Nevertheless, implied volatility forecasts are more accurate rather than *Riskmetrics*TM (0.576138) as well as random walk forecasts (0.335578). It is noticeable that although implied volatility has the lower R^2 value than GJR (1, 1) model, its coefficient is the highest among these competitors, which is 0.960649, quite close to one. Furthermore, the author wonders to know whether the coefficient β is equal to one. Wald test is employed to test this restricted hypothesis. Table 7 reports that the hypothesis of the unit coefficient is accepted with p-values of F-statistic and Chi-square equal to 0.5794 and 0.5783 respectively. This result can be interpreted as implied volatility is an efficient forecast, while it still commits larger forecasting errors rather than GJR (1, 1) model produces. For MBFs, regression based evaluation reports the same ranking order as loss functions do, which further confirms that more complicated models are more competent than simple historical models.

Considering the results of loss functions and regression based evaluation, the author concludes that GJR (1, 1) model is the most accurate forecast, even superior to implied volatility. Hypothesis 1 is rejected.

Table 7. Wald test for the unit coefficient of implied volatility.

Wald Test:			
Test Statistic	Value	df	Probability
F-statistic	0.308955	(1, 118)	0.5794
Chi-square	0.308955	1	0.5783
Null Hypothesis Summary:			
Normalized Restriction (=0)	Value		Std. Err
- 1+C(2)	- 0.039351		0.070795

The second target of this study is to test whether the efficiency of U.S option market has been improved after 2007 financial crisis. Table 8 and 9 presents separately the information content of implied volatility in pre-crisis (2005-2006) and after-crisis

period (2008-2009). The empirical result finds that R^2 value in pre-crisis period is quite low (0.095400), which is close to zero. It means that implied volatility forecast almost has no forecasting power during 2005-2006. Meanwhile, the coefficient of VIX is even not statistically significant (p-value=0.1420). The statistical measurement cannot reject the null hypothesis on which the coefficient of VIX equals to zero. That implies that implied volatility may have no relation with realized volatility in the pre-crisis sample. Oppositely, VIX index forecast performs much better in the after-crisis period. The R^2 value soars to 0.408334 with the much higher coefficient of VIX (0.913535). In this time, p-value equals to zero, which is strongly statistically significant. This obvious difference demonstrates that the predictive ability of implied volatility has increased a lot after 2007 financial crisis, which is in line with the previous studies on 1987 crash and 1995 Japanese crisis. With this positive evidence, hypothesis 2 is supported.

Table 8. Regression of VIX on the realized volatility from Jan2005 to Dec2006.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.756930	5.3122	0.3307	0.7440
VIX	0.634996	0.4169	1.5232	0.1420
R-squared	0.095400			
Adjusted R-squared	0.054282			
S.E. of regression	2.535481			
Sum squared resid	141.4306			
Log likelihood	- 55.339587			

Table 9. Regression of VIX on the realized volatility from Jan 2008 to Dec 2009.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.703746	8.0725	0.0872	0.9313
VIX	0.913535	0.2344	3.8965	0.0008
R-squared	0.408334			
Adjusted R-squared	0.381441			
S.E. of regression	14.23910			
Log likelihood	-96.754184			

7. SUMMARY AND LIMITATIONS

The purpose of this study is to compare the predictive ability between implied volatility and MBFs in the U.S. stock market. VIX index proposed by Chicago Board Options Exchange in 2003 is used as the representative of implied volatility. The over-lapping problem is noticed and avoided in the VIX index sampling procedure. The econometric models implemented here include random walk, *Riskmetrics*TM, GARCH (1, 1) and GJR (1, 1), which capture main characteristics of the volatility process. The parameters of GARCH and GJR models are re-estimated by rolling samples. In order to find the best forecasting method, four different evaluation criteria are chosen. MSE, MAPE and R^2 are common used measurements. As a relative new method, QLIKE loss function is used because of its outstanding reliability and efficiency, especially for volatility forecasting comparison. Since the results highly depend on the evaluation criteria chosen, the adoption of QLIKE loss function can increase the reliability of this study. This studying sample is from Jan 1990 to Dec 2009, which covers both long enough history and recent turbulent crisis stages. Monthly forecasting horizon is chosen for avoiding the mismatch problem. For decreasing noise, monthly actual volatility is produced by summing daily squared returns in each month. Based on the intensive literature reviews and common sense, two hypothesis are proposed, 1) Implied volatility performs better on the volatility forecasting than MBFs do. 2) The efficiency of option market improved after 2007 financial crisis.

The first hypothesis is rejected by three out of four evaluation criteria. MSE, QLIKE and regression based evaluation all find that implied volatility underperforms relative to GJR (1, 1) and GARCH (1, 1) models. Under these three criteria, GJR (1, 1) model is highlighted as the most accurate method. Even under MAPE measurement, GJR (1, 1) model is just slightly inferior to implied volatility. This result proves that GARCH group model is much powerful on volatility forecasting. This study gives controversial evidence relative to the research conducted by Blair, Poon and Taylor (2001), which asserts that implied volatility yields GJR using both low- and high-frequency data. However, in line with Lamoureux and Lastrapes (1993), Vasilellis and Meade (1996),

Christensen and Prabhala (1998), the author find that implied volatility dominates simple historical volatility models under all evaluation criteria.

