

VOLATILITY FORECASTING WITH RANGE MODELS. AN EVALUATION OF NEW ALTERNATIVES TO THE CARR MODEL

José Luis Miralles Quirós¹

miralles@unex.es

Julio Daza Izquierdo

juliodaza@unex.es

Department of Financial Economics, University of Extremadura

Av. Elvas s/n 06071 Badajoz (Spain)

Abstract

The aim of this paper is to analyze the forecasting ability of the CARR model proposed by Chou (2005) using the S&P 500. We extend the data sample, allowing for the analysis of different stock market circumstances and propose the use of various range estimators in order to analyze their forecasting performance. Our results show that there are two range-based models that outperform the forecasting ability of the GARCH model. The Parkinson model is better for upward trends and volatilities which are higher and lower than the mean while the CARR model is better for downward trends and mean volatilities.

JEL Classification: G10, G11, G14.

Keywords: CARR; GARCH; Range Estimators; Forecasting Performance

¹ Corresponding author

Volatility Forecasting with Range models. An evaluation of new alternatives to the CARR model

1 Introduction

In the years following the publication of the ARCH model proposed by Engle (1982) and its generalization (GARCH model) proposed by Bollerslev (1986), modeling and forecasting volatility has been the subject of vast empirical and theoretical investigation. As a result, many different studies have focused on evaluating different volatility measures that might improve volatility forecasts and, if possible, identify a preferred technique.

For forecasting monthly US stock index volatility, Akgiray (1989) finds the GARCH model superior to ARCH, as well as the exponentially weighted moving average and the historical means models. Lamoureux and Lastrapes (1993) show that implied volatility tends to underpredict realized volatility while forecasts of variance from past returns contain relevant information not contained in the forecast constructed for implied volatility.

Moreover, Brailsford and Faff (1996) find GJR and GARCH models slightly superior to a number of simpler models for predicting Australian monthly stock index volatility. In contrast, Tse (1991) and Tse and Tung (1992) use data from Japan and Singapore and find that an Exponentially Weighted Moving Average model produces better volatility forecasts than do ARCH models.

Martens (2001) analyzes the improvement of forecasting on different GARCH models by including additional intraday information and finds that the higher the frequency used the better the volatility forecast.

The use of ranges is another alternative to measure the variability of a share, an index or a stock market, which makes sense because that is what traders perceive volatility to be. The application of ranges in finance started with Parkinson (1980) who showed the superiority of his proposal when compared with the standard methods of volatility

estimations. This initial study was followed by others¹ where the range properties were analyzed.

More recently, Brandt and Jones (2006) compare a range-based EGARCH model with the return-based volatility model and find that the former produces better predictions for out of sample forecasts, while Chou (2005), in a very interesting paper, uses Standard and Poors 500 index data and proposes a range-based model, the Conditional Autoregressive Range (CARR), suggesting that it outperforms the forecasting ability of the GARCH model.

Our study furthers the line of research initiated by Chou (2005) by mainly discussing the results of the forecasting power of the CARR model, and suggesting the use of range estimators in order to improve the original model of Chou (2005).

We improve the previous literature in various ways. Firstly, we propose to extend the original sample of Chou (2005) to the week which begins on September 27, 2010 in order to analyze the performance of the models employed by Chou (2005) in different situations. Secondly, we analyze the performance of both models to forecast volatility depending on the day of the week the variables are calculated. Finally, we suggest substituting the original range employed by Chou (2005) with other range estimators with the aim of finding an alternative to the CARR model.

This paper shows that when forecasts are made on an upward trend and in a low volatility environment the Parkinson range estimator provides better forecasting results while the original CARR model is better on downward trends and mean volatility.

The remainder of the paper is organized as follows. Section 2 describes the original methodology and the initial results, Section 3 presents the alternatives, Section 4 shows the main results and Section 5 provides the main conclusions.

2. INITIAL METHODOLOGY AND PRELIMINARY RESULTS

Different applications and methodologies have been developed in recent years to analyze the dynamics of volatility. Among them, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models proposed initially by Engle (1982) and Bollerslev (1986), and the stochastic volatility (SV) models advocated by Taylor (1986) are two popular and useful alternatives for estimating and modeling time-varying

¹ See Garman and Klass (1980), Beckers (1983), Rogers and Satchell (1991), Kunitomo (1992) or Yang and Zhang (2000) among others.

conditional financial volatility. However, as pointed out by Alizadeh, Brandt, and Diebold (2002), Brandt and Diebold (2006) and Chou (2005) among others, both models are inaccurate and inefficient, because they are based on the closing prices of the reference period and fail to use the information contents in between the reference points.

We initially follow the Chou approach (2005), which proposed the Conditional Autoregressive Range (CARR) model as an alternative for the modeling of financial volatility.

The CARR model of order (p,q) is shown as

$$R_t = \lambda_t \varepsilon_t$$

$$\lambda_t = \omega + \sum_{i=1}^p \alpha_i R_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j} \quad (1)$$

$$\varepsilon_t | I_{t-1} \sim f(1, \xi_t)$$

where, $R_t = \text{Max}(P_t) - \text{Min}(P_t)$, is the range measure calculated as the difference between the highest and the lowest logarithms of the prices of a speculative asset observed at time t and λ_t is the conditional mean of the range based on all information up to time t. As pointed out by Chou (2005), the distribution of the disturbance term ε_t , or the normalized range $\varepsilon_t = R_t / \lambda_t$, is assumed to be distributed with a density function $f(\cdot)$ with a unit mean. Additionally, the coefficients in the conditional mean equation are all positive to ensure positivity of λ_t .

Assuming that the distribution follows an exponential distribution with unit mean then the log likelihood function can be written as:

$$L(\alpha_i, \beta_j; R_1, R_2, \dots, R_T) = - \sum_{t=1}^T \left[\ln(\lambda_t) + \frac{R_t}{\lambda_t} \right] \quad (2)$$

Chou (2005) also shows that the unconditional (long term) mean of range can be calculated as $\omega / \left(1 - \left(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right) \right)$. Finally, for the process to be stationary, the

coefficients α and β must follow the condition: $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$.

