Modeling and Forecasting Volatility in Indian Capital Markets

Ajay Pandey[*]

[*] Author is a member of faculty at Indian Institute of Management, Ahmedabad-380015, India and is thankful to the National Stock Exchange, Mumbai for providing high-frequency data for this and an earlier research work.

Abstract

Various volatility estimators and models have been proposed in the literature to measure volatility of asset returns. In this paper, we compare empirical performance of various unconditional volatility estimators and conditional volatility models (GARCH and EGARCH) using time-series data of S&PCNX Nifty, a value-weighted index of 50 stocks traded on the National Stock Exchange (NSE), Mumbai. The estimates computed by various estimators and conditional volatility models over non-overlapping one-day, five-day and one-month periods are compared with the "realized volatility" measured over the same period. We use three years' (1999-2001) high-frequency data set of five-minute returns to construct measures of realized volatility. In order to test the ability of the estimators and models to forecast volatility, we compare the estimates of unconditional estimators with the realized volatility measured in the next period of same length. For conditional volatility models, the forecasts for the same periods are obtained by estimating models from the time-series prior to the forecast period. Our results indicate that while conditional volatility models provide less biased estimates, extreme-value estimators are more efficient estimators of realized volatility. As far as forecasting ability of models and estimators is concerned, conditional volatility models fare extremely poorly in forecasting five-day (weekly) or monthly realized volatility. In contrast, extreme-value estimators, other than the Parkinson estimator, perform relatively well in forecasting volatility over these horizons.

1. Introduction

Modeling and forecasting stock market volatility is of considerable interest to the practitioners and researchers alike. This has led to considerable research in this area in the past decade or so. The ARCH model, introduced by Engle (1982) and later generalized by Bollerslev (1986) spawned numerous empirical studies modeling volatility in developed markets[1]. Later, there have been quite a few studies focussed on emerging stock markets[2] as well. Researchers have increasingly used conditional volatility models such as ARCH, GARCH and their extensions, as these models have helped them to model some of the empirical regularities. Starting with the pioneering work of Mandelbrot (1963) and Fama (1965), the following features of stock returns have been extensively documented[3] in the literature-

1. Positive serial correlation in volatility or Volatility clustering. Mandelbrot (1963) noticed that "large changes in stock prices tend to be followed by large changes of either sign, whereas small changes tend to be followed by small changes of either sign". This implies that volatility of returns changes with time and that the changes in volatility are non-random.
2. Thick-tailed marginal distribution of returns. Mandelbrot (1963) and Fama (1965) found that the asset returns tend to be leptokurtic.
3. Leverage effect, first noted by Black (1976). The changes in stock prices tend to be negatively correlated with changes in stock volatility. Black (1976) argued that the changes in stock volatility are too large in response to changes in return direction, to be explained by the leverage effect alone. The works of Christie (1982) and Schwert (1989) later supported this conclusion.

[1] For a review of empirical applications of ARCH models on low frequency data, see Bollerslev et al. (1992,1994)

[2] For example, Varma (1999) evaluates the performance of GARCH-GED and EWMA models, in term of goodness of fit, in the context of Indian capital markets.

[3] See Bollerslev et al. (1994).

4. Low volatility during non-trading periods. Fama (1965) and French and Roll (1986) noted that the volatility of returns during trading periods tend to be considerably higher than during non-trading periods.
5. Predictability of volatility. As noted by Cornell (1978) and Patell and Wolfson (1981), the volatility of individual firms' stock returns is high during earning announcement. Predictable changes in volatility have also been found within the trading period. Volatility is typically much higher at the beginning and close of trading period than the rest of the trading period (Harris 1986, Baillie and Bollerslev 1991)
6. Co-movements in volatility. Black (1976) observed that volatility seems to change across stocks. Later, Diebold and Nerlove (1989) and Harvey et al. (1992) in the context of exchange rate volatility movements and Engle et al. (1990) in the context of US bond markets for volatilities across maturities found similar results. Engle and Susmel (1993) and Hamao et al. (1990) and other later studies have also found close links in volatility movements across countries.

The autoregressive conditional heteroskedasticity (ARCH) model proposed by Engle (1982) and its various extensions such as GARCH (Bollerslev 1986), A-GARCH (Engle and Ng 1993), EGARCH (Nelson 1991), ARCH-M (Engle et al. 1987), Components ARCH (Engle and Lee 1993) etc. have been developed to model some of the above-mentioned characteristics of financial time series. In particular, different models have tried to capture time varying second moments of return distributions, time varying and mean reverting second moments, leverage effect, varying first moment, time varying "baseline" second moment, undefined second moment etc.

Despite the ability of the ARCH/GARCH-type models to capture the stylized facts about volatility and return distribution characteristics, their usefulness ultimately depends on their ability to forecast volatility as pointed out by Engle and Patton (2001) recently. Moreover, as pointed out by Poon and Granger (2003) in a recent article reviewing volatility forecasting findings, there are at least three stylized facts in the volatility modeling literature, which have not been captured by ARCH type models. They are-

1. The standardized residuals from ARCH/GARCH models tend to be leptokurtic, i.e., conditional heteroskedasticity alone is unable to explain the tail thickness of returns distribution (Bollerslev 1987, Hsieh 1989).
2. The hypothesis of a unit root in variance has not been rejected by several studies (French et al. 1987, Chou 1988, Pagan and Schwert 1990).
3. GARCH effect disappears once large shocks are controlled for (Aggarwal et al. 1999).

As far as volatility forecasting is concerned, Akigray (1989) in an early work found that GARCH (1, 1) outperformed models based on historical prices. Later works[4] across different countries using different data sets have reported different results. Pagan and Schwert (1990) compared GARCH, EGARCH, Markov switching regime and three non-parametric models for forecasting monthly US stock return volatilities and found that EGARCH followed by GARCH performed moderately, while other models had little prediction power. Comparing GARCH, QGARCH and GJR model for forecasting various European stock market indices, Franses and van Dijk (1996) found that non-linear GARCH models did not perform better than standard GARCH model. Brailsford and Faff (1996) find GJR and GARCH models slightly superior to various simpler models in predicting Australian stock index volatility. From the literature on forecasting volatility, it is clear that it is difficult to forecast volatility. Similarly, the evidence on the performance of various models is mixed. In Indian context, there have been not very many studies comparing forecasting ability of various volatility models, even though ARCH/GARCH type models have been used in various empirical works to model time-varying second moment and serial correlation in volatility.

[4] See Poon and Granger (2003) for a detailed review of 93 papers on volatility forecasting.

Given that the ability of various models to forecast is contingent on the empirical return distribution, the choice of an appropriate volatility model for its forecasting ability is essentially an empirical issue.

In addition to testing comparative performance of various conditional volatility models, another strand has evolved in the literature in the form of research on extreme-value volatility estimators, following work of Parkinson (1980). These estimators are historical unconditional estimators and are based on the range of prices observed during trading, unlike the classical or traditional volatility estimator, which uses closing price. These estimators though not popular, have been shown to be theoretically much more efficient compared to the traditional estimator. One reason for the lack of interest in these estimators is that they could be downward-biased compared to traditional estimator due to discreteness of prices and trading in the stock markets. However, recently Li and Weinbaum (2000) argued that the assumed unbiasedness of the traditional estimator is contingent on the validity of assumption of return generating process. They contend that both the bias and efficiency of extreme value estimators and the traditional estimator is more of an empirical issue. In their empirical work, they use the realized volatility measure using high frequency data, developed by Andersen et al. (2001a), as the benchmark to evaluate the empirical performance of extreme-value estimators vis-à-vis traditional estimators. Besides the issue of bias, extreme-value estimators, unlike conditional volatility models, do not explicitly incorporate the empirical features of returns' distribution discussed above, and are therefore, not as attractive as conditional volatility models. .

In this paper, we report the empirical performance of both historical, unconditional volatility estimators and of conditional volatility models using realized volatility measure as the benchmark. The motivation for comparing two different classes of volatility estimators and models in Indian context stems from the fact that ultimately their usefulness can only be determined empirically. In some ways, the work is similar to the study by Li and Weinbaum (2000) and its replication and extension in Indian capital markets by Pandey (2002). However, these studies did not include conditional volatility models for comparison. Besides including estimates from various conditional volatility models, we also extend the scope of previous studies by investigating the predictive power of the estimators and models. The latter part is similar to studies by Day and Lewis (1992) and Pagan and Schwert (1990). In this work however, we have used realized volatility measure, which has been shown to be model free by Andersen et al. (2001a, 2001b), as the "true" volatility to be forecasted by the estimators and models.

The paper is organized in five sections. In the next section, we review volatility models and estimators proposed in the literature. We also discuss theoretical and empirical issues related to these models and estimators. Our emphasis in this review is on extreme-value estimators as numerous papers and textbooks[5] provide comprehensive review of conditional volatility models. In section 3, we describe the methodology and the data used in this work. Section 4 covers our finding. In section 5, we conclude by summarizing the results and by discussing the directions for future research in Indian capital markets.

2. Review of Volatility Models

There are various classes of models and estimators, which have been proposed in the literature for measuring volatility of asset returns. Models and estimators, assuming volatility to be constant are the oldest ones among the models which have been used to estimate and forecast volatility. These models and estimators measure "unconditional volatility". With the recognition of empirical regularity that the volatility in financial markets is clustered in time and is time varying, these models gave way to models measuring "conditional volatility". In addition, volatility estimated from the value of options, in which typically volatility is the only unobservable parameter for valuation, allowed researchers and practitioners to use "implied volatility", i.e., the market forecast of

---

[5] For example, Bollerslev et al. (1992, 1994), Alexander (1998) etc.

volatility in valuing the traded options. Finally as shown by Andersen et al. (2001a, 2001b), volatility becomes observable and does not remain latent, if high-frequency data is available. The "realized volatility" estimated using high frequency data is model-free under very weak assumptions.

2.1 "Unconditional Volatility" Estimators and Models

2.1.1 Traditional Estimators

Traditionally, the unconditional volatility of asset returns has been estimated using close-to-close returns. The traditional close-to-close volatility (or, variance) estimator (σcc) for a driftless security is estimated using squared returns and is given by-

$$\sigma_{cc}^2 = 1/n \, \Sigma \, (c)^2 \qquad \ldots\ldots \quad (1)$$

where,

n = Number of days (or, periods) used to estimate the volatility

c = $\ln C_t - \ln C_{t-1}$

C_t = Closing price of day t

The mean-adjusted version of the close-to-close estimator (σ_{acc}) is estimated using sample standard deviation and is given by-

$$\sigma_{acc}^2 = 1/(n-1)*[\, \Sigma \, (c)^2 - n\underline{c}^2] \qquad \ldots\ldots \quad (2)$$

where,

$\underline{c}$ = $(\ln C_n - \ln C_0)/n$

While equation (2) provides an unbiased estimate of variance, the square root of the estimator is biased estimator of volatility due to Jensen inequality (Fleming 1998). The statistical properties of the sample mean make it a very inaccurate estimator, particularly for small samples (Figlewski 1997). He suggests taking deviations around zero, i.e., using equation (1), improves the volatility forecast accuracy.