Unlike hypothesis one, the second hypothesis is strongly supported. The U.S. option market achieved remarkable improvement after 2007 financial crisis. It is reasonable because huge losses in financial markets, like a ringing alarm, alert both institutional and individual investors to pay more attention on the risk control. Investors tend to modify their risk-management tools and develop their forecasting ability. Comparatively speaking, in tranquil time, profits are usually given more attention. Investors are prone to chase higher and higher returns while neglect the risk behind the abnormal profits. Implied volatility as the market expectation of volatility is easily influenced by investors' behaviors. So the information contents of implied volatility are entirely different in turbulent and steady periods.

Although clear conclusions have been obtained in this study, there are still some points needed to be studied further. The empirical evidence prefers MBFs to implied volatility here. Whereas it is well known that VIX index bases on the risk-neutral assumption, which is widely used by the implied volatility calculation but not consistent with the real markets. Relatively econometric models can be adjusted to capture the risk premium with the past actual data. This can be interpreted as some extent of unfairness. Therefore, the risk premium should be taken into account in further study to examine whether the risk premium exists and it affects the predictability of implied volatility or not. Additionally, when testing the second hypothesis, at this study time there are just two years data (2008-2009, 24 observations) available for the after-crisis sample. It is not sufficient to generate powerful statistical results. With the passing of time, more observations should be included to increase the statistical reliability.

REFERENCES

- Alexander, Carol (2001). *Market Models: A Guide to Financial Data Analysis*. Chichester: John Wiley & Sons Ltd.
- Alexander, Carol (2008). *Market Risk Analysis Volume II: Practical Financial Econometrics*. Chichester: John Wiley & Sons Ltd.
- Amin, Kaushik & Victor Ng (1997). Inferring future volatility from the information in implied volatility in Eurodollar options: a new approach. *Review of Financial Study* 10, 333-367.
- Andersen, Torben & Tim Bollerslev (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39:4, 885-905.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold & Paul Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 1:2, 579-625.
- Aretz, Kevin., Sohnke M. Bartram & Peter F. Pope (2009). Asymmetric Loss Functions and the Rationality of Expected Stock Returns. *International Journal of Forecasting*. Forthcoming. Available at internet: <URL: <http://ssrn.com/abstract=889323>>.
- Armstrong, J. Scott., Fred Collopy (1992). Error measures for generalizing about forecasting methods: empirical comparisons. *International Journal of Forecasting* 8, 69-80.

- Baillie, Richard T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5-59.
- Bali, Turan G. & K. Ozgur Demirtas (2008). Testing mean reversion in financial market volatility: evidence from S&P 500 index futures. *The Journal of Futures Markets* 28:1, 1–33.
- Barclay, Michael J., Robert H. Litzenberger & Jerold B. Warner (1990). Private information, trading volume, and stock-return variances. *The Review of Financial Studies* 3, 233-253.
- Beckers, Stan (1981). Standard deviations implied by options prices as predictors of future stock price variability. *Journal of Banking & Finance* 5:3, 363-382.
- Becker, Ralf & Adam E. Clements (2008). Are combination forecasts of S&P 500 volatility statistically superior? *International Journal of Forecasting* 24, 122-133.
- Becker, Ralf., Adam E. Clements & Andrew McClelland (2009). The jump component of S&P 500 volatility and the VIX index. *Journal of Banking & Finance* 33, 1033-1038.
- Becker, Ralf., Adam E. Clements & Christopher A. Coleman-Fenn (2009). Forecast performance of implied volatility and the impact of the volatility risk premium. *Working Paper*. Available from internet: <URL: <http://www.ncer.edu.au/papers/documents/WPNo45.pdf>>.
- Becker, Ralf., Adam E. Clements & Scott I. White (2006). On the informational efficiency of S&P 500 implied volatility. *The North American Journal of Economics and Finance* 17,139-153.

- Becker, Ralf., Adam E. Clements & Scott I. White (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking & Finance* 31, 2535-2549.
- Black, Fisher & Myron Scholes (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance* 2, 399-417.
- Black, Fisher & Myron Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637-659.
- Black, Fischer (1976). Studies of stock price volatility changes. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section. 177-181.
- Blair, Bevan J., Ser-Huang Poon & Stephen J. Taylor (2001). Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics* 105, 5-26.
- Bodie, Zvi & Robert C. Merton (1995). The informational role of asset prices: the case of implied volatility. *Working Paper 95-063*. Available at internet: <URL: <http://ssrn.com/abstract=6568>>.
- Bollerslev, Tim (1986). Generalized autoregressive conditional heteroskedasticity, In: *ARCH selected readings*, 42-60. Robert F. Engle. New York ect.: Oxford University Press.
- Bollerslev, Tim & Eric Ghysels (1996). Periodic autoregressive conditional heteroskedasticity. *Journal of Business & Economic Statistics* 14, 139-151.