Chou (2005) performs out of sample forecasts and makes comparisons with a GARCH (1,1) model with conditional normal distribution. He chooses the forecast horizons ranging from 1 week to 50 weeks and makes rolling sample estimations to estimate the parameters of both models. In each case, 972 weeks of data prior to the forecast interval are used and 100 out of sample forecast are made for each forecast horizon. Four measures are used as the benchmark of the ex post volatility: the sum of squared daily returns (SSDR), weekly return squared (WRSQ), weekly range (WRNG) and absolute weekly return (AWRET).

In order to evaluate the performance of both models, Chou (2005) uses two symmetric error statistics, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE):

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{\sigma}_m - \sigma_m)^2} \quad (3)$$

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |\hat{\sigma}_m - \sigma_m| \quad (4)$$

where $T=100$ and $\hat{\sigma}_m$ and σ_m denote the volatility forecast and the realized volatility in week m , respectively.

After computing those symmetric statistics over 50 forecast horizons, Chou (2005) considers that both criteria give almost unanimous support for the CARR model over GARCH. Chou (2005) also points out that a closer analysis of the results shows that the differences in the performance of the two models are more evident when the horizons are shorter and, in particular, for the SSDR and WRNG measures, because both of them use more information (daily) than WRSQ and AWRET and, therefore, contain less noise.

The results in Chou (2005) are not shown for a horizon longer than 13 weeks in order to save space², but if we analyze those results we find that the performance of the GARCH model on forecasting volatility is significantly better than the CARR model, specially for the RMSE statistics, and for most of the measures of volatility used as benchmarks, as reported in Tables 1 and 2.³

INSERT TABLE 1 ABOUT HERE

Table 1 shows the results for the RMSE statistics. It is interesting to observe that in the last 30 horizons most of the smaller stats are for the GARCH model, which means that its forecasting ability is better than the CARR model proposed by Chou (2005). The difference between these models is specially significant when the Weekly Range (WRNG) is used to measure volatility because 24 out of the 30 cases are smaller for the GARCH model.

This fact is in conflict with the original hypothesis of Chou (2005) who considers that the CARR model should be good at forecasting the Weekly Range (WRNG) because it is the variable used in the variance equation of the CARR model.

On the other hand, from the results of the Mean Absolute Error statistics reported in Table 2, we see that the CARR model performs better forecasts than the GARCH model for all measures of volatility. There is just one exception because, once again, the WRNG provides better results for the GARCH model (18 out of 30 stats are smaller for the GARCH model).

INSERT TABLE 2 ABOUT HERE

3. COMMENTS AND PRICE RANGE ESTIMATORS

From the previous results, and after analyzing in depth the methodology, we find some weaknesses in the model proposed by Chou (2005). The first one is related to the way in which the data is collected. Chou (2005) collects daily data from the Standard

² Chou (2005) points out that the results are available in a previous working paper.

³ Following the same reason of saving space only the results for a 13 weeks horizon and longer are reported. The remainders are partially in Chou (2005) or are available upon request.

and Poors 500 for the period from April 26, 1982 to October 17, 2003, which was downloaded from the website “Yahoo.com”. Daily and weekly data (obtained from the daily one) were considered, but only weekly estimations were shown because the results were basically the same and some weekday seasonal effects were found for the daily range data.

All the dates that are referred to in Chou (2005) are Mondays, which suggests us to think that all the weekly estimations are from Monday to Monday, but the fact is that the weekly format that can be downloaded from the aforementioned website is from Monday to Friday, which led us to check that the returns (close to close natural logarithm difference) used to calculate the GARCH models and, consequently, the GARCH forecasts are not from Monday to Monday but from Friday to Friday.⁴

We agree with the fact of using weekly returns in the analysis because they should not be subject to potential bias such as the bid–ask effect, non-trading days, etc, that might arise when daily returns are used. However, we must consider that the existence of calendar anomalies such as the Monday effect, the Friday effect or the day of the week effect could lead to an irregular behavior of the proposed models according to the day of the week on which the variables are calculated.

Secondly, the sample used by Chou (2005) was from April 26, 1982 to October 13, 2003, the forecast period⁵ being characterized by a mix of downward and upward trend in the Standard and Poors 500. Since the publication of Chou’s paper there have been different trends in the S&P500, with a higher maximum and a lower minimum than those that were considered by Chou (2005). In our opinion, it would be interesting to analyze the forecasting ability of the GARCH and the CARR models in different trends in order to check the performance of each one in special situations.

Thirdly, the classical estimator of volatility is based on the close to close prices but it has been demonstrated that the daily squared return is an unbiased estimator of the realized daily volatility, however, Andersen and Bollerslev (1998) show that it is also extremely noisy. Furthermore, it must be pointed out that by only looking at opening and closing prices we may wrongly conclude that volatility on a given day is small if both prices are similar, despite large intraday price fluctuations. For those reasons, more

⁴ It was also checked with the Chou (2005) data.

⁵ The first date is December 4, 2000.

sophisticated estimators using additional information such as high, low and open prices are needed to estimate volatility.

Taylor and Xu (1997) use the standard deviation of the intraday returns, while Martens (2001) uses the sum of squared intraday returns, in both cases provide better results for the conditional variance. However, we consider the extreme value methods to be more effective. In order to explain them we adopt the notation of Garman and Klass (1980) and Yang and Zhang (2000).

C_t = closing price on day t;

O_t = opening price on day t;

H_t = high price on day t;

L_t = low price on day t;

$c_t = \ln C_t - \ln O_t$, the normalized close price;

$o_t = \ln O_t - \ln C_{t-1}$, the normalized open price;

$u_t = \ln H_t - \ln O_t$, the normalized high price;

$d_t = \ln L_t - \ln O_t$, the normalized low price;

n = number of daily periods (five in our case).

Parkinson (1980) provides a simple way to measure daily volatility given the daily range of the high/low prices by suggesting the measurement of the daily volatility as follows:

$$\hat{\sigma}_{\text{PARK}}^2 = \frac{1}{n4\ln 2} \sum_{i=1}^n (u_i - d_i)^2 \quad (5)$$

It has been demonstrated that the efficiency of this estimator is very high, about 4.91 in comparison with the standard simple variance estimator and could be as much as 8.5 times more efficient than log-squared returns.