2.1.2 Extreme-Value Estimators

Parkinson (1980), following the work of Feller (1951) on the distribution of the trading range of a security following geometric Brownian motion (GBM), was first to propose an extreme-value volatility estimator for a security following driftless[6] GBM, which is theoretically 5 times more efficient compared to traditional close-to-close estimator. His estimator (σ_p) is given by-

$$\sigma^2_p \quad = \quad 1/(4n \ln 2)^* \Sigma (\ln H_t / L_t)^2 \qquad \ldots\ldots \quad (3)$$

where,

H_t = Highest price observed on day t

L_t = Lowest price observed on day t

Extending his work, Garman and Klass (1980) constructed an extreme-value estimator incorporating the opening and closing prices in addition to the trading range, which is theoretically 7.4 times more efficient than its traditional counterpart. Their estimator (σ_{gk}) is given by-

---

[6] Driftless means that log price process is driftless, i.e., $\mu = \sigma^2 /2$. The process is specified as $dS_t = \mu S_t dt + \sigma S_t dW_t$, where W_t is a standard Brownian motion and S_t is the price of asset at time t.

$$\sigma^2_{gk} = 1/n * \Sigma [0.511(\ln H_t/L_t)^2 - 0.019(\ln (C_t/O_t)*\ln (H_t L_t/O_t^2) - 2 \ln (H_t/O_t)* \ln (Lt/O_t)) - 0.383(\ln C_t/O_t)^2] \quad \ldots\ldots. \quad (4)$$

where,

O_t = Opening price of day t

Both the Parkinson and Garman-Klass estimators despite being theoretically more efficient are based on assumption of driftless GBM process. Rogers and Satchell (1991) relaxed this assumption and proposed an estimator (σ_{rs}), which is given by-

$$\sigma^2_{rs} = 1/n * \Sigma [\ln (H_t/C_t) \ln (H_t/O_t) + \ln (L_t/C_t) \ln (L_t/O_t)] \quad \ldots\ldots. \quad (5)$$

Kunitomo (1992) also proposed an extreme-value estimator based on the range of a Brownian Bridge process constructed from price process, which is 2 times more efficient than Parkinson estimator. His estimator however, cannot be computed directly from the daily data. Later, Spurgin and Schneeweis (1999) proposed an estimator based on the distribution of the range of Binomial Random walk. Their estimator (σ_{ss}) is given by-

$$\sigma^2_{ss} = 1/n^2 * 0.3927*\Sigma (\ln H_t/L_t)^2 - 0.4986 S \quad \ldots\ldots. \quad (6)$$

where,

S = The tick-size of the trades

Recently, Yang and Zhang (2000) proposed an estimator independent of drift, which also takes into account an estimate of closed market variance. The estimators proposed earlier, including the Rogers-Satchell estimator, do not take in to account the closed market variance. This means that the prices at the opening of the market are implicitly considered same as that of closing price on the previous day. The Yang-Zhiang estimator is based on the sum of estimated overnight variance and estimated open market variance. The estimated open-market variance in turn is based on weighted average sum of the open-market returns' sample variance and the Rogers-Satchell estimator with the weights chosen to minimize the variance of estimator. The Yang-Zhiang estimator (σ_{yz}) is given by-

$$\sigma^2_{yz} = 1/(n-1) * \Sigma (\ln O_t / C_{t-1} - \underline{o})^2 + k/(n-1) * \Sigma (\ln C_t / O_t - \underline{c})^2 + (1-k) * \sigma^2_{rs} \quad \ldots\ldots \quad (7)$$

where,

$\underline{o} = 1/n * \Sigma (\ln O_t / C_{t-1})$

$\underline{c} = (\ln C_n - \ln C_0)/n$ or, $1/n * \Sigma (\ln C_t / O_t)$

$k = 0.34 / [1.34 + (n+1)/(n-1)]$

The extreme-value estimators proposed in the literature have been usually derived under strong assumptions. As pointed out earlier, attempts have been made to relax the assumption of driftless price process and closed market variance by Rogers and Satchell (1991) and Yang and Zhang (2000) respectively. Besides these, it is argued that the observed extreme values may reflect certain liquidity-motivated trades (Li and Weinbaum 2000). This could make them less representative of "true" prices as compared to the closing prices.

Besides extreme values being potentially less representative of true prices, extreme values observed are in markets, where the trading is discrete, whereas, extreme-value estimators are derived under assumption of continuous trading. This can induce downward "discrete trading" bias in extreme-value estimators, as the observed highest prices are lower than the "true" highest price and the observed lowest price is higher than the "true" lowest price (Rogers and Satchell 1991, Li and Weinbaum 2000). Rogers and Satchell (1991) addressed this issue by proposing adjustment in their extreme-value estimator by taking into account the number of steps (trades) explicitly. The adjusted Rogers-Satchell estimator (σ_{ars}) is positive root of the following equation-

$$\sigma^2_{ars} = (0.5594/N_{obs})* \sigma^2_{ars} + (0.9072/ N^{1/2}_{obs})* \ln (H_t/ L_t)* \sigma_{ars} + \sigma^2_{rs} \quad \ldots\ldots \quad (8)$$

where,

N_{obs} = Number of observations/ transactions

σ_{rs} = Unadjusted Rogers-Satchell Estimator

Rogers and Satchell also proposed similar correction to the Garman and Klass (1980) estimator. The adjusted Garman-Klass estimator (σ_{agk}) is positive root of the following equation-

$$\sigma^2_{agk} = 0.511*[(\ln H_t/L_t)^2 + (0.9079/ N_{obs})* \sigma^2_{agk} + (1.8144/ N^{1/2}_{obs})*\ln H_t/L_t*\sigma_{agk}] + 0.038*[\ln H_t/O_t* \ln L_t/O_t - (0.2058/ N_{obs})* \sigma^2_{agk} - (0.4536/ N^{1/2}_{obs})* \ln (H_t/L_t)*\sigma_{agk}] - 0.019* \ln (C_t/ O_t)*\ln (H_t L_t/ O_t^2) - 0.383*(\ln C_t/ O_t)^2 \quad \ldots\ldots \quad (9)$$

While theoretically extreme value estimators are shown to be more efficient (5 to 14 times), yet they have not been very popular. This is mainly because these estimators are derived under strong assumptions about underlying returns generating process in the asset markets. It is assumed that the asset prices follow geometric Brownian motion (GBM) and are observable in a market trading continuously. While extreme-value estimators of volatility could be biased if the returns generating process is mis-specified, Li and Weinbaum (2000) point out that the assumed "unbiasedness" of the traditional estimator itself, is contingent on the validity of assumed return generating process. In particular, they show that the traditional estimator based on the sample standard deviation/variance of returns is not an unbiased estimator of the true instantaneous volatility/ variance for the trending Ornstein-Uhlenbeck process having predictable returns and constant volatility. They argue that the bias in the traditional or extreme-value estimators is more of an empirical issue, more so, when it is possible to assess the efficiency and/or bias of the traditional and extreme-value estimators of volatility using realized volatility measured from high frequency data.

Extreme-value estimators proposed in the literature have been tested using simulated stock prices, actual stock prices and recently, using realized volatility measures. Garman and Klass (1980) using simulated data with discrete price changes, show that extreme-value estimators are downward biased. Beckers (1983) using actual data also found downward bias in extreme-value estimators. Studies by Wiggins (1991, 1992) also reached similar conclusions. However, Spurgin and Schneeweis (1999) found that the binomial estimator developed by them outperformed traditional and other extreme-value estimators on daily and intra-day day data of two futures - CME SP500 and CBT Treasury Bonds contracts. Li and Weinbaum (2000) using intra-day high frequency data to measure realized volatility, found overwhelming support for extreme-value estimators for stock indices (S&P 500 and S&P 100) data set, but confirmed the bias of extreme-value estimators for currencies and S&P 500 futures data set despite efficiency gains. Li and Weinbaum investigated the performance of extreme value estimators for two stock indices (S&P 500 and S&P 100), a stock index futures (on S&P 500) and three exchange rates (Deutsche Mark: US$, Yen: US$ and UK Pound: US$).

Though a plausible reason for relatively less research on and application of extreme value estimators could be the time varying characteristic of volatility, yet the use of extreme-value estimators may still be preferred if they are as efficient empirically as implied by the theory. In that case, conditional volatility models, efficient extreme value volatility estimators and high frequency data based realized volatility model could possibly compete for modeling and forecasting volatility for various applications.

2.2 Conditional Volatility Models

Conditional volatility models, unlike the traditional or extreme-value estimators, incorporate time varying characteristics of second moment/volatility explicitly. Following the pioneering work of Engle (1982), various models have been proposed in the literature. The specification of an ARCH (q) model (Engle 1982) is given by-

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \; \varepsilon^2_{t-i} \qquad \text{……… (10)}$$

where,

$\omega, \alpha_1, \ldots, \alpha_q$ = parameters to be estimated

σ_t^2 = conditional variance at period t

q = number of lags included in the model

ε_t = innovation in return at time t

In the ARCH (q) model, the volatility at time t is a function of q past squared returns. For the ARCH model to be well-defined, the parameters should satisfy $\omega>0$ and $\alpha_1 \geq 0, \ldots, \alpha_q \geq 0$. Equation (10) gives the conditional variance equation. In the ARCH/GARCH type models, standard conditional mean equation is usually modeled as r_t = constant + ε_t. Since empirical application of ARCH(q) model required long lag length and a large number of parameters to be estimated, Bollerslev (1986) proposed GARCH (p,q) model in which volatility at time t is also affected by p lags of past estimated volatility. The specification of a GARCH (p,q) is given by-

q p

$$\sigma_t^2 = \omega + \sum_{i=1} \alpha_i \varepsilon^2_{t-i} + \sum_{j=1} \beta_j \sigma^2_{t-j} \quad \text{......... (11)}$$

where,

$\omega, \alpha_1,..., \alpha_q, \beta_1,.., \beta_p$ = parameters to be estimated

q = number of return innovation lags included in the model

p = number of past volatility lags included in the model

The coefficients of the model should satisfy certain conditions for the conditional variance in the GARCH (p,q) model to be well-defined. β_j's in the model capture GARCH coefficients, whereas α_i's capture ARCH coefficients. For the GARCH (1,1) model, these conditions are- $\omega>0$, $0\leq\alpha\leq0$, $0\leq\beta\leq0$, and $\alpha+\beta\leq1$. As pointed out elsewhere in the paper, the basic ARCH/GARCH models have been extended and new models proposed to model returns distribution better. EGARCH is one such model. In this work, we have used only GARCH and EGARCH models to model volatility in Indian stock market. In Exponential GARCH (EGARCH) model, proposed by Nelson (1991), the conditional variance depends upon both the size and the sign of lagged residuals. EGARCH as well as other asymmetric volatility models have been developed to incorporate the "leverage effect" and observed "asymmetric volatility changes with the change in return sign". The specification of EGARCH (1,1) model are given by-

$$\text{Log}\, \sigma_t^2 = \omega + \beta\, \sigma^2_{t-1} + \gamma\, \varepsilon_{t-1}/\sigma_{t-1} + \alpha \left[\, |\, \varepsilon_{t-1}\, | / \sigma_{t-1} - (2/\Pi)^{1/2} \right] \quad \text{....... (12)}$$