- Bollerslev, Tim., Gibson Michael S. & Zhou Hao. (2008). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *FEDS Working Paper* No. 2004-56; AFA 2006 Boston Meetings Paper; *Journal of Econometrics*, Forthcoming. Available at internet: <URL: <http://ssrn.com/abstract=614543>>.
- Braun, Phillip., Daniel Nelson & Alain Sunier (1995). Good news, bad news, volatility and betas. *Journal of Finance* 50, 1575-1603.
- Braisford, Tomthy J. & Robert W. Faff (1996). An evaluation of volatility forecasting techniques. *Journal of Banking & Finance* 20, 419-438.
- Britten-Jones, M. & A. Neuberger (2000). Option prices, implied price processes, and stochastic volatility. *Journal of Finance* 55, 839-866.
- Brownlees, Christian T., Robert F. Engle & Bryan T. Kelly (2009). A practical guide to volatility forecasting through calm and storm. *Working Paper*. Available at Internet: <URL: <http://ssrn.com/abstract=1502915>>.
- Canina, Linda & Stephen Figlewski (1993). The information content of implied volatility. *The Review of Financial Studies* 6:3, 659-681.
- Chernov. Mikhall (2001). Implied volatilities as forecasts of future volatility, time-varying risk premia, and returns variability. *AFA 2002 Atlanta Meetings*. Available at internet: <URL: <http://ssrn.com/abstract=292979> or doi:10.2139/ssrn.292979>.
- Chicago Board Options Exchange, Incorporated (2009). The *CBOE volatility index*[®]-*VIX*[®]. Available at internet: <URL: <http://www.cboe.com/micro/vix/vixwhite.pdf>>.

- Chou, Ray Yeutien (1988). Volatility persistence and stock valuations: some empirical evidence using GARCH. *Journal of Applied Econometrics* 3, 279-294.
- Christie, Andrew (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics* 3, 407-432.
- Christensen, B. J. & N. R. Prabhala (1998). The relation between implied and realized volatility. *Journal of Financial Economics* 50, 125-150.
- Christensen, B. J. & Charlotte Strunk Hansen (2002). New evidence on the implied-realized volatility relation. *The European Journal of Finance* 8, 187-205.
- Christoffersen, Peter F. & Francis X. Diebold (1996). Further results on forecasting and model selection under asymmetric loss. *Journal of Applied Econometrics* 11, 561-571.
- Christoffersen, Peter F. & Francis X. Diebold (1997). Optimal prediction under asymmetric loss. *Econometric Theory* 13, 808-817.
- Chiras, D.P. & S. Manaster (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics* 6:2/3, 213-234.
- Claessen, Holger & Stefan Mittnik (2002). Forecasting stock market volatility and the informational efficiency of the DAX-index options market. *The European Journal of Finance* 8, 302-312.
- Cont, Rama & José da Fonseca (2002). Dynamics of implied volatility surfaces. *Quantitative Finance* 2, 45-60.

- Cootner, Paul H (1964). *The Random Character of Stock Market Prices*. Cambridge: M.I.T. Press.
- Corrado, Charles J. & Thomas W. Miller, JR (2005). The forecast quality of CBOE implied volatility indexes. *The Journal of Futures Markets* 25, 339-373.
- Costas, Siriopoulos & Fassas Athanasios (2009). Implied volatility indices- a review. *Working Paper*. Available at internet: <URL: <http://ssrn.com/abstract=1421202>>.
- Cox, John C., Stephen A. Ross & Mark Rubinstein (1979). Option pricing: a simplified approach. *Journal of Financial Economics* 7:3, 229-263.
- Cumby, Robert., Stephen Figlewski & Joel Hasbrouck (1993). Forecasting volatilities and correlations with EGARCH models. *Journal of Derivatives* 1, 51-63.
- Day, Theodore E. & Craig M. Lewis. (1992). Stock Market volatility and the information content of stock index options. In: *ARCH selected readings*, 332-352. Robert F. Engle. New York ect.: Oxford University Press.
- Dowling. Sean & Jayaram Muthuswamy (2003). The implied volatility of Australian index options. Available at internet: <URL: <http://ssrn.com/abstract=500165>>.
- Edey, Macolm., Graham Elliot (1992). Some evidence on option prices as predictors of volatility. *Oxford Bulletin of Economics & Statistics* 54, 567-578.
- Ederington, Louis H. & Wei Guan (2005). Forecasting volatility. *The Journal of Futures Markets* 25. 465–490.

- Engle, Robert. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the UK inflation. *Econometrica* 50, 987-1008.
- Engle, Robert. F. & Gary. G. J. Lee (1999). A long-run and short-run component model of stock return volatility. In *Cointegration, Causality, and Forecasting: a Festschrift in Honour of Clive W. J. Granger*.475-497. R.F. Engle, H. White, Oxford: Oxford University Press.
- Engle, Robert. F. & Patton, Andrew J. (2001). What good is a volatility model? *Quantitative Finance* 1, 237-245.
- Engle, Robert. F. & Tim Bollerslev (1986). Modeling the persistence of conditional variances. *Econometric Reviews* 5:1, 1-50.
- Engle, Robert. F. & Victor K. Ng (1993). Time-varying volatility and the dynamic behavior of the term structure. In: *ARCH selected readings*, 145-175. Robert F. Engle. New York ect.: Oxford University Press.
- Eun, Cheol S. & Sangdal Shim (1989). International transmission of stock market movements. *Journal of Financial and Quantitative Analysis* 24:2, 241-256.
- Fama, Eugene. F. (1965). The behavior of stock-market prices. *Journal of Business* 38, 34-105.
- Figlewski, Stephen (1997). Forecasting volatility. *Financial Markets, Institutions and Instruments* (New York University Salomon Center) 6, 1-88.
- Figlewski, Stephen and Wang, Xiaozu (2000). Is the 'leverage effect' a leverage effect? *Working Paper*. Available at internet: <URL: <http://ssrn.com/abstract=256109> or doi:10.2139/ssrn.256109>.

