ESTIMATING HISTORICAL VOLATILITY

Michael W. Brandt,
The Fuqua School of Business
Duke University
Box 90120
One Towerview Drive
Durham, NC 27708-0120

Phone: (919) 660-1948
Fax: (919) 660-8038

Email: mbrandt@duke.edu
WWW: www.duke.edu/~mbrandt

Jonathan Kinlay
Investment Analytics
350 Albany Street
New York, NY 10280
Tel: 212 786 1781
Fax: 212 786 1782
Email: jkinlay@investment-analytics.com

ABSTRACT

The research considers the properties of a number of statistical measures of volatility, extending from the common standard deviation metric to less widely used range-based measures. Prior research in this field, which has typically featured the use of data series generated by Monte Carlo simulation within the theoretical framework of Geometric Brownian Motion, has tended towards the conclusion that alternate volatility estimators offer substantial efficiency improvements relative to the standard deviation estimator. This research indicates, however, that such findings are critically dependent on the assumptions made with regard to the nature of the underlying process of interest. The research considers the effect that departures from the behavior of the idealized Geometric Brownian Motion process may have on the performance of a variety of volatility estimators. More specifically, we evaluate the impact that sample size and frequency, process drift, opening gaps and time-varying volatility may have on the performance of both the standard deviation metric and its various alternatives. Under each of the scenarios considered, the integrated volatility estimator represents the "gold standard" in terms of bias and efficiency. All other estimators, with the possible exception of the Alizadeh-Brandt-Diebold estimator, produce biased estimates of the true process volatility unless

observations are taken at very high frequencies. The performance of these estimators further deteriorates in the presence of other departures from geometric Brownian motion such as process drift, stochastic volatility and opening gaps. As bid-offer spreads are almost certain to introduce a further source of substantial bias at high frequencies, these findings call into question the usefulness of many of the alternate estimators despite their superior efficiency. However, when the research moves on to consider empirical data derived from the S&P 500 Index process, the Alizadeh-Brandt-Diebold estimator is amongst the worst performers, while the range-based Parkinson estimator comes out ahead. None of the estimators achieves anything close to the levels of efficiency expected from theory or seen in simulation studies. One common factor, however, is that the standard deviation estimator is generally the worst performing metric on every criteria, for both simulated and empirical data series.

Introduction

Volatility estimation is of central importance to risk management, pricing and portfolio construction and a number of attempts have been made in the last 25 years to improve upon the classical standard deviation of daily returns as an estimator of asset volatility. Many of these estimators, such as those developed by Parkinson [1980], Garman and Klass [1980], Rogers and Satchell [1991], Alizahdeh, Brandt and Diebold [2001] and Yang and Zhang [2002] use information on daily trading ranges – the intraday high and low prices in the asset. Compared to the classical close to close estimator, these estimators have a theoretical efficiency that is 5 times greater in the case of the Parkinson estimator, while in the case of the Garman and Klass and Yang and Zang estimators the efficiency is typically as much as 7 or 8 times, depending on the estimation period and the assumed characteristics of the underlying asset process. However, some of the assumptions do not realistically apply to asset processes: for instance, it is typically assumed that the asset process is continuous and follows a geometric Brownian motion, that volatility is constant, and that the drift in the underlying process is zero. Empirical evidence suggests that none of these assumptions is tenable when applied to asset processes and the question at issue is to what extent the theoretical advantages of the alternate estimators endure in the real world application of estimating asset volatility.

As the true volatility of an asset process if unobservable, prior studies have tended to focus on comparison of the performance of various estimators for a process generated by Monte Carlo simulation. While these have tended to support the claimed efficiency improvements of the various estimators, in most cases they have ignored inconvenient departures from GBM behavior such as non-zero drift, non-constant variance and discontinuities in the underlying process. By contrast, Rogers, Satchell and Yoon [1994] do consider this issue and show that the Rogers and Satchell estimator significantly outperforms other estimators when the asset process includes a time-varying drift. The more recent Yang and Zhang estimator is likewise theoretically better equipped to cope with underlying processes that have non-zero drift and in this study we compare the performance of this and other estimators for a simulated GBM process incorporating non-zero drift of varying levels.

Another problem encountered in estimating volatility for asset processes is the existence of opening price jumps. Rather than being continuous, most asset markets are closed overnight and for certain holidays. Information arriving during periods when the market is closed often results in opening prices that differ significantly from the closing price of the prior trading session. Price gaps of this kind are particularly evident in markets that are affected by trading in other, non-contemporaneous, markets (for instance, the opening of a trading session in European markets often reflects activity in related US markets in its prior trading session), or in markets in which non-trading related news frequently arrives when the market is closed (for instance, in agricultural commodity markets, where news is often weather-related). Hence it is important to test the effect of opening price gaps on the performance of the various volatility estimators. Finally, it is important to consider how estimators perform for processes in which volatility is not constant. The expectation here is that some of the more efficient estimators that incorporate price range information are likely to perform better as they require fewer data points to arrive at volatility estimates and hence will be influenced to a lesser degree by observations from earlier periods in which a different volatility regime may have pertained.

While for a simulated process with known drift and volatility the procedure for assessing estimator performance is straightforward, the same is not true for real market processes where both drift and volatility are unobservable. Here we take the approach developed by Anderson and Bollerslev [1997], using high frequency data at five minute intervals to arrive at daily volatility estimates.

Data and Methodology

Volatility Estimation

A geometric Brownian motion process for an asset with price S_t evolves as follows:

$$dS_t = \mu S_t dt + \sigma S_t dZ_t \quad [1.1]$$

Where μ is the asset drift, σ is the volatility (assumed constant) and Z_t is a Weiner process. From Ito's lemma thef the log asset price is

$$d \ln S_t = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dZ_t \qquad [1.2]$$

In the rest of the paper the notation developed by Garman and Klass is adopted:

σ= asset volatility, to be estimated;

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.

