STOCKHOLM SCHOOL OF ECONOMICS
MASTER'S THESIS IN FINANCE

FORECASTING VOLATILITY:

EVIDENCE FROM THE SWEDISH STOCK MARKET

Yu Meng* Natalia Rafikova**

ABSTRACT

This study evaluates the performance of alternative models for predicting stock price volatility on Swedish market. The model set contains various methods for producing volatility forecasts ranging from simple ones (Random Walk, Moving Average and Exponentially Moving Average) to non-linear group of models (GARCH and EGARCH) and Implied volatility from OMX S30 option prices (IV). Overall model performance is evaluated using RMSE and MAE measures. The main results are the following: (1) Forecasts based in implied volatility produce the most accurate results under both measures, while GARCH (1,1) model gain the highest overall error statistics. (2) Allowing asymmetry in variance and non-normal error distribution, the EGARCH (1,1)-GED models perform much better than GARCH (1,1), especially for 20- and 40-day forecasts. (3) Further tests applied to IV confirm that indeed it is an unbiased and efficient estimate of future volatility. The results suggest that the findings in the OMX S30 are mainly in line with the majority of recent evidence, although the study can be extended by the inclusion of other models.

Tutor : Andrei Simonov
Presentation Date : June 07, 2006
Venue : Room 342
Discussants: Jens Ewert and Ellinor Hult

*80299@student.hhs.se
**80295@student.hhs.se

Acknowledgements
Foremost, we would like to express gratitude to our tutor, Andrei Simonov, for providing helpful feedback and support in writing this thesis. In particular, we would also like to thank Jan Engvall at OMX Exchanges for providing us with OMX S30 index option data for conducting this study. Finally, we would like to thank Yangguoyi Ou and Zhengfang Zhao for giving us valuable advice on econometric analysis part.

Table of Contents

1 Introduction

Forecasting financial market volatility has received extensive attention in the literature by academicians and practitioners in recent time. This can be easily explained by crucial importance of volatility forecasts in such arias as investment decision-making, derivatives pricing, risk management and financial market regulation.

Future volatility forecast is an important input for investment decision, portfolio selection and static hedging, where volatility is typically used as a quantitative representation of risk. Volatility does not equal to risk, however. And a clear understanding of volatility is extremely important for making financial decisions based on return fluctuations.

Volatility also plays a central role in dynamic trading and the pricing of derivative securities. Modern option price theory, beginning with Black-Scholes, usually uses volatility of the underlying asset over the option life as a basic input in determining the fair value of an option. Among the parameters in Black-Scholes option pricing model, volatility is the only one that is not observed directly in the market, which magnifies the importance of effectively forecasting volatility of underlying asset returns. A naive and common approach is simply to assume constant volatility and to project the observed past volatility into the future. However, this is only one of the several possible methods, and may not be accurate enough. Particularly, as the maturity of available derivative instruments has lengthened dramatically in the recent years, volatility over the option life can change greatly and thus assuming constant volatility over the option life is unrealistic.

Also risk management that became a compulsory procedure for financial institutions after the Basel Accord I and turned out to be much more complex after the introduction of the Basel II put a central role into the assumptions about future volatility. Furthermore, volatility receives the great deal of concern from public policy makers whose prime duty is stability of financial markets and economy as a whole.

Given that important role of volatility this paper aims to provide an analysis of various techniques that can be used for creating volatility forecasts in stock market by evaluating the performance of the chosen models.

Applying two different approaches, namely the group of historical information based approaches and efficient market approach (option based forecast) suggesting that all information about the future volatility has been condensed by the market and is reflected in option prices, we will try to define whether volatility is predictable, to what extent and which model is superior to others. In other words, we will examine and contrast different procedures from the perspective of their accuracy in producing out-of-sample forecasts.

The object of our empirical study is the Swedish Stock Market and two instruments that are widely traded there: the OMXS30 index and options on this index.

As the return data does not show the same pattern through different markets and time spans, we hope that this study will contribute to the existent evidence on the volatility modeling. Also, to our knowledge no exact study on Swedish stocks has been performed before and this makes us believe that this analysis can add some facts to current knowledge on volatility predicting or open a number of topics for further discussion.

While there vast amount of literature on studying the volatility across different markets and variety of instruments exists, it is hardly possible to name one superior model. Although relying in those works that utilize different methods while testing the same sample, implied volatility is expected to perform better than other methods. As results showed these expectations are correct. Furthermore, implied volatility not only contains predictive power but also produces unbiased and efficient forecasts.

The remaining sections are organized as follows. Section 2 briefly summarizes the concept of volatility and stylized facts about volatility. Section 3 introduces the volatility forecasting models as well as previous literature review. Model evaluation methods are discussed in section 4. Section 5 provides data description and section 6 brings up results and analysis. Finally, section 7 concludes and in section 8 further research suggestions are proposed.

2 Preliminary Discussion about Volatility

2.1 Volatility definition and measurement

Just before starting the analysis we will briefly comment on the difference between such issues as volatility, standard deviation and risk measure as it is not always clearly defined throughout financial literature. Volatility arises from random price movements which occur naturally in every market. It has no exact definition, however in finance, it is a measure of the dispersion of security returns over time and is usually referred to variance σ^2 or standard deviation σ of a sample set of observations according to the formula

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{t=1}^{N}(R_t - \overline{R})^2 \tag{1}$$

Where R_t is the asset return at time t, calculated by $R_t = \ln(S_t / S_{t-1})$, $\overline{R}$ is the average return over this sample period, and N is the number of days.

Throughout this paper, we use sample variance as our volatility measure, and this also helps to avoid the problem of the square root of $\hat{\sigma}^2$ being a biased estimate of the real standard deviation due to Jensen inequality, even though this effect is often negligible (Fleming, 1998).

As mentioned above, volatility is calculated according to the formula (1). But some contradictions arise with respect to the estimate of the mean return since the accuracy of the volatility assessment depends directly on the accuracy of the mean estimation. In Figlewsky 2004, the author points out that the sample average return is a very noisy estimate of the true parameter μ. He suggests a number of approaches that allow avoiding the extreme value of the mean, referring to the previous experience. For example, it is possible to impose a value for the mean as it has been done in Black (1976), where author sets it equal to zero. The same approach was proposed by Perry (1982) who calculates the volatility in a month simply as the sum of squared daily returns in that month. One more approach that was advocated by Ding et al. (1993) is to use the absolute values of daily stock returns. In this paper we rely on the first approach, namely simple average return. With regard to OMXS30 behavior which was relatively smooth in comparison to the samples that were studied in other works (for example S&P500 Index in Figlewsky, 2004) we assume it to be close to the true mean.

It is important to distinguish between volatility and risk. While sample variance $\hat{\sigma}^2$ is a distribution free parameter representing a second moment characteristic of the sample, it is meaningless to use it as a risk measure unless it is attached to a distribution or a price dynamics (see Poon and Granger 2003 for detailed discussion). Meanwhile variance is a correct dispersion measure for the normal distribution and some other distributions, but not all.

2.2 Stylized facts about volatility

Financial time series such as stock returns and exchange rates exhibit certain patterns which are crucial for correct model specification, estimation and forecasting. The statistical properties of financial time series data have been widely reviewed throughout numerous pieces of research and certain interesting stylized facts are revealed, which seem to be common to a wide variety of markets, instruments and periods.

1. *Fat tail distribution*: the (unconditional) distribution of the financial asset return R_t is found to deviate consistently from normality (also referred as excess kurtosis). This is because that there are more very large changes and (consequently) more very small ones than a normal distribution calls for (see Paretian and Levy (1925) for modeling excess kurtosis and distributions that have fatter tails than normal distribution).
2. *Volatility clustering*: it is observed that large price variations are more likely to be followed by large price variations, and the inverse is also true. This is an indication of persistence in shocks. This kind of time variation in the returns distribution suggests that volatility is predictable.
3. *Asymmetry:* gains and losses in stock and index returns are not symmetric (also known as leverage effect). One observes large drawdowns but not equally large upward movements. This makes the distribution of stock returns to be negatively skewed.
4. Slow decay of autocorrelation in absolute and squared returns: it is observed that (linear) autocorrelation in returns series (except high-frequency return series) are typically insignificant. On contrary, positive, significant and slow

decaying autocorrelation is commonly present in absolute and squared returns. The autocorrelation often remains significant for even very long lags. This serves as a quantitative manifestation of volatility clustering and is also an evidence of conditional heteroskedasticity. This finding may further imply long range dependence in volatility.

Apart from the stylized facts discussed above, financial asset returns also exhibit other characteristics such as mean reversion, co-movements of volatilities across assets and financial markets, and long memory property (the existence of unit root). It is important to be aware of these stylized facts as optimal modeling and forecasting should take into account these statistical properties.

In the next section, we will introduce the most frequently used volatility forecasting models (by both practitioners and researchers).

3 Volatility Forecasting Models

In this section we describe various types of models that are used for volatility forecasting, limiting our discussion to the most popular methods. Relevant literature is reviewed in each subsection. Generally all models that practitioners and academicians use to forecast volatility can be classified as either time series forecasting models or options based forecasting approach. Time series volatility forecasts make use of historical information set and can be further classified as naïve models, such as random walk and moving average models, and the more sophisticated models such as GARCH type models and stochastic volatility models. All models discussed in this section take into account volatility persistence or clustering, and some models also capture asymmetry.

3.1 Naïve models

This group of models bases forecasts of future volatility on past information, usually variances or standard deviations. In our paper, we choose some most frequently used naïve models including random walk (RW), moving average (MA) and exponentially weighted moving average (EMWA) model.

Although given that the actual security prices do not come from a constant volatility lognormal diffusion process, computing historical volatility using those models is no longer theoretically optimal. But, in spite of this, it is quite common among option traders and academic researchers to calculate historical volatility estimates by the most basic methods. The normal (though not necessarily optimal) way most traders deal with the fact that volatility changes stochastically over time is to use only recent observations in the calculation and discard data from the distant past. It then becomes necessary to decide how much past data to include in a historical sample.

Detailed specification of the selected naïve models is provided further as well as the review of the previous findings that were based on those approaches.

Random walk (RW)

The well-known random walk model is the simplest possible model and it assumes that the best forecast of next period's volatility is this period's volatility, i.e.

$$\hat{\sigma}_{T+1}^2 = \sigma_T^2 \qquad (2)$$

Moving Average (MA)

According to the historical moving average model, the most accurate forecasts are the ones based on the most recent data, and this model is defined as

$$\hat{\sigma}_{T+1}^2 = \frac{1}{n}\sum_{j=1}^{n} \sigma_{T+1-j}^2 \qquad (3)$$

The choice of the period that the forecast is based on is an arbitrary issue. Too few data contains too little information, while including too many data points is also not appropriate as the very old ones are obsolete and have no explanatory power for the volatility in recent time. In the paper two moving average models namely based on previous 5 and 20 volatilities are used for making predictions. In contrast to Figlewski (1997) and other earlier papers we do not use overlapping data[1]. The same approach (non-overlapping data) is utilized in consequent EWMA, GARCH and IV modeling.

Exponentially Weighted Moving Average (EWMA)

Another model that belongs to the group of the simple historical based forecasts is exponentially-weighted moving average (EWMA). Conceptually it is very similar to simple moving average except that it places more weight into latest data, thus making the forecasts more relied on recent past.

$$\hat{\sigma}_{T+1}^2 = (1-\lambda)\hat{\sigma}_T^2 + \lambda \frac{1}{N}\sum_{i=1}^{N} \sigma_{T+1-j}^2 \qquad (4)$$

[1] Although using overlapping data gives more observations at hand, it produces autocorrelation for regression-based results.

The selection of the smoothing parameter value (λ) is an empirical issue. The optimal value of λ is chosen by a search of values between zero and one. In our tests we selected it to be equal to 0.94, an industry standard[2].

The next question is how to choose the time span. The 12- and 26-day EWMAs are the most popular short-term averages in the industry and they are used to create indicators like the moving average convergence divergence (MACD) and the percentage price oscillator (PPO). In general, the 50- and 200-day EWMAs are used as signals of long-term trends. In line with the time span picked for moving average model, we make forecasts based on EWMA employing 5 and 20 previous observations[3].

All models described above capture volatility clustering. And due to the calculation simplicity, this group of models is frequently used either in the literature or in practice. In Brainsford et al (1996) all of the models mentioned above have been used for monthly volatility forecasting of Statex-Actuaries Accumulation Index (this index comprises the 50 most actively traded companies listed on the Australian Stock Exchange). The best results were obtained by utilizing MA (12-year) model following by EWMA and MA (5-year). The RW model was the least successful. Similar results have been obtained in the study of the German market volatility by Claessen and Mitnik (2002).

If we rank the models by their RMSE[5] in the study by Yu (2002), RW will take the last place again. The best way to forecast volatility of the New Zealand Stock Market is to use GARCH-type models (this type of models will receive the great deal of attention in the next sections), while MA models work slightly better than EWMA. Similar ranks are obtained for MAE[6] measure.

In contrast to those studies, Dimson and Marsh (1990) in their paper on UK FT All Share Index found that RW approach is better than MA model when predicting the next quarter volatility using the daily returns as an input.

---

[2] The value of λ is chosen according to The RiskMetrics database (J.P. Morgan) and Yu Meng's previous working experience as an investment analyst in Chinese market.
[3] For both MA and EWMA models, we have also tried other time span, such as previous 40 and 100, and find that shorter time span (5 and 20) leads to slightly better results, even though the difference is quite small.
5 Root Mean Squared Error, see Section 4.1.
6 Mean Absolute Error, see Section 4.2

Interesting results are given in Figlewsky (2004), where the author does not evaluate the range of historical based volatility models, but performs the study on different specification of the same model. In general, he finds that historical volatility computed over many past periods provides the most accurate forecasts for both long and short horizons. Although the error statistics is substantially lower for long term than for short term volatility forecasts.

