A realised volatility measurement using quadratic variation and dealing with microstructure effects

C du Toit[*]   WJ Conradie[†]

Abstract

A volatility measurement that overcomes the respective problems encountered when implementing the realised and Discrete Sine Transform volatility measurements is defined and discussed in this paper. First the shortcomings of these measurements are briefly discussed. Then a modified realised volatility measurement is defined and relevant theoretical results are derived. Finally simulation results are used to evaluate these three volatility measurements.

1 Introduction

In Anderson, *et al.* (2001a, 2001b), Barndorff-Nielsen and Shepard (2001, 2002a, 2002b, 2002c) and Comte and Renault (1998), a model-free (non-parametric) volatility measurement is specified and studied. This measurement, termed *realised volatility*, satisfies the criteria of a good measurement. Given this measurement, volatility may be seen as "observable." For this measurement, no model is specified and it is relatively error-free (in the sense that it approximates the second moment of returns very well). Hence it has considerable advantages relative to other methods. Since it makes it possible for volatility to be observed relatively error-free, its characteristics, known as *stylised facts,* are easily observable and examinable. This leads to the construction of an appropriate volatility model that can be tested for forecasting purposes.

The problem with the realised volatility measurement is that as the sample frequency increases, market microstructure effects increasingly undermine the assumption that returns are independently distributed. Bid-ask spread and price discreteness are the main causes of microstructure effects (Cohen *et al.*, 1981; Glosten, 1994; Roll, 1984; Glottlieb and Kalay, 1985). Owing to these microstructure effects, it no longer holds that the variance

---

[*]ISS International, PO Box 12063, Die Boord, 7613, Stellenbosch, South Africa

[†]Corresponding author: Department of Statistics and Actuarial Science, University of Stellenbosch Private Bag X1, Matieland, 7602, South Africa, email wjc@sun.ac.za

of the sum of returns is the sum of the variances of returns. Corsi and Cursi (2003) demonstrated how to measure volatility in the presence of non-zero autocorrelation of returns based on the Discrete Sine Transform (DST) approach. This method is appropriate if the instantaneous volatility is constant, but it is well-known that in practice this is not the case. Nevertheless, this measurement still gives good estimates of annualised volatility, if the instantaneous volatility does not change too much as a result of a weighted average instantaneous volatility over the time horizon being used. In practice, this method cannot be employed to obtain volatility estimates over short time intervals, *e.g.* daily or weekly estimates. Although this measurement may yield good annualised volatility estimates for any given year, it produces constant daily volatility estimates, which is not satisfactory in most real practical situations. One way of overcoming this problem is to sample returns over a short enough time interval for the instantaneous volatility over that period to be approximately constant. In reality though, too short time intervals are then used, resulting in noisy volatility estimates. Hence there is still a need for a volatility measurement that can simultaneously manage microstructure effects and varying instantaneous volatility.

The realised volatility measurement based on quadratic variation martingale theory handles varying instantaneous volatility effectively, but fails to deal with the microstructure effects. On the other hand, the volatility measurement based on the DST approach manages the microstructure effects efficiently, but breaks down under a changing instantaneous volatility environment.

A modified realised volatility measurement, which we term *microstructure realised volatility*, is subsequently defined in §3. It attempts to address some of the problems of realised volatility based on the quadratic variation and the DST approaches respectively. This method is also based on quadratic variation theory, but the underlying return model is more realistic and incorporates microstructure effects. This measurement is therefore better equipped to deal with both microstructure effects and non-constant volatility. Only the derivation for the first-order autocorrelation version of the model is given, but the model can easily be extended for higher order lags.

2   Definitions, notation and background

Let S_t be the market value of the security or index at time t, with $p_t = \log(S_t)$. Assume that the security or index of interest is very marketable, so that at any instant a trade occurs. We may therefore model the return process and the volatility process in continuous time. The continuous return process is defined as $r_t = p_t - p_0$, with $t \in [0, T]$, *i.e.* $S_t = S_0 e^{r_t}$. A general return model is

$$r_t = A_t + M_t, \tag{1}$$

where A_t denotes the predictable process, and where M_t denotes the unpredictable process (local martingale).