Using this notation the classical variance estimator σ^2 is given by

$$\hat{\sigma}_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}[(o_i + c_i) - \frac{1}{n}\sum_{i=1}^{n}(o_i + c_i)]^2 \qquad [1.3]$$

The Parkinson [1980] estimator uses instead the high and low values to estimate the variance:

$$\hat{\sigma}_P^2 = \frac{1}{4n\ln 2}\sum_{i=1}^{n}(u_i - d_i)^2 \qquad [1.4]$$

Garman and Klass provide an estimator with superior efficiency, having minimum variance on the assumption that the process follows a geometric Brownian motion with zero drift:

$$\hat{\sigma}_{GK}^2 = \frac{0.511}{n}\sum_{i=1}^{n}(u_i - d_i)^2 - \frac{0.019}{n}\sum_{i=1}^{n}[c_i(u_i + d_i) - 2u_i d_i] - \frac{0.383}{n}\sum_{i=1}^{n}c_i^2 \qquad [1.5]$$

Here the term efficiency is used to refer to the ratio

$$\frac{Var(\hat{\sigma}_{GK}^2)}{Var(\hat{\sigma}_c^2)} \qquad [1.6]$$

The no-drift assumption is a good approximation when the quantity $\mu\sqrt{n}/\sigma$ is small, which is usually the case for daily series. However there are often periods during which an asset process trends strongly, the drift being large relative to the volatility. The Parkinson and Garman-Klass estimators will tend to overestimate volatility during these periods.

Rogers and Satchell [1991] derive an estimator that allows for nonzero drift:

$$\hat{\sigma}^2_{RS} = \frac{1}{n}\sum_{i=1}^{n}[u_i(u_i - c_i) + d_i(d_i - c_i)] \qquad [1.7]$$

Yang and Zhang [2002] devise an estimator that combines the classical and Rogers-Satchell estimator, showing that it has the minimum variance and is both unbiased and independent of process drift and opening gaps. Their estimator is given by

$$\hat{\sigma}^2_{YZ} = \hat{\sigma}^2_o + k\hat{\sigma}^2_c + (1-k)\hat{\sigma}^2_{RS} \qquad [1.8]$$

Where the constant k takes the form

$$k = \frac{0.34}{1 + \frac{m+1}{m-1}}$$

Yang and Zhang show that the efficiency of the estimator is 1 + 1/k, and that k is at a minimum of 0.0783 for m = 2. Under these conditions, the efficiency has a peak value of 14, meaning that using only two days' data for this estimator gives the same accuracy as the classical estimator using three week's data. However, where the process is dominated by opening jumps the efficiency reduces to almost one, which means there is no improvement over the classical estimator.

Alizadeh, Brandt and Diebold [2002] show that the log range $\ln(u_t - d_t)$ is to a very good approximation normally distributed with mean $(0.43 + \ln\sigma_t)$, and variance 0.29^2. They test their estimator under a wide variety of simulated market conditions, finding it to be robust to microstructure noise, in contrast to other popular volatility estimators including realized volatility. This estimator has approximately constant variance and hence its efficiency with vary according to the characteristics of the asset process.

Anderson [2000] introduces the concept of integrated volatility, being the sums of squares of high frequency intra-day returns, and demonstrates the superior accuracy of the volatility metric compared to the classical standard deviation measure. Anderson's estimator has become the accepted benchmark for estimating volatility using market data and we adopt it as such in this research.

Data

The test dataset used to evaluate the volatility estimators comprised observations of the S&P 500 Index from 4-Jan-1988 to 31-Dec-2003, some 4,037 trading days. The realized variance is calculated as the sum of squares of five minute intra-day returns. Summary statistics for daily returns reveal the long term upward drift in the S&P 500 index process, which equates to an annual drift rate of 8.1% (Exhibit 1).

Summary statistics for realized volatility show that the average volatility of the process is just over 12%, but with some considerable variation and high levels of kurtosis (Exhibit 2).

Asset Process Simulation

If the log asset price follows a geometric Brownian motion we can use a discrete approximation to generate a simulated asst process using

$$\ln S_t - \ln S_{t-1} = (\mu - 0.5\sigma_t^2)/N + \sigma_t X_{t,} \quad t = 1,\ 2, \ldots N \qquad [1.9]$$

Where X_t is N(0,1/N), normally distributed with zero mean and variance 1/N, N being the number of price movements per day.

We consider a number of simulation scenarios. If the baseline scenario we test the relative efficiency of the various estimators assuming constant volatility and zero drift. Here the focus of interest is on the relative performance of the various estimators as the sample size and frequency of observation vary. Monte Carlo simulation is used to generate 100 instances of the process which are sampled at frequencies of 1-minute, 5-minute and 15-minute in a trading day lasting 6.5 hours. The various estimators are calculated using sample sizes of 5, 25 and 50 days.

In the second scenario volatility again remains constant at 14%, but we allow for a time-varying annual drift of 5%, 10% and 15%. The question of interest here is how volatility estimators are affected by levels of non-zero drift which typically characterize asset processes.

It is well known empirical finding that asset volatility changes over time and this is evidenced in the case of the S&P 500 index, where the process volatility itself has an annual volatility of 4%. The process itself is approximately follows a log-normal distribution as illustrated in Exhibit 3 for the log-volatility process.

In the third scenario we simulate this behavior by allowing volatility to fluctuate each day, while remaining constant during the day. Volatility for a given day is generated by sampling from a LogNormal distribution with the same characteristics as that found empirically for the S&P 500 index. This scenario is likely to favor alternatives to the classical volatility estimator, which assumes that volatility remains constant, and is therefore unlikely to be able to reflect short-term changes in the level of the volatility process.

The final simulation scenario is designed to assess the impact of opening price gaps on the performance of the volatility estimators. Under this scenario the assumption of continuous trading breaks down and one-period estimators such as the Parkinson, Garman-Klass, Rogers and Satchell and Alizadeh-Brandt-Diebold estimators may perform poorly. The Yang and Zhang estimator is a multi-period estimator and should therefore able to capture the effect of the opening jump correctly. The classical close-close estimator is likely to give inflated estimates of volatility because of the discontinuities in the process and its inability to capture anything but close – close movements. In this scenario we assume as before that we can observe the process at 1-minute intervals during a 6.5 hour trading day (390 steps). However we also now assume that prices continue to move outside out trading hours and we examine the outcome if we simulate 50, 100, 150 and 200 price movements when the market is closed.