Now, after the brief review of the previous works that contain simple methods we will precede the discussion by introducing more sophisticated techniques for producing volatility forecasts. Also we try to assess the performance of different naïve models, we emphasize that the main purpose is to evaluate the whole set of models against their performance.

3.2 Regression based forecasts (GARCH type models)

This group consists of various types of time series forecast models such as simple regression, ARMA type models, ARCH class conditional volatility models, etc. The first two types of models predict future volatility based on historical volatility data, while ARCH type models forecast volatility from historical return data. Simple regression model is principally an autoregressive (AR) process, where volatility is expressed as a function of its own lagged values and an error term. Further, when past volatility errors are added, we get an ARMA model, and by introducing a differencing order I (d), we get ARIMA when $d = 1$ and ARFIMA when $d < 1$ (see, for example, Poterba and Summers, 1986, and French, Schwert, and Stambaugh, 1987).

Like the naïve models discussed above, AR and ARMA type models involve the calculation of a series of, for example, monthly sample volatilities, which implicitly assumes that volatility is constant within a month and becomes variable for only longer horizons. In addition, according to Chou (1988), the parameter estimates are extremely sensitive to the sampling frequency for which the time series of volatility estimates is calculated. Given these shortcomings of the simple regression and ARMA models, we will not include them in this study but focus on the more popular ARCH type models.

The autoregressive conditional heteroscedastic (ARCH, see Engle, 1982) and the generalized ARCH (GARCH, see Bollerslev, 1986) models have been widely employed since their introduction. The ARCH family models make use of only historical return data and involve volatility as an integrated aspect of the return behavior. These non-linear time series models were designed to capture volatility clustering and unconditional return distributions with fatter tails which are commonly associated with macro-economic series such as stock market returns.

A GARCH (p,q) model is represented as

$$r_t = \mu + \varepsilon_t \text{ with } \varepsilon_t = \sigma_t z_t \tag{5}$$

where z_t is assumed to be normally distributed with mean zero and variance of one. This equation is usually called mean equation, which tries to capture the dynamic of returns. σ_t^2 is the conditional variance of returns based on information available up to time t-1, being given by

$$\sigma_t^2 = \lambda + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{6}$$

Thus the estimation of GARCH model involves joint estimation of the mean and conditional variance equation. In practice, GARCH (1,1) is normally found to be sufficient to model time-varying variance for most financial time series data parsimoniously (that low-order GARCH models describe stock return volatility behavior very well is shown by Akgiray (1989), Pagan and Schwert (1990), etc.). For GARCH (1,1) the conditional variance equation is the following

$$\sigma_t^2 = \lambda + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{7}$$

In estimating the model, we impose restrictions of $\alpha + \beta < 1$ on the parameter to ensure that ε_t is stationary. The sum of α and β measures the persistence of conditional variance to shocks.

Some researchers often combine a GARCH (1,1) model with AR(1), i.e. $r_t = \mu + \beta_0 r_{t-1} + \varepsilon_t$, for the mean equation to transform each original daily return series into an uncorrelated new series. The rational is that by removing the first-order autocorrelation from the return series, the noise in volatility resulting from microstructural factors can also be largely reduced. However, in the case of OMX S30, we find that the $\hat{\beta}_0$ is insignificant in almost all the sub-sample estimations, and thus do not include the autoregressive term in the mean equation.

The GARCH models can be estimated by maximum likelihood method and the parameters can be obtained by numerical maximization of the log-likelihood function.

After the model parameters have been estimated, the h-day ahead forecast of volatility can be calculated by iterating on

$$E_T\left[\sigma^2_{T+h}\right] = \hat{\lambda}\frac{1-\left(\hat{\alpha}+\hat{\beta}\right)^{h-1}}{1-\left(\hat{\alpha}+\hat{\beta}\right)} + \left(\hat{\alpha}+\hat{\beta}\right)^{h-1}\hat{\sigma}^2_{T+1}$$

Where $\hat{\sigma}^2_{T+1} = \hat{\lambda} + \hat{\alpha}\varepsilon^2_T + \hat{\beta}\hat{\sigma}^2_T$ (8)

Another GARCH type model that also has wide application in dealing with financial time series data is the exponential GARCH model (EGARCH, see Nelson, 1991), where mean equation is the same but the conditional variance in EGARCH (1,1) is defined as

$$\ln\sigma^2_t = \lambda + \alpha\ln\sigma^2_{t-1} + \beta_1\varepsilon_{t-1} + \beta_2\left(\left|\varepsilon_{t-1}\right| - \sqrt{\frac{2}{\pi}}\right) \quad (9)$$

The log-transformation guarantees that volatility is always nonnegative, even if the parameter values are negative. Another important improvement of EGARCH models is that it allows for an asymmetric reaction to positive and negative shocks. As for the error distribution, instead of assuming normal distribution, we assume an iid (independently and identically distributed) generalized error distribution (GED) for z_t, which is defined as

$$f(z) = \frac{\upsilon\exp\left(-\frac{1}{2}\left|z/\lambda\right|^{\upsilon}\right)}{\lambda 2^{(1+1/\upsilon)}\Gamma(1/\upsilon)}, \quad -\infty \prec z \prec \infty, \upsilon \succ 0 \quad (10)$$

where $\Gamma(\cdot)$ is the gamma function, and $\lambda^2 = 2^{-2/\upsilon}\Gamma(1/\upsilon)/\Gamma(3/\upsilon)$. When υ = 2, we obtain a special case of GED, the standard normal distribution. In other cases, the distribution has either fatter tails (when $\upsilon > 2$) or thinner tails (when $\upsilon < 2$) than the normal distribution. This type of behavior in daily returns is specifically modeled by the EGARCH model. Stationarity is obtained by imposing the value of α to be less than 1.

According to Figlewski (2004), all (G)ARCH-type models have at least two significant shortcomings as forecasting tools. First, a large number of data points are needed for robust estimation[7]. The second problem is that all these models essentially focus on variance one step ahead, which means that they seem not to be designed for very long-term forecasting.

The predictive power of GARCH type models has been investigated in a large amount of papers and many empirical studies show supportive results. Akigray (1989) is one of the first researchers that test GARCH model. He finds that GARCH (1,1) beats EWMA and HIS (historical volatility derived from standard deviation of past returns over a fixed interval) in all subperiods in case of CRSP VW and EW[8] indices. Pagan and Schwert (1990) find EGARCH to be best especially compared with nonparametric methods.

Other studies show that volatility prediction performance of different models depends on the specific asset class, error statistics, sampling schemes or time periods. For instance, Heynen and Kat (1994) find that GARCH (1,1) is superior to EGARCH (1,1) and stochastic volatility for currencies, but not for stock indexes.

3.3 Option based forecasts

Apart from the methodology of forecasting volatility from historical prices, the option based volatility measures are also commonly used in practice. Volatility is often an input for option price calculation, and thus given the option prices, we can derive the volatility having been used, namely implied volatility. In contrast to time series models, forecasting volatility from option prices does not involve historical information. It is believed that implied volatility is superior to historical volatility since it is the market

[7] By convention, around 2000 data points can be seen as sufficient.

[8] Center of Recearch in Security Prices Value WEIGHTED Index and Equaly Weighted Index, http://www.gsb.uchicago.edu/research/crsp/

participants' best "guess" of the future average volatility during the option life and thus contains more information. Option prices are highly related to market expectations about the asset's future value movements. Assuming the rational market "behavior", market should use all the information available to form its expectations about future volatility. Hence, the market option price reveals the market's true volatility estimate. Furthermore, if the market is efficient, the market's estimate, the implied volatility, is the best possible forecast given the currently available information. That is, all information necessary to explain future realized volatility generated by all other explanatory variables in the market information set should be subsumed in implied volatility.

In order to test whether implied volatility is a valuable estimate of the future volatility, we need first to derive it from the option prices. Traditionally there are two types of options, European options and American options. And Black-Scholes model is commonly used to price European options while binomial trees and Monte-Carlo simulation are usually employed to obtain the theoretical value of American options. Since OMX S30 index option is European style, we assume that Black-Scholes Model gives efficient option pricing, and by solving the B-S formula, we can obtain the implied volatility from option prices.

Under the Black-Scholes framework, stock price follows a geometric Brownian motion (GBM)

$$dS_t = \mu S_t dt + \sigma S_t dW \tag{11}$$

where dW is Wiener process, and μ and σ are the drift and diffusion (volatility), respectively.

And from Ito lemma, the logarithmic of stock price has the following dynamics

$$d \ln S = \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dW \tag{12}$$

which implies that stock price is lognormally distributed and that its return is normally distributed with constant variance. For a non-dividend paying stock, values for a call price c or put price p are:

$$c = s\Phi(d_1) - Xe^{-rT}\Phi(d_2) \text{ and } p = Xe^{-rT}\Phi(-d_2) - s\Phi(-d_1) \qquad (13)$$

where

$$d_1 = \frac{\ln(s/X) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \text{ and } d_2 = d_1 - \sigma\sqrt{T}$$

σ is average (annualized) standard deviation over the option life

s is the underlying stock price

X is the option strike price

r is continuous (annualized) risk-free rate

T is time to maturity in calendar days[9]

Black-Scholes option price depends on five arguments, i.e. underlying stock price, standard deviation of the underlying stock price, option strike price, time to maturity, and continuous risk-free interest rate[10]. Although there is a closed form solution for Black-Scholes call and put price in terms of the five arguments, the solution for volatility σ as a function of call or put price and the other arguments is done by approximation. In practice, trial and error in spreadsheet, bisection and Newton-Raphson iterative algorithm are commonly used in solving Black-Scholes implied volatility. For large set of data, iterative algorithm (bisection and Newton-Raphson) is more efficient. Newton-Raphson is employed for deriving implied volatility in this paper[11] (see appendix A for details).

In the Black-Scholes framework, volatility is assumed to be constant over the option life across all the exercise prices. This means that given different exercise prices, the implied volatility from Black-Scholes model should be the same. This, however, does not hold in the real world. Usually implied volatility is higher for out-of-the-money (OTM) and for in-the-money (ITM) options than for at-the-money (ATM) options. This phenomenon is known as volatility smile and is illustrated in the figure 1.

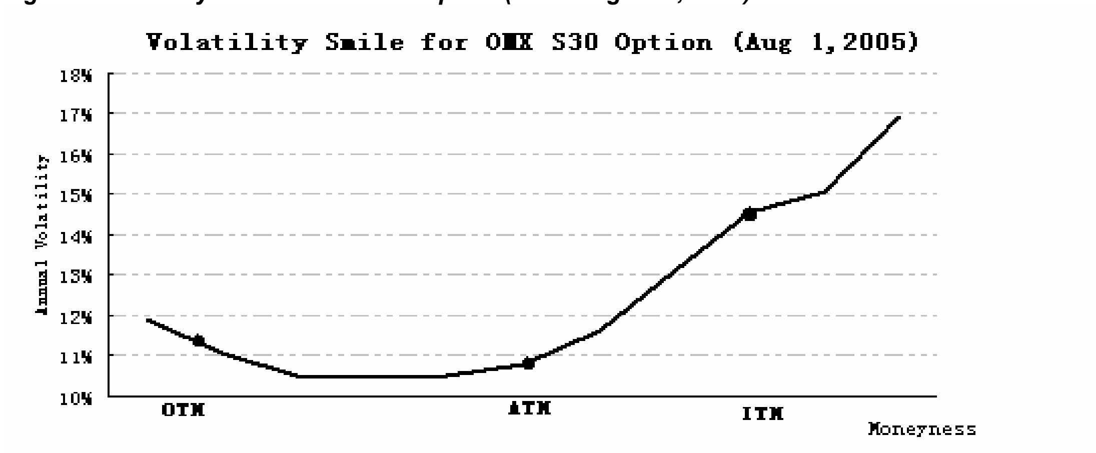
[9] The denotion *T* in option based approach here is time to maturity, and this only applies to formula (13)

[10] In this study, we ignore dividend yield of the components of OMX S30 Index.

[11] Implied standard deviation is calculated by statistical software R and VBA.

[13] Time to maturity is denoted by *TTM* here, rather than by *T* as in formula (13), to avoid confusion.

Figure 1 Volatility Smile of OMXS30 Option (as of August 1, 2005)



There are several explanations for volatility smile. And the most common arguments are violation of distribution assumptions and stochastic volatility. As mentioned before, under Black-Scholes framework, stock price is assumed to be lognormally distributed and thus stock returns are normally distributed, while empirical evidence shows that stock returns have leptokurtic tails (which has been illustrated above). A leptokurtic right tail will lead to a higher call price and a higher Black-Scholes implied volatility at both high strike (deep-out-of-the-money) and low strike (deep-in-the-money). A much more concrete explanation can be found in Poon & Granger (2003).

In face of volatility smiles, many researchers have adopted the approaches of combining implied volatilities from all or several available options with the same expiration date to obtain a single measure of volatility forecast over the option life. By calculating the average implied volatility from these options using certain weighting scheme, pricing bias and measurement errors are reduced and better forecasting performance may be achieved.

Lots of empirical work has been done in order to find the most suitable weighing scheme for certain data set. And particularly, Ederington & Guan (2000) tests the most frequently used weighting methods and find that an average (and often an equally weighted average) of just a few at-the-money implied volatilities performs slightly better than the broader weighted average.

In this paper, we make use of all the at-the-money call and put options. We define at-the-money to be +/-5% out-of-the-money or in-the-money since in most cases there are no calls or puts that are exactly at-the-money in the option markets. Instead of using equally weighted average, we weight the implied volatilities obtained from these at-the-money options by trading volume. The rational is that options with strike prices closest to the index prices are usually the most heavily traded and by doing so we give these relatively closer to at-the-money implied volatilities more weights.