In equation (12), ω, α, β, γ are the parameters to be estimated, while other symbols are same as in equation (10) and (11). Besides EGARCH, there have quite a few extensions of basic ARCH/GARCH model proposed in the literature for modeling volatility. In addition, there is a separate class of conditional volatility model called stochastic volatility models, in which the conditional variance specification contains two error terms. In case of ARCH/GARCH models, the conditional variance equation is determined by the information available at that time, with only one error term associated with the past return. Because of computation difficulties, stochastic volatility models are not as popular as ARCH/GARCH type models.

2.3 Realized Volatility

If high-frequency data is available, the volatility becomes observable and does not remain latent. The realized volatility measure developed by Andersen et al. (2001a) therefore, can be used to directly compare performance various volatility models and estimators[7]. The realized volatility measure for day t is given by-

$$\sigma^2_t = \Sigma r^2_{j,t} \qquad \text{……. (13)}$$

where,

$r^2_{j,t}$ = Squared return series of intra-day data

j = Intra-day interval over which returns are being measured

It is possible to annualize the realized volatility so measured, by scaling it up with an annualizing factor. The annualizing factor is simply square root of number of trading days in a year. Measuring realized volatility requires choosing appropriate interval over which the squared returns are used to measure the realized volatility. While shorter time intervals reduce the measurement error, they are also likely to be biased by the microstructure effects (Andersen and Bollerslev 1998,

---

7 Li and Weinbaum (2000) use the realized volatility measure to evaluate empirical performance of extreme-value estimators. In a similar vein, the study by Day and Lewis (1992) use variance of daily returns multiplied by the number of trading days to compute weekly variance for evaluating out-of-sample predictive power of various volatility models.

Andersen et al. 1999). Andersen et al. (2001a, 2001b) and Li and Weinbaum (2000) found that sampling the returns over 5-minute interval is optimal. Without investigating the desirability of using 5-minute returns series on our data set, we have used 5-minute returns to compute the realized volatility.

3. Characteristics of Nifty Daily Returns Time-series and Methodological Issues

In this study, our objective is to empirically investigate the performance of some of the popular volatility models and estimators proposed in the literature. With the availability of high-frequency data being compiled by the National Stock Exchange, a direct comparison of estimates with the model-free realized volatility estimates is possible and hence the realized volatility estimates have been used in the study to assess the bias and efficiency of various volatility models estimators. Traditional close-to-close estimators, various extreme-value estimators and two popular conditional volatility models are estimated and compared with the realized volatility estimates. We also test the ability of these estimators and models to forecast one-day, five-day (approx. weekly) and monthly volatility. In this section, we describe the data set used, the characteristics of daily returns of the index chosen to represent Indian stock market and discuss the performance criteria used to assess the performance of various models and estimators.

3.1 Nifty daily returns characteristics

We use high-frequency data on S&P CNX Nifty, a value-weighted stock index of National Stock Exchange, Mumbai, derived from prices of 50 large capitalization stocks. As the National Stock Exchange started compiling the high-frequency data for research purposes since 1999, our data set covers the period of January 1999- December 2001, i.e., three years. NSE records the data on the index for each day separately. Since forecasting out-of-the-sample volatility required estimating conditional volatility models before the period for which we have high frequency data available, we use daily data on S&P CNX Nifty for the period 1st January 1996 to 31st December 2001 for forecasting volatility on a rolling basis, as well as on the basis of model fitted using data from 1st January 1996 to 31st December 1998. For estimation of conditional volatility similarly, we estimate models using daily data on S&P CNX Nifty for the period 1st January 1996 to 31st December 2002, some part of which falls outside the period for which realized volatility has been measured and compared. We also use daily data for period 1st January 1999 to 31st December 2001 to estimate conditional volatility models, the period which coincides with the period for which the realized volatility estimates are computed. The descriptive statistics of the entire returns series used, i.e., 1st Jan 1996-31st Dec 2002, and its parts, 1st Jan 1996- 31st Dec 1998, 1st Jan 1999- 31st Dec 2001, are given in Table 1a.

Insert Table 1a about here.

As Table 1a shows, the index had a small negative average return in the first sub-period and a small positive average return during the second period. The standard deviation of daily return is of the order of 1.7% in both the sub-periods, implying average annualized volatility of around 27%. The kurtosis of daily returns in each of the period is higher than 3, the kurtosis of Gaussian distribution. It is however, closer to normal in the second sub-period. The Jarque-Bera test for normality of returns distribution yield statistics much greater than any critical value at conventional confidence levels in both the sub-periods.

3.2 Time Dependence in Daily Index Returns and Volatility

Presence of serial correlation in the returns time-series is inconsistent with weak form of market efficiency and also poses issues in modeling volatility directly from daily returns. In Table 1a, we also report autocorrelation coefficients for five lags for each of the three series and associated Ljung-Box Q^* statistic. Except for autocorrelation coefficient associated with first lag for the entire data set, all others are insignificant at conventional confidence levels. The first-order correlation coefficient too is significant at 5% but not at 1% significance level (in terms of Ljung-Box and Box-Pierce tests). As pointed out earlier, we use the entire data set only for one set of estimates of conditional volatility models, the other sets are based on the period for which the high-frequency data was available, i.e., between Jan'1999 to Dec'2001.

A more general test for time-based dependence in the returns series, due to Brock et al. (1996), viz. BDSL test statistics is also reported for epsilon ranging from 0.5 to 2 times of standard deviation and embedding dimensions up to 10, as suggested by Hsieh (1991,1993). This test can detect a variety of departures from randomness including non-linear dependence and deterministic chaos. The BDS test statistic follows standard normal distribution under the null hypothesis. The associated z-statistics on the daily Nifty returns on each of the sub-periods and the entire data set are reported in Table 1b. As can be seen from the table, BDS test strongly rejects the null hypothesis of randomness in the Nifty return series in each of the data set. One possible reason for the non randomness in the returns series is attributed to predictability of volatility, or autocorrelation in volatility.

Insert Table 1b about here.

In order to check the presence of volatility clustering, we report the autocorrelation of squared returns in Table 1c. As can be seen from the table, the Ljung-Box Q^* statistic is significant at 1% level for up to five lags in each of the sub-periods and the entire data set. This confirms volatility clustering in the Indian markets, just as it has been found and reported in case of other markets. In the first sub-period of Jan'1996-Dec'1998, the first order autocorrelation of squared returns though significant at 5%, is not as high as in the remaining data set.

Insert Table 1c about here.

The use of ARCH/GARCH-type conditional volatility models is motivated by the presence of volatility clustering and time-varying characteristics of volatility. In order to test the presence of "ARCH effect", we compute and report the F-statistics and the LM-statistics associated with ARCH-LM test on each data set of Nifty daily returns in Table 1d. While computing these, we use the residuals of OLS residuals of the daily returns regressed on a constant. The number of lags included is five. The results in Table 1d indicate presence of "ARCH effect" in the Nifty daily returns series in each of the data set.

Insert Table 1d about here.

To sum up, our analysis indicates that the daily return series of the index is non-normal and exhibits "ARCH effect".

3.3 Realized Volatility: Descriptive Statistics

In computing volatility measures for the chosen index (S&P CNX Nifty), we faced measurement problems due to trading breaks. On quite a few days during the period, trading was stopped (and later resumed) at NSE because of communication and operational reasons. Since the extreme-value estimators and the traditional estimator are based on extreme values and closing prices are reported for the entire day, we use the squared return series even if there are breaks. In other words, the returns between the breaks are treated as if they are 5-minute returns. This is likely to introduce measurement errors in the realized volatility measure and make them slightly downward biased.

Insert Table 2 about here.

The descriptive statistics of daily realized volatility during the period Jan'1999-Dec'2001 is given in Table 2. For making comparisons with Table 1c easy, the descriptive statistics in Table 2 is reported for the variance rather than volatility. As is clear from the comparison, the mean of realized daily variance is slightly higher than squared returns. On the other hand, the standard deviation of realized daily variance is lower. The auto-correlation at first lag are somewhat similar in magnitude in both the cases, but for lags between 2 to 5, the autocorrelation coefficients drop considerably in case of squared returns, whereas, they drop extremely gradually in case of realized daily variance series indicating greater "volatility clustering or persistence". On closer examination, we find that partial autocorrelation for lags 2 to 5 are considerable lower in case of squared return series. The mean daily realized variance implies annualized volatility of around 31%. The volatility during this period was slightly higher than the average long run volatility, as the capital markets in India were volatile during this period (driven by boom in technology, telecom and media stocks), as is evident from Table 1c.