Since then, different methods have been proposed for estimating the volatility parameter. Garman and Klass (1980) incorporate the opening and closing prices and suggest the following measure (V_{GK}):

$$V_{GK} = V'_O - 0.383V'_C + 1.364V_{PARK} + 0.019V_{RS} \quad (6)$$

where:

$$V'_O = \frac{1}{n} \sum_{i=1}^n o_i^2 \quad \text{and} \quad V'_C = \frac{1}{n} \sum_{i=1}^n c_i^2$$

and V_{RS} is another alternative measure of volatility proposed by Rogers and Satchell (1991) which is calculated as:

$$V_{RS} = \frac{1}{n} \sum_{i=1}^n [u_i(u_i - c_i) + d_i(d_i - c_i)] \quad (7)$$

Finally, Yang and Zhang (2000) propose a new estimator which is, in their opinion, the minimum-variance unbiased variance estimator and is independent of both the drift and opening jumps of the underlying price movements. This estimator, V_{YZ} , is given by the equation:

$$V_{YZ} = V_O + kV_C + (1 - k)V_{RS} \quad (8)$$

where:

$$V_O = \frac{1}{n-1} \sum_{i=1}^n (o_i - \bar{o})^2 \quad \text{and} \quad V_C = \frac{1}{n-1} \sum_{i=1}^n (c_i - \bar{c})^2$$

$$k = \frac{0.34}{1.34 + \frac{n+1}{n-1}}$$

Considering all these comments, we propose to extend the sample to the week which begins on September 27, 2010 in order to analyze the performance of the original models (GARCH and CARR) employed by Chou (2005) in different situations along this time (upward and downward trends) but always keeping the number of observations

(1120) used by Chou (2005); we also analyze the performance of both models on forecasting volatility depending on the day of the week the variables are calculated and, additionally, we suggest to substitute the original range employed by Chou (2005) with the volatility measures previously mentioned (Parkinson, Garman-Klass, Rogers-Satchell and Yang-Zhang) with the aim of finding an alternative to the CARR model.

4. DATA AND RESULTS

The data consists of daily data from the Standard and Poors 500 for the sample period from April 26, 1982 to September 27, 2010. As in Chou (2005), open, high, low and close prices are collected. Consistent with reviewed literature, the previous day of trading data was taken to calculate the different estimators in those cases when a holiday occurred. Weekly series were constructed for each day of the week so for example in the case of Monday, data from Tuesday to the following Monday (including it) was collected.

Having observed that there are different trends in the sample, we decide to divide the full sample in four sub-samples with the aim of analyzing the forecasting ability of the different range estimators in various periods.

Following those reasons, the first sub-sample (April 26, 1982-October 13, 2003) is the sample used by Chou (2005). The second sub-sample (April 14, 1986-October 9, 2007) is the period which ends with the historical maximum quote of the Standard and Poors 500. The phase which ends with the minimum after the technological bubble crash is the period analyzed in the third sub-sample (September 7, 1987-March 9, 2009). Finally, the period from March 13, 1989 to September 27, 2010 is the fourth sub-sample.

The results for the first subsample are shown in Table 3. They denote a better forecasting performance for the CARR model proposed by Chou (2005), specially when the Mean Absolute Error (MAE) criteria is considered. We observe that in most of the cases, and the CARR model provides the best forecasting results often for more than a 90% of the horizons estimated.

INSERT TABLE 3 ABOUT HERE

However, we find that for the Root Mean Squared Errors (RMSE) the results of forecasting ability are not so favorable for the CARR model. Considering the RMSE we show that the GARCH model provides better forecasting results in the central part of the week, 3 out of 4 cases on Wednesday and 4 out of 4 on Thursday. The results for Friday samples show that both models performance is very similar and with just one exception, when the Weekly Range (WRNG) is used as the measure of ex post volatility, both models provide the best forecasting ability in 25 of the 50 cases (50% each).

In this case the proposed models based on different range estimators are not significant. There is just one case, for the MAE estimator and the Weekly Return squared (WRSQ), where the Parkinson model fits volatility better than the others (on Friday for 6 times), and the Garman-Klass and Yang-Zhang models for Monday (twice and once respectively).

We find some interesting results for the second subsample which are reported in Table 4. Depending on the choice of the error estimator either the Parkinson or the GARCH model are better at forecasting volatility.

INSERT TABLE 4 ABOUT HERE

The Parkinson model is best for forecasting volatility, specially for the WRSQ and AWRET volatility measures, when the Mean Absolute Error estimator is considered. In those cases the Parkinson model performs better than the others for 46 out of 50 forecast horizons for Friday (WRSQ measure) and for 41 out of 50 forecast horizons on Monday, (AWRET measure), having also a high percentage of better results on the rest of the days and for the other two measures of volatility (SSDR and WRNG).

On the other hand, the GARCH model provides the best forecasting performance when the RMSE is considered. However, in three cases the model with the Parkinson volatility estimator is better (on Friday when SSDR is used and on Monday for WRNG and AWRET measures).

In this sample, the CARR model proposed by Chou provides the better forecasting performance in just 2 out of the 40 possibilities, being equal to the GARCH model in

one case (for the Thursday estimations when the Weekly Return Squared, WRSQ, is used as the measured volatility).

The rest of the models proposed fit better results than in the first sample but their relevance is minor when compared with the Parkinson or the GARCH model. The most interesting results are provided by the Rogers-Satchell model when the MAE estimator and SDDR measure are considered (13 out of 50 forecast horizons on Monday are better suited to this model) and for Monday and Wednesday when WRSQ is used (6 out of 50 forecast horizons). Furthermore, the Yang-Zhang model is better on Thursday when the MAE estimator and WRSQ measure are considered (12 out of 50 forecast horizons are better suited to that model).

The results for the third sub-sample (September 7, 1987-March 9, 2009), which are reported in Table 5, show smaller differences among the proposed models than in the previous samples. As well as in the previous sample, the best forecasting results for the Parkinson model are obtained when the Mean Absolute Error (MAE) estimator is calculated and, in this case, for the WRSQ and AWRET measures, specially on Tuesday, Wednesday and Thursday.