The final test is an empirical test of the performance of the various estimators on a data sample comprising observations of the S&P 500 index at 5-minute intervals from 4-Jan-1988 to 31-Dec-2004.

The measures used to assess the performance of the various estimators in the simulation scenarios are as follows:

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i - \sigma_i) \qquad [1.10]$$

$$MAD = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\sigma}_i - \sigma\right| \qquad [1.11]$$

$$Eff = \frac{Var(StDev)}{Var(M)} \qquad [1.12]$$

The efficiency test compares the relative uncertainty of the various estimators, using the variance of the classical close-close standard deviation as a benchmark. If the ratio is larger than 1, the variance of the current estimator, M, is smaller than that of the classical estimator. The larger the ratio, the more efficient the estimator is.

Simulation Testing

Exhibit 4(a) in the Appendix shows the results of the first simulation, a base case in which volatility is constant at 14% and the process contains no drift term. In terms of bias, the integrated volatility measure outperforms all of the other estimators, as expected. The classical estimator consistently tends to overestimate the true process volatility, although the bias is small and moderates as the sample size increases. With the exception of the Brand-Diebold estimator, all of the other volatility metrics show a marked tendency to underestimate the true process volatility, a bias that increases as the sampling frequency diminishes. Increasing the sample size from five to fifty days diminishes the bias, but the downward bias induced by low sampling frequency dominates. The findings show how sensitive these estimators are to the assumption that the process is continuously observed. Although the bias is in most cases acceptably small for high frequency sampling, this disguises a further difficulty: at high frequencies, bid-offer bounce will significantly impact observed returns and hence volatility estimates. Consequently in practice, in order to mitigate bid-offer effects, volatility estimates should be based on observations made at 5-minute intervals or more. At these levels the bias in the various non-classical estimators, with the exception of the Alizadeh-Brandt-Diebold estimator, becomes unacceptably large. The performance behavior of the Alizadeh-Brandt-Diebold estimator is unusual. At higher sampling frequencies it exhibits a tendency to over-estimate volatility, to a greater degree than the classical estimator. The bias diminishes as the sampling frequency is reduced to 5-minutes, to less than half that of the classical estimator. As the observation frequency falls still further to 15 minute intervals, the bias again reappears, this time in the form of a tendency to underestimate the true process volatility.

Exhibit 4(b) shows a clear improvement in the accuracy of the estimators, in terms of Mean Absolute Deviation, as the sample size increases. This is true of all of the estimators, including the realized volatility metric, but is most pronounced for the classical estimator. The Alizadeh-Brandt-Diebold estimator dominates the classical estimator in terms of MAD at all sampling frequencies and sample sizes.

As Exhibit 4(c) indicates, the non-standard estimators all show significant efficiency gains relative to the classical estimator, in line with our expectations from theory. Apart from the realized volatility metric, the most efficient estimators are the Garman-Klass and Yang-Zhang estimators. The Alizadeh-

Brandt-Diebold estimator is the least efficient of the non-standard estimators, but its efficiency is not much lower than the Parkinson estimator and its variance is five times smaller than that of the classical estimator. Relative efficiency is not much improved by higher sampling frequency or larger sample size for any of the volatility estimators.

In Exhibit 4(d) we examine the correlations between the volatility estimators for the simulated data series comprising 25 days observations at 5-minute intervals. The most striking finding is the low level of correlation between the realized volatility estimator and its alternatives. This is in sharp contrast with the very high correlations between the estimators when applied to market data, examined later in Exhibit 12(b) in the Appendix, which suggests that there are fundamental behaviors of the market process which are not captured by the simulated process. Further points of interest are the perfect correlation between the Parkinson and Alizadeh-Brandt-Diebold estimators, which stems from their use of the same log-range metric; the low level of correlation between the classical estimator and the Garman-Klass, Rogers-Satchell and Yang-Zhang estimators; and the high correlations between the Yang-Zhang, Garman-Klass and Rogers-Satchell estimators, which again is due to similarities in the metrics employed by these estimators. Exhibit 5 represents the inter-relationships between the various estimators, in the form a cluster diagram.

In Exhibit 6(a) we examine the impact of non-zero drift on the performance of the volatility estimators. In terms of bias, none of the estimators appears much affected by drift levels below the 15% threshold. At higher levels of drift all of the estimators, with the exception of the realized volatility estimator, show a tendency to under-estimate volatility, and this is typically more marked than at lower drift levels. Overall, however, process drift does not appear to be a major influence on estimator bias. In Exhibit 6(b) the Alizadeh-Brandt-Diebold estimator shows very stable MAD performance at all drift levels, while the MAD for both the classical and realized volatility estimators actually improves as drift levels rise. The reverse relationship is generally seen for other estimators. In terms of estimator efficiency, Exhibit 6(c) indicates that the performance of the Parkinson and Alizadeh-Brandt-Diebold estimators is unaffected by drift levels. By contrast, the efficiency of the Garman-Klass, Rogers-Satchell and Yang-Zhang estimators at first increases with the drift level and then declines as it reaches higher drift level of 15% per annum.

As far as correlations between the estimators is concerned, Exhibit 6(d) indicates that the principal effect of process drift is to induce noticeably higher levels of correlation between the realized volatility measure and the other estimators. This finding suggests the hypothesis that process drift is one of the factors which induces high levels of correlation between the volatility estimators seen in the market process (Exhibit 12(b) in the Appendix).

Exhibit 7(a) sets out the results from the scenario in which the asset process variance is a random variable, assumed to follow a LogNormal distribution of the form:

$$f(x) = \frac{1}{x\sqrt{2\pi\xi}} e^{-\frac{1}{2}\left[\frac{\ln x - \sigma^2}{\xi}\right]^2} \qquad [1.13]$$

Where ξ is the volatility of volatility.