3.4 Conditional Volatility Models

3.4.1 Symmetric Conditional Volatility Model: GARCH

As pointed out earlier, the return series of the index exhibits ARCH effect in all the periods studied. We use therefore, GARCH (p, q) model, the most popular member of the ARCH class of models, to model volatility of Nifty returns. We use EViews software for model estimation. EViews uses maximum likelihood procedure to estimate the model under the assumption that the errors are conditionally normally distributed. For initialization[8] of variance, by default, EViews first uses the

[8] Different initial variances for maximum likelihood procedure in conditional volatility models could lead to different estimates affecting model performance.

coefficient values to compute the residual of mean equation and then computes an exponential smoothing estimator of the initial values with smoothing parameter, $\lambda=0.7$. Even though the software provides for options for initialization, we have used the default initialization procedure in this work throughout. Using the Schwarz Information Criterion, we find that the best model in the GARCH (p, q) class for $p\in$ [1, 5] and $q\in$ [1, 5] is GARCH (1, 1). The results from the model estimated for different periods are reported in Table 3a. The sum of ARCH and GARCH coefficients (α and β respectively) estimated by the model is close to 0.9 in both the sub-periods as well for the entire data set. However during the 1999-2001 period, the volatility in Indian Capital markets was spikier (higher α) and less persistent (lower β) than the 1996-1998 period and the entire data set. The sum of coefficients being significantly less than one indicates that volatility is mean reverting. The coefficients of the estimated GARCH (1, 1) models are significant as can be seen from the z-statistic reported in Table 3a. This inference from the z-statistics as reported in the table is valid only if errors are conditionally normally distributed. Table 3a also reports the descriptive statistics of residuals from the estimated models. Standardized residuals from estimated GARCH models in each of the period are not normally distributed as indicated by the Jarque-Bera statistic. The standard errors (and therefore associated z-statistics) computed under the assumption of conditionally normally distributed error terms, are not consistent if the errors are not normally distributed. However, Bollerslev and Wooldridge (1992) provided a method for obtaining consistent and robust estimates of the standard errors. The robust standard errors and associated z-statistics computed following Bollerslev and Wooldridge procedure (not reported here), for each of these models are significant, though lower compared to the ones reported in Table 3a.

Insert Table 3a about here.

In order to test whether the GARCH (1, 1) model has adequately captured the persistence in volatility and there is no ARCH effect left in the residuals from the models, ARCH LM test was conducted for lags up to five. The tests (not reported here) indicate that the standardized residuals do not exhibit any ARCH effects.

3.4.2 Asymmetric Conditional Volatility Model: EGARCH

Conditional volatility of returns may not only be dependent on the magnitude of error terms or innovations, but also on its sign. In order to test for asymmetries in volatility, we compute cross-correlation between the squared residuals of the GARCH (1, 1) model and the lagged residuals. In the presence of the asymmetry of conditional volatility, these correlations should be negative. As shown in Table 3b, the cross-correlation for the entire data set as well as for the period 1999-2001 are significant for up to lags of three.

Insert Table 3b about here.

Since there is asymmetry in volatility in the period used for comparing performance of various estimators and models with the realized volatility, i.e., Jan'1999 to Dec'2001, we estimate EGARCH (1,1) models for each of the three periods. The results for estimated EGARCH (1,1) model are reported in Table 3c. The results are consistent with the test for asymmetry in conditional volatility as reported In Table 3b. The asymmetry term, γ, is insignificant in the period Jan'1996 to Dec'1998, but is significant in the other sub-period as well as for the entire data set. The other coefficients are significant at conventional significance levels. The Bollerslev-Wooldridge robust standard errors (not reported here) are higher and z-statistic lower than under the assumption of conditionally normally distributed error terms. Like in case of GARCH model, the standardized residuals from the estimated models are not normally distributed. The ARCH-LM test on residual (not reported here) indicates that there is no ARCH effect left after estimating the model. While the insignificance of asymmetry term, γ, for the period 1996-1998 does not affect the evaluation of EGARCH model for estimation, it does affect its evaluation for its forecasting ability as the period between 1999-2001 over which the forecasts are made, this term is large and significant. For evaluation of models for their forecasting ability, we have used the period of 1996-1998 for estimation of the model, which then has been used to predict out-of-sample (1999-2001) volatility. This constraint was due to availability of high frequency data only after 1999.

Insert Table 3c about here.

In this paper, we use only these two commonly used conditional volatility models from the class of ARCH/GARCH type models to test their performance vis-à-vis traditional and extreme-value unconditional volatility estimators.

3.5 Performance Criteria for Evaluation of Estimators and Models

In order to compare bias and efficiency of various estimators and models for estimation, we use following finite sample criteria-

1. Bias of the Estimator
2. Mean Square Error of the Estimator
3. Relative Bias of the Estimator
4. Mean Absolute Error of the Estimator

The first and the second criterion measure bias and efficiency respectively and are standard measures. The third criterion is to assess the magnitude of bias with respect to the true parameter (realized volatility measure, in this case) as the first criterion gives only absolute amount. The fourth criterion is another measure of efficiency like the second one but is less likely to be affected by the presence of outliers in the data set.

If true volatility (realized volatility) on day t is σ_t and the estimated volatility given by an estimator or model is σ_{est}, then the five performance criteria are computed as under-

Bias = E ($\sigma_{est} - \sigma_t$)

Mean Square Error (MSE) = E [($\sigma_{est} - \sigma_t)^2$]

Relative Bias = E [($\sigma_{est} - \sigma_t)/ \sigma_t$]

Mean Absolute Error (MAE) = E [Abs ($\sigma_{est} - \sigma_t$)]

For forecasting, we use h-period volatility estimates of "unconditional volatility" estimators for forecasting volatility h-period ahead. In case of conditional models, we forecast based on model parameters estimated from the period out side the period of study, i.e., of 1st January 1996 to 31st December 1998. In case of GARCH model, we also report result based on estimation of model on a rolling basis. For example, for forecasting volatility on 1st January 1999, we use model estimated on daily data from 1st January 1996 to 31st December 1998. In case of forecast for 2nd January 1999, we re-estimate the model parameters from data of period- 1st January 1996 to 1st January 1999, and so on. For evaluating the forecasts given by the models, we use the same criteria, as we do for estimation. We use term "forecast error" in place of "bias" in the context of evaluating predictive power of estimators and models. In addition to forecast error, mean square forecast error, relative forecast error and mean absolute forecast error, we also report the results of OLS regressions of the realized volatility on a constant and forecast value given by the various models and estimators following one of the approach used by Day and Lewis (1992) to test out-of-sample predictive power of volatility models. We also report the results of the other approach followed by them, i.e., forecast encompassing regressions based on a procedure due to Fair and Shiller (1990). Following this approach, we regress realized volatility on a constant and the forecast values obtained from different models and estimators.

3.6 Close-to-close Market Variance Estimates and Extreme-Value Estimators

While using conditional volatility models and traditional volatility estimators, using closing daily prices does not pose any problems. However, as pointed out elsewhere in the paper, extreme-value estimators prior to Yang and Zhiang (2000) did not take the closed-market variance (between the closing prices of the previous day and opening prices) into account. Similarly, the realized volatility measure, in the absence of continuous trading markets, is essentially a measure of open-market variance of volatility. In order to compare therefore, some of the extreme-value estimators and realized volatility measure need to be modified for estimating close-to-close market variance. In the absence of any observation during the period during which market is closed, treatment of the closed-market variance however, has to be alike for all the estimators. For incorporating the closed-market variance in such estimators, we use traditional unadjusted estimator, as given in equation (1), for one-day period and traditional mean-adjusted estimator, as given in equation (2), for longer periods. The closed-market variance is computed using close-to-open returns.

4. Empirical Results

In this work, we analyzed empirical performance of the volatility models and estimators vis-à-vis the realized volatility measure for the S&P CNX Nifty stock index, in terms of- (a) estimation, and (b) predictive power. Accordingly, we report the results separately on bias and efficiency of these models and estimators for estimation and for their out-of-sample predictive power.

4.1 Volatility Models and Estimation

Estimation of volatility from data over a given horizon is important for researchers and practitioners alike. In order to test empirical performance for estimation, we compute volatility estimates from unconditional estimators and conditional volatility models and compare them with the realized volatility in Table 4. For comparisons, we use three time periods of one-day, non-overlapping five-days and calendar months. We chose these horizons partly because we had only three years' high frequency data and partly because shorter horizons are more likely to be used in case of time-varying volatility, particularly in case of unconditional estimators. While estimating conditional volatility estimates using GARCH and EGARCH models, we report the results based on the estimates from the sub-period (in sample) of Jan'1999-Dec'2001 as well as from the estimates from the entire data set (complete data set from Jan'1996-Dec'2002 including the sample period). In case of traditional estimators, two estimates are reported. The first one is based on separate adjustment for the closed market variance, while the second one is more commonly used one and is based on close-to-close daily returns.

In panel A of table 4, we report the result for one-day period. While volatility estimates given by the GARCH/EGARCH models exhibit lower bias than traditional and extreme-value estimators, the extreme-value estimators given by Garman-Klass and Rogers-Satchell exhibit lower relative bias and higher efficiency in terms of both MSE and MAE criteria. In case of five-day period as reported in panel B, the results are similar. Garman-Klass, Rogers-Satchell and Yang-Zhiang extreme-value estimators perform well compared to the GARCH/EGARCH estimates on Relative Bias, MSE and MAE criteria despite exhibiting higher bias. For results on one-month (calendar month) period, as reported in panel C, these three estimators perform well on both efficiency criteria but exhibit higher absolute and relative bias than conditional volatility models. The extent of bias in case of these three extreme-value estimators as well as conditional volatility models increases with the increase in horizon. Even then, the bias in terms of annualized volatility is less than 1% for one-day period, around 2% for five-day periods and less than 3% for one-month period. In case of conditional volatility models, it is about 1% less than extreme-value estimators for all horizons. The bias exhibited by Parkinson estimator is exceptionally high.

Higher efficiency of extreme-value estimators compared to traditional estimators and conditional volatility models in Indian Capital markets is in line with the earlier findings of Pandey (2002), wherein extreme-value estimators were compared with traditional estimators. Higher efficiency of extreme-value estimators in comparison with conditional volatility models is somewhat surprising though. This could be because while conditional volatility models have been estimated using daily closing prices, the extreme-value estimators take into account intra-day information on prices, similar to realized volatility measure, the benchmark used for making comparisons. These results are also similar to some of the recent works discussed earlier (Li and Weinbaum 2000, Spurgin and Schneeweis 1999).

4.2 Predictive Power of Volatility Models

Besides estimation, the other important application and use of unconditional and conditional volatility models is for forecasting volatility. In case of unconditional estimators, generally h-period volatility estimates are used to forecast h-period volatility ahead. For evaluating the predictive power, we use estimates over a given a horizon, as the forecast for next horizon of equal length and compare it with the realized volatility next period. In case of volatility models, the forecasts are based on the model estimated on out-of-sample data and forecasts are obtained for different length of periods. We use data from the period 1st Jan'1999-31st Dec'1998 to estimate GARCH (1, 1) and EGARCH model for forecasting. As pointed out earlier, we also forecast on the basis of rolling estimation of GARCH model, by successively estimating model to include the data just prior to the forecast period.

The results for one-day ahead forecast performance are reported in panel A of Table 5. Among the various estimators, conditional volatility models (GARCH and EGARCH) perform well for one-day period on all parameters except relative forecast errors. However as can be seen from results for five-day and one-month period in panel B and C, extreme-value estimator perform as well for these horizons on both efficiency criteria (Mean Square Forecast Error and Mean Absolute Forecast Error). The relative forecast error for these estimators is also lower than conditional volatility models; even though mean forecast errors are higher.