For X_t, a semi martingale, the quadratic variation process $[X, X]_t$ is given by

$$[X, X]_t = X_t^2 - 2 \int_0^t X_{s-} \mathrm{d}X_s. \tag{2}$$

Anderson, *et al.* (2001a) showed that

$$p \lim_{n \to \infty} \sum_{k=1}^{n} \left(r_{t_k} - r_{t_{k-1}} \right)^2 = [r, r]_t, \tag{3}$$

where $t_o = 0$ and $t_n = t$, and hence

$$[r, r]_t = p \lim_{n \to \infty} \sum_{k=1}^{n} \left(r_{t_k} - r_{t_{k-1}} \right)^2 = p \lim_{n \to \infty} \sum_{k=1}^{n} \left(r_{t_k}, r_{t_{k-1}} \right)^2 \tag{4}$$

with $r_{t_k, t_{k-1}} = \log S_{t_k} / S_{t_{k-1}}$. The quadratic variation process $[r, r]_t$ may thus be approximated by $\sum_{k=1}^{n} [r_{t_k, t_{k-1}}]^2$, where $t_0 = 0$, $t_n = t$, and n is as large as possible. For the model (1) it may easily be shown that

$$[r, r]_t = [M, M]_t. \tag{5}$$

Anderson, *et al.* (2001a) defined *realised volatility* as

$$RV_{t, t-h} = \sum_{k=1}^{n} r_{t_k, t_{k-1}}^2, \tag{6}$$

where $t_0 = t - h$, $t_n = t$, and n is as large as possible. The time interval considered is therefore $[t - h, t]$ with length h. Since

$$[r, r]_t - [r, r]_{t-h} = p \lim_{n \to \infty} \sum_{k=1}^{n} r_{t_k, t_{k-1}}^2 = p \lim_{n \to \infty} RV_{t, t-h}, \tag{7}$$

$[r, r]_t - [r, r]_{t-h}$ may be approximated by the realised volatility measurement $RV_{t, t-h}$. Let $\mathcal{F}_t$ be the σ field that reflects the information at time t. Then the important result

$$\mathrm{Var}\left(r_t | \mathcal{F}_{t-h} \right) \approx \mathrm{E}\left([r, r]_t - [r, r]_{t-h} \, | \mathcal{F}_{t-h} \right) \tag{8}$$

(over relatively small time intervals) derived by Anderson, *et al.* (2001a) suggests an estimator for the variation of the return process between times $t - h$ and t. The expression in (6) is thus an approximate unbiased estimator for $\mathrm{Var}\left(r_t | \mathcal{F}_{t-h} \right)$ for small h. In practice $\mathrm{Var}\left(A_t | \mathcal{F}_{t-h} \right) \approx \mathrm{Cov}(A_t M_t | \mathcal{F}_{t-h}) \approx 0$ and differs from 0 only over long-time horizons.

Although $\mathrm{Var}\left(r_t \mid \mathcal{F}_{t-h} \right)$ is unobservable, it may be estimated via (7). By studying the stylised facts of $[r, r]_t - [r, r]_{t-h}$ as an estimate of $\mathrm{Var}\left(r_t | \mathcal{F}_{t-h} \right)$, a model may be built for the latter for forecasting purposes. Recall that $\sum_{k=1}^{n} [r_{t_k, t_{k-1}}]^2$, where $t_0 = t - h$, and $t_n = t$ converges to $[r, r]_t - [r, r]_{t-h}$ and not to $\mathrm{E}\left([r, r]_t - [r, r]_{t-h} \, | \mathcal{F}_{t-h} \right)$. We would like to have

$$\mathrm{Var}\left(r_t | \mathcal{F}_{t-h} \right) - \mathrm{E}\left(\sum_{k=1}^{n} [r_{t_k, t_{k-1}}]^2 | \mathcal{F}_{t-h} \right) = o_p(1), \tag{9}$$

but only have $\mathrm{Var}\left(r_t \mid \mathcal{F}_{t-h} \right) \approx \mathrm{E}\left([r, r]_t - [r, r]_{t-h} \, | \mathcal{F}_{t-h} \right)$. However, under some weak conditions (9) will be true.