Once a single weighted implied volatility over the option life is obtained, the next step is to calculate the n-day ahead volatility forecast by interpolation or extrapolation methods using the formula

$$\sigma_{T+n} = \sigma_T \times \frac{\sqrt{n}}{\sqrt{TTM}} \tag{14}$$

where *TTM* is the time to maturity[13].

Specifically, on the first day of each week in out-of-sample period, we calculate the implied standard deviation from option prices. As the options can only be exercised on the fourth Friday each month, this method leads to different time to maturity for different dates. We utilize all chosen call and put options on the first day of every week, with time to maturity varying between 16 and 45 calendar days, the most of which concentrate around 25 to 35. After calculating the implied standard deviations for these selected calls and puts, we obtain the final value of implied standard deviation on each specific date by taking the average weighted by volume.

After obtaining the (average) implied standard deviation for the whole option life at the beginning of each week, we calculate the (average) standard deviation for every week according to formula (14). Thus we can get the weekly forecasted volatility series, i.e. the squared implied standard deviation, as option implied standard deviation is by nature forward looking and can be used directly as forecasts. To reduce computation burden without losing accuracy, we utilize the weekly predicted volatility series to obtain the biweekly, monthly, and bimonthly volatility forecasts. The biweekly volatilities are calculated simply as the average of two non-overlapping weekly predicted volatilities in sequence. And for monthly and bimonthly forecasts, the method is the same except that the average of four and eight forecasted volatilities is used respectively.

The hypothesis that implied volatility is a rational forecast of subsequently realized volatility has been frequently tested in the literature. Several studies have cast doubt about the rationality hypothesis in the context of the most active option market, namely, the market for OEX options on the S&P 100 stock market index (see Day and Lewis (1992), Harvey and Whaley (1992), and Canina and Figlewski (1993)). More recently, however, Christensen and Prabhala (1998) find that implied volatility from at-the-money one-month OEX call options in fact is an unbiased and efficient forecast of ex-post realized index volatility after the 1987 stock market crash. The remaining forecast errors cannot be explained by simple ARCH or GARCH specifications (see, for example, Fleming (1998) on the OEX case).

There are also many studies that provides opposite or no clear-cut results. For instance, Aguilar (1999) finds that option based approach provides better forecasts for future volatility than GARCH, at least for shorter forecast horizons, but GARCH is superior for currencies.

3.4 Other models and general comments

Besides the models described above, there are some other approaches for volatility prediction, most of which do not appear frequently in the literature. Some of these models fail to gain much popularity due to their complexity. This is the case of neural networks models (see for example, Donaldson and Kamstra (1997). Other models are proposed relatively late and are getting more and more attention among academicians and practitioners. For example, volatility forecast based on stochastic volatility (SV) model was developed in the mid 1990’s and its high performance has been documented in several research papers. Heynen (1995) finds that SV forecast is the best for a number of stock indices across several continents. Other studies of SV volatility forecasts includes Yu (2002), Lopez (2001), Dunis, Laws and Chanvin (2000), etc.

In addition to SV models, some other classes of models are proposed with the aim of capturing certain features of volatility. Long memory (LM) volatility model takes into account the long memory characteristics of volatility and are studied in, for example, Vilasuso (2002) and Zumbach (2002). Furthermore, the regime switching models are developed to model changing volatility persistence (for detailed discussion, see Hamiton (1989), Pagan and Schwert (1990), Hamilton and Susmel (1994), etc.)

Albeit the variety of the modeling technique and pieces of research exist, it is still hard to underline the priority of one to another. Probably the most extensive comparative analysis on existent works is done by Poon and Granger (2003). In their paper, authors outline the main findings of 93 papers that were written on the topic of volatility forecasting and show the statistic of models' performance. Although in the majority of works results of option based approach outweigh all others, still no preferable model for this kind of analysis exist and the choice will depend on the features of the data (frequency, accuracy, instruments and so on) and current situation on the market.

In this study not all of the above models will be tested, the research will be focused on the following models: Random Walk, Moving Average, Exponentially Weighted Moving Average (EWMA), GARCH (1,1), EGARCH (1,1)-GED and Implied Volatility (IV) based forecasts.

4 Forecast Evaluation

Following the models described above, the performance evaluation methods applicable to those tests are presented in this section. Our focus in this part is on predicting short and middle horizon volatility. It would not be statistically meaningful to perform long–term volatility forecast which would let us avoid any effects from market microstructure noise and other self-correcting short run phenomena. For this kind of study monthly data is necessary and the time span available would be insufficient (see Figlewsky 2004).

Each approach produces 5, 10, 20 and 40-day forecasts, then performance of the models is tested by comparing results to the ex-post realized volatility over the same periods, which is calculated as

$$RV_t = \sqrt{\frac{1}{m}\left(\sum_{j=1}^{m} r_{t+j} - \mu\right)^2 \times 250} \qquad (15)$$

where m is the forecast horizon, and μ is the estimated average of returns.

4.1 Root Mean Square Error (RMSE)

The first category of forecast performance evaluation is to measure forecasting errors, which involves the calculation of difference between predicted values and realized one. Previous papers have used a variety of statistics to evaluate and compare forecast errors. Among them, the root mean squared error (RMSE) is frequently used and consistent with this research, the forecast errors generated from each model are compared by this measure. RMSE is calculated according to the following formula:

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{t=T}(\hat{\sigma}_t^2 - \sigma_t^2)^2} \qquad (16)$$

4.2 Mean Absolute Error (MAE)

The other type of forecasting errors measurement that is used in this study as a supplement for comparison purpose is MAE. It is calculated as

$$MAE = \frac{1}{T}\sum_{i=1}^{T}\left|\hat{\sigma}_i^2 - \sigma_i^2\right| \quad (17)$$

The result of forecast performance evaluation is sometimes quite sensitive to the error measurements. Thus different accuracy test measures may produce different rankings of models, although following previous works, consistency over different estimates should be maintained.

4.3 Regression based forecast efficiency test

In this subsection the study of volatility forecasting is extended by applying the test of market efficiency. Under the assumption of the efficient market we expect that implied volatility is the best possible forecast given the currently available information. Typically, this can be tested by running the following regression:

$$RV_t = \alpha + \beta_1 IV_t + \varepsilon_t \quad (18)$$

Where RV_t is ex-post realized volatility of OMX S30 index over the specific period and IV is the volatility forecast over the same period implied from the option price at time t.

According to the previous literature, the following major inferences can be drawn from those regressions. First, by estimating β_1 one can assess the information content of the implied volatility. In this case we will test the hypothesis that β_1 is significantly different from zero.

Second, it is argued that option prices contain biased forecast if the joint hypothesis of α being equal to 0 and β_1 being equal to 1 is rejected.

This framework can be extended to comparison of the predictive power of the IV and time-series models.

$$RV_t = \alpha + \beta_1 IV_t + \beta_2 HisV_t + \varepsilon_t \,, \quad (19)$$

Where $HisV_t$ is the average volatility forecast based on past historical information available at time t. Generally, in most studies, the term $HisV_t$ is presented by GARCH family estimates. The similar approach is given in this paper, where the EGARCH (1,1)-GED forecasts are added into the initial regression.

By testing β_2 it is possible to estimate the efficiency of implied volatility. In case it is statistically distinguishable from 0, we can argue that option prices contain informationally inefficient forecasts of future volatility.

5 Data description

5.1 The Swedish market for OMX-stock index options

In this section, the brief description of data set (OMX S30 Index and Option on index prices) will be introduced.

Presently, OMX is the Nordic derivatives market and Europe's third largest marketplace for derivatives by volume. As the world's leading exchange and clearing house for Nordic derivatives, OMX provides the optimal marketplace for liquidity and transparency in Nordic derivatives[14].

The OMX Index was first introduced in September 1986. It is a value-weighted index based on the 30 largest capitalized shares at Stockholm Stock Exchange (SSE). The purpose of the introduction was to use the index as an underlying asset for trading in standardized European options and futures. Since the introduction, the market for OMX-derivatives has grown substantially.

The OMX-index option market consists of European and American call and put options, as well as futures contracts, with different time to expiration. At any time throughout a calendar year when the exchange is open, trading is possible in at least three series of option contracts, with up to one, two and three months left to expiration respectively. On the fourth Friday each month, if the exchange is open for trading, one series of contracts expires and another with time to expiration equal to three months is initiated. If the exchange is closed the expiration date moves to the previous trading day.

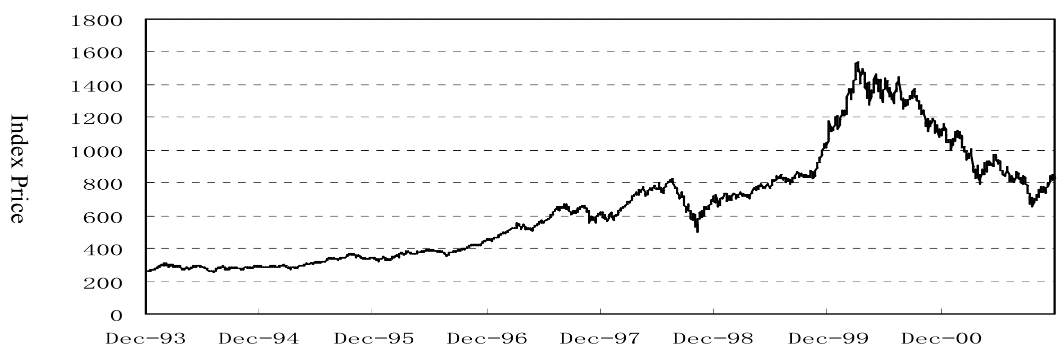
When new options series are introduced, strike prices are chosen so that they are centered round the current OMX-stock index value. Further, as the stock index value increases or decreases with a considerable amount during the time to expiration, new strikes are introduced. Thus, the prevailing strike price range depends on the development of the index during the time to expiration.

[14] See http://www.omxgroup.com/omxcorp/ for detailed information.

5.2 Data

The data is taken from OMX closing quotes for OMXS30 Index and options on the index. The time span that has been chosen for our study is the period Dec 01, 1993 through November 30, 2005. Although the data for longer time period was available we decided not to include the earlier observations. The trading volumes and frequencies on the late eighties – early nineties were not sufficient to perform any reliable analysis.

Figure 2 Time series plot of OMXS30 index prices, Dec 1, 1993 – Nov 30, 2001



The data from the 1st of December 1993 till the 30th of November 2001 will be used for creating appropriate models and the latter for years up to the end of November 2005 will serve for the evaluation of forecasts. In total we have 2012 in sample and 995 out of sample observations. Figure 2 shows in sample time series plot of OMX S30 index prices. During this period of time, the index price increases from 261.94 on Dec 01, 1993 until the burst of IT bubble in the first quarter of 2000, followed by subsequent long term downwards adjustments.

The out-of-sample data set is then divided to as many as possible non-overlapping 5-, 10-, 20-, and 40-day sub forecasting periods. Thus for a given model, the amount of predicted volatilities for the four forecasting horizons are 199, 99, 49, and 24, respectively.

Specifically, in the naïve model forecast, we first utilize the weekly, biweekly, monthly[15], and bimonthly volatility series of in-sample data set to calculate the first volatility forecast

[15] Unless specified, one month corresponds to 20 work days throughout this paper.

for each forecasting horizon. After that the sample of every volatility series is rolled over one horizon ahead to get the next forecast until all the out-of-sample data set is used up.

As for GARCH and EGARCH forecast, in-sample return series is used to fit the model and make the first set of forecasts. Then starting with the second month in out-of-sample data, we re-estimate the model at the beginning of every month using the previous 2012 data and obtain the sequential set of forecasts.

In option implied volatility forecast, for the first day of each week in out-of-sample period, the values of the five variables, i.e. option price, index price, strike price, risk-free interest rate, and time to maturity, are collected. Option prices are the closing prices of OMX S30 index options for the specific dates[16]. As mentioned in Section 3.3, we utilized all +/- 5% in- and out-of-the-money call and put options, and time to maturity varies between 16 and 45 calendar days. We use Swedish 30-day Treasury bill[17] rate as a proxy for risk-free interest rate. In accordance with Black-Scholes model, risk-free interest rate is in continuous time setting, so we convert the 30-day Treasury bill to continuous rate before using it.