In order to test the ability of the estimators and models to forecast volatility, we also regress realized volatility on the forecasted value given by each model and a constant, following Day and Lewis (1992). The specification of the OLS regression is given by-

$$\sigma^2_{t+1} = \beta_0 + \beta_1 \sigma^2_{ft} + \varepsilon_{t+1} \qquad \ldots\ldots (14)$$

where,

σ^2_{t+1} = actual value of "realized variance" at time t+1

σ^2_{ft} = value forecasted for the realized variance of time t+1 at time t

In case the forecasts are accurate, we would expect value of β_0 to be 0 and that of β_1 to be equal to 1. The sign and magnitude of coefficients and R-squared values of these regressions therefore, can be interpreted to assess the predictive power of various models and estimators. We report the values of coefficients, associated t-statistics and R-squared values of these regressions in Table 6. The forecasted values in regressions based on equation (14) are as given by different models and estimators. In panel A, B and C of the table, we report regressions for one-day period, five-day period and one-month period forecasts respectively. As can be seen from panel A, the values of β_0 in all the regressions are significantly different from zero. In terms of R-squared values, conditional volatility models perform the best, though only slightly better than the traditional estimator. In case of five-day and monthly forecasts however, the extreme-value estimators (Garman-Klass, Rogers-Satchell and Yang-Zhiang) perform well. In contrast, the conditional volatility models perform extremely poorly on monthly forecasts. This is expected to some extent, as the forecasts by conditional volatility models are extremely sensitive to recent volatility and errors term by construction. Volatility forecasts by conditional volatility models far out into the future, are likely to result in considerable error if the volatility is not as persistent as estimated by them. Despite significant volatility persistence observed in the squared Nifty returns series and in realized volatility series, latter being more persistent than the former, the forecasting power of conditional volatility models is not vastly greater than the traditional estimator.

We also performed forecast encompassing regression, similar to Day and Lewis (1992), wherein different set of forecasts given by different models and estimators are used as independent variables to test whether they contain different sets of information from each other. The results for forecasts of one-day, five-day and one-month periods are given in panel A, B and C of Table7. For forecast encompassing regression, we chose forecasts of the best traditional estimator from among the traditional estimators, two best performing extreme-value estimators for each horizon from the class of extreme-value estimators, and the GARCH and EGARCH forecasts. The specification used in the regression is given by-

$$\sigma^2_{t+1} = \beta_0 + \beta_1 \sigma^2_{Tt} + \beta_2 \sigma^2_{E1t} + \beta_3\sigma^2_{E2t} + \beta_4 \sigma^2_{Et} + \beta_5 \sigma^2_{EGt} + \varepsilon_{t+1} \qquad \text{....... (15)}$$

where,

σ^2_{t+1} = actual value of "realized variance" at time t+1

σ^2_{Tt} = value forecasted for the realized variance of time t+1 at time t by the best traditional estimator

σ^2_{E1t} = value forecasted for the realized variance of time t+1 at time t by the best extreme-value estimator

σ^2_{E2t} = value forecasted for the realized variance of time t+1 at time t by the second-best extreme-value estimator

σ^2_{Gt} = value forecasted for the realized variance of time t+1 at time t by GARCH (1, 1) model

σ^2_{EGt} = value forecasted for the realized variance of time t+1 at time t by EGARCH model

As is evident from panel A of Table 7, the results from forecast encompassing regressions are in line with the results discussed earlier. In case of one-day period, the GARCH forecasts and traditional estimators have significant coefficients and put together, forecast realized volatility better. Both the extreme-value estimators (Parkinson and Garman-Klass) for one-day period perform poorly. In case of five-day period however, only extreme-value estimators (Yang-Zhiang and Rogers-

Satchell) have significant coefficients and predictive power. In the result of regression using both the extreme-value estimators and other forecasts, the coefficients for both are insignificant, as the forecasts from both the extreme-value estimators are highly correlated (0.99+). For one-month period also, extreme-value estimators (Rogers-Satchell and Garman-Klass) forecast volatility much better than others and are the only ones to have significant coefficients. In the regression involving all the forecasts, the coefficients are once again insignificant due to high correlation (0.99+) of the forecasts given by two extreme-value estimators. Another interesting aspect of one-month period results is that, unlike five-day period, traditional and conditional volatility model forecasts add to the predictive power somewhat. The poor performance of EGARCH model, in term so incremental predictive ability compared to GARCH, can be understood as the model estimated for forecasting over 1996-1998 data did not have significant asymmetric term whereas during 1999-2001 period, the asymmetric term was found to be significant and large. On examining the forecasts given by conditional volatility models closely, it was clear that the variance in their forecasts was much lower than other estimators' forecasts as well as that of realized volatility for each of the three horizons. This implies that the realized volatility is not as persistent as forecasted by conditional volatility models.

Besides relatively superior performance of extreme-value estimators in forecasting volatility for five-day and one-month period ahead, the other striking aspect of our result is that the R-square values are of much higher order than the results of Day and Lewis (1992). Most of the explanatory power in volatility prediction comes from extreme-value estimators, which have mostly not been used in other studies. To that extent, it would be interesting to replicate the study on different samples and contexts.

4. Summary and Conclusions

Modeling and forecasting volatility of capital markets been an important area of inquiry and research in financial economics with the recognition of time-varying volatility, volatility clustering and asymmetric response of volatility to market movements. This stream of research has been aided by various conditional volatility (ARCH/GARCH type) models proposed to handle these empirical regularities. Nonetheless, researchers have found that forecasting volatility is difficult. In this paper, we model the volatility of S&P CNX Nifty, an index of 50 stocks of the National Stock Exchange, Mumbai, using different class of estimators and models.

Our results indicate that while conditional volatility models perform well in estimating volatility for the past in terms of bias, extreme-value estimators based on observed trading range perform well on efficiency criteria. As far as forecasting is concerned, the extreme-value estimators are able to forecast volatility five-day (approx. a week) and one-month volatility ahead much better than conditional volatility models.

In this paper, we have not used "implied volatility" forecasts, used extensively elsewhere, for two reasons. Firstly, the options in Indian Capital Markets have been introduced only recently and therefore, long enough time-series is not available. Secondly, as pointed out by Varma (2002), Indian market seem to underprice (by implication, underestimate) volatility of index options in its short history of pricing options. Nonetheless, comparisons incorporating "implied volatility" forecasts remains potentially an area worth investigating at a future date. Similarly, even though we have used five-minute returns for computing realized volatility, the optimality of use of lower or higher frequency returns needs to be verified empirically. Another interesting area requiring further work in Indian context is to model volatility explicitly for non-trading days and for any plausible "day-of-the-week" effect in returns and volatility.

References:

1. Aggarwal, R., C. Inclan and R. Leal, 1999, "Volatility in Emerging Stock Markets", *Journal of Financial and Quantitative Analysis 34(1),* 33-55.
2. Akgiray, V., 1989, "Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts", *Journal of Business 62,* 55-80.
3. Alexander, C., 1998, "Volatility and Correlation: Measurement, Models and Applications", *Risk Management and Analysis, Vol.I: Measuring and Modeling Financial Risk,* Ed C.Alexander, John Wiley and Sons, New York, 125-171.
4. Andersen, T.G. and T. Bollerslev, 1998, "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts", *International Economic Review 39,* 885-905.
5. Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 1999, "Microstructure Bias and Volatility Signatures", Unpublished Manuscript, New York University.
6. Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 2001a, "The distribution of Realized Exchange Rate Volatility", *Journal of American Statistical Association 96,* 42-55.
7. Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens, 2001b, "The distribution of Stock Return Volatility", *Journal of Financial Economics 61(1),* 43-76.
8. Baillie, R.T. and T. Bollerslev, 1990, "Intra-day and Inter-Market Volatility in Foreign Exchange Rates", *Review of Economic Studies 58*, 567-585.
9. Beckers, S., 1983, "Variances of Security Price Returns based on High, Low and Closing Prices", *Journal of Business 56,* 97-112.
10. Black, F., 1976, "Studies of Stock Market Volatility Changes", *Proceedings of 1976 Meetings of American Statistical Association, Business and Economics Statistics Section,* 177-181.
11. Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity" *Journal of Econometrics 31,* 307-328.
12. Bollerslev, T., 1987, "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *Review of Economics and Statistics 69(3),* 542-547.

13. Bollerslev, T., R.Y. Chou and K.F. Kroner, 1992, "ARCH Modeling in Finance: A Review of Theory and Empirical Evidence", *Journal of Econometrics 52,* 5-59.

14. Bollerslev, T., R.F. Engle and D.B. Nelson, 1994, "ARCH models", *Handbook of Econometrics vol. IV,* Ed R.F. Engle and D. McFadden, North-Holland, Amsterdam, 2959-3038.

15. Brailsford, T.J. and R.W. Faff, 1996, "An Evaluation of Volatility Forecasting Techniques", *Journal of Banking and Finance 20(3),* 419-438.

16. Brock, W.A., W.D. Dechert, J.A. Scheinkman and Le Baron, 1996, " A Test for Independence Based on Correlation Dimension", *Econometric Review 15,* 197-235.

17. Cho, D.C. and E.W. Frees, 1988, "Estimating the Volatility of Discrete Stock Prices", *Journal of Finance 43,* 451-466.

18. Chou, R.Y., 1988, "Volatility Persistence and Stock Valuations: Some Empirical Evidence using GARCH", *Journal of Accounting and Economics 3,* 279-294.

19. Christie, A.A., 1982, "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest rate Effects", *Journal of Financial Economics 10,* 407-432.

20. Cornell, B., 1978, "Using the Options Pricing Models to Measure the Uncertainty Producing Effect of Major Announcements", *Financial Management 7,* 54-59.

21. Day, T.E. and C.M. Lewis, 1992, "Stock Market Volatility and the Information Content of Stock Index Options", *Journal of Econometrics 52*, 267-287.

22. Diebold, F.X. and M. Nerlove, 1989, "The Dynamic of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model", *Journal of Applied Econometrics 4,* 1-21.

23. Engle, R.F., 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica 50,* 987-1008.

24. Engle, R.F., D.M. Lilien and R.P. Robins, 1987, "Estimating Time-varying Risk Premia in the Term Structure: The ARCH-M Model", *Econometrica 55(2),* 391-407.

25. Engle, R.F., V. Ng and M. Rothschild, 1990, "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills", *Journal of Econometrics 45,* 213-238.