Given (1), these conditions are (Barndorff-Nielsen and Shepard, 2002b):

- A_t must be continuous,

- $\text{Var}(A_t \mid \mathcal{F}_{t-h}) = Cov(A_t M_t \mid \mathcal{F}_{t-h}) = 0$,

- $\text{E}\left(\sum_{k=1}^n \left(A_{t_k} - A_{t_{k-1}} \right)^2 \mid \mathcal{F}_{t-h} \right) \longrightarrow 0$, and

- $\text{E}\left(\sum_{k=1}^n \left(A_{t_k} - A_{t_{k-1}} \right) \left(M_{tk} - M_{t_{k-1}} \right) \mid \mathcal{F}_{t-h} \right) \longrightarrow 0$ as $n \to \infty$.

3   A modified realised volatility measurement

The model (1) is not realistic when the sampling frequency is very small, because it assumes that returns are independent. A model that can handle dependency of returns of one or more lags is required. A more applicable return model is

$$r_t = A_t + M_t + \phi M_{t-s}, \tag{10}$$

where s is the time between consecutive trades and where A_t denotes the predictable process. Here M_t and M_{t-s} are unpredictable single processes (local martingales) at different times as before with $0 \le \phi \le 1$. The model (10) assumes first order autocorrelation of returns. Let $h > 2s$. Then

$\text{Cov}\{(r_t - r_{t-s})(r_{t-s} - r_{t-2s})|\mathcal{F}_{t-h}\}$

$$
\begin{aligned}
=\ & \text{E}\{(A_t - A_{t-s} + M_t - M_{t-s} + \phi\left(M_{t-s} - M_{t-2s}\right)) \cdot \\
& (A_{t-s} - A_{t-2s} + M_{t-s} - M_{t-2s} + \phi\left(M_{t-2s} - M_{t-3s}\right))|\mathcal{F}_{t-h}\} \\
& - \text{E}\left(A_t - A_{t-s}|\mathcal{F}_{t-h}\right)\text{E}\left(A_{t-s} - A_{t-2s}|\mathcal{F}_{t-h}\right) \\
=\ & \text{E}\{(A_t - A_{t-s})\, A_{t-s} - A_{t-2s} \mid \mathcal{F}_{t-h}\} + \phi\text{E}\{(M_{t-s} - M_{t-2s})^2|\mathcal{F}_{t-h}\} \\
& - \text{E}\left(A_t - A_{t-s}|\mathcal{F}_{t-h}\right)\text{E}\left(A_{t-s} - A_{t-2s}|\mathcal{F}_{t-h}\right) \\
=\ & \phi\text{E}\{(M_{t-s} - M_{t-2s})^2|\mathcal{F}_{t-h}\} + \text{E}\left(A_t - A_{t-s} \mid \mathcal{F}_{t-h}\right)\text{E}\left(A_{t-s} - A_{t-2s}|\mathcal{F}_{t-h}\right) \\
& - \text{E}\left(A_t - A_{t-s}|\mathcal{F}_{t-h}\right)\text{E}\left(A_{t-s} - A_{t-2s}|\mathcal{F}_{t-h}\right) \tag{11}
\end{aligned}
$$

because the $\{A_t\}$ are independent in a time window of length s. Therefore

$$\text{Cov}\{(r_t - r_{t-s})(r_{t-s} - r_{t-2s})|\mathcal{F}_{t-h}\} = \phi\text{E}\{(M_{t-s} - M_{t-2s})^2 \mid \mathcal{F}_{t-h}\} \tag{12}$$

and if we assume M_t to follow a Brownian motion process, it follows that

$$\text{Cov}\{(r_t - r_{t-s})(r_{t-s} - r_{t-2s})|\mathcal{F}_{t-h}\} = \frac{\phi}{1 + \phi^2} \ \left(MA(1) - \text{process}\right). \tag{13}$$

The same path as was taken for model (1) has to be taken for model (10) in order to obtain an ex-post measurement for the volatility of model (10). In the following theorem the variance of the return process is expressed in terms of M_t.