This distribution has mean

$$e^{\sigma^2 + \frac{\xi^2}{2}} \qquad [1.14]$$

We allow ξ to take values 0%, 85%, and 116% giving expected values of process volatility of 10%, 12% and 14%. Samples from the distribution are drawn for each day and, once sampled, volatility remains at that level throughout the day. Volatility estimators are calculated for 5-minute sample frequencies and sample length of 25 days. The results indicate how estimator bias increases with kurtosis. For the Parkinson, Rogers-Satchell and Yang-Zhang estimators the downward bias becomes very marked at the highest levels of kurtosis (corresponding to a process volatility of 14%). By contrast, the bias of the classical estimator is quite moderate, while that of the realized volatility and Alizadeh-Brandt-Diebold metrics is minor. As might be expected, Exhibit 7(b) shows demonstrates that the MAD of all of the estimators increases substantially with the level of kurtosis, in some cases by two- or three-fold. At the highest level of kurtosis, the MAD of the classical estimator is at least 1.5 times greater than that of other estimators. The effect of kurtosis can be identified more precisely by comparing the results in Exhibit 7(b) with those of Exhibit 4(b), in which volatility was fixed deterministically at 14% (sample frequency 5 minutes, sample length 25 days), as shown in Exhibit 8.

From Exhibit 9 it is apparent that kurtosis has by far the greatest effect on the classical estimator. The increase in MAD is least for the Garman-Klass and Rogers-Satchell estimators, but all of the estimators are outperformed by the realized volatility and Alizadeh-Brandt-Diebold estimators, whether volatility is deterministic or random.

A similar comparison shows that the impact of kurtosis on estimator bias is quite modest. As Exhibit 10 demonstrates, it is estimator efficiency that is affected most by kurtosis, which compresses the efficiency of all of the volatility estimators by around 50% or more (dramatically more so in the case of the realized volatility estimator). Nonetheless, estimator variance remains in all cases less than half that of the classical estimator.

Another effect of random volatility is to induce much higher levels of correlation between the volatility estimators than results from process drift alone. As Exhibit 9(d) in the Appendix shows, the most important correlations, those between the realized volatility estimator and its alternatives, approaches the levels seen in empirical tests on market data (see Exhibit 12(b) in the Appendix for comparison).

The final simulation scenario considers the performance of the volatility estimators in the presence of opening gaps. Note that in computing realized volatility we omit the first observation of each day as otherwise inflated volatility estimates would result from inclusion of the overnight 'gap' return. As might be expected (since they are single period estimators and/or only consider daily price ranges during periods when the market is open) the bias of the realized volatility, Parkinson, Rogers-Satchell and Alizadeh-Brandt-Diebold estimators are unaffected in any material way by the introduction of opening gaps, as the results in Exhibit 11(a) confirm. Unsurprisingly, the classical, Garman-Klass and Yang-Zhang estimators, which incorporate close-close information, tend to produce increasingly inflated volatility estimates as the average size of the opening gaps increases. At the same time, the efficiency of the latter estimators reduces in proportion to the number of time steps outside market hours, ultimately falling by more than half.

The impact of opening gaps on the correlations between volatility estimators is two-fold. Firstly, as can be seen in Exhibit 11(d), there is a rise in the correlations between the realized volatility estimator and the volatility estimators which use a range-based metric: the Parkinson, Rogers-Satchell and Alizadeh-Brandt-Diebold estimators. These appear to correlate more highly with realized volatility than do other estimators, which are strongly affected by opening gaps. Secondly, there is a breakdown in the very high levels of correlation between the Yang-Zhang, Paarkinson and Rogers-Satchell estimators, again because of the impact of opening gaps on estimators which are not primarily range-based. Note that the Yang-Zhang and Garman-Klass estimators remain highly correlated, as might be deduced from this analysis.

In an attempt to reproduce the very high levels of correlation seen between the volatility estimators in empirical tests (see next section) a final simulation study was conducted in which the various departures from geometric Brownian motion were combined. Volatility was sampled from a LogNormal distribution producing an expected process volatility of 14%; process drift was set to be 8% per annum; and finally we allowed for opening gaps comprising 200 steps outside market hours. The results, shown in Exhibit 11(e) in the Appendix, indicate that this combination of features produces still higher levels of correlation between the various estimators which approaches, but does not yet attain, the very high levels seen in empirical tests. This suggests that the simulated process does not yet incorporate approximations for all of the "anomalies" which typify real asset processes in the way that they depart from idealized GBM behavior. In particular, we conjecture that it is the long term serial autocorrelation (long-memory) in real asset volatility processes that induces very high levels of correlation between the volatility estimators. This important feature is absent from the simulated test series used in this analysis

Empirical Testing

For an empirical test of the volatility estimators 5-minute return data for the S&P 500 index from 4 Jan 1988 to 31 Dec 2003 were used to construct a series of 4,037 daily observations comprising open, high, low and close prices. The series was divided into 161 non-overlapping periods comprising 25 days of data and used to construct volatility estimates using each of the estimators. Of course, in this empirical test we do not know the true process volatility, so we take as our standard the integrated volatility comprising the sums of squares of 5-minute returns over each 25 day period. The results, summarized in Exhibit 12(a), paint a very different picture from that produced from simulation studies. Here it is the Parkinson estimators which outperforms all of the other estimators in terms of bias, Mean Absolute Deviation and Mean Absolute Percentage Error. The closely related Alizadeh-Brandt-Diebold measure, by contrast, appears to have produced inflated volatility estimates, and with larger MAD and variance than other estimators, save the classical estimator. The latter performs worst amongst all of the estimators, with a bias and MAPE that are large enough to lead to economically significant arbitrage opportunities if used as the basis for pricing index options. While the classical estimator produces inflated volatility estimates, the Garman-Klass, Rogers-Satchell and Yang-Zhang estimators all show signs of negative bias, as in the simulation tests. In general, MAD levels are 40% to 100% higher than in the baseline simulation analysis using the same sample length and frequency. The efficiency level of all of the estimators is well below theoretical expectations and even the most extreme of the simulation results. Noticeably, the realized volatility has the lowest relative efficiency of any of the estimators, an outcome not seen in any of the simulation tests.