- Fleming, Jeff., Barbara Ostdiek & Robert E. Whaley (1995). Predicting stock market volatility: a new measure. *The Journal of Futures Markets* 15, 265-302.
- Fleming, Jeff (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance* 5, 317-345.
- Fouque, Jean-Pierre., George Papanicolaou & K. Ronnie Sircar (2000). Mean-reverting stochastic volatility. *Working Paper*. Available from internet: <URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.50.237>>.
- Forbes, Kristin J. & Menzie D. Chinn (2004). A decomposition of global linkages in financial markets over time. *The Review of Economics and Statistics* 86, 705-722.
- Franses, Philip Hans & Dick Van Dijk (1996). Forecasting stock market volatility using non-linear GARCH models. *Journal of Forecasting* 15:3, 229-235.
- Fung, William K. H. & David A. Hiseh (1991). Estimating the dynamics of foreign currency volatility. *Working Paper*. Available from internet: <URL: <http://faculty.fuqua.duke.edu/~dah7/vitae.htm>>.
- Gemmill, Gordon (1986). The forecasting performance of stock options on the London traded options markets. *Journal of Business Finance & Accounting* 13:4. 535-546.
- Giot, Pierre & Sébastien Laurent (2006). The information content of implied volatility in light of the jump/continuous decomposition of realized volatility. *Working Paper*. Available from internet <URL: http://www.core.ucl.ac.be/econometrics/Giot/Papers/IMPLIED4_8.pdf>.

- Glosten, Lawrence., Ravi Jagannathan & David Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779-1801.
- Gonzalez, M. Teresa & Alfonso Novales (2007). Why a volatility index can be useful in the Spanish financial market? *Working Paper*. Available from internet <URL: http://www.uibcongres.org/imgdb/archivo_dpo3655.pdf>.
- Granger, C. W. J. & H. Pesaran (2000). Economic and statistical measures of forecast accuracy. *Journal of Forecasting* 19, 537–560.
- Granger, Clive W. J. & Namwon Hyung (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance* 11:3, 399-421.
- Granger, Clive W. J. & Roselyne Joyeux (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1:1, 15-29.
- Granger, Clive W. J. & Paul Newbold (1986). *Forecasting Economic Time Series*. 2nd Edition. London: Academic Press Inc. (London) Ltd.
- Haan, Wouter J. den. & Scott A. Spear (1998). Volatility clustering in real interest rates: theory and evidence. *Journal of Monetary Economics* 41, 413-453.
- Hagerud, Gustaf E. (1997). Modeling nordic stock returns with asymmetric GARCH models. *Working paper series in Econometrics and Finance No.164*.

- Hamao, Yasushi., Ronald W. Masulis & Victor Ng (1990). Correlations in price changes and volatility across international stock markets. *The Review of Financial Studies* 3, 281-307.
- Hamilton, James. D. & Raul Susmel (1994). Autoregressive conditional heteroskedasticity and change in regime. *Journal of Econometrics* 64, 307-333.
- Hansen, Charlotte Strunk (2001). The relation between implied and realized volatility in the Danish option and equity markets. *Accounting & Finance* 41, 197-228.
- Hilliard, J (1979). The relationship between equity indices on world exchanges. *Journal of Finance* 26, 89-94.
- Hong, Yongmiao (2001). A test for volatility spillover with application to exchange rates. *Journal of Econometrics* 103,183-224.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika* 68, 165-176.
- Hull, John & Alan White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42:2, 281-300.
- Hull, John C. (2006). *Options, Futures, and Other Derivatives*. 6th edition. Upper Saddle River, New Jersey: Pearson Education, Inc.
- Jackwerth, Jens Carsten & Mark Rubinstein (1996). Recovering probability distributions from option prices. *Journal of Finance* LI, 1611-1631.
- Jaffe, Jeffrey & Randolph Westerfield (1985). Pattern in Japanese common stock returns: day of the week and turn of the year effects. *Journal of Financial and Quantitative Analysis* 20, 261-272.

Jarrow, Robert A. & James B. Wiggins (1989). Option pricing and implicit volatilities. *Journal of Economic Survey* 3:1, 59-81.

Jiang, George J. & Yisong S. Tian (2005). The model-free implied volatility and its information content. *Review of Financial Studies* 18, 1305-1342.

Jorion, Phillip (1995). Predicting volatility in the foreign exchange market. *Journal of Finance* 50, 507-528.

Jorion, Phillip (1996). Risk and turnover in the foreign exchange market. In: *The Microstructure of Foreign Exchange Markets*, 19- 40. J. A. Frankel, G. Galli and A. Giovanini, eds. Chicago: University of Chicago Press

Karolyi, G Andrew (1995). A multivariate GARCH model of international transmissions of stock returns and volatility: the case of the United States and Canada. *Journal of Business & Economic Statistics* 13, 11-25.

Knif, Johan & Seppo Pynnönen (1998). Common short-term volatility on international stock markets. *Meddelanden från Svenska Handelshögskolan*. 389

King, Mervyn., Enrique Sentana & Sushil Wadhvani (1994). Volatility and links between national stock markets. *Econometrica* 62, 901-933.

Koutmos, Gregory & G. Geoffrey Booth (1995). Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance* 14, 747-762.

Krylova, Elizaveta., Jussi Nikkinen & Sami Vähämaa (2009). Cross-dynamics of volatility term structures implied by foreign exchange options. *Journal of Economics and Business* 61, 355-375.