Theorem 1 *Over short time horizons,*

$$\operatorname{Var}(r_t|\mathcal{F}_{t-h}) = \operatorname{E}(M_t^2|\mathcal{F}_{t-h}) - M_{t-h}^2 + \phi\,(\phi + 2)\,(\operatorname{E}(M_t^2|\mathcal{F}_{t-h}) - M_{t-h}^2). \qquad (14)$$

Proof

$$
\begin{aligned}
\operatorname{Var}(r_t|\mathcal{F}_{t-h}) &= \operatorname{E}\{(r_t - \operatorname{E}(r_t|\mathcal{F}_{t-h}))^2|\mathcal{F}_{t-h}\}\\
&= \operatorname{E}\{(A_t + M_t + \phi M_{t-s} - \operatorname{E}(A_t + M_t + \phi M_{t-s}|\mathcal{F}_{t-h}))^2|\mathcal{F}_{t-h}\}\\
&= \operatorname{E}\{(A_t + M_t + \phi M_{t-s} - \operatorname{E}(A_t|\mathcal{F}_{t-h}) - M_{t-h} - \phi M_{t-h})^2|\mathcal{F}_{t-h}\}\\
&= \operatorname{E}\{(A_t - \operatorname{E}(A_t|\mathcal{F}_{t-h}) + (M_t - M_{t-h}) + \phi(M_{t-s} - M_{t-h}))^2|\mathcal{F}_{t-h}\}\\
&= \operatorname{E}\{(M_t - M_{t-h})^2\,|\mathcal{F}_{t-h}\} + \phi^2 \operatorname{E}\{(M_{t-s} - M_{t-h})^2\,|\mathcal{F}_{t-h}\}\\
&\quad + 2\phi \operatorname{E}\{(M_t - M_{t-h})\,(M_{t-s} - M_{t-h})\,|\mathcal{F}_{t-h}\} + \operatorname{Var}(A_t|\mathcal{F}_{t-h})\\
&\quad + 2\phi\,\operatorname{Cov}(A_t M_t|\mathcal{F}_{t-h}) + 2\phi \operatorname{Cov}(A_t M_{t-s}|\mathcal{F}_{t-h})\\
&= \operatorname{E}(M_t^2|\mathcal{F}_{t-h}) - M_{t-h}^2 + \phi^2(\operatorname{E}(M_{t-s}^2|\mathcal{F}_{t-h}) - M_{t-h}^2)\\
&\quad + 2\phi \operatorname{E}\{(M_t - M_{t-s} + M_{t-s} - M_{t-h})\,(M_{t-s} - M_{t-h})\,|\mathcal{F}_{t-h}\}\\
&\quad + \operatorname{Var}(A_t|\mathcal{F}_{t-h}) + 2\phi\,\operatorname{Cov}(A_t M_t|\mathcal{F}_{t-h}) + 2\phi\,\operatorname{Cov}(A_t M_{t-s}|\mathcal{F}_{t-h})\\
&= \operatorname{E}(M_t^2|\mathcal{F}_{t-h}) - M_{t-h}^2 + \phi^2(\operatorname{E}(M_{t-s}^2|\mathcal{F}_{t-h}) - M_{t-h}^2)\\
&\quad + 2\phi \operatorname{E}\{(M_{t-s} - M_{t-h})^2\,|\mathcal{F}_{t-h}\} + \operatorname{Var}(A_t|\mathcal{F}_{t-h})\\
&\quad + 2\phi\,\operatorname{Cov}(A_t M_t|\mathcal{F}_{t-h}) + 2\phi\,\operatorname{Cov}(A_t M_{t-s}|\mathcal{F}_{t-h})\\
&= \operatorname{E}(M_t^2|\mathcal{F}_{t-h}) - M_{t-h}^2 + \phi\,(\phi + 2)\,(\operatorname{E}(M_{t-s}^2|\mathcal{F}_{t-h}) - M_{t-h}^2)\\
&\quad + \operatorname{Var}(A_t|\mathcal{F}_{t-h}) + 2\phi\,\operatorname{Cov}(A_t M_t|\mathcal{F}_{t-h}) + 2\phi\,\operatorname{Cov}(A_t M_{t-s}|\mathcal{F}_{t-h}).(15)
\end{aligned}
$$

The last three terms in (15) are usually negligible and will only have an influence over long time horizons. This proves the theorem. ∎

For Y_t, a stochastic process, the quadratic variation process $[Y, Y]_t$ may be approximated by

$$p\lim_{n \to \infty} \sum_{k=1}^{n} \left(Y_{t_k} - Y_{t_{k-1}}\right)^2$$

if $\int Y_{s-}\mathrm{d}Y_s$ exists, *i.e.* if $\int Y_{s-}\mathrm{d}Y_s$ may be written in terms of Ito integrals.