Another significant departure from the results seen in simulation studies is the very high levels of correlation between all of the volatility estimators, as shown in Exhibit 12(b) in the Appendix. Once again, a cluster diagram is used to illustrate the interrelationships between the estimators and shows a grouping very similar to that seen in simulation analysis (Exhibit 13). This suggests that the factor(s) inducing higher levels of correlation between the estimators affects each one approximately equally, since we do not see a substantial re-alignment of the inter-estimator correlations, but rather a rise in correlations across the board.

Taken together, the findings of significant disparities between the performance characteristics of the volatility estimators in simulated vs. empirical analysis suggest that the departures from the idealized geometric Brownian motion process are rather greater than envisaged in this or other parallel research and may involve a combination of all of the behaviors considered in this analysis (random volatility, opening gaps, process drift), together with other factors such as market microstructure effects not considered here. One conjecture is that long memory effects which empirical studies have repeatedly demonstrated to exist within volatility processes may result in still higher levels of correlation between the volatility estimators than are likely to result from process drift, opening gaps, and random volatility effects alone.

Conclusion

In this study we compare the performance characteristics of a number of alternatives to the classical volatility estimator under a variety of simulated market conditions and in an empirical test on S&P 500 index data. Numerical tests with Monte Carlo simulation show that estimator performance is strongly dependent on the assumptions made about the underlying process, and to the chosen sampling frequency and sample size. With the exception of the realized volatility and Alizadeh-Brandt-Diebold estimators, all of the estimators produce biased estimates of true process volatility unless observations are taken at high frequencies of 1-minute intervals. The performance of these estimators further deteriorates in the presence of other departures from geometric Brownian motion such as process drift, stochastic volatility and opening gaps. As bid-offer spreads are almost certain to introduce a further source of substantial bias at high frequencies, these findings call into question the usefulness of many of the alternate estimators despite their superior efficiency.

Other than the realized volatility estimator, whose performance dominates that of every other estimator in every scenario, the only alternative estimator with consistently superior performance characteristics to the classical estimator is the Alizadeh-Brandt-Diebold log-range estimator. This metric shows itself to be a robust and accurate volatility estimator under many different types of departure from idealized GBM conditions, and in addition performs well with low frequency data and small sample sizes.

The results from empirical research differ significantly from those seen in simulation studies in a number of respects. Here it is the Parkinson estimator that has the best performance, while the Alizadeh-Brandt-Diebold estimator is amongst the worst performers against a number of criteria. None of the estimators achieves anything close to the levels of efficiency expected from theory or seen in simulation analysis. Finally, the levels of correlation between volatility estimators in far higher in the empirical test than in any of the simulation studies. The only common finding between simulation and empirical testing is that the classical estimator performs significantly worse than any of the other estimators on every criterion. Further research is required to devise asset process models that produce simulation results closer to those seen in this empirical test before a definitive assessment can be made of the performance characteristics of the various volatility estimators considered in this study.

APPENDIX

Exhibit 1 Summary Statistics for Daily Returns.

Daily Returns	
Mean	0.0322%
Standard Error	0.0160%
Median	0.0423%
Mode	0
Standard Deviation	0.010144
Sample Variance	0.000103
Kurtosis	4.26
Skewness	-0.26
Range	13.13%
Minimum	-7.11%
Maximum	6.02%
Count	4037

Exhibit 2 Summary Statistics for Realized Volatility.

Realized Volatility	
Mean	12.09%
Standard Error	0.12%
Median	10.17%
Mode	7.11%
Standard Deviation	7.77%
Sample Variance	0.60%
Kurtosis	13.63
Skewness	2.46
Range	114.88%
Minimum	2.14%
Maximum	117.02%
Count	4037

Exhibit 3 Log-Volatility Distribution.

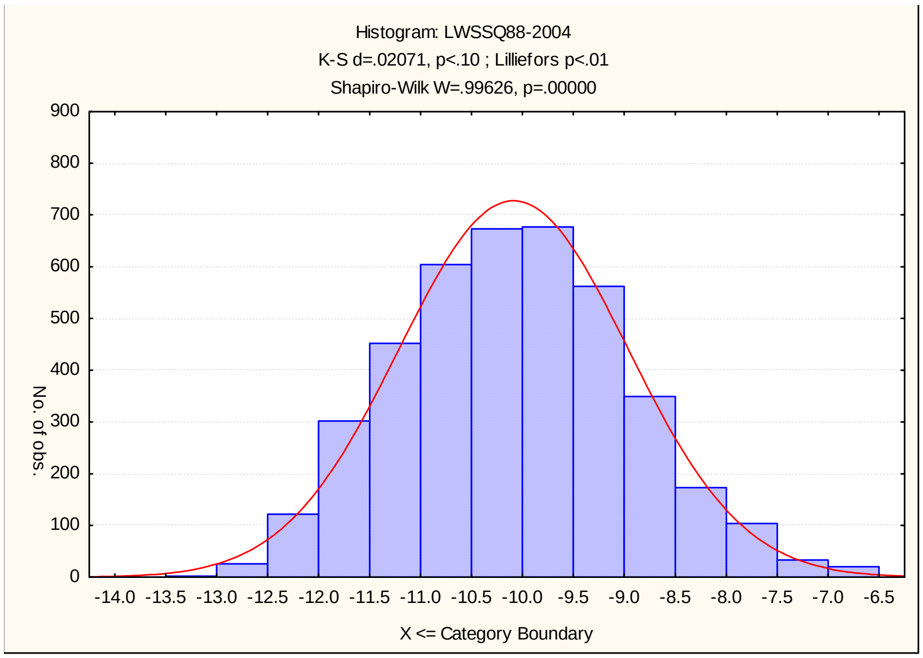


Exhibit 4 Performance statistics for simulated geometric Brownian motion process with zero drift and constant annual volatility of 14%, with observations taken at 1-minute, 5-minute and 15-minute intervals and sample periods of 5, 25 and 50 days.