- Lamoureux, Christopher & William Lastrapes. (1993). Forecasting stock-return variance: towards an understanding of stochastic implied volatilities. *The review of Financial Studies* 6:2. 293-326.
- Latane, Henry & Richard Rendleman. (1976). Standard deviations of stock price ratios implied in option prices. *Journal of Finance* 31, No.2. 369-381.
- Li, Kai. (2002). Long-memory versus option-implied volatility prediction. *Journal of Derivatives* 9:3. 9-25.
- Lin, Bing-Huei., Ing-Jye Chang & Dean A. Paxson (2008). Smiling less at LIFFE. *Journal of Futures Markets* 28, 57–81.
- Makridakis, Spyros., Michele Hibon (1979). Accuracy of forecasting: an empirical Investigation (with discussion). In *The Forecasting Accuracy of Major Time Series Methods*.35-101. S. Makridakis., A. Andersen., R. Carbone., R. Fildes., M. Hibon., R. Lewandowski., J. Newton., E. Parzen., R. Winkler. Chichester: John Wiley & Sons.
- Makridakis, Spyros (1986). The art and science of forecasting: an assessment and future directions. *International Journal of Forecasting* 2, 15-39.
- Makridakis, Spyros (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting* 9,527-529.
- Mandelbrot, Benoit. (1963). The variation of certain speculative prices. *Journal of Business* 36:4, 394-419.
- Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4, 141-183.

- Mikosch, Thomas & Cătălin Stărică (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics* 86, 378-390.
- Nelson, Daniel (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* 59, 347-370.
- Nishina, Kazuhiko., Maghrebi Nabil. & Kim Moo-Sung (2006). Stock market volatility and the forecasting accuracy of implied volatility indices. *Discussion Papers in Economics and Business No. 06-09*. Graduate School of Economics and Osaka School of International Public Policy.
- Pagan, Adrian R. & G. William Schwert (1990). Alternative models for conditional stock volatility. *Journal of Econometrics* 45, 267-290.
- Patton, A. (2006). Volatility forecast comparison using imperfect volatility proxies. *Technical Report, University of Oxford*. Available from internet: <URL: http://www.business.uts.edu.au/qfrc/research/research_papers/rp175.pdf>.
- Patton, Andrew J. & Kevin Sheppard (2008). Evaluating volatility and correlation forecasts. *Working Paper*. Available from internet: <URL: http://www.kevinsheppard.com/images/c/c6/Patton_Sheppard.pdf>.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53, 61-65.
- Pong, Shiu-Yan., Mark B. Shackleton., Stephen J. Taylor & Xinzhong Xu. (2002). Forecasting currency volatility: a comparison of implied volatilities and AR(FI)MA models. *Working Paper*. Available from internet <URL: <http://ssrn.com/abstract=301981> or doi:10.2139/ssrn.301981>.

- Poon, Ser-Huang (2005). *A Practical Guide for Forecasting Financial Market Volatility*. Chichester: John Wiley & Sons, Ltd.
- Poon, Ser-Huang & Clive W. J. Granger. (2003). Forecasting volatility in financial markets: a review. *Journal of Economic Literature* XLI. 478-539.
- RiskMetricsTM* (1996). *RiskMetricsTM – Technical Document*. 4th ed. JP Morgan & Reuters: New York.
- Rogers, L. C. G. & Satchel S. E. (1991). Estimating variance from high, low, and closing prices. *Annals of Probability* 1, 504-512.
- Rubinstein, Mark (1994). Implied binomial trees. *Journal of Finance* LXIX, 771-817.
- Schmalensee, Richard & Robert R. Trippi. (1978). Common stock volatility expectations implied by option premia. *Journal of Finance* XXXIII:I. 129-147.
- Schwert, G. William (1989). Why does stock market volatility change over time? *Journal of Finance* 44, 1115–53.
- Sentana, Enrique (1995). Quadratic ARCH models. *The Review of Economic Studies* 62: 4, 639-661.
- Shephard, Neil (2005). *Stochastic Volatility Selected Readings*. New York ect.: Oxford University Press.
- Szakmary, Andrew., Evren Ors, Jin Kyoung Kim & Wallace N. Davidson III (2003). The predictive power of implied volatility: evidence from 35 futures markets. *Journal of Banking & Finance* 27, 2151-2175.

- Sørensen, Carsten (1997). An equilibrium approach to pricing foreign currency options. *European Financial Management* 3, 63-84.
- Taylor, Stephen (1986). *Modeling Financial Time Series*. Chichester: John Wiley & Sons.
- Taylor, Stephen J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton: Princeton University Press.
- Tse, Yiuman & G. Geoffrey Booth (1996). Common volatility and volatility spillovers between U.S. and Eurodollar interest rates: evidence from the futures market. *Journal of Economics and Business* 48, 299-312.
- Vasilellis, George A. & Nigel Meade (1996). Forecasting volatility for portfolio selection. *Journal of Business Finance & Accounting* 23. 125-143.
- Wong, H.Y. & Lau, K.Y. (2008). Path-dependent currency options with mean reversion. *The Journal of Futures Markets* 28, 275–293.
- Wei, Shang Jin & Jaffrey A. Frankel (1991). Are option-implied forecasts of exchange rate volatility excessively variable? *NBER Working Paper*. Available from internet <URL: <http://www.nber.org/papers/w3910.pdf>>.
- Whaley, Robert E. (1993). Derivatives on market volatility: hedging tools long overdue. *Journal of Derivatives* 1, 71—84.
- Whaley, Robert E. (2008). Understanding VIX. *Working Paper*. Available from internet<URL: <http://ssrn.com/abstract=1296743>>.