26. Engle, R.F., and G. Lee, 1993, "A Permanent and Transitory Component Model of Stock Return Volatility", *Mimeo.*

27. Engle, R.F., and R. Susmel, 1993, "Common Volatility in International Equity Markets", *Journal of Business and Economic Statistics 11,* 167-176.

28. Engle, R.F. and V. Ng, 1993, "Measuring and Testing the Impact of News on Volatility", *Journal of Finance 48,* 1749-1778.

29. Engle, R.F., and A.J. Patton, 2001, "What Good is a Volatility Model?", *Quantitative Finance 1,* 237-245.

30. Fair, R.C. and R.J. Shiller, 1990, "Comparing Information in Forecasts from Econometric Models, *American Economic Review 80*, 375-389.

31. Fama, E.F., 1965, "The Behavior of Stock-market Prices", *Journal of Business 38,* 34-105.

32. Feller, W., 1951, "The Asymptotic Distribution of the Range of sums of Independent Random Variables", *Annals of Mathematical Statistics 22,* 427-432.

33. Figlewski, S., "Forecasting Volatility", *Financial Markets, Institutions and Instruments,* New York University Salomon Center 6(1), 1-88.

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

35. Franses, P.H. and D. Van Dijk, 1996, "Forecasting Stock Market Volatility using (Non-Linear) GARCH Models", *Journal of Forecasting 15(3),* 229-235.

36. French, K.R. and R.Roll, 1986, "Stock Return Variances: The Arrival of Information and the Reaction of Traders", *Journal of Financial Economics* 17, 5-26.

37. French, K.R., G.W. Schwert and R.F. Stambaugh, 1987, "Expected Stock Returns and Volatility", *Journal of Financial Economics 19(1),* 3-30.

38. Garman, M.B. and M.J. Klass, 1980, "On the estimation of Security Price Volatilities from Historical data", *Journal of Business 53,* 67-78.

39. Hamao, Y., R.W. Masulis and V.K. Ng, 1990, "Correlations in Price Changes and Volatility Across International Stock Markets", *Review of Financial Studies 3,* 281-307.

40. Harris, L., 1986, "A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns", *Journal of Financial Economics 16,* 99-117.

41. Harvey, A.C., E. Ruiz and E. Sentana, 1992, "Unobserved Component Time Series Models with ARCH Disturbances", *Journal of Econometrics 52*, 129-158.

42. Hsieh, D.A., 1989, "Modeling Heteroskedasticity in Daily Foreign Exchange Rates", *Journal of Business and Economics Statistics 7(3),* 307-317.

43. Hsieh, D.A., 1991, "Chaos and Nonlinear Dynamics: Application to Financial Markets", *Journal of Finance 46,* 1839-1877.

44. Hsieh, D.A., 1993, "Implication of Nonlinear Dynamic for Financial Risk Management", *Journal of Financial and Quantitative Analysis 28,* 41-64.

45. Kunitomo, N., 1992, "Improving Parkinson Method of estimating Security Price Volatilities", *Journal of Business 65,* 295-302.

46. Li, K. and D. Weinbaum, 2000, "The Empirical Performance of Alternative Extreme Value Volatility Estimators", W*orking Paper,* Stern School of Business, New York (at www.stren.nyu.edu)

47. Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices", *Journal of Business 36,* 394-419.

48. Marsh, T. and E. Rosenfeld, 1986, "Non-Trading, Market Making, and Estimates of Stock Price Volatility", *Journal of Financial Economics 15,* 359-372.

49. Nelson, D.B., 1991, "Conditional Heteroskedasticity in Asset Returns: A new Approach", *Econometrica 59,* 347-370.

50. Pagan, A.R. and G.W. Schwert, 1990, "Alternative Models for Conditional Models for Conditional Stock Volatility", *Journal of Econometrics 45(1-2),* 267-290.

51. Parkinson, M., 1980, "The extreme value method for estimating the variance of the rate of return", *Journal of Business 53,* 61-65.

52. Pandey, A., 2002, "Extreme-Value Volatility Estimators and their Empirical Performance in Indian Capital Markets", *Working Paper #14*, National Stock Exchange, Mumbai (at www.nseindia.com)

53. Patell, J.M. and M.A. Wolfson, 1981, "The Ex-Ante and Ex-Post Price Effects of Quarterly Earnings Announcement Reflected in Option and Stock Price", *Journal of Accounting Research 19*, 434-458.

54. Poon, S.H. and C. Granger, 2003, "Forecasting Financial Market Volatility: A Review", *Journal of Economic Literature 41(2)*, 478-539.

55. Rogers, L.C.G. and S.E. Satchell, 1991, "Estimating Variance from High, Low and Closing Prices", *Annals of Applied Probability 1,* 504-512.

56. Schwert, G.W., 1989, "Why does Stock Market Volatility Change over time?", *Journal of Finance 44,* 1115-1153.

57. Spurgin, R.B. and T. Schneeweis, 1999, "Efficient Estimation of Intra-day Volatility: A Method-of-Moments approach incorporating Trading Range", in *Financial Markets Tick by Tick,* Ed: P. Lequeux, John Wiley & Sons, England.

58. Varma, J.R., 1999, "Value at Risk Models in the Indian Stock Market", *Working Paper 99-07-05,* Indian Institute of Management, Ahmedabad.

59. Varma, J.R., 2002, "Mispricing of Volatility in the Indian Index Options Market", *Working Paper 2002-04-01,* Indian Institute of Management, Ahmedabad.

60. Wiggins, J.B., 1991, "Empirical Tests of the Bias and Efficiency of the Extreme-Value Variance Estimator for Common Stocks", *Journal of Business 64,* 417-432.

61. Wiggins, J.B., 1992, "Estimating the Volatility of S&P Futures Prices Using the Extreme-Value Method", *Journal of Futures Markets 12,* 265-273.

62. Yang, D. and Q. Zhang, 2000, "Drift Independent Volatility Estimation based on High, Low, Open and Close Prices", *Journal of Business 73,* 477-492.

Table 1a: Descriptive Statistics: S&P CNX Nifty Daily Returns

	1st Jan 1996- 31st Dec 1998	1st Jan 1999- 31st Dec 2001	1st Jan 1996- 31st Dec 2002
Observations	743	753	1747
Mean	-3.57E-05	0.000240	0.000106
Median	-0.000625	0.000625	0.000207
Maximum	0.099339	0.075394	0.099339
Minimum	-0.088405	-0.077099	-0.088405
Std. Dev.	0.017043	0.018416	0.016901
Skewness	0.154152	-0.134935	0.003877
Kurtosis	6.902943	4.928598	6.114683
Jarque-Bera	474.5297	118.9840	706.1745
Probability	0.000000	0.000000	0.000000
Autocorr.-(1) Ljung-Box Statistic	0.050 (1.8996)	0.059 (2.6429)	0.050* (4.3963*)
Autocorr. (2) Ljung-Box Statistic	-0.003 (1.9083)	-0.039 (3.7762)	-0.023 (5.3323)
Autocorr. (3) Ljung-Box Statistic	0.033 (2.7446)	-0.014 (3.9284)	0.010 (5.5226)
Autocorr. (4) Ljung-Box Statistic	0.015 (2.9056)	0.035 (4.8800)	0.033 (7.3954)
Autocorr. (5) Ljung-Box Statistic	0.023 (3.3157)	0.015 (5.0419)	0.020 (8.0639)

* Significant at 5% level.

Table 1b: BDS Test's z-statistics for daily Nifty returns

m	1st Jan 1996- 31st Dec 1998	1st Jan 1999- 31st Dec 2001	1st Jan 1996- 31st Dec 2002
2	2.551501*	5.296226**	6.658036**
3	3.176285**	7.558023**	9.169510**
4	3.897283**	9.179223**	10.91606**
5	3.513997**	10.84801**	12.05665**
6	3.121982**	13.50464**	13.87616**
7	3.007875**	17.54613**	16.19647**
8	5.187132**	23.65487**	20.38272**
9	7.227700**	33.13304**	25.77901**
10	8.048009**	50.29900**	35.09079**
2	2.403108*	5.680864**	6.949092**
3	3.058822**	6.861229**	8.806732**
4	3.847627**	7.897389**	10.21271**
5	3.544891**	8.206487**	10.60246**
6	4.047078**	8.912416**	11.65575**
7	4.256483**	10.24643**	13.05368**
8	4.494213**	11.75281**	14.56064**
9	4.533683**	13.13159**	16.04893**
10	4.476401**	14.56015**	17.49432**
2	2.455383*	6.058242**	7.609784**
3	2.989136*	6.791372**	8.989942**
4	3.723518**	7.401422**	10.03836**
5	3.602098	7.563355**	10.26141**
6	4.171508**	7.935225**	10.94753**
7	4.525305**	8.598299**	11.77826**
8	4.777024**	9.248246**	12.53487**
9	4.742911**	9.758905**	13.20222**
10	4.767014**	10.23080**	13.70829**
2	2.776471*	6.237481**	8.080689**
3	3.419191**	6.589254**	9.223566**
4	4.106672**	6.939119**	9.965322**
5	4.036499**	7.136335**	10.14337**
6	4.568144**	7.367190**	10.59916**
7	4.849485**	7.803029**	11.10080**
8	5.008651**	8.184174**	11.51581**
9	4.828123**	8.462457**	11.81751**
10	4.807102**	8.660134**	11.96142**

* Significant at 5% level.

** Significant at 1% level.

Table 1c: Squared S&P Nifty Returns

	1st Jan 1996- 31st Dec 1998	1st Jan 1999- 31st Dec 2001	1st Jan 1996- 31st Dec 2002
Mean	0.000290	0.000339	0.000285
Median	8.56E-05	0.000106	8.37E-05
Maximum	0.009868	0.005944	0.009868
Minimum	0.000000	0.000000	0.000000
Std. Dev.	0.000705	0.000671	0.000646
Autocorr.-(1) Ljung-Box Statistic	0.088 (5.7139*)	0.247 (46.247**)	0.174 (52.718**)
Autocorr. (2) Ljung-Box Statistic	0.085 (11.169**)	0.084 (51.623**)	0.097 (69.218**)
Autocorr. (3) Ljung-Box Statistic	0.018 (11.418**)	0.090 (57.744**)	0.064 (76.471**)
Autocorr. (4) Ljung-Box Statistic	-0.010 (11.488*)	0.046 (59.330**)	0.030 (78.063**)
Autocorr. (5) Ljung-Box Statistic	0.124 (23.050**)	0.085 (64.869**)	0.115 (101.32**)

* Significant at 5% level.