Exhibit 4(a)

Frequency (mins)	Days	Bias: Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
1	5	0.12%	0.79%	-0.26%	-0.54%	-0.71%	-0.44%	0.88%
	25	0.11%	0.43%	-0.28%	-0.38%	-0.50%	-0.28%	0.86%
	50	0.11%	0.20%	-0.38%	-0.45%	-0.56%	-0.35%	0.75%
5	5	0.12%	0.79%	-0.85%	-1.30%	-1.58%	-1.11%	0.24%
	25	0.11%	0.43%	-0.89%	-1.15%	-1.37%	-0.95%	0.20%
	50	0.11%	0.20%	-0.99%	-1.21%	-1.41%	-1.00%	0.09%
15	5	0.12%	0.79%	-1.67%	-2.24%	-2.77%	-1.98%	-0.65%
	25	0.11%	0.43%	-1.73%	-2.06%	-2.54%	-1.74%	-0.71%
	50	0.11%	0.20%	-1.86%	-2.15%	-2.63%	-1.83%	-0.85%

Exhibit 4(b)

Frequency (mins)	Days	MAD: Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
1	5	0.42%	3.84%	1.62%	1.39%	1.61%	1.43%	1.73%
	25	0.20%	1.70%	0.74%	0.66%	0.85%	0.67%	1.03%
	50	0.14%	1.19%	0.59%	0.55%	0.67%	0.51%	0.82%
5	5	0.42%	3.84%	1.79%	1.75%	2.05%	1.66%	1.69%
	25	0.20%	1.70%	1.04%	1.19%	1.44%	1.03%	0.77%
	50	0.14%	1.19%	1.03%	1.22%	1.43%	1.02%	0.52%
15	5	0.42%	3.84%	2.21%	2.49%	3.04%	2.29%	1.85%
	25	0.20%	1.70%	1.76%	2.08%	2.58%	1.76%	0.98%
	50	0.14%	1.19%	1.87%	2.16%	2.65%	1.84%	0.91%

Exhibit 4(c)

Frequency (mins)	Days	Efficiency: Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
1	5	84.1	1.0	6.0	8.8	6.6	7.9	5.1
	25	80.0	1.0	5.8	8.2	5.3	7.3	4.9
	50	88.0	1.0	6.1	9.4	6.4	8.7	5.2
5	5	84.1	1.0	5.9	9.4	6.8	8.3	5.0
	25	80.0	1.0	5.6	8.2	4.8	7.0	4.8
	50	88.0	1.0	6.0	8.8	5.7	8.1	5.1
15	5	84.1	1.0	5.6	8.2	5.6	7.3	4.8
	25	80.0	1.0	5.5	8.8	4.4	7.5	4.7
	50	88.0	1.0	5.9	9.4	5.7	8.4	5.0

Exhibit 7(d)

Correlations are for the simulated process with sampling at 5-minute frequencies for a 25-day sample period.

Correlation	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Realized	1.000	0.032	0.094	0.156	0.110	0.144	0.094
StDev		1.000	0.786	0.209	-0.164	0.228	0.786
Parkinson			1.000	0.735	0.400	0.715	1.000
Garman-Klass				1.000	0.891	0.970	0.735
Rogers-Satchell					1.000	0.917	0.400
Yang-Zhang						1.000	0.715
Brandt-Diebold							1.000

Exhibit 5 Cluster Diagram for Volatility Estimators.

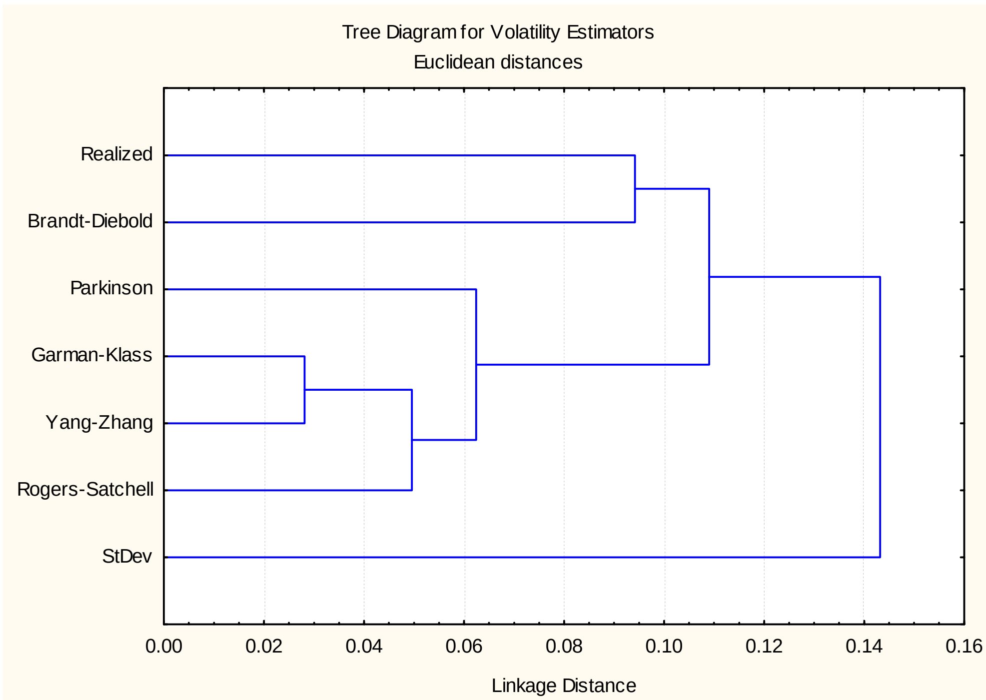


Exhibit 6 Performance statistics for simulated geometric Brownian motion process with varying drift and constant annual volatility of 14%, with observations taken at 5-minute intervals and sample periods of 25 days.