The quadratic variation process of the model (10) follows as

$$
\begin{aligned}
[r, r]_t &= r_t^2 - 2\int_0^t r_{q-}\mathrm{d}r_q\\
&= (A_t + M_t + \phi M_{t-s})^2 - 2\int_0^t (A_{q-} + M_{q-} + \phi M_{q-s-})\mathrm{d}r_q\\
&= A_t^2 + M_t^2 + \phi^2 M_{t-s}^2 + 2A_t M_t + 2\phi A_t M_{t-s} + 2\phi M_t M_{t-s}\\
&\quad - 2\int_0^t A_{q-}\mathrm{d}A_q - 2\int_0^t M_{q-}\mathrm{d}A_q - 2\phi\int_s^t M_{q-s-}\mathrm{d}A_q - 2\int_0^t A_{q-}\mathrm{d}M_q\\
&\quad - 2\int_0^t M_{q-}\mathrm{d}M_q - 2\phi\int_s^t M_{q-s-}\mathrm{d}M_q - 2\phi\int_0^{t-s} A_{q-}\mathrm{d}M_{q-s}\\
&\quad - 2\phi\int_0^{t-s} M_{q-}\mathrm{d}M_{q-s} - 2\phi^2\int_0^t M_{q-s-}\mathrm{d}M_{q-s}
\end{aligned}
$$

$$
\begin{aligned}
&= [M, M]_t + \phi^2 [M, M]_{t-s} \\
&\quad + 2\phi \left(M_t M_{t-s} - \int_s^t M_{q-s-} \mathrm{d}M_q - \int_0^{t-s} M_{q-} \mathrm{d}M_{q-s} \right)
\end{aligned}
\tag{16}
$$

because the terms involving $\{A_t\}$ are all zero. Define

$$
[M, M_{-s}]_t = M_t M_{t-s} - \int_s^t M_{q-s-} \mathrm{d}M_q - \int_0^{t-s} M_{q-} \mathrm{d}M_{q-s}.
$$

In this notation $[r, r]_t$ may be written as

$$
[r, r]_t = [M, M]_t + \phi^2 [M, M]_{t-s} + 2\phi [M, M_{-s}]_t.
\tag{17}
$$

The above definition has a significant implication for ex-post volatility measurements if the return process is given by (10). Note that