- Xu, Xinzhong & Stephen J. Taylor (1995). Conditional volatility and the informational efficiency of PHLX currency options market. *Journal of Banking & Finance* 19, 803-821.
- Yu, J (2002). Forecasting volatility in the New Zealand stock market. *Applied Financial Economics* 12, 193-202.
- Yu, Wayne W., Evans C.K. Lui & Jacqueline W. Wang (2010). The predictive power of the implied volatility of options traded OTC and on exchanges. *Journal of Banking & Finance* 34, 1-11.
- Zumbach, Gilles O. (2007). The Riskmetrics 2006 Methodology. Available from internet: <URL: <http://ssrn.com/abstract=1420185>>.

APPENDIX 1

Summary of previous empirical evidence pertaining to the comparative performance of implied volatility and MBFs on future volatility forecasting.

Author	Asset (s)	Data Freq	Forecasting Methods	Forecasting Horizon	Results
Latane and Rendleman (1976)	24 stock options from CBOE	W	$Implied_{vega\ weighted}$ $HIS_{4\ years}$		ISD outperforms HIS.
Schmalensee and Trippi (1978)	6 CBOE stock options	W	$Implied_{BS\ call}$	1 week ahead	ISD outperforms HIS. Positive correlation among different stock options.
Chiras and Manaster (1978)	All stock options from CBOE	M	Implied (weighted by price elasticity) $HIS_{20\ mont\ hes}$	20 month ahead	IV outperforms HIS.

Lamoureux and Lastrapes (1993)	Stock options for 10 non-dividend paying stocks (CBOE)	D	<i>Implied</i> _{Hull-White call} <i>HIS</i> _{updated expanding} GARCH	90 to 180 days matching option maturity	IV outperforms HIS. BS model and market imperfect.
Gemmill (1986)	13 UK stocks LTOM options	M	<i>Implied</i> _{ITM} <i>Implied</i> _{ATM, vega WLS} <i>Implied</i> _{OTM, equal} <i>HIS</i> _{20 weeks}	13-21 non-overlapping option maturity	In-the-money call ISD adds predictive power on HIS. Out-the-money options contain no information.
Vasilellis and Meade (1996)	Stock option 12 UK stocks	W	Combine (Implied + GARCH) Implied (various) GARCH EWMA <i>HIS</i> _{3 months}	3month ahead	Implied volatility is the better predictor than historical volatility. Combination between implied and GARCH is best.

Day and Lewis (1992)	S&P 100 OEX option	W	<p><i>Implied</i>_{BS call} (shortest but >7 days, volume WLS)</p> <p><i>HIS</i>_{1 week} GARCH EGARCH</p>	1 week ahead	Short-run market volatility is difficult to predict. GARCH, EGARCH contain incremental information relative to implied volatility.
Canina and Figlewski (1993)	S&P 100 (OEX)	D	<p><i>HIS</i>_{60 calendar days} <i>Implied</i>_{binominal call}</p>	7 to 127 calendar days matching option maturity	Implied volatility has no correlation with future realized volatility.
Christensen and Prabhala (1998)	S&P 100 (OEX)	M	<p><i>Implied</i>_{BS ATM 1-month call} <i>HIS</i>_{18 days}</p>	Non-overlapping 24 calendar days.	Implied volatility dominates HIS. HIS has no additional information.

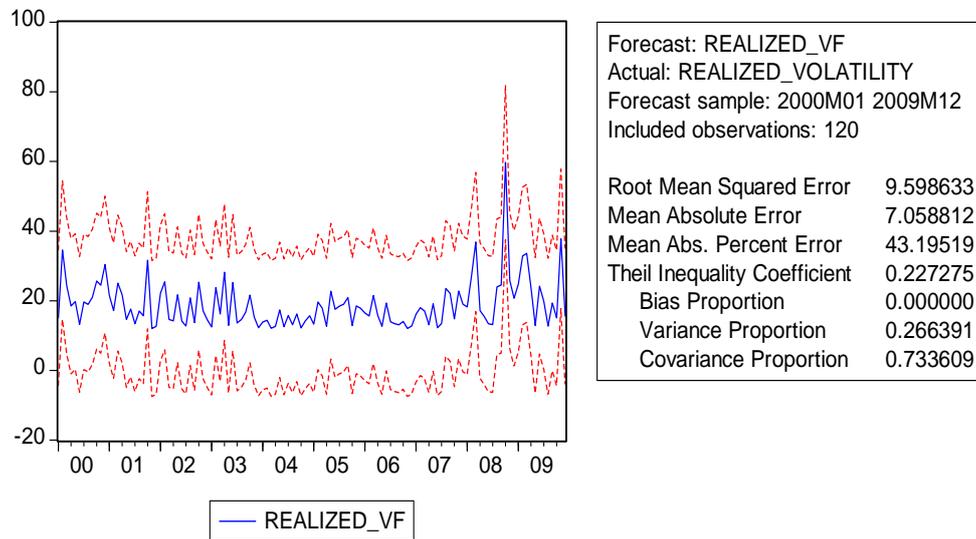
Christensen and Strunk (2002)	S&P 100 (OEX)	M	$Implied_{BS\ ATM,ITM,OTM}^{1-month\ call}$ $HIS_{18\ days}$	Non-overlapping 24 calendar days.	Implied volatility dominates HIS. Implied volatility is unbiased and efficient.
Fleming, Osdiak and Whaley (1995)	VIX (BS model) S&P 100	D	$Implied_{ATM\ call}$ $Implied_{ATM\ put}$ $HIS_{H-L\ 28\ days}$ ARCH GARCH	Option maturity (shortest but >15 days, average 30 days), 1 and 28 days ahead	Implied volatility dominates.
Blair, Poon and Taylor (2001)	VIX (BS model) S&P 100	Tick	$Implied_{BSVIX}$ GJR HIS_{100} (low- and high-frequency data)	1, 5, 10 and 20 days ahead using rolling sample	Implied volatility yields GJR and HIS by using both low- and high-frequency data.