Exhibit 6(a)

	Bias						
Drift	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
0%	0.11%	0.43%	-0.89%	-1.15%	-1.37%	-0.95%	0.20%
5%	0.10%	0.11%	-1.12%	-1.34%	-1.58%	-1.16%	-0.05%
10%	0.08%	0.25%	-1.03%	-1.24%	-1.44%	-1.04%	0.05%
15%	0.04%	-0.24%	-1.30%	-1.41%	-1.55%	-1.20%	-0.24%

Exhibit 6(b)

	MAD						
Drift	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
0%	0.20%	1.70%	1.04%	1.19%	1.44%	1.03%	0.77%
5%	0.21%	1.81%	1.29%	1.38%	1.61%	1.22%	0.85%
10%	0.17%	1.63%	1.15%	1.27%	1.49%	1.09%	0.71%
15%	0.18%	1.54%	1.41%	1.46%	1.59%	1.26%	0.77%

Exhibit 6(c)

	Efficiency						
Steps	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
50	81.4	1.0	7.0	10.9	8.0	10.4	6.0
100	132.8	1.0	6.5	7.5	6.4	6.6	5.5
150	87.0	1.0	6.4	5.0	6.3	5.0	5.4
200	129.9	1.0	7.0	4.4	8.3	4.4	6.0

Exhibit 6(d)

Correlations are for the simulated process with sampling at 5-minute frequencies for a 25-day sample period, with annual drift of 15%.

Correlation	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Realized	1.000	0.228	0.304	0.328	0.222	0.332	0.304
StDev		1.000	0.812	0.373	-0.059	0.323	0.812
Parkinson			1.000	0.812	0.454	0.754	1.000
Garman-Klass				1.000	0.857	0.964	0.812
Rogers-Satchell					1.000	0.916	0.454
Yang-Zhang						1.000	0.754
Brandt-Diebold							1.000

Exhibit 7 Performance statistics for simulated geometric Brownian motion process with zero drift and stochastic volatility LogNormally distributed with mean 10% and standard deviation of 0%, 85%, and 116%, with observations taken at 5-minute intervals and sample periods of 25 days. The results in expected process volatility of 10%, 12% and 14% respectively.

Exhibit 7(a)

	Bias						
Volatility	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
10.00%	0.08%	-0.11%	-0.86%	-0.99%	-1.14%	-0.86%	-0.10%
12.00%	0.17%	-0.08%	-0.97%	-1.09%	-1.27%	-0.94%	-0.05%
14.00%	0.13%	0.58%	-0.96%	-1.31%	-1.61%	-1.11%	0.13%

Exhibit 7(b)

	MAD						
Volatility	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
10.00%	0.13%	1.08%	0.93%	1.02%	1.17%	0.90%	0.49%
12.00%	0.73%	1.89%	1.38%	1.44%	1.61%	1.35%	1.12%
14.00%	1.28%	3.31%	2.08%	1.95%	2.22%	1.89%	1.89%

Exhibit 7(c)

	Efficiency						
Volatility	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
10.00%	87.66	1.00	5.23	7.53	5.48	7.42	4.46
12.00%	4.30	1.00	2.87	3.15	2.53	2.96	2.45
14.00%	4.49	1.00	2.68	3.56	3.69	3.44	2.28

Exhibit 7(d)

Correlations are for the simulated process with sampling at 5-minute frequencies for a 25-day sample period, with mean volatility of 14% and zero drift.

Correlation	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Realized	1.000	0.616	0.751	0.756	0.721	0.783	0.751
StDev		1.000	0.875	0.675	0.502	0.688	0.875
Parkinson			1.000	0.940	0.839	0.940	1.000
Garman-Klass				1.000	0.968	0.992	0.940
Rogers-Satchell					1.000	0.970	0.839
Yang-Zhang						1.000	0.940
Brandt-Diebold							1.000

Exhibit 8 Comparison of Estimator MAD for Deterministic and Stochastic Volatility.

	MAD						
Volatility	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
Fixed	0.20%	1.70%	1.04%	1.19%	1.44%	1.03%	0.77%
Variable	1.28%	3.31%	2.08%	1.95%	2.22%	1.89%	1.89%

Exhibit 9 Comparison of Estimator Bias for Deterministic and Stochastic Volatility.

	Bias						
Volatility	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
Fixed	0.11%	0.43%	-0.89%	-1.15%	-1.37%	-0.95%	0.20%
Variable	0.13%	0.58%	-0.96%	-1.31%	-1.61%	-1.11%	0.13%

Exhibit 10 Comparison of Estimator Efficiency for Deterministic and Stochastic Volatility.

	Efficiency						
Volatility	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Fixed	80.0	1.0	5.6	8.2	4.8	7.0	4.8
Variable	4.5	1.0	2.7	3.6	3.7	3.4	2.3

Exhibit 11 Performance statistics for simulated geometric Brownian motion process with zero drift and volatility of 14%, with observations taken at 1-minute intervals within each trading day and for a variable number of steps when the market is closed. The sample length is 50 days.

Exhibit 11(a)

	Bias						
Steps	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
50	-0.08%	0.89%	-0.59%	0.22%	-0.76%	0.31%	0.52%
100	-0.10%	1.75%	-0.48%	1.17%	-0.67%	1.26%	0.65%
150	-0.10%	2.35%	-0.66%	1.87%	-0.80%	1.96%	0.45%
200	-0.09%	3.26%	-0.58%	2.71%	-0.74%	2.79%	0.53%

Exhibit 11(b)

	MAD						
Steps	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
50	0.15%	1.36%	0.69%	0.40%	0.83%	0.43%	0.64%
100	0.14%	1.86%	0.67%	1.16%	0.80%	1.25%	0.73%
150	0.15%	2.33%	0.74%	1.86%	0.87%	1.94%	0.63%
200	0.13%	3.19%	0.68%	2.69%	0.78%	2.77%	0.70%

Exhibit 11(c)

	Efficiency						
Steps	**Realized**	**StDev**	**Parkinson**	**Garman-Klass**	**Rogers-Satchell**	**Yang-Zhang**	**Brandt-Diebold**
50	81.44	1.00	7.00	10.88	8.05	10.43	5.96
100	132.76	1.00	6.46	7.49	6.40	6.59	5.51
150	87.02	1.00	6.36	5.02	6.25	4.96	5.42
200	129.85	1.00	7.03	4.44	8.35	4.36	5.99

Exhibit 11(d) Correlations are for the simulated process with sampling at 5-minute frequencies for a 25-day sample period, with constant volatility of 14%, zero drift and 200 steps between market close and open.