$$
\begin{aligned}
[M, M_{-s}]_t &= M_t M_{t-s} - \int_s^t M_{q-s-} \mathrm{d}M_q - \int_0^{t-s} M_{q-} \mathrm{d}M_{q-s} \\
&= M_t M_{t-s} - p \lim_{n \to \infty} \sum_{k=s}^n M_{t_{k-s-1}} [M_{t_k} - M_{t_{k-1}}] \\
&\quad - p \lim_{n \to \infty} \sum_{k=s}^n M_{t_{k-1}} [M_{t_{k-s}} - M_{t_{k-s-1}}] \\
&= M_t M_{t-s} - p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_{k-s-1}} [M_{t_k} - M_{t_{k-1}}] + M_{t_{k-1}} [M_{t_{k-s}} - M_{t_{k-s-1}}] \right) \\
&= p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_k} M_{t_{k-s}} - M_{t_{k-1}} M_{t_{k-s-1}} \right) \\
&\quad - p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_{k-s-1}} [M_{t_k} - M_{t_{k-1}}] + M_{t_{k-1}} [M_{t_{k-s}} - M_{t_{k-s-1}}] \right) \\
&= p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_k} M_{t_{k-s}} - M_{t_{k-1}} M_{t_{k-s-1}} - M_{t_{k-s-1}} [M_{t_k} - M_{t_{k-1}}] \right. \\
&\quad \left. - M_{t_{k-1}} [M_{t_{k-s}} - M_{t_{k-s-1}}] \right) \\
&= p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_k} M_{t_{k-s}} - M_{t_{k-1}} M_{t_{k-s-1}} - M_{t_{k-s-1}} [M_{t_k} - M_{t_{k-1}}] \right. \\
&\quad \left. - M_{t_{k-1}} [M_{t_{k-s}} - M_{t_{k-s-1}}] \right) \\
&= p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_k} M_{t_{k-s}} - M_{t_k} M_{t_{k-s-1}} - M_{t_{k-1}} M_{t_{k-s}} + M_{t_{k-1}} M_{t_{k-s-1}} \right) \\
&= p \lim_{n \to \infty} \sum_{k=s}^n \left(M_{t_k} [M_{t_{k-s}} - M_{t_{k-s-1}}] - M_{t_{k-1}} [M_{t_{k-s}} + M_{t_{k-s-1}}] \right) \\
&= p \lim_{n \to \infty} \sum_{k=s}^n (M_{t_k} - M_{t_{k-1}})(M_{t_{k-s}} - M_{t_{k-s-1}}).
\end{aligned}
$$

Therefore $[r, r_{-s}]_t$ may be approximated by

$$\sum_{k=s}^{n} (r_{t_k, t_{k-1}})(r_{t_{k-1}, t_{k-2}}),$$

where n is made as large as possible and $s = 1$. Hence the variance of the model (10) may be written in terms of (17) and $[r, r_{-s}]_t$. But first note that

$$
\begin{aligned}
[r, r_{-s}]_t &= r_t r_{t-s} - \int_s^t r_{q-s-} \mathrm{d}r_q - \int_0^{t-s} r_{q-} \mathrm{d}r_{q-s} \\
&= (A_t + M_t + \phi M_{t-s})(A_{t-s} + M_{t-s} + \phi M_{t-2s}) \\
&\quad - \int_s^t (A_{q-s-} + M_{q-s-} + \phi M_{q-2s-})\mathrm{d}r_q - \int_0^{t-s} (A_{q-} + M_{q-} + \phi M_{q-s-})\mathrm{d}r_{q-s} \\
&= A_t A_{t-s} + A_t M_{t-s} + \phi A_t M_{t-2s} + M_t A_{t-s} + M_t M_{t-s} + \phi M_t M_{t-2s} \\
&\quad + \phi M_{t-s} A_{t-s} + \phi M_{t-s}^2 + \phi^2 M_{t-s} M_{t-2s} - \int_s^t A_{q-s-} \mathrm{d}A_q - \int_s^t A_{q-s-} \mathrm{d}M_q \\
&\quad - \phi \int_s^t A_{q-s-} \mathrm{d}M_{q-s} - \int_s^t M_{q-s-} \mathrm{d}A_q - \int_s^t M_{q-s-} \mathrm{d}M_q - \phi \int_s^t M_{q-s-} \mathrm{d}M_{q-s} \\
&\quad - \phi \int_s^t M_{q-2s-} \mathrm{d}A_q - \phi \int_s^t M_{q-2s-} \mathrm{d}M_q - \phi^2 \int_s^t M_{q-2s-} \mathrm{d}M_{q-s} \\
&\quad - \int_s^t A_{q-} \mathrm{d}A_{q-s} - \int_s^t A_{q-} \mathrm{d}M_{q-s} - \phi \int_s^t A_q \mathrm{d}M_{q-2s} - \int_s^t M_{q-} \mathrm{d}A_{q-s} \\
&\quad - \int_s^t M_{q-} \mathrm{d}M_{q-s} - \phi \int_s^t M_q \mathrm{d}M_{q-2s} - \phi \int_s^t M_{q-s-} \mathrm{d}A_{q-s} \\
&\quad - \phi \int_s^t M_{q-s-} \mathrm{d}M_{q-s} - \phi^2 \int_s^t M_{q-s} \mathrm{d}M_{q-2s} \\
&= [M, M_{-s}]_t + \phi[M, M_{-2s}]_t + \phi[M, M]_{t-s} + \phi^2[M, M_{-s}]_{t-s}
\end{aligned}
$$
(18)

because the terms involving A_t are zero.