Jiang and Tian (2005)	options on S&P 500 index	Tick	$Implied_{BS}$ $Implied_{model\ free}$ HIS	30, 60, 120, 180 days ahead	Model-free implied volatility yields both BS implied volatility and historical volatility.
Becker, Clemants & White (2006)	VIX (model-free method) S&P 500	D	$Implied_{model\ -free\ VIX}$ RV	22 trading days ahead	Significant positive correlation between the VIX index and future volatility. But implied volatility is not efficient
Becker, Clemants & White (2007)	VIX (model-free method) S&P 500	D	$Implied_{model\ -free\ VIX}$ GJR SV AR ARF RV GARCH	1-, 5-, 10-, 15-, 22-, trading days ahead	Implied volatility does not contain any information that is not captured by model based forecasts.

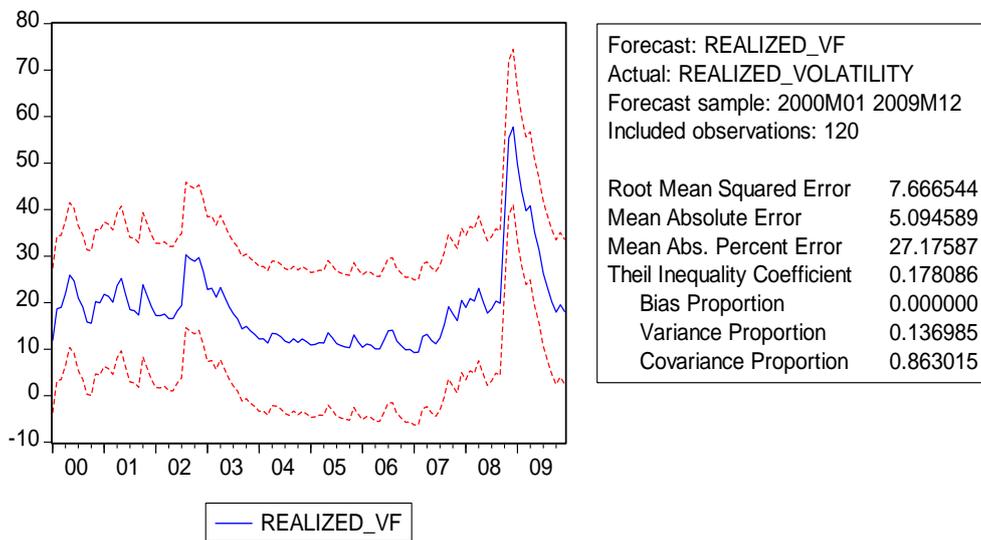
<p>Becker and Clements (2008)</p>	<p>VIX (model-free method) S&P 500</p>	<p>D</p>	<p><i>Implied_{model-free VIX}</i> ARMA ARFIMA GARCH GJR RV SV ALL ALLMBF Combination between any two of MBFs.</p>	<p>22trading days ahead</p>	<p>Combination model is the best estimation. VIX does not only contain the additional information, it also cannot efficiently reflect the information incorporated in MBF.</p>
<p>Becker, Clements and Coleman-Fenn (2009)</p>	<p>VIX (model-free method) S&P 500</p>	<p>D</p>	<p><i>Implied_{model-free VIX}</i> GARCH SV RV ARMA BGZ GJR GJRVIX GJRRV GVIX</p>	<p>22trading days ahead</p>	<p>Risk-adjusted implied volatility provided the better prediction rather than the risk-neutral implied volatility.</p>

Corrado and Miller (2005)	VIX (model-free method) VNX VOX S&P 500 S&P 100 NASDAQ100	D	$Implied_{model-free VIX}$ $Implied_{VNX}$ $Implied_{VOX}$ $HIS_{22\ days}$	22trading days ahead	VIX and VOX are biased but more efficient than HIS.
Giot and Laurent (2006)	VIX (model-free method) VOX S&P 500 S&P 100	D	$Implied_{model-free VIX}$ $Implied_{VOX}$ HIS_{jump} $HIS_{continuous}$	10- and 22-days	Implied volatility contains high information. Decomposed measure of realized volatility does not bring valuable additional information.
Becker, Clements and McClelland (2009)	VIX (model-free method) S&P 500	D	$Implied_{model-free VIX}$ HIS_{jump} $HIS_{continuous}$ GJR GJR + RV SV SV - RV ARMA ARFIMA	22trading days ahead	VIX does reflect the past jump activity in the S&P 500 and provides incremental information.