INSERT TABLE 5 ABOUT HERE

With respect to the Root Mean Squared Error (RMSE) estimator, the GARCH and the CARR models provide similar forecasting ability, when SDDR and WRSQ measures are considered. The GARCH model is better on Wednesday and Friday for SDDR and Tuesday and Wednesday for WRSQ but the CARR model is better on Tuesday and Thursday for SDDR and Monday and Thursday for WRSQ. They are equal on Monday for SDDR and Friday for WRSQ.

For the rest of the volatility measures used as benchmarks (WRNG and AWRET) the CARR model performance is better than the GARCH model one. However, the improvement is insignificant because in most of the cases the GARCH model provides a high number of instances in which performs better than the CARR model.

Once again the Rogers-Satchell and the Yang-Zhang models are the only alternative to the Parkinson, GARCH and CARR models. In this case, it is significant to point out the fact that the best forecasting ability of the Rogers-Satchell is produced on

Wednesday when the MAE estimator and the AWRET measure are considered (13 out of 50 forecast horizons are better fitted for that model). On the other hand, Table 5 shows that the best results for the Yang-Zhang model are on Thursday when RMSE estimator and WRNG are considered (in that case 20 out of the 50 forecast horizons are better fitted for that model).

Finally, the results for the last sample (March 13, 1989 to September 27, 2010) are shown in Table 6. In this case there is no debate about which is the best model for forecasting volatility because the CARR model is clearly better than the other ones for both the error estimators as well as the volatility measures. According with Chou (2005), the differences in the performance of the models are more obvious when SDR and WRNG are used to measure volatility.

INSERT TABLE 6 ABOUT HERE

The main explanation of the results we obtained is that the choice of the best forecasting model depends on three factors: the trend, the level of volatility in the analysis period and the error estimator that is used to analyze the forecasting ability of each model. On the other hand, the day of the week on which the estimations are made is insignificant.

Figure 1 presents the plot of the weekly closes on Friday⁶ of the Standard and Poors 500 for the whole sample. The A_F , A_L and A denote the first one-step forecast, the last one-step forecast and the end of the sample respectively (by substituting B, C and D for A we would get the same variables for the second, third and fourth samples).

INSERT FIGURE 1 ABOUT HERE

The CARR model proposed by Chou (2005) shows better results for forecasting volatility in the first and fourth samples, denoted by A and D on Figure 1, where the forecasting period is a mix of downtrend⁷ (most of the time) and uptrend. In both cases,

⁶ We choose Friday in order to maintain the methodology of Chou (2005).

⁷ The downtrend is most significant in the fourth sample because it contains the technological bubble crash.

the mean of the different volatility measures used as benchmarks on each weekday is approximately the mean of the four samples as reported in Table 7.

INSERT TABLE 7 ABOUT HERE

However, the CARR model is no longer adequate for forecasting volatility in other cases. That is the point of the second sample (denoted by B in Figure 1), where the whole forecasting period is entirely included in an upward trend and the volatility values are significantly lower than the mean as show in Table 7. In that case, the proposed Parkinson model forecast volatility better than the rest of the models when the MAE estimator and the four volatility measures are considered. However, for the same forecasting period the GARCH model is better at forecasting when the RMSE estimator is used.

On the other hand, the Parkinson and the GARCH models perform better than the CARR model for both error estimators when the WRSQ and AWRET measures are used and when the volatility is higher than the mean and the trend is mixed, being in an upward trend most of the time, which is the case of the third sample (denoted by C on Figure 1). The GARCH model also shows good results when the SDDR measure is used while the CARR model reports the best forecasting ability only when the WRNG is used (precisely the variable used in the variance equation of the CARR model).

The best performance of the Parkinson estimator and, therefore, the worst of the Garman-Klass, Rogers-Satchell and Yang-Zhang models is in accordance with the results of Brandt and Kinlay (2005) who show that the latter ones were downward biased. They demonstrate that the Parkinson estimator outperforms all of the other estimators in terms of bias and that the Garman-Klass, Rogers-Satchell and Yang-Zhang estimators show signs of negative bias.

In spite of the theoretical heterogeneity of the results, they agree with the findings of Poon and Granger (2003) who, in a review about forecasting volatility in financial markets, provide some useful insights into comparing different studies about this topic. The authors say that the conclusions of these studies depend strongly on the error statistics used, the sampling schemes employed (e.g. rolling fixed sample estimation or recursive expanding sample estimation), as well as the period and assets studied.

To sum up, we agree with the fact that the CARR model proposed by Chou (2005) is a good model, but it is also very important to point out that the Parkinson model is clearly an alternative for the CARR and the GARCH models, specially when the volatility is low and we run the forecast analysis on an upward trend.

5. CONCLUSIONS

In this paper we analyze the forecasting ability of the Conditional Autorregressive Range (CARR) model, proposed by Chou (2005), by extending the sample of analysis of the Standard and Poors 500 till the last week of September 2010 and allowing for the analysis of different stock market circumstances (like upward or downward trends). Additionally, we analyze the volatility forecasting ability for all the weekdays and we propose to use various range estimators in order to analyze their forecasting performance.

The results show that the original CARR model can be improved depending on three factors: the trend, the level of volatility in the analysis period and the error estimator that is used to analyze the forecasting ability of each model.

For that reason, in those samples where the whole forecasting period is entirely included in an upward trend and the volatility values are significantly lower than the mean, the use of the Parkinson estimator instead of the range used by Chou (2005) leads to an improvement in the forecasting ability of the model.