Theorem 2 $\mathrm{E}[M, M_{-s}]_t = 0$.

Proof

$$
\begin{aligned}
&\mathrm{E}([M, M_{-s}]_t - [M, M_{-s}]_{t-h} | \mathcal{F}_{t-h}) \\
&= \mathrm{E}\left(M_t M_{t-s} - \int_s^t M_{q-s-} \mathrm{d}M_q - \int_0^{t-s} M_{q-} \mathrm{d}M_{q-s} - M_{t-h} M_{t-s-h} + \int_s^{t-h} M_{q-s-} \mathrm{d}M_q \right. \\
&\quad \left. + \int_0^{t-s-h} M_{q-} \mathrm{d}M_{q-s} \Big| \mathcal{F}_{t-h} \right) \\
&= \mathrm{E}(M_t M_{t-s} | \mathcal{F}_{t-h}) - \mathrm{E}\left(\int_{t-h}^t M_{q-s-} \mathrm{d}M_q | \mathcal{F}_{t-h} \right) - \mathrm{E}\left(\int_{t-s-h}^{t-s} M_{q-} \mathrm{d}M_{q-s} | \mathcal{F}_{t-h} \right) \\
&\quad - M_{t-h} M_{t-s-h} \\
&= \mathrm{E}(M_t - M_{t-s} + M_{t-s})M_{t-s} | \mathcal{F}_{t-h}) - \mathrm{E}\left(\int_{t-h}^t M_{q-s-} \mathrm{d}M_q \right) \\
&\quad - \mathrm{E}\left(\int_{t-s-h}^{t-s} (M_{t-s} - M_{t-s} + M_{q-})\mathrm{d}M_{q-s} | \mathcal{F}_{t-h} \right) - M_{t-h} M_{t-s-h}
\end{aligned}
$$

$$
\begin{aligned}
&= 0 + \mathrm{E}(M_{t-s}^2|\mathcal{F}_{t-h}) - 0 - \mathrm{E}\left(\int_{t-s-h}^{t-s} M_{t-s}\mathrm{d}M_{q-s}|\mathcal{F}_{t-h}\right) \\
&\quad + \left(\mathrm{E}\int_{t-s-h}^{t-s} (M_{t-s} - M_{q-})\mathrm{d}M_{q-s}|\mathcal{F}_{t-h}\right) - M_{t-h}M_{t-s-h} \\
&= \mathrm{E}(M_{t-s}^2|\mathcal{F}_{t-h}) - \mathrm{E}(M_{t-s}(M_{t-s} - M_{t-s-h})|\mathcal{F}_{t-h}) + 0 - M_{t-h}M_{t-s-h} \\
&= \mathrm{E}(M_{t-s}^2|\mathcal{F}_{t-h}) - \mathrm{E}(M_{t-s}^2|\mathcal{F}_{t-h}) + M_{t-h}M_{t-s-h} - M_{t-h}M_{t-s-h} \\
&= 0.
\end{aligned}
$$

This completes the proof. ∎

From (16), (18) and Theorem (2) the expected values of $[r,r]_t$ and $[r,r_{-s}]_t$ are

$$
\begin{aligned}
\mathrm{E}\left([r,r]_t - [r,r]_{t-h}|\mathcal{F}_{t-h}\right) &= \mathrm{E}([M,M]_t + \phi^2\,[M,M]_{t-s} - [M,M]_{t-h} \\
&\quad + \phi^2\,[M,M]_{t-s-h}\,|\mathcal{F}_{t-h}) \\
&= \mathrm{E}\left(M_t^2|\mathcal{F}_{t-h}\right) - M_{t-h}^2 \\
&\quad + \phi^2 E\left(M_{t-s}^2|\mathcal{F}_{t-h}\right) - \phi^2 M_{t-s-h}^2 \qquad (19)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}\left([r,r_{-s}]_t - [r,r_{-s}]_{t-h+s}|\mathcal{F}_{t-h}\right) &= \mathrm{E}\left(\phi[M,M]_{t-s} - \phi[M,M]_{t-h}|\mathcal{F}_{t-h}\right) \\
&= \phi\mathrm{E}\left(M_{t-s}^2|\mathcal{F}_{t-h}\right) - \phi M_{t-h}^2 \qquad (20)
\end{aligned}
$$