[16] Option data are obtained directly from the Stockholm Office of OMX AB
[17] The 30-day Swedish Treasury bill rate data is obtained from www.scb.se

6 Empirical Results and Analysis

6.1 In-sample results

Table 1: Sample Statistics for OMX S30 Index Returns, December 1, 1993 - November 30, 2001

	OMX S30 Index Returns
# of observations	2012
Mean (%)	0.02570
St.Dev. (% per year)	10.5096
Skewness	0.02035
Excess Kurtosis	3.1451
JB p-value	<0.0001
LB Q(10)	14.3205 [0.15887]
LB Q(20)	51.1601 [0.00015]
LM ARCH 1-50 test	6.5694 [0.0000]
ADF1	-32.61 ***
ADF2	-23.28 ***

JB: Jarque-Bera (1980)'s normality test

LB Q (10): Ljung-Box Q-statistic for serial correlation in return series (10 lags)

LB Q (20): Ljung-Box Q-statistic for serial correlation in return series (20 lags)

LM ARCH 1-50 test: Engle's Lagrange multiplier test for presence of ARCH effects

ADF1: Augmented Dickey-Fuller test for unit root in return series

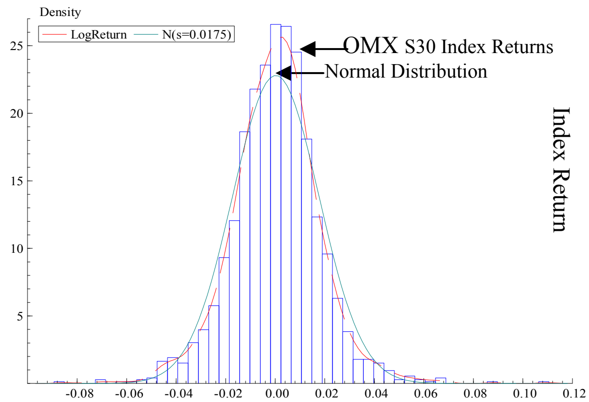
ADF2: Augmented Dickey-Fuller test for unit root in squared return series

St.Dev: Annualized standard deviation for returns assuming 250 days per year

*** implies the rejection of null hypothesis of unit root presence in the data at 1% level.

Table 1 displays some sample statistics for the 2012 daily OMX index returns from Dec 01, 1993 through Nov 30, 2001. Assuming 250 trading days per year, the annualized standard deviation for the daily returns is 10.51%. Our empirical results also reveal that OMX S30 index returns match the stylized facts discussed in section 2.2. Firstly, the sample excess kurtosis and histogram of return series in *Figure 3* provide the evidence of non-normality. This is also supported by the Jarque-Bera normality test. The null hypothesis of normality is rejected at 1% significance level. Secondly, the plot of daily return series (*Figure 4*) shows the existence of volatility clustering. One can see from this figure that large changes tend to be followed by large changes, and small changes tend to be followed by small changes. Thirdly, from *Figure 5*, the autocorrelation function for OMX S30 index returns, squared and absolute returns, we can see that serial correlations for index returns are insignificant for most lags while the autocorrelations in squared and absolute returns are significant and persistent. This is consistent with the results of Ljung-Box Q-statistics and Engle's Lagrange multiplier ARCH effects test as shown in Table 1. Thus the null hypothesis of no serial correlation up to lag 10 cannot be rejected for returns, but there exists non-linear serial correlation, which indicates the presence of conditional heteroskedsticity.

Figure 3 Density function of OMX S 30 Index Returns, Dec 1, 1993 – Nov 30, 2001



The normal distribution is the one with the same mean and standard deviation as OMX S30 index returns.

Figure 4 Time Series Plot of OMX S 30 Index Returns, Dec 1, 1993 – Nov 30, 2001

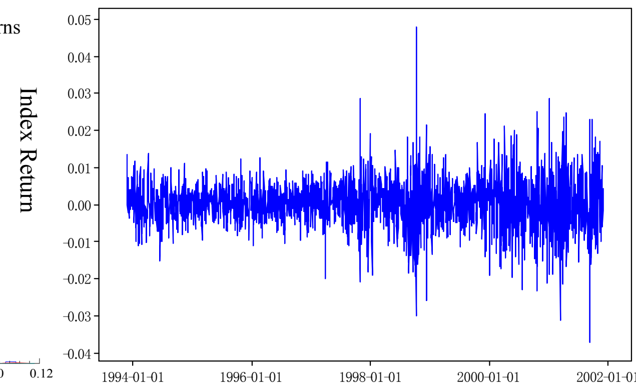
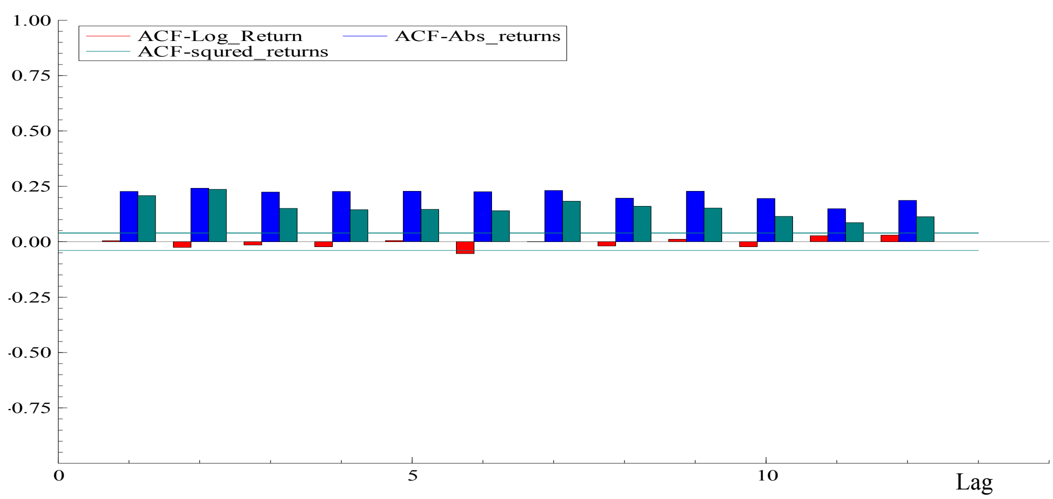


Figure 5. Autocorrelation function of OMX S30 index returns, the squared and absolute returns



Also the results of the augmented Dickey-Fuller (ADF) unit root test show that there are no unit roots in both return series and squared return series. The null hypothesis of a unit root is rejected at 1% significance level in both cases.

In summary, the OMX S30 index return series seems to have a non-normal distribution with excess kurtosis and possess significant conditional heteroskedasticity. These characteristics make the GARCH type models with student-t or the more generalized GED distribution a natural candidate for volatility forecasts.

However, it is too early to draw any conclusion about the forecasting performance of the alternative models at the stage. A model that does not show very good in-sample performance may also produce good out-of-sample forecasts. It is important to keep in mind that effective modeling does not equal to effective forecasting. For instance, non-normal distribution in return series violates the assumption of Black-Scholes models but this does not necessarily mean that option-based approach will produce worse volatility forecasts than GARCH type models. Also our sample does not show high degree of deviation from normal distribution. It behaves not that 'badly' in comparison with other studies, where markets have been much more unstable (for example the kurtosis in the New Zealand data set was almost 78). In addition the Skewness value of 0.02035 shows that the asymmetry in return distribution is not quite obvious. Worthy to mention also is that a model may have different forecasting ability for different forecast horizons or different financial assets due to its specific features.

Table 2 shows the in-sample estimation of GARCH (1,1) and EGARCH (1,1)-GED models. All the coefficients are significant at 1% confidence level for both models. And both models are stationary, as $\alpha + \beta$ is less than 1 for GARCH and α is less than 1 for EGARCH. The standardized residuals from GARCH (1,1) fail the Jarque-Bera normality test, i.e. the null hypothesis of normal distribution is rejected. This is further supported by the negative skewness value for standardized residuals. The results show contradiction on the normality assumption of GARCH models and indicate a sub-optimal fit of data.

Unlike the case in GARCH (1,1), whether the standardized residuals are normally distributed or not is not important for EGARCH (1,1)-GED, as the latter assumes generalized error distribution (GED), and normal distribution is only one special case of GED. As shown in Table 2, even though the estimated value for υ is 1.896958, close to 2, Jarque-Bera test result shows that standardized residuals are not normal at 5% significance level.

As for serial correlation, the null hypothesis of no serial correlation for standardized residuals and squared standardized residuals up to lag 10 cannot be rejected at even 15% significance level for both models, which is in line with the model assumption. In addition, performing Engle's Lagrange multiplier ARCH test on standardized residuals, we do not discover significant ARCH effects.

Table 2: Empirical Estimation of GARCH Models and GED-EGARCH (1,1) models

	GARCH (1,1)	EGARCH (1,1)
λ	5.81E-7 (-0.002)	-0.223591 (0.002)
α	0.109971(0.000)	0.978391(0.000)
β	0.880181(0.000)	
$\beta 1$		-0.0592288(0.001)
$\beta 2$		0.197989(0.000)
υ		1.896958
LogL	7484.548	7501.1206
AIC	-7.4359	-7.4504
$\alpha+\beta$	0.990151	
Skewness	-0.12981	-0.078413
Excess Kurtosis	0.49239	0.36658
JB	25.976(p<0.0001)	13.3270(p=0.00127)
LB1 Q (10)	10.098(p=0.4319)	10.3827(p=0.40758)
LB2 Q (10)	4.8001(p=0.7787)	8.95587(p=0.53630)
ARCH 1-10 test	0.4890(p=0.8982)	0.90634 [p=0.5263]

JB: Jarque-Bera (1980)'s normality test

LB1 Q (10): Ljung-Box Q-statistic for serial correlation in standardized residuals (10 lags)

LB2 Q (10): Ljung-Box Q-statistic for serial correlation in squared standardized residuals (10 lags)

LM ARCH 1-10 test: Engle's Lagrange multiplier test on standardized residuals for ARCH effects (10 lags)

Compared with GARCH (1,1), the EGARCH (1,1)-GED has two improvements: it takes into account asymmetry effects and allows for non-normal distribution. This coincides with the higher log likelihood and the lower AIC value of EGARCH (1,1)-GED, which indicates a better fit.

A final comment on the in-sample performance is that the results of misspecification tests above (Jarque-Bera's normality test, Ljung-Box test for serial correlation, and Engle's Lagrange Multiplier ARCH effects tests) sometimes do not give much insight into the out-of-the-sample forecasting performance. Nelson (1991) proves that even if the misspecification is quite severe for both conditional mean and the dynamic of the conditional variance, (G)ARCH type models could still generate consistent one-step-ahead conditional variance forecasts and short term estimates. Nelson and Foster (1995) further point out that (G)ARCH type models can produce consistent medium-and long-term variance forecasts.

6.2 Out-of-sample Comparison

In this section, we discuss the out-of-sample predictive power of the alternative forecasting models described in section 3.

The main results of model evaluation are presented in the table 3 and 4. In Table 3 the value and ranking of all eight competing models under RMSE and MAE are reported, while table 4 contains the second part of evaluation, the efficiency and unbiasedness tests of option based predictions.

Table 3 : Evaluation of Predictive Power of Random Walk, Moving Average EWMA, GARCH (1,1), EGARCH (1,1), and Implied Volatility models: RMSE and MAE*

	MODEL	FORECAST HORIZON (DAYS)							
	RMSE	Rank	**5**	Rank	**10**	Rank	**20**	Rank	**40**
1	RW	8	0,01801	7	0,01365	7	0,01141	5	0,00945
2	MA (5)	4	0,01502	3	0,01149	4	0,01028	4	0,00870
3	MA (20)	6	0,01530	6	0,01221	6	0,01069	7	0,01104
4	EWMA (5)	3	0,01499	2	0,01145	3	0,01016	3	0,00861
5	EWMA (20)	1	0,01488	5	0,01179	5	0,01033	6	0,01007
6	GARCH (1,1)	7	0,01758	8	0,01479	8	0,01333	8	0,01451
7	EGARCH (1,1)-GED	5	0,01509	4	0,01162	2	0,00942	2	0,00839
8	OPTION IV	2	0,01493	1	0,01049	1	0,00602	1	0,00442
	MAE**	Rank	**5**	Rank	**10**	Rank	**20**	Rank	**40**
1	RW	7	0,80848	6	0,62160	5	0,59667	3	0,56204
2	MA (5)	3	0,67365	2	0,52929	4	0,56929	5	0,63285
3	MA (20)	5	0,69840	7	0,67305	8	0,71785	8	0,99164
4	EWMA (5)	1	0,66872	1	0,52862	3	0,56175	4	0,61990
5	EWMA (20)	2	0,67294	5	0,61978	6	0,66912	7	0,88771
6	GARCH (1,1)	8	0,84054	8	0,73129	7	0,686359	6	0,85349
7	EGARCH (1,1)-GED	4	0,67823	4	0,55839	2	0,48545	2	0,50270
8	OPTION IV	6	0,73490	3	0,53294	1	0,39321	1	0,29842

* The calculation of RMSE and MAE is based on annualized predicted variance and annualized realized variance for each respective horizon. RW is random walk forecast. MA (5) is moving average forecast based on previous 5 volatilities, while MA (20) is based on previous 20. Similarly, EWMA (5) is based on previous 5 volatilities and EWMA (20) is based on previous 20. ** MAE measure is multiplied by 100

According to the results of Table 3 it is noted that following the RMSE statistic IV model in most cases provides the most accurate forecasts (only for 5-day horizon, IV model is placed second, performing 0,3 % worse then a winner, EWMA (20) model). The discrepancy of forecast accuracy among models ranges for different forecast horizons. For example, for weekly horizon it equals 21%, while with the raise of horizon it steadily increases and reaches 228% for bimonthly forecasts. Although many previous findings report smaller RMSE for IV than for historical volatility models, some papers found controversial results. In Lamuoreux and Lastrapes (1993), (the analysis of 10 individual stock options traded on CBOE from 1982 to 1984) it was documented that implied variance has smaller RMSE for only two companies out of ten, for remaining companies, the GARCH and the naïve forecasts have lower RMSE then the IV. However, such difference can be attributed to many factors: different markets, time spans, data frequencies, methods of IV derivation and forecasts horizons.

The GARCH (1,1) model produces the highest overall RMSE values. This is quite in line with the mediocre performance of this model that was documented in other papers. In the extensive study by Poon and Granger 2003, where the results of 93 papers were compared from the perspective of competing models that were divided into 4 categories (simple historical models, GARCH, Stochastic volatility and Implied volatility models), GARCH model performs worse than any other. For example, simple historical based forecasts outperformed GARCH model in 56% of all cases. Similar results can be found in papers by Ederington and Guan (2002), where S&P Futures were studied, and Taylor (2001) where the range of different indexes was reviewed.

The MAE statistics favors the exponentially weighted moving average based on previous 5 volatility values for weekly and biweekly horizons, while for longer horizons implied volatility outperforms other models. Extension of the model to 20 previous volatilities worsens the results emphasizing the short memory of the volatility series.