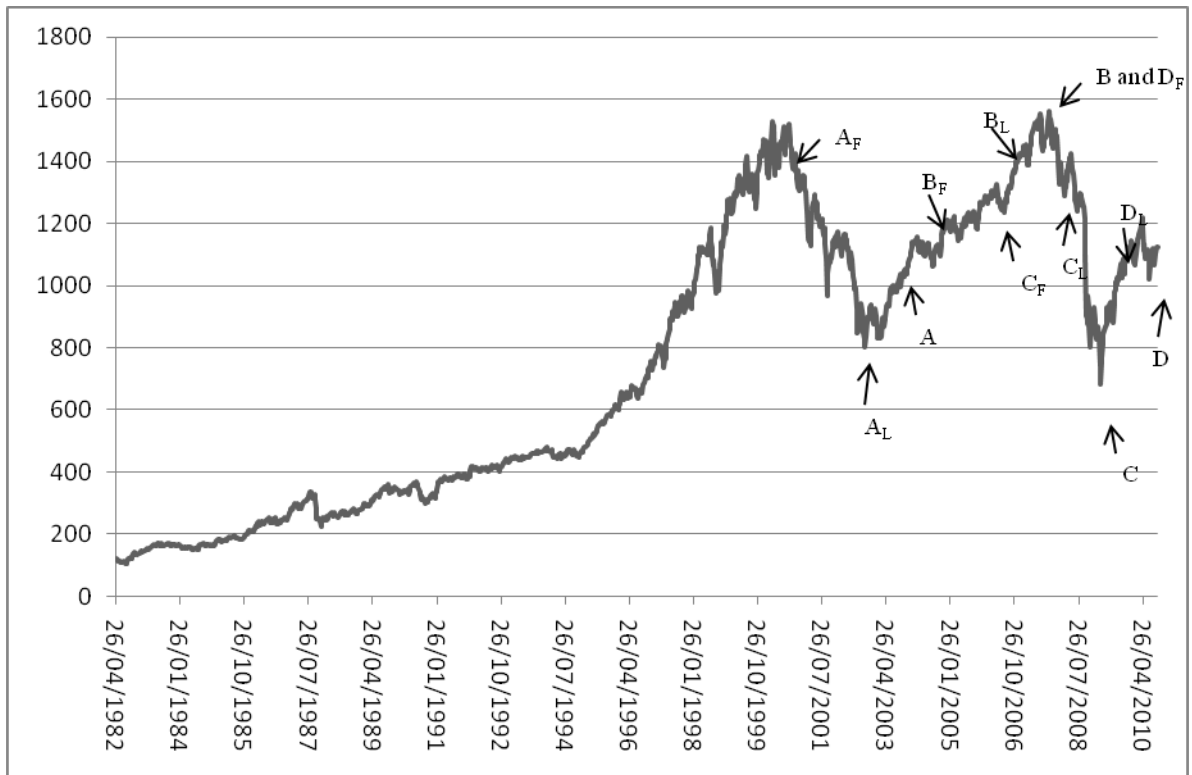
Finally, we must point out that these results are of greatest relevance when applied to option pricing or market risk management where an under or over- estimation of risk can be disastrous.

REFERENCES

- Akgiray V. (1989). "Conditional heteroskedasticity in time series of stock returns: Evidence and Forecasts." *Journal of Business* 62, 55-80.
- Alizadeh, S., M. Brandt and F. Diebold (2002). "Range-based estimation of stochastic volatility models." *Journal of Finance* 57, 1047-1091.
- Andersen T.G. and T. Bollerslev (1998). "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts." *International Economic Review* 39, 885-905.
- Beckers, S. (1983). "Variances of security price returns based on highest, lowest, and closing prices." *Journal of Business* 56, 97-112.
- Bollerslev, T. (1986). "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* 31, 307-327.
- Brailsford, T.J. and R.W. Faff. (1996). "An evaluation of volatility forecasting techniques." *Journal of Banking and Finance* 20, 419-438.
- Brandt, M. and F. Diebold (2006). "A no-arbitrage approach to range-based estimation of return covariances and correlations." *Journal of Business*. 79, 61-74.
- Brandt, M. and C. Jones. (2002). "Volatility Forecasting with Range-based EGARCH Models." Manuscript, Wharton School, University of Pennsylvania.
- Brandt, M. and J. Kinlay (2005). "Estimating historical volatility." *Investment Analytics*.
- Chou, R. (2005). "Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model." *Journal of Money, Credit and Banking* 37, 561-582.
- Engle, R.F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation." *Econometrica* 50, 987-1008.
- Garman, M. and M. Klass. (1980). "On the estimation of security price volatilities from historical data." *Journal of Business* 53, 67-78.
- Kunitomo, N. (1992). "Improving the Parkinson method of estimating security price volatilities." *Journal of Business* 65, 295-302.

- Lamoureux, C.G. and W.D. Lastrapes. (1993). "Forecasting stock return variances: Toward an understanding of stochastic implied volatilities." *Review of Financial Studies* 6, 293-326.
- Martens, M. (2001). "Forecasting daily exchange rate volatility using intraday returns." *Journal of International Money and Finance* 20, 1-23.
- Parkinson, M. (1980). "The extreme value method for estimating the variance of the rate of return." *Journal of Business* 53, 61-65.
- Poon, S. and C. Granger. (2003). "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41, 478-539.
- Rogers, C. and S. Satchell. (1991). "Estimating variance from high, low and closing prices." *Annals of Applied Probability* 1, 504-512.
- Taylor, S. J. (1986). "Modelling Financial Time Series." New York: John Wiley & Sons.
- Taylor, S. and X. Xu (1997). "The incremental volatility information in one million Foreign exchange quotations." *Journal of Empirical Finance* 4, 317– 340.
- Tse, Y.K. (1991). "Stock Returns Volatility in the Tokyo Stock Exchange." *Japan and the World Economy* 3, 285-298.
- Tse, S.H. and K.S. Tung. (1992). "Forecasting Volatility in the Singapore Stock Market." *Asia Pacific Journal of Management* 9, 1-13.
- Yang D. and Q. Zhang, (2000), "Drift-independent volatility estimation based on high, low, open, and closing prices." *Journal of Business* 73, 477-491.

Figure 1: Weekly closes of the Standard and Poors 500



Sample April 26, 1982 to September 27, 2010

The A_F , A_L and A denote the first one-step forecast, the last one-step forecast and the end of the sample respectively (by substituting B, C and D for A we would get the same variables for the second, third and fourth samples).