** Significant at 1% level.

Table 1d: ARCH-LM Test statistics on Nifty daily returns

	1st Jan 1996- 31st Dec 1998	1st Jan 1999- 31st Dec 2001	1st Jan 1996- 31st Dec 2002
F-Statistics	4.491289**	11.34598**	17.10076**
LM-Statistics	21.96662**	53.12680**	81.77179**

* Significant at 5% level.

** Significant at 1% level.

Table 2: Descriptive Statistics: Daily Realized Variance of S&P CNX Nifty

	Jan'99- Dec'2001
Observations	737
Mean	0.000383
Median	0.000208
Maximum	0.007979
Minimum	1.89E-05
Std. Dev.	0.000589
Autocorr.-(1) Ljung-Box Statistic	0.261 (50.248)
Autocorr. (2) Ljung-Box Statistic	0.245 (94.743**)
Autocorr. (3) Ljung-Box Statistic	0.237 (136.25**)
Autocorr. (4) Ljung-Box Statistic	0.255 (184.67**)
Autocorr. (5) Ljung-Box Statistic	0.192 (211.98**)

** Significant at 1% level.

Table 3a: Results from GARCH (1, 1) model

	1st Jan 1996- 31st Dec 1998	1st Jan 1999- 31st Dec 2001	1st Jan 1996- 31st Dec 2002
Constant	-9.20E-05 (-0.139151)	0.001097 (1.783237)	0.000386 (1.134965)
ω	2.84E-05 (4.124338**)	3.25E-05 (3.857869**)	1.55E-05 (7.086951**)
α	0.056814 (4.117929**)	0.153048 (5.602553**)	0.102514 (9.857610**)
β	0.846745 (33.39316**)	0.756544 (19.11235**)	0.847406 (84.94265**)
Annualized long-run Volatility implied by ω,α & β	27.13%	29.98%	27.82%
Standardized Residuals: Descriptive Statistics			
Mean	9.43E-05	-0.049790	-0.023647
Std. Dev.	0.999903	0.999320	0.999791
Skewness	0.011157	-0.044657	-0.043485
Kurtosis	7.262294	4.692043	6.204649
Jarque-Bera	562.4400	90.07721	748.1046
Probability	0.000000	0.000000	0.000000

Figures in parenthesis are z-statistics associated with coefficients

* Significant at 5% level.

** Significant at 1% level.

Table 3b: Cross-correlation of Squared Residuals from GARCH (1,1) Models with the Lagged Residuals

	1st Jan 1996- 31st Dec 1998	**1st Jan 1999- 31st Dec 2001**	**1st Jan 1996- 31st Dec 2002**
Lag=1	-0.04226	-0.10034**	-0.07240**
Lag=2	-0.06436	-0.07691*	-0.06780**
Lag=3	-0.01743	-0.08322*	-0.05107*

* Significant at 5% level.

** Significant at 1% level.

Table 3c: Results from EGARCH (1, 1) model

	1st Jan 1996- 31st Dec 1998	**1st Jan 1999- 31st Dec 2001**	**1st Jan 1996- 31st Dec 2002**
Constant	-0.000339 (-0.499085)	0.000424 (0.686293)	4.56E-05 (0.117468)
ω	-0.943289 (-2.982314**)	-1.158396 (-4.752127**)	-0.756000 (-7.896799**)
α	0.133435 (5.769960**)	0.272820 (6.661729**)	0.212420 (13.76127**)
β	0.896355 (23.06698**)	0.882940 (31.20140**)	0.927577 (80.11228**)
γ	-0.030392 (-1.699621)	-0.118988 (-4.542789**)	-0.064712 (-5.346620**)
Standardized Residuals: Descriptive Statistics			
Mean	0.015696	-0.008306	0.003416
Std. Dev.	0.999373	1.001506	0.999910
Skewness	0.005987	0.015417	-0.024480
Kurtosis	7.412521	4.560810	6.420721
Jarque-Bera	602.7737	76.46337	851.9340
Probability	0.000000	0.000000	0.000000

Figures in parenthesis are z-statistics associated with coefficients

* Significant at 5% level.

** Significant at 1% level.

Table 4: Performance of Volatility Models (and Estimators) for Estimation

Panel A: One-day Period

(Number of Observations- 737)

Model/Estimator	Bias	Relative Bias	Mean Square Error	Mean Absolute Error
Traditional Cl-O-Cl♦	-0.003258	-0.193098	0.000103	0.007680
Traditional Cl-Cl♦	-0.003505	-0.193992	0.000126	0.008435
Parkinson	-0.005915	-0.340108	0.000062	0.006031
Garman-Klass	-0.001502	-0.065399	0.000028	0.003543
Rogers-Satchell	-0.001475	-0.064936	0.000038	0.003962
GARCH(1,1)- In sample	0.000681	0.228127	0.000071	0.005854
GARCH (1,1)- Complete	0.001092	0.257060	0.000070	0.005964
EGARCH-In sample	0.000450	0.213318	0.000069	0.005718
EGARCH-Complete	0.000903	0.242548	0.000070	0.005840

Panel B: Five-day Period*

(Number of Observations- 147)

Model/Estimator	**Bias**	**Relative Bias**	**Mean Square Error**	**Mean Absolute Error**
Traditional Cl-Cl♦	-0.002295	-0.132712	0.000026	0.003847
Traditional Cl-O-Cl♦	-0.001673	-0.099923	0.000021	0.003443
Traditional Adj. Cl-Cl♦	-0.002190	-0.132206	0.000032	0.004329
Trad. Adj. Cl-op-Cl♦	-0.002009	-0.122123	0.000031	0.004104
Parkinson	-0.006173	-0.340889	0.000052	0.006173
Garman-Klass	-0.001582	-0.076001	0.000011	0.002247
Rogers-Satchell	-0.001311	-0.060231	0.000012	0.002349
Yang-Zhiang	-0.001399	-0.067837	0.000010	0.002118
GARCH(1,1)- In sample	-0.000162	0.090241	0.000032	0.004098
GARCH (1,1)- Complete	0.000280	0.117163	0.000029	0.004028
EGARCH-In sample	-0.000397	0.078539	0.000032	0.004087
EGARCH-Complete	0.000097	0.106139	0.000028	0.004057

* The volatility estimates for these comparisons are based on average daily volatility estimated over the relevant period and have not been annualized. For converting them in % annualized volatility, the volatility needs to be multiplied with $(N)^{1/2}$ * 100 where N is approx. 250. The same factor will also scale up the reported Bias and Mean Absolute Error while Relative Bias will remain unaffected. The Mean Square Error needs to be scaled up by multiplying with N instead of its square root.

Panel C: Calendar Month*

(Number of Observations- 36)

Model/Estimator	**Bias**	**Relative Bias**	**Mean Square Error**	**Mean Absolute Error**
Traditional Cl-Cl♦	-0.001371	-0.067712	0.000010	0.002362
Traditional Cl-O-Cl♦	-0.001328	-0.067828	0.000009	0.002244
Traditional Adj. Cl-Cl♦	-0.001294	-0.063739	0.000010	0.002282
Trad. Adj. Cl-op-Cl♦	-0.001205	-0.061844	0.000009	0.002144
Parkinson	-0.006420	-0.343027	0.000049	0.006420
Garman-Klass	-0.001717	-0.085808	0.000005	0.001843
Rogers-Satchell	-0.001440	-0.071387	0.000005	0.001716
Yang-Zhiang	-0.001413	-0.070288	0.000004	0.001553
GARCH(1,1)- In sample	-0.000620	0.023725	0.000015	0.002762
GARCH (1,1)- Complete	-0.000131	0.054742	0.000014	0.002718
EGARCH-In sample	-0.000858	0.014670	0.000016	0.002893
EGARCH-Complete	-0.000300	0.046885	0.000014	0.002933

♦ Traditional Cl-op-cl estimator is based on sum of closed market and open market squared returns, whereas traditional Cl-Cl estimator is based on close-to-close squared returns. Similarly, traditional adjusted Cl-Cl estimator is estimated using close-to-close returns, whereas in case of traditional adjusted Cl-op-Cl estimator, open and closed variances are separately measured and added to arrive at daily variance/volatility.

* The volatility estimates for these comparisons are based on average daily volatility estimated over the relevant period and have not been annualized. For converting them in % annualized volatility, the volatility needs to be multiplied with $(N)^{1/2}$ * 100 where N is approx. 250. The same factor will also scale up the reported Bias and Mean Absolute Error while Relative Bias will remain unaffected. The Mean Square Error needs to be scaled up by multiplying with N instead of its square root.

Table 5: Predictive Power of Volatility Models and Estimators

Panel A: One-day Period

(Number of Observations- 736)

Model/Estimator	Mean Forecast Error	Relative Forecast Error	Mean Square Forecast Error	Mean Absolute Forecast Error
Traditional Cl-O-Cl♦	-0.003495	-0.158813	0.000157	0.009154
Traditional Cl-Cl♦	-0.003743	-0.156675	0.000173	0.009755
Parkinson	-0.006151	-0.296215	0.000120	0.007392
Garman-Klass	-0.001739	-0.005943	0.000102	0.006353
Rogers-Satchell	-0.001713	0.004317	0.000121	0.007058
GARCH(1,1)	0.000274	0.243228	0.000083	0.006269
GARCH (1,1)- Rolling	0.000634	0.259618	0.000081	0.006297
EGARCH	0.000253	0.239605	0.000083	0.006196

Panel B: Five-day Period*

(Number of Observations- 146)

Model/Estimator	Mean Forecast Error	Relative Forecast Error	Mean Square Forecast Error	Mean Absolute Forecast Error
Traditional Cl-Cl♦	-0.002295	-0.077617	0.000076	0.006309
Traditional Cl-O-Cl♦	-0.001669	-0.038811	0.000074	0.006304
Traditional Adj. Cl-Cl♦	-0.002166	-0.060767	0.000088	0.007045
Trad. Adj. Cl-op-Cl♦	-0.001985	-0.051693	0.000087	0.006814
Parkinson	-0.006181	-0.302628	0.000084	0.006947
Garman-Klass	-0.001589	-0.028514	0.000050	0.005045
Rogers-Satchell	-0.001320	-0.012849	0.000050	0.005071
Yang-Zhiang	-0.001404	-0.017729	0.000050	0.005076
GARCH(1,1)	-0.000465	0.122342	0.000052	0.005242
GARCH (1,1)- Rolling	-0.000034	0.144828	0.000051	0.005295
EGARCH	-0.000407	0.125388	0.000052	0.005256

* The volatility estimates for these comparisons are based on average daily volatility estimated over the relevant period and have not been annualized. For converting them in % annualized volatility, the volatility needs to be multiplied with $(N)^{1/2}$ * 100 where N is approx. 250. The same factor will also scale up the reported Bias and Mean Absolute Error while Relative Bias will remain unaffected. The Mean Square Error needs to be scaled up by multiplying with N instead of its square root.