The second best model is EGARCH for monthly and bimonthly forecasts and for 5- and 10-day horizons are EWMA (20) and MA (5) consequently. Taking into account asymmetry and releasing the normality assumption greatly improve the GARCH model performance in terms of mid-term forecasts. Although the gap between IV and EGARCH models is quite noticeable.

Another finding with EGARCH model is that it seems to perform better for mid-term forecast then for short-term forecasts. Under both RMSE and MAE, EGARCH (1,1)-GED ranks the second place for 20 and 40 days forecasts, while for 5 and 10 days horizons it only gains middle scores. This does not surprise us as several previous studies, such as Heynen and Kat (1994) find that volatility is more predictable over longer horizon. Also Poon and Granger (2003) argue that GARCH type model can provide adequate, or even very good forecast for long periods, but not all the time. Even though GARCH type models only produce one variance on each step, they have considered many forms of specifications, and this may be the reason why EGARCH (1,1)-GED model can forecast volatility well in the case of OMX S30 index. Also 40-day forecasting horizon is not that long, and how EGARCH model works for, e.g. 6 month horizon is not studied in this paper.

And as mentioned before, we believe that the pronounced difference in performance of EGARCH (1,1)-GED and GARCH (1,1) model is due to the fact that the former captures the asymmetric effect in variance as well as has more generalized assumption for error distribution, while the latter does not. This indicates that the performance of GARCH type models depends largely on the specific characteristics of the data set in question.

Table 4. Results of predictive power, unbiasedness and efficiency tests of OMX S30 implied vlatility estimates.

Parameter Hypothesis	Constant (H_0: α=0)	Coefficient IV (H_0: β_1=0)	Coefficient EGARCH (H_0: β_2=0)	t-stat IV (H_0: β_1=1)	F-stat (H_0: α=0; β_1=1)	Model
5-day	0,002	0,675**		-3,85**	0,99	18
	0,001**		1,013**			18
	0,001	0,558**	0,316	-3,85**		19
10-day	0,001	0,799**		-2,13***	0,09	18
	0,001		1,018**			18
	0,001	0,795**	0,008***	-1,39		19
20-day	-0,001	0,993**		-0,08	0,66	18
	0,001		1,050**			18
	0,000	1,218**	-0,442***	1,65		19
40-day	-0,001	0,993**		-0,85	0,16	18
	0,001		0, 929**			18
	0,000	0.932**	0.020	-0,59		19

** significant at 5% level, i.e. reject the null hypothesis at 5% significance level
***significant at 1% level, i.e. reject the null hypothesis at 1% significance level.
Model 18 and 19 are regression based efficiency test models shown on page 23.
Applying formula 18 while estimating the EGARCH coefficients, means that in that particular case we replace the IV in the formula by EGARCH forecasts.

In Table 4 the results of 2 regressions (see formulas 18, 19) are presented. We first test the information content of implied volatility (see column Coefficient IV in the table). All the slope coefficients are above zero and rather close to 1 with p-values less then 1%. These results allow us to conclude that indeed implied volatility has a predictive power and this is applied for all forecasting horizons.

The next step is to test whether implied volatility is biased estimate of future volatility. Testing separately for IV coefficient being equal to unit gives different results for various forecast horizons. This null hypothesis of $\beta_1 = 1$ is rejected for weekly and biweekly forecasts but cannot be rejected for 1- and 2-month forecasts, indicating that that option IV is an unbiased forecast of future volatility over mid-term horizon. When testing jointly for alpha being equal to zero and beta being equal to one we found that our results can not be rejected at any conventional level. This joint test statistics produces very strong and important results, implying in our particular case that IV is not only has a predictive power (the results that are basically achieved in other studies) but also that IV gives unbiased forecasts. Majority of previous findings failed to gain such outcome. For example, Lamuoreux and Lastrapes (1993) along with Flemmimg (1998) and Jorion (1995) found that option prices provide biased predictions of future volatility.

Interesting results are obtained by introducing the EGARCH forecast into the model (according to its performance, the EGARCH model was a natural candidate to be put into encompassing regression, (see the table above)). While single parameter regression (containing only EGARCH forecasts as a dependent variable) produces the coefficients that are significantly different from zero, adding EGARCH estimates into the initial regression (see equation 19) shows that in almost all cases it does not contain any additional information: as for 5- and 40-day forecasts, the results are not statistically different from 0, for weekly forecast the coefficient is significant at 5% level, but the value of the coefficient itself is close to zero (0,008). Only monthly forecasts give surprising results. The IV coefficient is significant along with EGARCH beta. While IV beta equals 1,2, EGARCH produces negative slope coefficient (-0,44). According to such results implied volatility forecasts for the next 20 days overperform the market, while EGARCH estimates underperform.

These tests results are in line with Flemming (1998) who documents that the forecast power of implied volatility from S&P 100 options dominates that of historical volatility.

Also Christensen and Prabhala (1998) find that volatility implied by S&P 100 index option prices outperforms past volatility in forecasting future volatility and even subsumes the information content of past volatility in some cases. However according to the evidence from the overlapping analysis of time series observations by Day and Lewis (1992) (S&P 100 index options) and Lamoureux and Lastrapes (1993), in addition to IV that contains some useful information in forecasting volatility, time series models contain information incremental to the implied volatility. Canina and Figlewsky (1993) conclude that S&P implied volatility is such a poor forecast that it is dominated by the historical volatility estimate.

Mainly the biased estimates of IV and failure to confirm efficiency can be explained by measurement errors in data. In almost all papers that studied implied volatility, those errors were found to be the main cause of bias in producing volatility forecasts.

There are several sources of measurement errors in implied volatility. First of all, measurement errors may stem from limitations of the Black-Scholes option pricing formula. The Black-Scholes option pricing formula assumes that the price of the underlying index evolves according to a lognormal diffusion process. Even if the theoretical option pricing formula is correct, market microstructure effects may cause additional measurement errors. Furthermore, it is unlikely that all underlying asset prices reflect trades which are simultaneous with the option trade. Lastly, such market imperfections as transaction costs, taxation, etc. exist.

All in all, the results indicate that, for OMX S30 index, in all cases the implied volatility subsumes most of the forecasting information and that in most cases the coefficient of the EGARCH forecast is not significantly larger than zero.

7 Conclusion

In this paper we examine the predictive ability of a number of models, namely Random Walk, Moving Average, Exponentially weighted moving average, GARCH type models and implied volatility in producing the volatility forecasts. The relative forecasting accuracy of the various volatility models is evaluated using both error statistics and regression based evaluation methods. The latter is applied while proving the efficiency and unbiasedness of implied volatility forecasts.

Our results do not contradict many earlier findings in the literature, where recent research has indicated that implied volatility provides the most accurate volatility forecasts as it contains the “market’s” expectation of the asset’s future volatility. The naïve models and GARCH class of models make the forecasts based on backward looking information and thus tomorrow’s volatility is predicted to be similar to today’s volatility. Among these models, GARCH (1,1) provides the worst overall performance by both RMSE and MAE error measurement. The main problem with GARCH (1,1) is that the normality assumption is violated and also it does not take into account the asymmetric effects in variance, as the performance of EGARCH (1,1)-GED model is much better, specially for 20 and 40 days forecasting horizons.

Generally, the naïve models give better forecasts for short-term horizon than mid-term horizon. Particularly the EWMA model performs well for 5- and 10-day forecast and EWMA (5) ranks first under MAE for these two short horizons. As all the naïve models make use of no longer than 20 previous volatility values, we can see from the forecasting results that the volatility series of OMX S30 index have relatively short memory.

In regression-based efficiency tests, we find that both of the two best performers from RMSE and MAE measures have significant predictive power for all forecasting horizons. Further, we find strong support that option implied volatility provides unbiased forecasts for future volatility, especially for 20- and 40-day forecasting horizons. In fact, implied volatility remains significant even in the multiple regression where historical volatility is included, it subsumes the information content of this, and the bias in the implied volatility forecast is insignificant.

Finally, again it is hard to conclude which model is the best. The performance depends on the specific assets, data frequency, forecasting horizons, and also error measurements. The results of the empirical studies for OMX S30 index and during this specific time period may differ significantly from the results of other studies and cannot be applied directly to other markets.

8 Future Research Suggestion

The topic of volatility forecasts is far from being fully researched within this thesis and while writing this work, we have encountered many questions that to our opinion need a separate study or more extensive further research.

For example, while much literature is written on high frequency data, nothing is done on this direction in Swedish market. This can be particular interesting with further application of the results to some trading strategies.

Second moment that can find a broad review within the Swedish market is a search of the optimal weighting scheme for derivation of implied volatility out of the option prices. Other techniques besides the one used in this thesis exist and involve complex calculations and theories.

Third issue that is still questionable is the appropriated measure of volatility (this is discussed in the beginning of section 2). Finally, it is quite interesting to investigate the predictive power of some other approaches not studied in this paper, for instance, the stochastic volatility and long memory models.

9 References

Aguilar, J. (1999). GARCH, implied volatilities and implied distributions: An evaluation for forecasting purposes. Sveriges Risksbank Working Paper Series. No.88.

Akgiray V. (1989) Conditional heteroskedasticity in time series of stock returns:evidence and forecasts, Journal of Business, 62, 55-80.

Black, Fischer. (1976b) The Pricing of Commodity Options. Journal of Financial Economics 3 March/June, pp. 167-79.

Björk, T(2004). Arbitrage Theory in Continuous Time, Second edition, Oxford Press…

Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics, 31, 307-27.

Brailsford, T. J. and Faff, R. W. (1996) An evaluation of volatility forecasting technique, Journal of Banking and Finance, 20, pp. 419-38.

Canina L., Figlewski S. (1993), The Informational Content of Implied Volatility, Rev. Financial Stud. . 6/3 pp. 659-81.

Cho D.C. Frees W. (1988) Estimating the Volatility of Discrete Stock Prices, J Finance, 43/2, pp. 451-66

Chou, R.Y. (1988) Volatility persistence and stock valuation: some empirical evidence using GARCH. Journal of Applied Econometrics 3, 279-294.

Christensen B.J., Prabhala N.R., (1998), The Relation between Implied and Realised Volatility, J. Finan. Econ., 50/2, pp. 125-50

Claessen H. Mittnik S (2002) Forecasting Stock Market Volatility and the Informational Efficiency of the DAXindex Options Market, Center for Financial Studies No. 2002/04

Cont, R. (2000) Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance.

Cont, R. (2005) Volatility clustering in financial markets: Empirical facts and agent–based models.

Day T., Lewis G.M. (1992) Stock Market Volatility and the Information Content of Stock Indez Options, J Econometrics 52, pp. 267-87

Ding Z. et al (1993) A Long Memory Property of Stock Market Returns and A New Model. J. Empirical Finance, 1 pp. 83-106

Dimson, E. and Marsh, P. (1990) Volatility forecasting without data-snooping, Journal of Banking and Finance, 14, 399-421.

Donaldson, R. G. and Kamstra, M. (1997). "An artificial neural network-GARCH model for international stock return volatility", Journal of Empirical Finance, 4(1), 1997, pp 17-46.

Dunis C., Laws J., Chauvin S. (2000), The Use of Market Data and Model Combination to Improve Forecast Accuracy, work. paper, Liverpool BS.

Ederington,L. & Guan,W.(2000) Measuring Implied Volatility: Is an Average Better?, work paper, U. Oklahoma

Engle R. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Vriance of UK Inflation, Econometica 50/4, pp. 987-1007.

Engle R. Bollerslev T. (1986), Modeling the Persistence of Conditional Variances, Econometric Rev., 5, pp. 5-50

Figlewski, S. (1997). “Forecasting volatility" Financial Markets, Institutions & Instruments, Vol. 6 No. 2.

Figlewski, S. (2004) Forecasting volatility, work. paper,

Fleming J. (1998). The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices, J. of Empirical Finance, 5, 4 pp. 317-345

French K., Schwert G.W. and Stambaugh R.F. (1987) Expected Stock Returns and Volatility, J. Finan. Econ. 19/1, pp. 3-30

Hamilton J. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and Business Cycle, Econometrica 57, pp. 357-84.

Hamilton J., Susmel R. (1994), Autoregressive Conditional Heteroscedasticity and Changes in Regime, J. Econometrics 64/1-2, pp. 307-33.

Harvey C.R, Whaley R.E. (1992), Market Volatility Prediction and the Efficiency of the S&P Index Option Market, J. Finan. Econ, 31/1, pp. 43-74

Harvey A.C., Ruiz E., Shepard N., (1994) Multivatiate Stocastic Variance Models, Rev. Econ. Stud. 61, pp. 247-64.

Heynen R. C. (1995), Essays on Derivatives Pricing Theory, Amsterdam: Thesis Publishers.

Heynen R. C.and Kat, H. M. (1994), Volatility Prediction: A Compatison of Stochastic Volatility, GARCH (1,1) and EGARCH (1,1) Models, J. Derivatives, pp. 50-65: Thesis Publishers.

Jarque, C.M., and Bera, A.K. (1980) Efficient tests for normality, homoscedasticity, and serial independence of regression residuals. Economics Letters, 6, 255-9.

Levy, Paul (1925) Calcul des Probabilities, Paris: Gauthier-Villars.

Lopez J. (2001), Evaluating the Predictive Accuracy of Volatility Models, J. Forecast 20/2, pp. 87-109

Lamoureuex D, Lastrapes W, 1993, Forecasting Stock Return Variance: Toward an Understanding of Stochastic Implied Volatilities, Review of Financial Studies 6, pp. 293-326

Martens M, Zein J., Predicting financial volatility: High-frequency time -series forecasts vis-à-vis implied volatility

Nelson, D. (1991), Conditional Heteroscedasticity in Asset Returns: A New Approach, Econometrica 59/2, pp. 347-70

Nelson, D.B. and D. Foster (1995), "Filtering and Forecasting with Misspecified ARCH Models II: Making the Right Forecast with the Wrong Model," Journal of Econometrics, 67, 303-335.