Table 1: Root Mean Squared Errors (RMSE) estimations

Horizon	SSDR		WRSQ		WRNG		AWRET	
	GARCH	CARR	GARCH	CARR	GARCH	CARR	GARCH	CARR
13	12.674	11.604	19.760	19.603	2.595	2.482	2.131	2.104
14	12.916	11.776	19.806	19.591	2.616	2.493	2.136	2.099
15	12.708	11.603	19.407	19.152	2.586	2.487	2.089	2.047
16	12.537	11.609	19.346	19.130	2.521	2.433	2.073	2.042
17	12.420	11.637	19.241	19.131	2.500	2.444	2.064	2.048
18	12.006	11.274	19.160	19.115	2.443	2.414	2.042	2.042
19	11.939	11.329	19.100	19.008	2.441	2.419	2.028	2.018
20	12.022	11.294	19.420	19.394	2.446	2.428	2.050	2.048
21	11.920	11.197	19.313	19.307	2.444	2.405	2.036	2.036
22	11.867	11.087	19.171	19.184	2.440	2.385	2.014	2.008
23	12.041	11.146	19.256	19.173	2.445	2.392	2.024	2.000
24	12.085	11.158	19.388	19.219	2.443	2.398	2.042	2.011
25	12.109	11.228	19.485	19.336	2.456	2.420	2.068	2.035
26	12.041	11.211	19.380	19.262	2.427	2.402	2.053	2.028
27	11.805	11.221	19.322	19.347	2.393	2.416	2.054	2.042
28	11.751	11.272	19.453	19.418	2.402	2.431	2.059	2.046
29	11.614	11.243	19.439	19.424	2.368	2.426	2.058	2.051
30	11.388	11.178	19.313	19.387	2.320	2.407	2.032	2.037
31	11.241	11.175	19.315	19.427	2.295	2.418	2.032	2.041
32	11.041	11.143	19.293	19.446	2.266	2.416	2.037	2.051
33	10.880	11.109	19.237	19.445	2.234	2.412	2.023	2.046
34	10.969	11.162	19.168	19.412	2.206	2.410	2.012	2.037
35	10.831	11.138	19.114	19.412	2.154	2.409	2.010	2.041
36	10.581	11.110	19.133	19.385	2.136	2.403	2.011	2.036
37	10.352	11.109	19.215	19.398	2.099	2.401	2.039	2.043
38	10.129	11.106	19.201	19.410	2.079	2.413	2.056	2.054
39	9.745	11.074	19.001	19.345	2.062	2.402	2.024	2.035
40	9.431	11.068	18.809	19.318	2.022	2.390	2.003	2.026
41	9.110	11.036	18.689	19.321	2.015	2.394	1.973	2.023
42	8.380	10.459	11.677	12.442	1.726	2.127	1.682	1.727
43	8.201	10.336	11.179	11.348	1.762	2.093	1.639	1.639
44	8.469	10.369	11.677	11.379	1.817	2.092	1.669	1.636
45	9.231	10.425	11.594	11.386	1.882	2.107	1.661	1.641
46	9.742	10.473	11.689	11.417	1.917	2.109	1.674	1.646
47	10.111	10.507	11.877	11.473	2.001	2.122	1.706	1.658
48	10.210	10.497	11.827	11.483	2.027	2.125	1.705	1.664
49	10.245	10.515	11.748	11.477	2.010	2.127	1.692	1.660
50	10.187	10.531	11.716	11.483	2.033	2.135	1.704	1.667

This table computes the Root Mean Squared Error (RMSE) where T=100. In all cases, the smaller the error, the better the forecasting ability. SDDR, WRSQ, WRNG and AWRET are the sum of squared daily returns, weekly return squared, weekly range and absolute weekly return, respectively.

Table 2: Mean Absolute Errors (MAE) estimations

Horizon	SSDR		WRSQ		WRNG		AWRET	
	GARCH	CARR	GARCH	CARR	GARCH	CARR	GARCH	CARR
13	8.853	7.441	10.061	8.915	1.919	1.752	1.496	1.420
14	9.046	7.438	10.046	8.919	1.938	1.752	1.490	1.412
15	8.907	7.242	9.601	8.431	1.910	1.724	1.438	1.362
16	8.776	7.161	9.526	8.424	1.828	1.668	1.421	1.357
17	8.666	7.062	9.401	8.475	1.839	1.685	1.442	1.375
18	8.422	6.764	9.471	8.367	1.802	1.642	1.425	1.366
19	8.428	6.810	9.350	8.219	1.789	1.651	1.411	1.340
20	8.423	6.921	9.575	8.364	1.806	1.679	1.431	1.354
21	8.340	6.895	9.405	8.273	1.798	1.680	1.416	1.351
22	8.154	6.733	9.301	8.094	1.796	1.676	1.402	1.329
23	8.268	6.778	9.238	8.022	1.777	1.677	1.418	1.313
24	8.371	6.728	9.155	8.021	1.721	1.652	1.419	1.322
25	8.292	6.757	9.410	8.104	1.727	1.652	1.462	1.344
26	8.316	6.800	9.305	8.108	1.748	1.635	1.449	1.344
27	8.197	6.787	9.205	8.139	1.718	1.641	1.454	1.353
28	8.138	6.898	9.238	8.087	1.719	1.636	1.457	1.350
29	8.071	6.898	9.226	8.124	1.683	1.631	1.440	1.359
30	7.840	6.845	9.097	8.111	1.632	1.611	1.420	1.349
31	7.770	6.805	9.112	8.075	1.603	1.621	1.426	1.343
32	7.627	6.726	9.168	8.122	1.577	1.625	1.434	1.355
33	7.323	6.596	9.142	8.087	1.558	1.608	1.422	1.342
34	7.528	6.671	9.103	8.008	1.537	1.622	1.397	1.323
35	7.421	6.665	9.216	8.058	1.524	1.636	1.403	1.329
36	7.278	6.624	9.252	8.058	1.541	1.641	1.401	1.334
37	7.239	6.665	9.394	8.078	1.515	1.644	1.420	1.339
38	7.136	6.715	9.528	8.128	1.489	1.662	1.453	1.352
39	6.945	6.722	9.267	7.978	1.497	1.663	1.417	1.327
40	6.747	6.697	9.068	7.836	1.482	1.642	1.399	1.314
41	6.426	6.671	8.962	7.814	1.449	1.636	1.381	1.305
42	6.190	6.374	7.545	6.338	1.362	1.546	1.289	1.208
43	6.137	6.266	7.409	5.918	1.362	1.518	1.276	1.165
44	6.405	6.282	7.455	5.870	1.384	1.495	1.279	1.156
45	6.681	6.252	7.334	5.863	1.414	1.495	1.272	1.162
46	6.882	6.284	7.397	5.868	1.444	1.496	1.285	1.162
47	7.084	6.292	7.595	5.959	1.489	1.503	1.306	1.176
48	7.193	6.206	7.495	5.995	1.530	1.503	1.308	1.187
49	7.228	6.226	7.437	5.957	1.494	1.500	1.290	1.178
50	7.175	6.241	7.368	5.972	1.525	1.506	1.298	1.194

This table computes the Mean Absolute Error (MAE) where $T=100$. In all cases, the smaller the error, the better the forecasting ability. SSSR, WRSQ, WRNG and AWRET are the sum of squared daily returns, weekly return squared, weekly range and absolute weekly return, respectively.