Correlation	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Realized	1.000	-0.009	0.243	0.106	0.267	0.102	0.243
StDev		1.000	0.483	0.398	-0.135	0.418	0.483
Parkinson			1.000	0.372	0.523	0.392	1.000
Garman-Klass				1.000	0.275	0.967	0.372
Rogers-Satchell					1.000	0.342	0.523
Yang-Zhang						1.000	0.392
Brandt-Diebold							1.000

Exhibit 11(e) Correlations are for the simulated process with sampling at 5-minute frequencies for a 25-day sample period, with random volatility with mean of 14%, annual drift of 8% and 200 steps between market close and open.

Correlation	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Realized	1.000	0.580	0.840	0.845	0.803	0.849	0.840
StDev		1.000	0.700	0.689	0.394	0.693	0.700
Parkinson			1.000	0.768	0.745	0.777	1.000
Garman-Klass				1.000	0.804	0.995	0.768
Rogers-Satchell					1.000	0.813	0.745
Yang-Zhang						1.000	0.777
Brandt-Diebold							1.000

Exhibit 1 Performance statistics for volatility estimators computed on S&P 500 Index data at 5-minute intervals from 4-Jan-1988 to 31-12-2003.

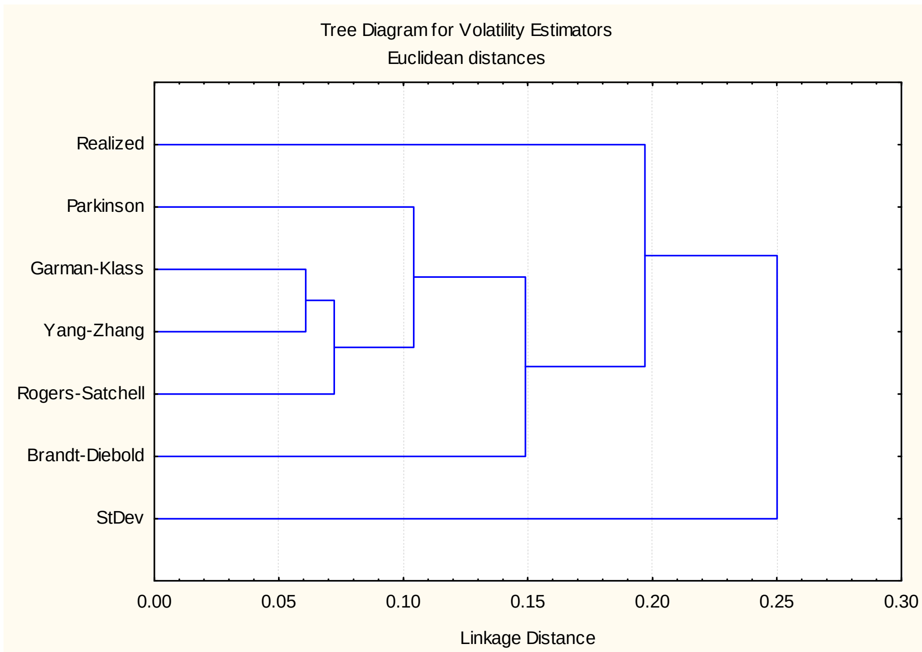
Exhibit 12(a)

	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Mean	13.03%	15.16%	13.05%	12.35%	12.21%	12.73%	14.13%
Variance	0.37%	0.46%	0.29%	0.26%	0.25%	0.28%	0.34%
Bias		2.12%	0.01%	-0.68%	-0.82%	-0.30%	1.10%
MAD		2.37%	1.21%	1.27%	1.51%	1.23%	1.48%
MAPE		20.26%	9.55%	8.75%	10.49%	9.17%	13.41%
Efficiency	1.24	1.00	1.58	1.77	1.83	1.67	1.35

Exhibit 12(b)

Correlation	Realized	StDev	Parkinson	Garman-Klass	Rogers-Satchell	Yang-Zhang	Brandt-Diebold
Realized	1.000	0.954	0.971	0.968	0.953	0.967	0.971
StDev		1.000	0.975	0.944	0.911	0.942	0.975
Parkinson			1.000	0.991	0.976	0.990	1.000
Garman-Klass				1.000	0.994	0.999	0.991
Rogers-Satchell					1.000	0.996	0.976
Yang-Zhang						1.000	0.990
Brandt-Diebold							1.000

Exhibit 2 Cluster Analysis for Volatility Estimators Using Market Data.



REFERENCES

Alizadeh ,S., Brandt, W. M., and Diebold, X.F., 2002. Range-based estimation of stochastic volatility models." *Journal of Finance* 57: 1047-1091

Andersen, T. G., and Bollerslev, T. 1997. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4: 115-158.

Andersen, T. G. 2000. Some reflections on analysis of high-frequency data. *Journal of Business & Economic Statistics* 18: 146-153.

Ball, C.A., and Torous, W. [1984], "The Maximum Likelihood Estimation of Security Price Volatility: Theory, Evidence, and Application to Option Pricing," *Journal of Busines*s, 57, 97-112.

Chang, E. C., Jain, P. C., and Locke, P. R. 1995. Standard & Poor's 500 index futures volatility and price changes around the New York Stock Exchange close. *Journal of Business* 68: 61-84.

Fleming, J. 1998. The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance* 5: 317-345.

Garman, M., and Klass, M. 1980. On the estimation of security price volatilities from historical data. *Journal of Business* 53: 67-78.

Nandi, S. 1998. How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the S & P 500 index options market. *Journal of Banking and Finance* 22: 589-610.

Parkinson, M. 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53: 61-68.

Rogers, L. C. G., and Satchell, S. E. 1991. Estimating variance from high, low and closing prices. *Annals of Applied Probability* 1: 504-512

Rogers L. C. G., and Satchell, S. E., and Yoon, Y. 1994. Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Applied Financial Economics* 4: 241-247.

Yang, D., and Zhang Q. 2000. Drift independent volatility estimation based on high, low, open and close prices. *Journal of Business* 73: 477-491.