Noh,.J, Engle, R. F and Kane, A. (1995). Forecasting volatility and option prices of the S&P 500 Index. The Journal of Derivatives. Volume 2, number 1.

Pagan, A. and Schwert G. W. (1990) Alternative models for conditional stock volatilities, Journal of Econometrics, 45, 267-90.

Perry, Ph. (1982), The Time-Variance Relationship of Security Returns: Implications for the Return-Generating Stochastic Process, J. Finance 37, pp. 857-70

Poon, S-H and Granger C. (2003), Forecasting Volatility in Financial Markets: A Review. Journal of Economic Literature, Vol. XLI (June 2003) pp. 478 – 539.

Poterba, J. M. and Summers, L.H (1986). "The Persistence of Volatility and Stock Market Fluctuations," American Economic Review 76, pp. 1141-1151.

Timo Teräsvirta, Statistical properties of volatility models and stylized facts of financial time series

Vilasus J. (2002), Forecasting Exchange Rate Volatility, Econ. Letters, 76, pp. 59-64

Yu, J. 2002 Forecasting Volatility in the New Zealand Stock Market, Applied Finan. Econ. 12, pp. 193-202

Zumbach, G. (2002), Volatility Process and Volatility Forecast with Long Memory, work. paper, Olsen Associates

10 Appendix A

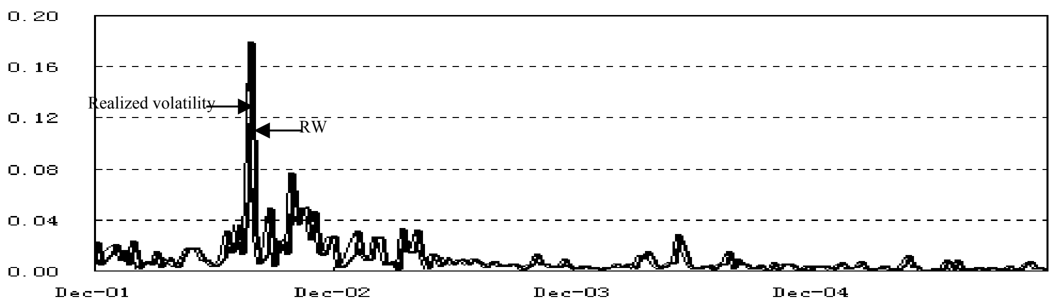
Newton-Raphson method is the most frequently used algorithm. It makes an initial guess for the iteration and then uses the Greek derivative of the option price relative to changes in volatility (the vega) to make a new guess if the initial guess is off the mark.

Unlike Newton-Raphson procedure, Bisection method does not involve derivative calculation, but utilizes linear interpolation, which requires two initial values, i.e. a minimum and a maximum in the iteration process. The number of steps it takes to convert depends greatly on the starting numbers. In general, this method takes more iterations in comparison to the Newton method.

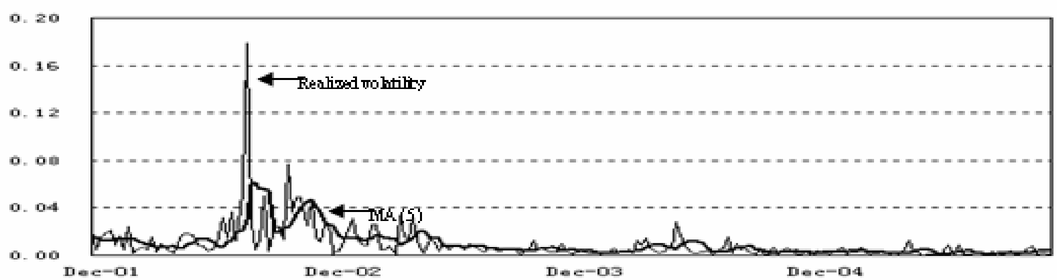
11 Appendix B

Graph 1-32 show out-of-the-sample performance of the models studied in this paper, compared with ex-post realized volatility. The y-axis is annualized variance.

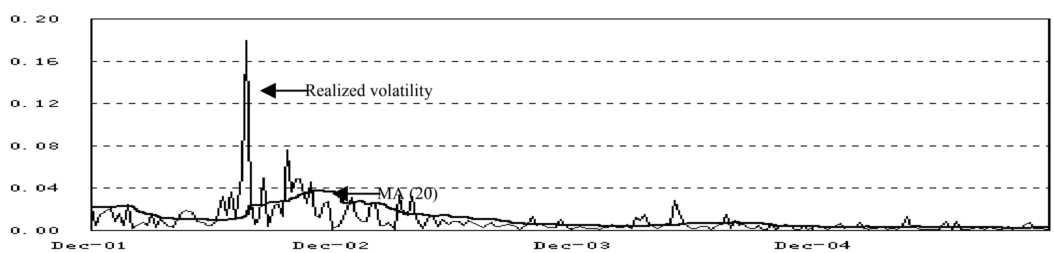
Graph 1: RW *and realized one-week volatility, Dec 2001-Nov 2005*



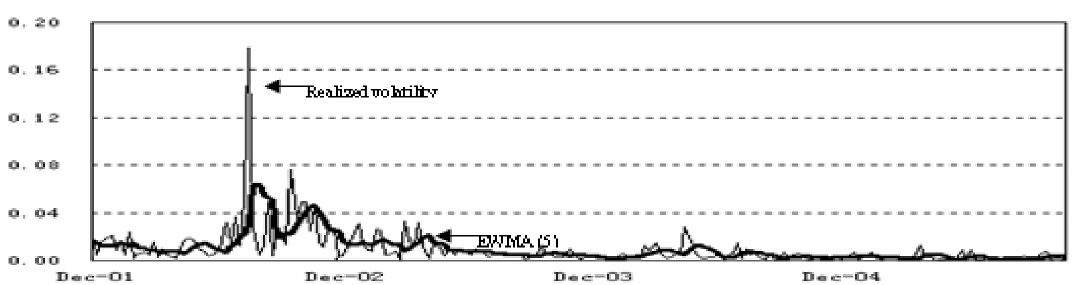
Graph 2: MA (5) *and realized one-week volatility, Dec 2001-Nov 2005*



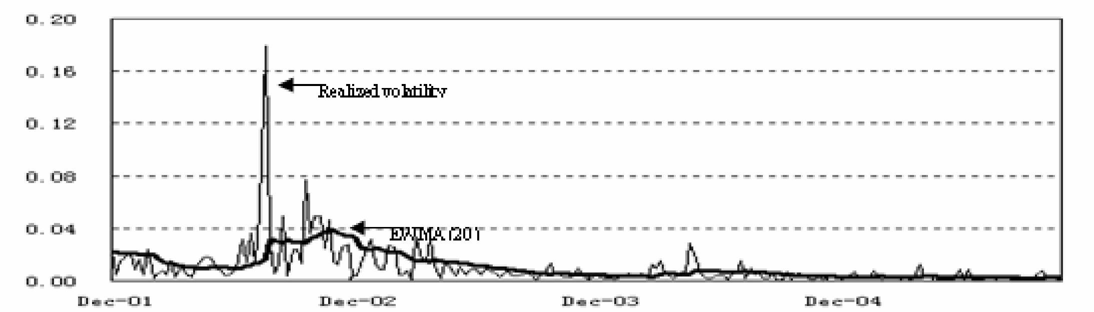
Graph 3: MA (20) *and realized one-week volatility, Dec 2001-Nov 2005*



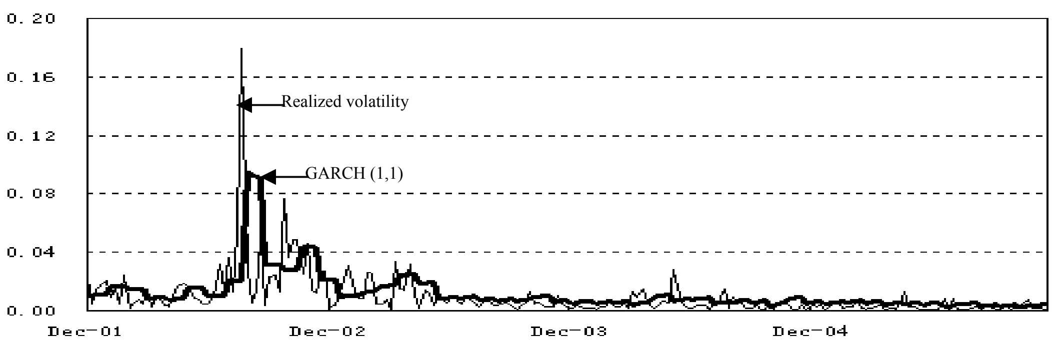
Graph 4: EWMA (5) *and realized one-week volatility, Dec 2001-Nov 2005*



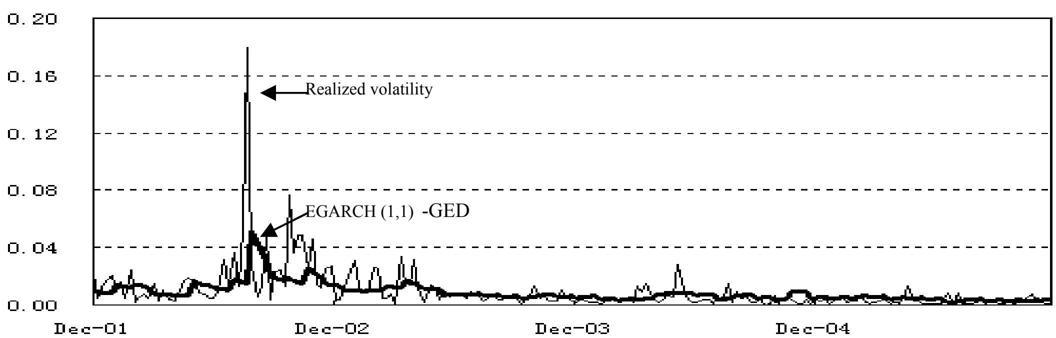
Graph 5: EWMA (20) *and realized one-week volatility, Dec 2001-Nov 2005*



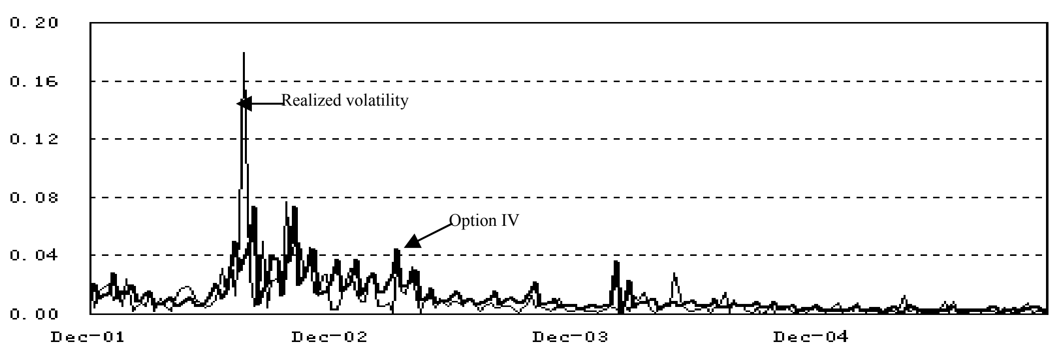
Graph 6: *GARCH (1,1) and realized one-week volatility, Dec 2001-Nov 2005*



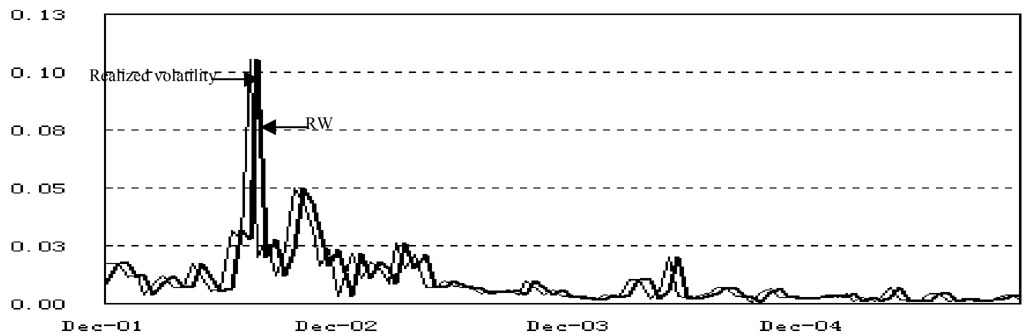
Graph 7: E*GARCH (1,1)-GED and realized one-week volatility, Dec 2001-Nov 2005*



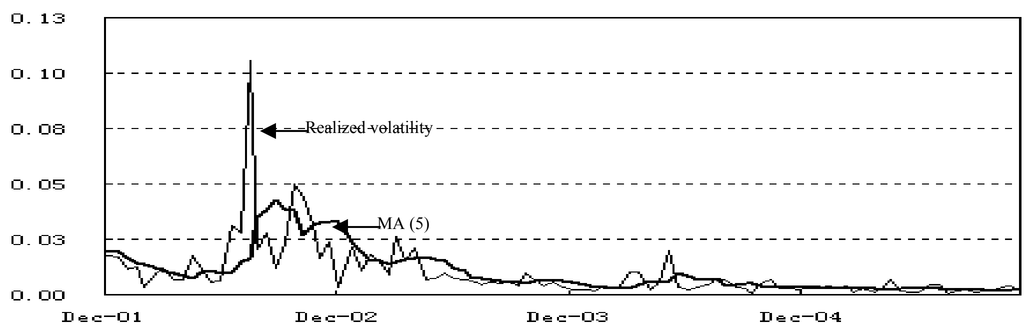
Graph 8: Implied *and realized one-week volatility, Dec 2001-Nov 2005*