**Table 3: Out of sample forecast comparison.
First Sample: April 26, 1982 to October 13, 2003**

RMSE							MAE						
SSDR							SSDR						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	0	50	0	0	0	0	MO	4	46	0	0	0	0
TU	0	50	0	0	0	0	TU	3	47	0	0	0	0
WD	26	24	0	0	0	0	WD	2	48	0	0	0	0
TH	34	16	0	0	0	0	TH	3	47	0	0	0	0
FR	20	30	0	0	0	0	FR	2	48	0	0	0	0
WRSQ							WRSQ						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	0	50	0	0	0	0	MO	28	19	0	0	2	1
TU	0	50	0	0	0	0	TU	0	50	0	0	0	0
WD	31	19	0	0	0	0	WD	3	47	0	0	0	0
TH	35	15	0	0	0	0	TH	0	50	0	0	0	0
FR	19	31	0	0	0	0	FR	0	44	6	0	0	0
WRNG							WRNG						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	16	34	0	0	0	0	MO	1	49	0	0	0	0
TU	0	50	0	0	0	0	TU	0	50	0	0	0	0
WD	28	22	0	0	0	0	WD	0	50	0	0	0	0
TH	36	14	0	0	0	0	TH	31	19	0	0	0	0
FR	25	25	0	0	0	0	FR	19	31	0	0	0	0
AWRET							AWRET						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	25	25	0	0	0	0	MO	13	37	0	0	0	0
TU	0	50	0	0	0	0	TU	0	50	0	0	0	0
WD	19	31	0	0	0	0	WD	0	50	0	0	0	0
TH	38	12	0	0	0	0	TH	11	39	0	0	0	0
FR	18	32	0	0	0	0	FR	1	49	0	0	0	0

This table reports a summary of the results for the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) estimations for all the horizons (50), all the measures of volatility (SSDR, WRSQ, WRNG and AWRET) and the six models considered. The row next to each day shows the number of times in which forecasting ability of each model is better than the others. MO, TU, WD, TH and FR are the reference of Monday, Tuesday, Wednesday, Thursday and Friday respectively.

Table 4: Out of sample forecast comparison.
Second Sample: April 14, 1986 to October 9, 2007

RMSE							MAE						
SSDR							SSDR						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	30	16	2	0	2	0	MO	22	0	15	0	13	0
TU	39	11	0	0	0	0	TU	47	3	0	0	0	0
WD	43	7	0	0	0	0	WD	22	0	26	0	2	0
TH	26	12	12	0	0	0	TH	10	5	24	3	0	8
FR	15	6	29	0	0	0	FR	2	6	42	0	0	0

WRSQ							WRSQ						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	41	0	0	0	3	6	MO	13	0	27	0	6	4
TU	46	4	0	0	0	0	TU	7	0	43	0	0	0
WD	42	7	1	0	0	0	WD	15	0	29	0	6	0
TH	25	25	0	0	0	0	TH	0	0	38	0	0	12
FR	19	31	0	0	0	0	FR	0	4	46	0	0	0

WRNG							WRNG						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	14	11	25	0	0	0	MO	14	0	36	0	0	0
TU	48	1	1	0	0	0	TU	41	0	9	0	0	0
WD	29	7	14	0	0	0	WD	19	0	31	0	0	0
TH	38	10	2	0	0	0	TH	17	4	29	0	0	0
FR	45	5	0	0	0	0	FR	22	6	22	0	0	0

AWRET							AWRET						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	14	8	28	0	0	0	MO	9	0	41	0	0	0
TU	43	0	7	0	0	0	TU	23	0	27	0	0	0
WD	30	8	12	0	0	0	WD	20	0	25	0	5	0
TH	28	10	12	0	0	0	TH	10	3	37	0	0	0
FR	40	10	0	0	0	0	FR	46	3	1	0	0	0

This table reports a summary of the results for the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) estimations for all the horizons (50), all the measures of volatility (SSDR, WRSQ, WRNG and AWRET) and the six models considered. The row next to each day shows the number of times in which forecasting ability of each model is better than the others. MO, TU, WD, TH and FR are the reference of Monday, Tuesday, Wednesday, Thursday and Friday respectively.

**Table 5: Out of sample forecast comparison.
Third Sample September 7, 1987 to March 9, 2009**

RMSE							MAE						
SSDR							SSDR						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	25	25	0	0	0	0	MO	31	19	0	0	0	0
TU	19	31	0	0	0	0	TU	27	23	0	0	0	0
WD	34	14	0	2	0	0	WD	38	12	0	0	0	0
TH	17	31	0	0	0	2	TH	18	32	0	0	0	0
FR	28	22	0	0	0	0	FR	46	4	0	0	0	0
WRSQ							WRSQ						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	21	28	0	0	0	1	MO	32	6	9	0	3	0
TU	27	20	0	2	0	1	TU	13	0	25	1	11	0
WD	27	18	0	2	0	3	WD	10	0	24	4	9	3
TH	15	24	0	0	0	11	TH	11	0	24	2	3	10
FR	24	24	0	1	0	1	FR	46	0	2	0	2	0
WRNG							WRNG						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	14	28	0	0	0	8	MO	10	40	0	0	0	0
TU	15	31	0	0	4	0	TU	10	40	0	0	0	0
WD	22	21	0	0	5	2	WD	22	27	0	0	1	0
TH	0	30	0	0	0	20	TH	10	38	0	0	0	2
FR	20	28	0	0	0	2	FR	16	34	0	0	0	0
AWRET							AWRET						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	11	30	0	0	0	9	MO	4	46	0	0	0	0
TU	21	23	0	0	3	3	TU	27	5	16	0	2	0
WD	26	17	0	0	5	2	WD	19	1	17	0	13	0
TH	11	19	0	0	2	18	TH	15	6	16	0	9	4
FR	18	28	0	1	0	3	FR	29	21	0	0	0	0

This table reports a summary of the results for the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) estimations for all the horizons (50), all the measures of volatility (SSDR, WRSQ, WRNG and AWRET) and the six models considered. The row next to each day shows the number of times in which forecasting ability of each model is better than the others. MO, TU, WD, TH and FR are the reference of Monday, Tuesday, Wednesday, Thursday and Friday respectively.