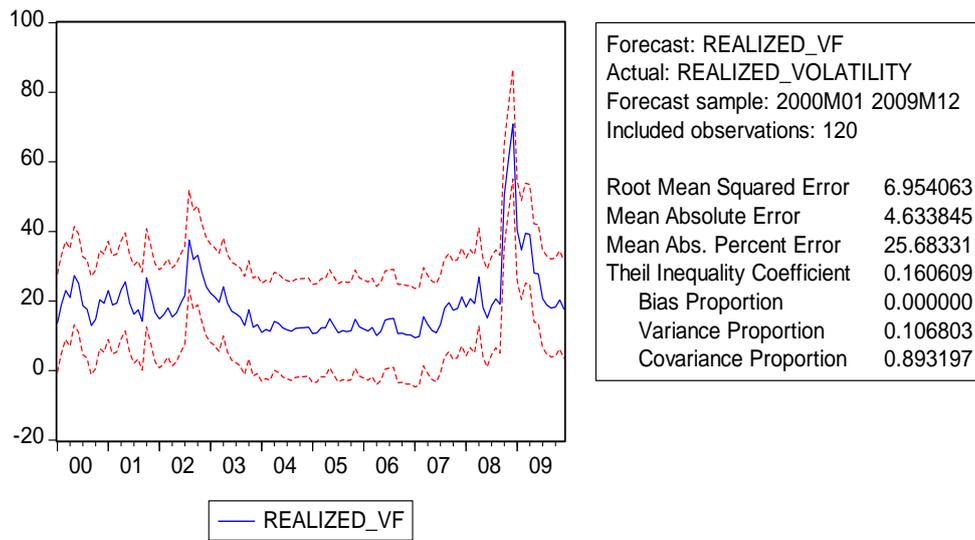
APPENDIX 2



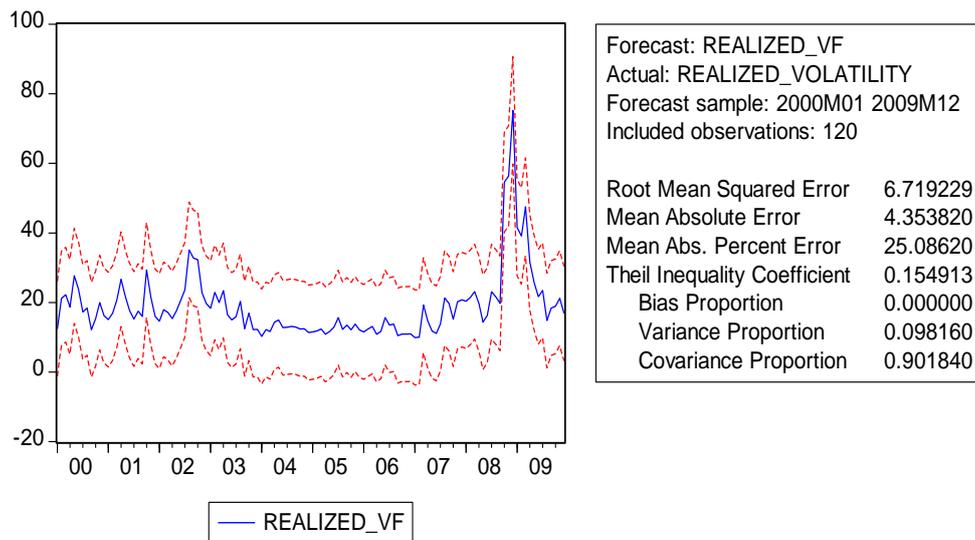
a) Monthly forecasts by Random walk model from Jan 2000 to Dec 2009.



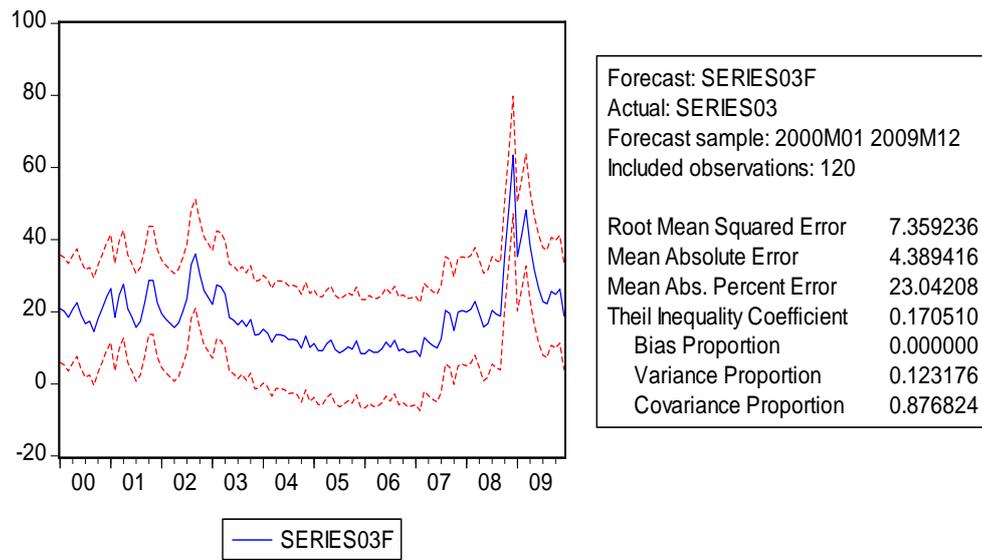
b) Monthly forecasts by *Riskmetrics*TM model from Jan 2000 to Dec 2009.



c) Monthly forecasts by GARCH (1, 1) model from Jan 2000 to Dec 2009.



d) Monthly forecasts by GJR (1, 1) model from Jan 2000 to Dec 2009.



e) Monthly forecasts by implied volatility from Jan 2000 to Dec 2009.