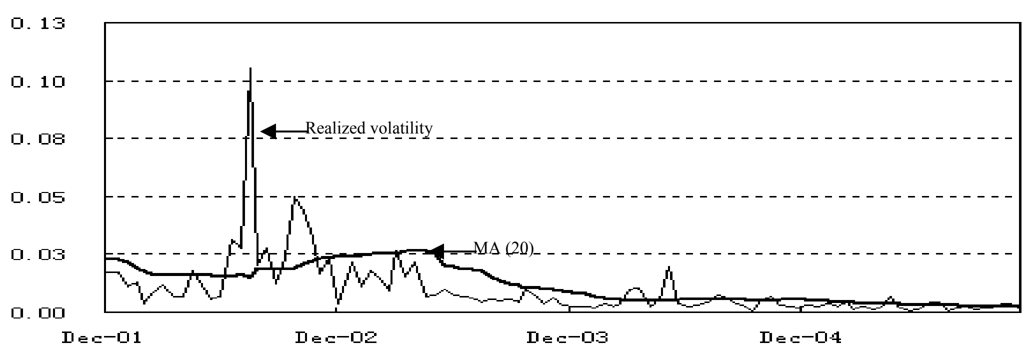
Graph 9: RW *and realized two-week volatility, Dec 2001-Nov 2005*



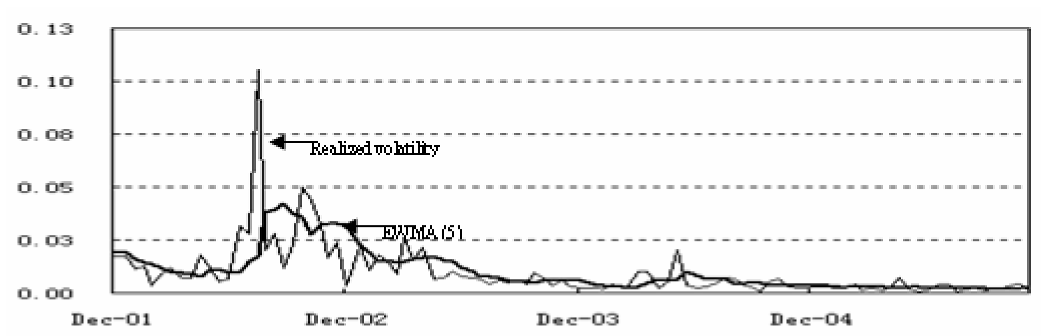
Graph 10: *MA (5) and realized two-week volatility, Dec 2001-Nov 2005*



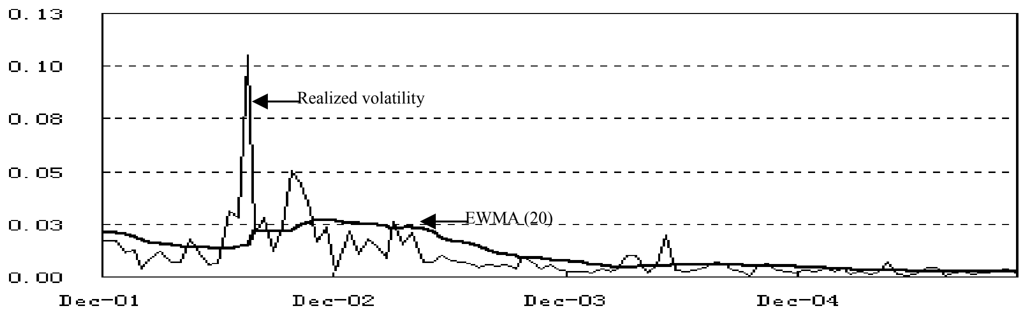
Graph 11: *MA (20) and realized two-week volatility, Dec 2001-Nov 2005*



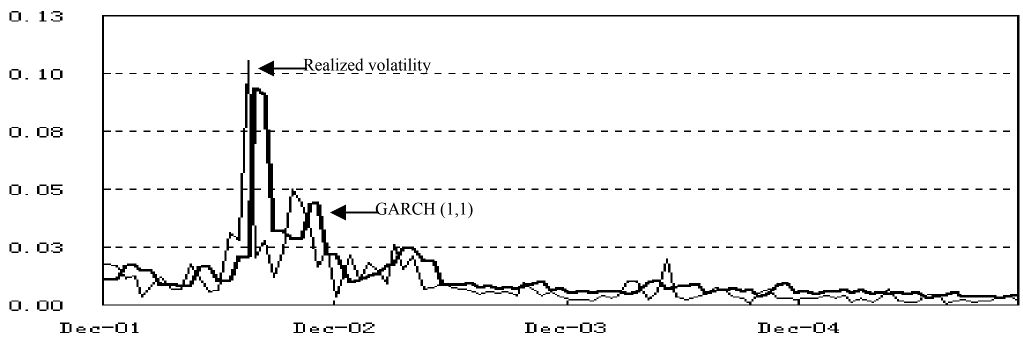
Graph 12: EW*MA (5) and realized two-week volatility, Dec 2001-Nov 2005*



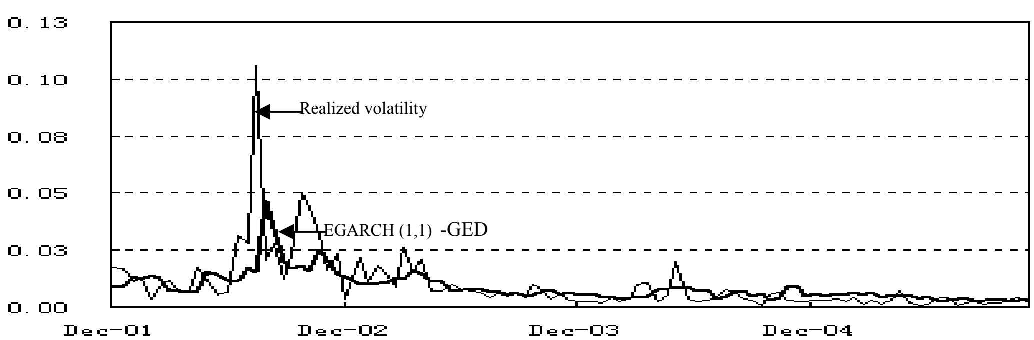
Graph 13: EW*MA (20) and realized two-week volatility, Dec 2001-Nov 2005*



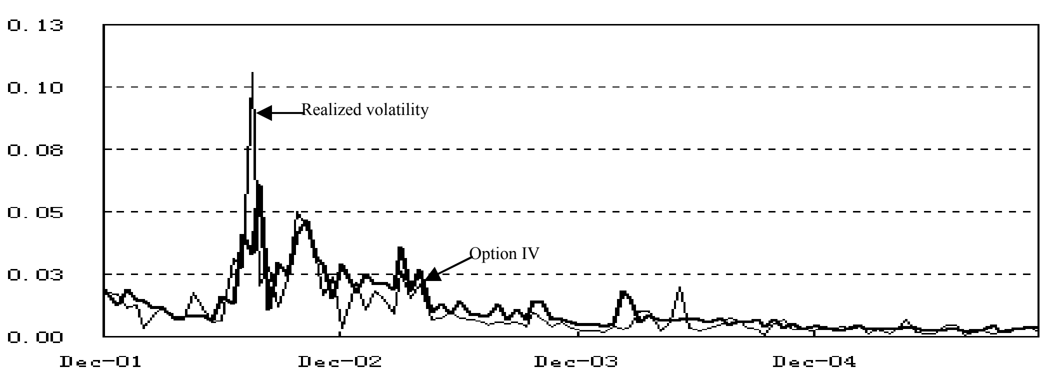
Graph 14: GARCH *(1,1) and realized two-week volatility, Dec 2001-Nov 2005*



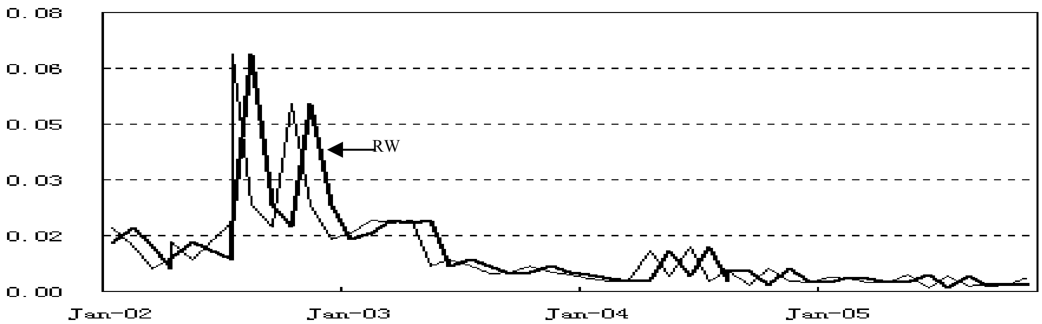
Graph 15: EGARCH *(1,1)-GED and realized two-week volatility, Dec 2001-Nov 2005*



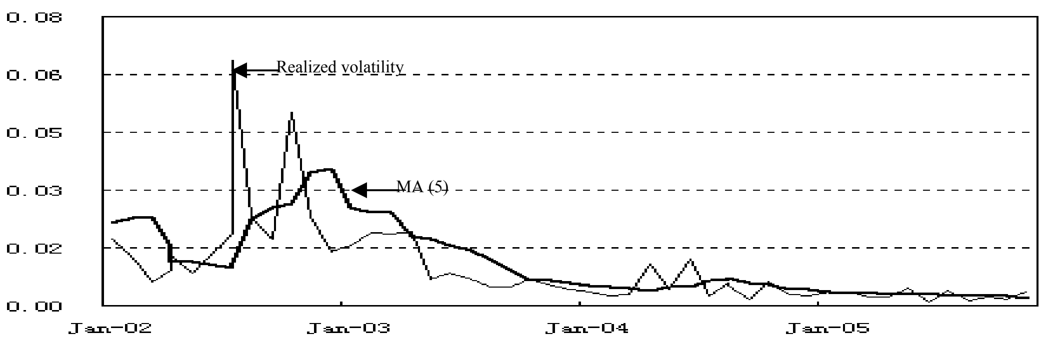
Graph 16: IV *and realized two-week volatility, Dec 2001-Nov 2005*



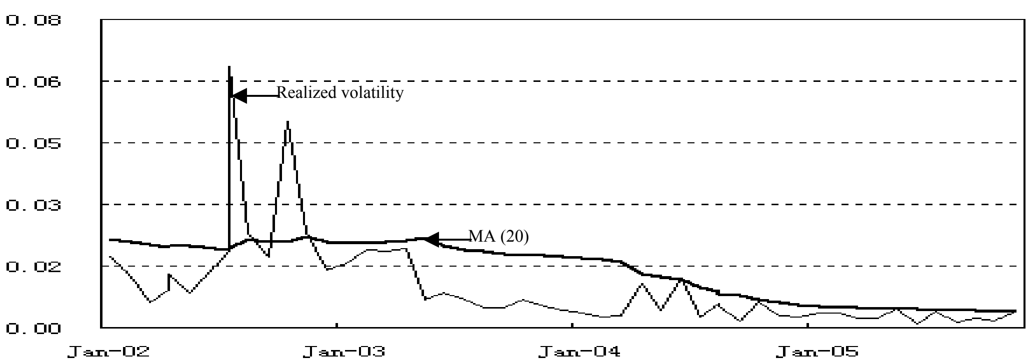
Graph 17: RW *and realized one-month volatility, Dec 2001-Nov 2005*



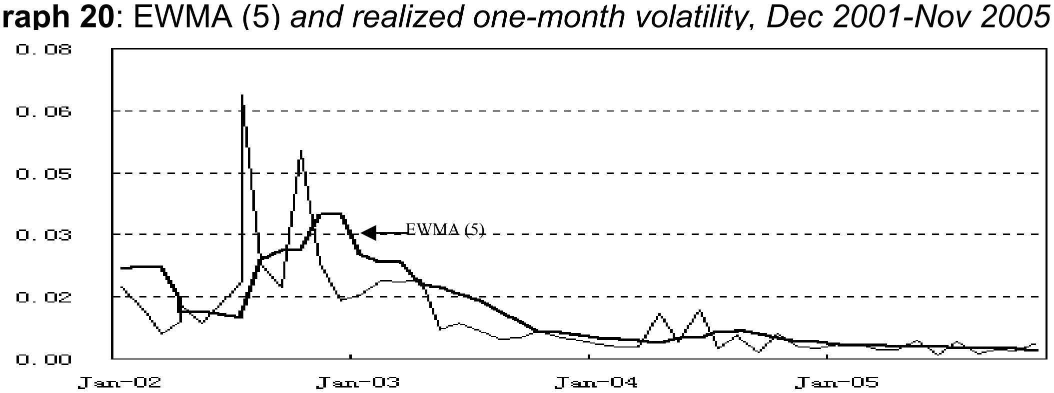
Graph 18: MA (5) *and realized one-month volatility, Dec 2001-Nov 2005*