**Table 6: Out of sample forecast comparison.
Fourth Sample March 13, 1989 to September 27, 2010**

RMSE							MAE						
SSDR							SSDR						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	0	50	0	0	0	0	MO	0	50	0	0	0	0
TU	0	50	0	0	0	0	TU	0	50	0	0	0	0
WD	0	50	0	0	0	0	WD	0	50	0	0	0	0
TH	0	50	0	0	0	0	TH	0	50	0	0	0	0
FR	0	50	0	0	0	0	FR	0	50	0	0	0	0
WRSQ							WRSQ						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	6	44	0	0	0	0	MO	0	50	0	0	0	0
TU	9	41	0	0	0	0	TU	6	44	0	0	0	0
WD	4	46	0	0	0	0	WD	6	44	0	0	0	0
TH	0	48	1	0	0	1	TH	0	49	0	0	0	1
FR	0	50	0	0	0	0	FR	0	50	0	0	0	0
WRNG							WRNG						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	0	50	0	0	0	0	MO	0	50	0	0	0	0
TU	0	50	0	0	0	0	TU	0	50	0	0	0	0
WD	0	50	0	0	0	0	WD	0	50	0	0	0	0
TH	0	50	0	0	0	0	TH	0	50	0	0	0	0
FR	0	50	0	0	0	0	FR	0	50	0	0	0	0
AWRET							AWRET						
	GARCH	CARR	PARK	GK	RS	YZ		GARCH	CARR	PARK	GK	RS	YZ
MO	2	48	0	0	0	0	MO	1	49	0	0	0	0
TU	1	49	0	0	0	0	TU	5	45	0	0	0	0
WD	0	50	0	0	0	0	WD	7	43	0	0	0	0
TH	0	50	0	0	0	0	TH	0	50	0	0	0	0
FR	0	50	0	0	0	0	FR	0	50	0	0	0	0

This table reports a summary of the results for the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) estimations for all the horizons (50), all the measures of volatility (SSDR, WRSQ, WRNG and AWRET) and the six models considered. The row next to each day shows the number of times in which forecasting ability of each model is better than the others. MO, TU, WD, TH and FR are the reference of Monday, Tuesday, Wednesday, Thursday and Friday respectively.

Table 7: Main statistics of the volatility measures

		SSDR					WRSQ					
		MO	TU	WD	TH	FR	MO	TU	WD	TH	FR	
1st	Mean	5,691	5,692	5,691	5,691	5,691	Mean	6,381	5,980	5,460	4,991	4,957
	Std. Dev.	18,599	19,195	18,966	19,362	18,637	Std. Dev.	32,599	26,366	21,292	13,324	11,098
2nd	Mean	5,468	5,469	5,466	5,466	5,472	Mean	5,863	5,575	5,062	4,552	4,573
	Std. Dev.	18,565	19,177	18,955	19,351	18,619	Std. Dev.	32,226	26,147	21,131	13,022	10,685
3rd	Mean	7,008	7,041	7,041	7,055	7,053	Mean	6,863	6,114	5,681	5,858	5,710
	Std. Dev.	21,941	22,578	22,614	22,679	22,145	Std. Dev.	33,723	27,610	23,157	22,759	17,477
4th	Mean	6,590	6,590	6,590	6,589	6,585	Mean	6,146	5,502	5,319	5,649	5,617
	Std. Dev.	13,790	13,969	14,344	13,844	13,904	Std. Dev.	14,972	13,923	13,406	20,558	16,642
Mean	Mean	6,189	6,198	6,197	6,200	6,200	Mean	6,313	5,793	5,381	5,263	5,214
	Std. Dev.	19,402	19,883	19,908	19,965	19,511	Std. Dev.	30,048	24,469	20,660	20,264	15,814

		WRNG					AWRET					
		MO	TU	WD	TH	FR	MO	TU	WD	TH	FR	
1st	Mean	3,236	3,240	3,219	3,173	3,196	Mean	1,750	1,754	1,697	1,661	1,675
	Std. Dev.	2,064	2,126	2,021	1,896	1,852	Std. Dev.	1,822	1,704	1,607	1,495	1,467
2nd	Mean	3,077	3,073	3,047	3,006	3,033	Mean	1,666	1,675	1,604	1,561	1,589
	Std. Dev.	2,031	2,100	2,005	1,866	1,824	Std. Dev.	1,758	1,665	1,579	1,455	1,432
3rd	Mean	3,293	3,257	3,233	3,213	3,240	Mean	1,775	1,716	1,654	1,653	1,696
	Std. Dev.	2,415	2,397	2,345	2,345	2,269	Std. Dev.	1,928	1,781	1,717	1,769	1,685
4th	Mean	3,298	3,256	3,240	3,230	3,250	Mean	1,780	1,708	1,667	1,653	1,703
	Std. Dev.	2,227	2,131	2,143	2,196	2,160	Std. Dev.	1,726	1,608	1,594	1,709	1,648
Mean	Mean	3,226	3,206	3,185	3,155	3,180	Mean	1,743	1,713	1,656	1,632	1,666
	Std. Dev.	2,269	2,234	2,188	2,193	2,126	Std. Dev.	1,866	1,710	1,655	1,690	1,618

This table reports a summary of the mean and standard deviation for all the measures of volatility (SSDR, WRSQ, WRNG and AWRET). MO, TU, WD, TH and FR refer to Monday, Tuesday, Wednesday, Thursday and Friday respectively. 1st, 2nd, 3rd and 4th refer to the first, second, third and fourth samples, while Mean refer to the mean values of each statistic for the four periods.