Panel C: Calendar Month*

(Number of Observations- 35)

Model/Estimator	**Mean Forecast Error**	**Relative Forecast Error**	**Mean Square Forecast Error**	**Mean Absolute Forecast Error**
Traditional Cl-Cl♦	-0.001225	0.008519	0.000062	0.005938
Traditional Cl-O-Cl♦	-0.001178	0.004903	0.000057	0.005687
Traditional Adj. Cl-Cl♦	-0.001156	0.011293	0.000061	0.005915
Trad. Adj. Cl-op-Cl♦	-0.001062	0.010622	0.000057	0.005711
Parkinson	-0.006333	-0.299390	0.000075	0.006540
Garman-Klass	-0.001608	-0.028542	0.000040	0.004855
Rogers-Satchell	-0.001345	-0.014075	0.000040	0.004831
Yang-Zhiang	-0.001311	-0.010940	0.000041	0.004914
GARCH(1,1)	-0.001126	0.050896	0.000044	0.004854
GARCH (1,1)- Rolling	-0.000534	0.086979	0.000044	0.004940
EGARCH	-0.000949	0.063854	0.000044	0.005043

♦ Traditional Cl-op-cl estimator is based on sum of closed market and open market squared returns, whereas traditional Cl-Cl estimator is based on close-to-close squared returns. Similarly, traditional adjusted Cl-Cl estimator is estimated using close-to-close returns, whereas in case of traditional adjusted Cl-op-Cl estimator, open and closed variances are separately measured and added to arrive at daily variance/volatility.

* The volatility estimates for these comparisons are based on average daily volatility estimated over the relevant period and have not been annualized. For converting them in % annualized volatility, the volatility needs to be multiplied with $(N)^{1/2}$ * 100 where N is approx. 250. The same factor will also scale up the reported Bias and Mean Absolute Error while Relative Bias will remain unaffected. The Mean Square Error needs to be scaled up by multiplying with N instead of its square root.

Table 6: Results of OLS Regression on Predictive Power of Volatility Models and Estimators

Panel A: One-day Period

(Number of Observations- 736)

Model/Estimator	**β_0**	**β_1**	**R-Squared**
Traditional Cl-O-Cl♦	0.000284 (11.76185)	0.340427 (10.85265)	0.138275
Traditional Cl-Cl♦	0.000284 (11.56929)	0.346103 (10.26028)	0.125434
Parkinson	0.000255 (9.414865)	0.864431 (9.254452)	0.10449
Garman-Klass	0.000257 (9.099892)	0.448345 (8.262659)	0.085098
Rogers-Satchell	0.000297 (10.5897)	0.308335 (6.183036)	0.049506
GARCH(1,1)	-0.00037 (-5.39779)	2.45939 (11.77057)	0.158784
GARCH (1,1)- Rolling	-0.00021 (-3.64259)	1.863083 (11.18503)	0.145622
EGARCH	-0.00042 (-5.59641)	2.631491 (11.36021)	0.149532

Figures in parenthesis are t-statistic associated with the coefficient.

Panel B: Five-day Period

(Number of Observations- 146)

Model/Estimator	β_0	β_1	**R-Squared**
Traditional Cl-O-Cl♦	0.000262 (6.727447)	0.366746 (4.873579)	0.141589
Traditional Cl-Cl♦	0.000266 (7.011435)	0.375357 (4.937798)	0.144801
Traditional Adj. Cl-Cl♦	0.000277 (7.053345)	0.332819 (4.261833)	0.112006
Trad. Adj. Cl-op-Cl♦	0.000293 (7.678961)	0.277347 (3.980774)	0.099136
Parkinson	0.000184 (4.449891)	1.219383 (6.615672)	0.233093
Garman-Klass	0.000158 (3.943864)	0.720093 (7.666167)	0.289836
Rogers-Satchell	0.00016 (4.04632)	0.687434 (7.741847)	0.293897
Yang-Zhiang	0.000153 (3.818194)	0.718399 (7.797907)	0.296900
GARCH(1,1)	-0.00024 (-2.05022)	2.024221 (5.496624)	0.173425
GARCH (1,1)- Rolling	-0.00009 (-0.9534)	1.469621 (5.037380)	0.149816
EGARCH	-0.00031 (-2.43191)	2.226531 (5.630326)	0.180424

Figures in parenthesis are t-statistic associated with the coefficient.

Panel C: Calendar Month

(Number of Observations- 35)

Model/Estimator	β_0	β_1	**R-Squared**
Traditional Cl-O-Cl♦	0.000270 (3.299997)	0.348629 (1.844389)	0.093451
Traditional Cl-Cl♦	0.000289 (3.440199)	0.294417 (1.492245)	0.063213
Traditional Adj. Cl-Cl♦	0.000288 (3.464967)	0.293773 (1.525348)	0.065862
Trad. Adj. Cl-op-Cl♦	0.000271 (3.349638)	0.338724 (1.847000)	0.093691
Parkinson	0.000171 (2.16143)	1.308826 (3.365789)	0.255559
Garman-Klass	0.000140 (1.87633)	0.778776 (4.072741)	0.334506
Rogers-Satchell	0.000139 (1.928762)	0.756023 (4.303069)	0.359427
Yang-Zhiang	0.000147 (1.961677)	0.730613 (3.974285)	0.323700
GARCH(1,1)	-0.00010 (-0.25220)	1.60115 (1.26143)	0.046000
GARCH (1,1)- Rolling	0.00007 (0.211446)	0.972845 (0.942551)	0.026216
EGARCH	-0.00030 (-0.60958)	2.234786 (1.397152)	0.055849

Figures in parenthesis are t-statistic associated with the coefficient.

Table 7: Results of Forecast Encompassing Regression

Panel A: One-day Period

(Number of Observations- 736)

Forecast Comparisons	β_0	β_1	β_2	β_3	β_4	β_5	R^2
Garch vs. EGARCH	-0.0004 (-5.27)	-	-	-	1.935 (2.96)	0.610 (0.85)	0.1596
Garch, EGARCH, Trad.	0.000 (-2.62)	0.198 (5.40)	-	-	1.566 (2.43)	0.176 (0.25)	0.1917
GARCH vs. Trad.	0.000 (-2.77)	0.200 (5.46)	-	-	1.714 (6.96)	-	0.1917
EGARCH vs. Trad.	0.000 (-2.74)	0.208 (5.67)	-	-	-	1.773 (6.50)	0.1852
Garch vs. EV (P)	0.000 (-3.77)	-	0.374 (3.39)	-	1.961 (7.72)	-	0.1718
EGARCH vs. EV(GK)	0.000 (-4.25)	-	-	0.151 (2.33)	-	2.239 (7.84)	0.1558
Garch, EGARCH, EV(P)	0.000 (-3.40)	-	0.371 (3.28)	-	1.906 (2.94)	0.069 (0.09)	0.1718
All	0.000 (-1.77)	0.341 (5.46)	-1.041 (-3.15)	0.504 (3.39)	1.916 (2.94)	-0.487 (-0.65)	0.2044

1. Figures in parenthesis are t-statistic associated with the coefficient.

2. For one-day period, the traditional estimator used for forecasts is traditional close-open-close estimator, extreme-value estimators used for forecasts are Parkinson and Garman-Klass Estimators. The choice was based on R^2 as reported in Table 6.

Panel B: Five-day Period

(Number of Observations- 146

Forecast Comparisons	β_0	β_1	β_2	β_3	β_4	β_5	R^2
Garch vs. EGARCH	0.000 (-2.39)	-	-	-	0.710 (0.65)	1.506 (1.29)	0.1829
Garch, EGARCH, Trad.	0.000 (-1.06)	0.130 (0.99)	-	-	0.104 (0.08)	1.596 (1.36)	0.1885
GARCH vs. Trad.	0.000 (-0.77)	0.117 (0.89)	-	-	1.552 (2.40)	-	0.1780
EGARCH vs. Trad.	0.000 (-1.09)	0.136 (1.19)	-	-	-	1.680 (2.77)	0.1885
Garch vs. EV (YZ)	0.000 (1.26)	-	0.735 (5.01)	-	-0.078 (-0.14)	-	0.2970
EGARCH vs. EV(RS)	0.000 (0.72)	-	-	0.647 (4.81)	-	0.223 (0.40)	0.2947
Garch, EGARCH, EV(YZ)	0.000 (1.09)	0.020 (0.18)	0.729 (4.69)	-	-	-0.144 (-0.21)	0.2972
All	0.000 (1.00)	0.087 (0.47)	0.046 (0.03)	0.597 (0.43)	-0.037 (-0.03)	-0.080 (-0.07)	0.2982

1. Figures in parenthesis are t-statistic associated with the coefficient.

2. For five-day period, the traditional estimator used for forecasts is traditional close-close estimator, extreme-value estimators used for forecasts are Yang-Zhiang and Rogers-Satchell Estimators. The choice was based on R^2 as reported in Table 6.

Panel C: Calendar Month

(Number of Observations- 35)

Forecast Comparisons	β_0	β_1	β_2	β_3	β_4	β_5	R^2
Garch vs. EGARCH	0.000 (-0.61)	-	-	-	0.610 (0.31)	1.645 (0.66)	0.0586
Garch, EGARCH, Trad.	0.000 (0.37)	0.337 (1.10)	-	-	0.048 (0.02)	-0.037 (-0.01)	0.0937
GARCH vs. Trad.	0.000 (0.55)	0.335 (1.30)	-	-	0.035 (0.02)	-	0.0937
EGARCH vs. Trad.	0.000 (0.39)	0.339 (1.16)	-	-	-	-0.002 (0.00)	0.0937
Garch vs. EV (RS)	0.000 (1.10)	-	0.829 (4.05)	-	-0.861 (-0.71)	-	0.3694
EGARCH vs. EV(GK)	0.001 (2.02)	-	-	1.104 (4.22)		-3.237 (-1.76)	0.3933
Garch, EGARCH, EV(RS)	0.001 (1.76)	-	0.984 (4.25)	-	0.596 (0.37)	-3.181 (-1.37)	0.4053
All	0.001 (1.07)	-1.780 (-2.97)	-5.723 (-2.10)	8.828 (2.55)	1.216 (0.80)	-2.793 (-1.26)	0.5441

1. Figures in parenthesis are t-statistic associated with the coefficient.

2. For one-month period, the traditional estimator used for forecasts is traditional adj. close-open-close estimator, extreme-value estimators used for forecasts are Rogers-Satchell and Garman-Klass Estimators. The choice was based on R^2 as reported in Table 6.