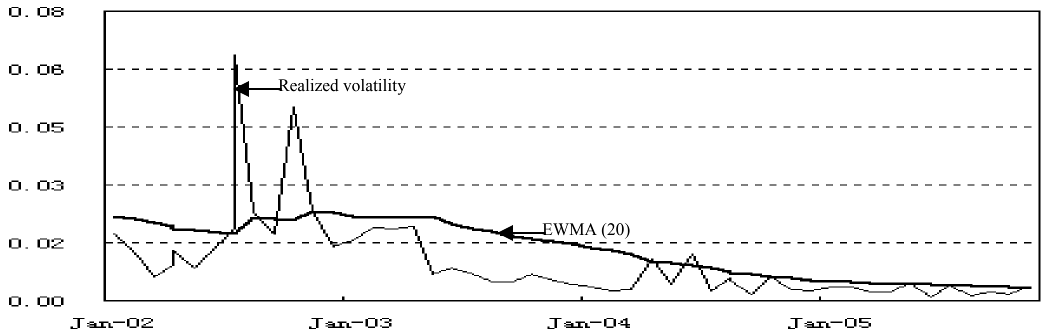
Graph 19: MA (20) *and realized one-month volatility, Dec 2001-Nov 2005*



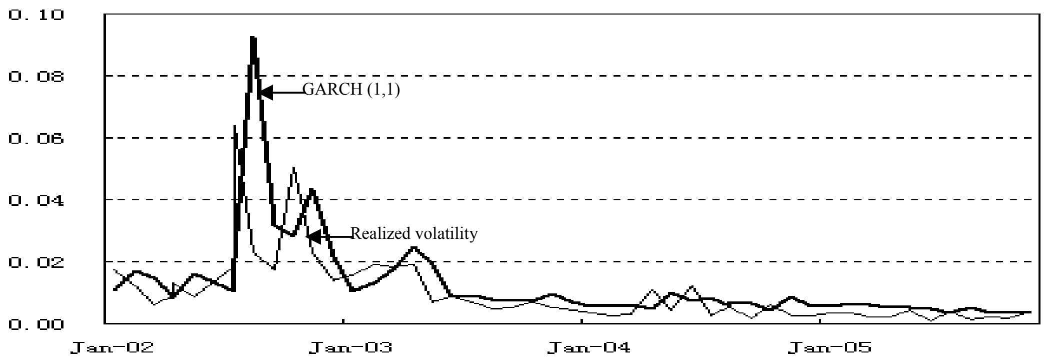
Graph 20: EWMA (5) *and realized one-month volatility, Dec 2001-Nov 2005*



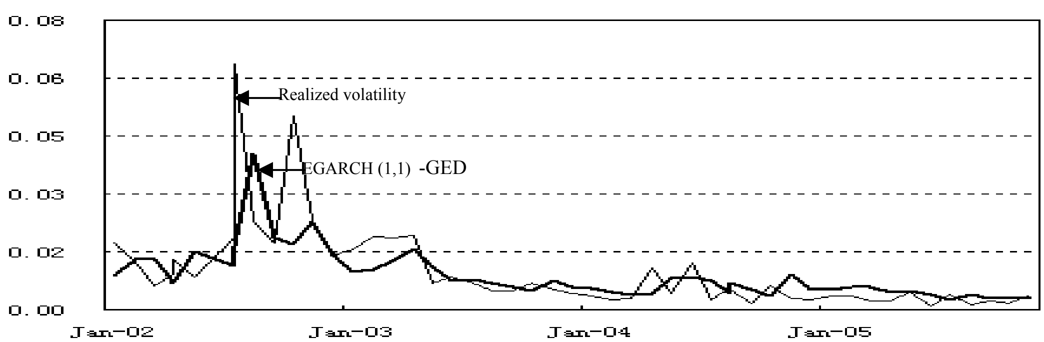
Graph 21: EWMA (20) *and realized one-month volatility, Dec 2001-Nov 2005*



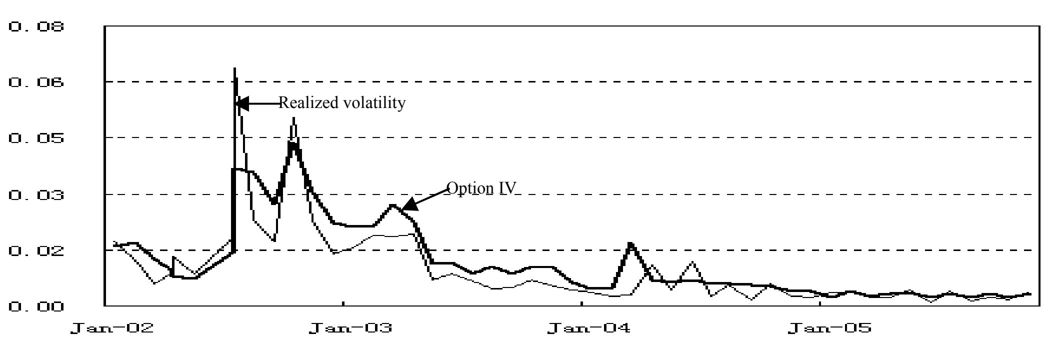
Graph 22: GARCH (1,1) *and realized one-month volatility, Dec 2001-Nov 2005*



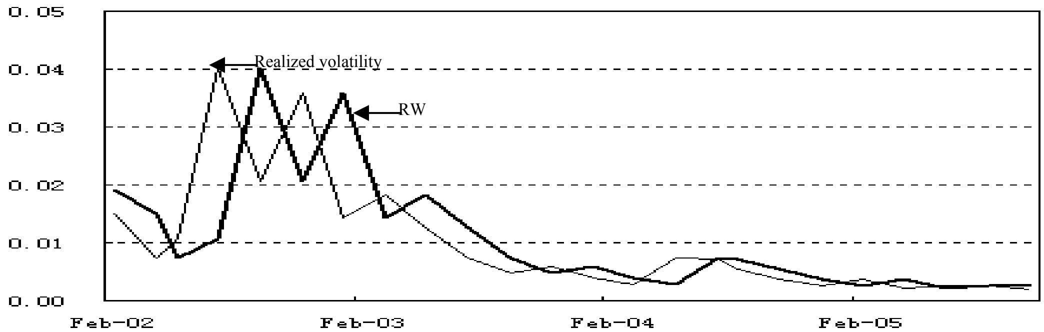
Graph 23: EGARCH (1,1)-GED *and realized one-month volatility,* Dec 2001-Nov 2005



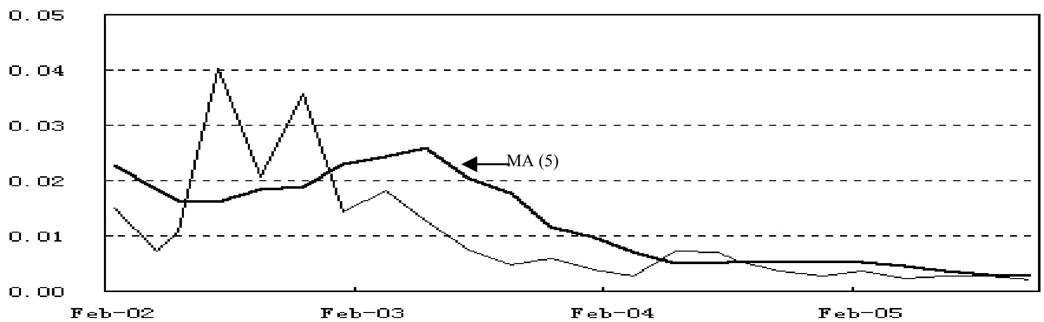
Graph 24: IV *and realized one-month volatility, Dec 2001-Nov 2005*



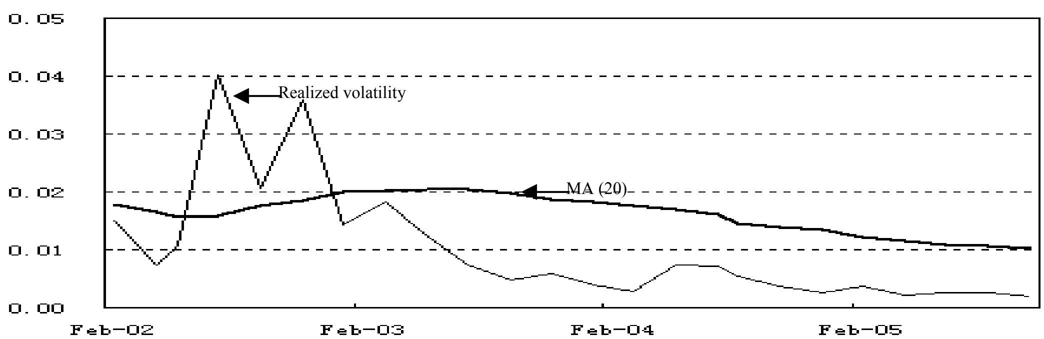
Graph 25: RW *and realized two-month volatility, Dec 2001-Nov 2005*



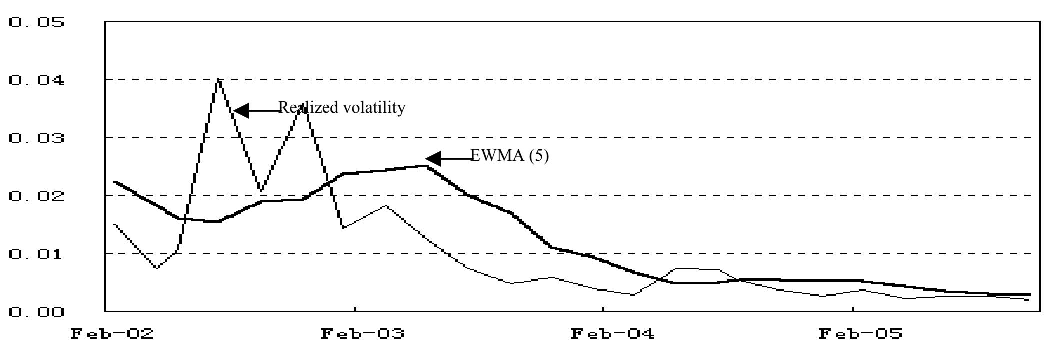
Graph 26: MA (5) *and realized two-month volatility, Dec 2001-Nov 2005*



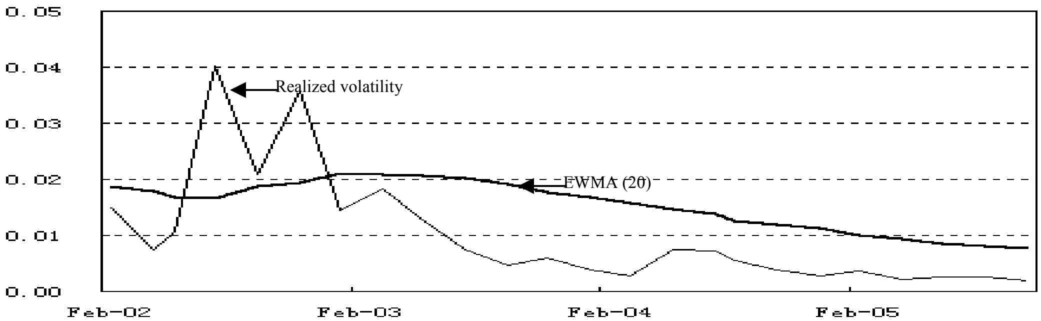
Graph 27: MA (20) *and realized two-month volatility, Dec 2001-Nov 2005*



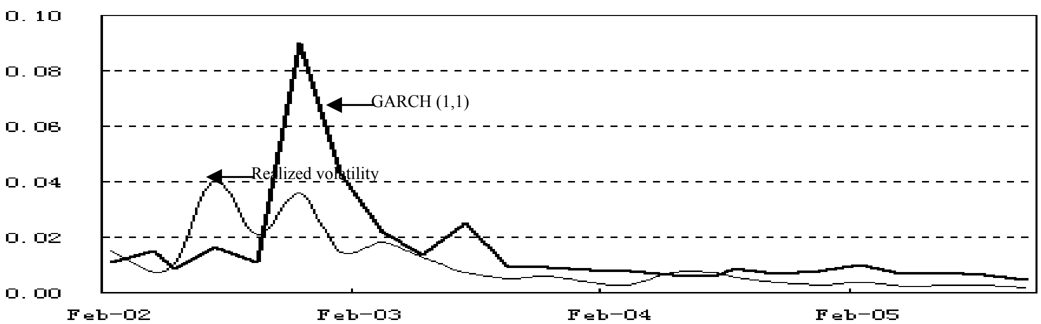
Graph 28: EWMA (5) *and realized two-month volatility, Dec 2001-Nov 2005*



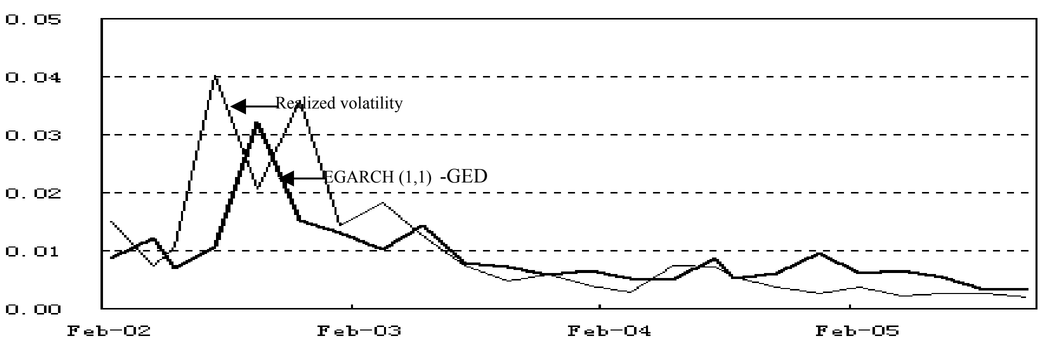
Graph 29: EWMA (20) *and realized two-month volatility, Dec 2001-Nov 2005*



Graph 30: GARCH (1,1) *and realized two-month volatility, Dec 2001-Nov 2005*



Graph 31: EGARCH (1,1) *and realized two-month volatility, Dec 2001-Nov 2005*



Graph 32: IV *and realized two-month volatility, Dec 2001-Nov 2005*