respectively. Combining (19) and (20) it follows that

$$
\begin{aligned}
&\mathrm{E}\left([r,r]_t - [r,r]_{t-h}\mid\mathcal{F}_{t-h}\right) + 2\mathrm{E}\left([r,r_{-s}]_t - [r,r_{-s}]_{t-h+s}\mid\mathcal{F}_{t-h}\right) \\
&= \mathrm{E}\left(M_t^2\mid\mathcal{F}_{t-h}\right) - M_{t-h}^2 + \phi^2\mathrm{E}\left(M_{t-s}^2\mid\mathcal{F}_{t-h}\right) - \phi^2 M_{t-s-h}^2 \\
&\quad + 2\left(\phi\mathrm{E}\left(M_{t-s}^2\mid\mathcal{F}_{t-h}\right) - \phi M_{t-h}^2\right) \\
&= \mathrm{E}\left(M_t^2\mid\mathcal{F}_{t-h}\right) - M_{t-h}^2 + \phi(\phi+2)\left(\mathrm{E}\left(M_{t-s}^2\mid\mathcal{F}_{t-h}\right) - M_{t-h}^2\right) \\
&\quad + \phi^2\left(M_{t-h}^2 - M_{t-s-h}^2\right) \\
&= \mathrm{Var}(r_t\mid\mathcal{F}_{t-h}) + \phi^2\left(M_{t-h}^2 - M_{t-s-h}^2\right).
\end{aligned}
$$

The bias of this estimator is $\phi^2\left(M_{t-h}^2 - M_{t-s-h}^2\right)$. However, in practice ϕ^2 should be small, and $M_{t-h}^2 - M_{t-s-h}^2$ should be of the order of s. So

$$
\mathrm{Var}(r_t\mid\mathcal{F}_{t-h}) \approx \mathrm{E}\left([r,r]_t - [r,r]_{t-h}\mid\mathcal{F}_{t-h}\right) + 2\mathrm{E}\left([r,r_{-s}]_t - [r,r_{-s}]_{t-h+s}\mid\mathcal{F}_{t-h}\right)
$$

and

$$
[r,r]_t - [r,r]_{t-h} \approx \sum_{k=1}^n [r_{t_k,t_{k-1}}]^2
$$

$$
[r,r_{-s}]_t - [r,r_{-s}]_{t-h+s} \approx \sum_{k=2}^n (r_{t_k,t_{k-1}})(r_{t_{k-1},t_{k-2}})
$$

with $s = 1$.

The variance of the model (10) may therefore be approximated by

$$\sum_{k=1}^{n}[r_{t_k,t_{k-1}}]^2 + 2\sum_{k=2}^{n}(r_{t_k,t_{k-1}})(r_{t_{k-1},t_{k-2}}).$$

The term $2\sum_{k=2}^{n}(r_{t_k,t_{k-1}})(r_{t_{k-1},t_{k-2}})$ is approximately the difference between the variances of returns for the models (1) and (10). If a lag m return model is used, *i.e.*

$$r_t = A_t + M_t + \phi_1 M_{t-s} + \phi_2 M_{t-2s} + \ldots + \phi_n M_{t-ms}, \tag{21}$$

then

$$\sum_{k=1}^{n}[r_{t_k,t_{k-1}}]^2 + 2\sum_{k=2}^{n}(r_{t_k,t_{k-1}})(r_{t_{k-1},t_{k-2}}) + 2\sum_{k=3}^{n}(r_{t_k,t_{k-1}})(r_{t_{k-2},t_{k-3}})$$
$$+ \ldots + 2\sum_{k=m+1}^{n}(r_{t_k,t_{k-1}})(r_{t_{k-n},t_{k-(m+1)}})$$

is an ex-post measurement of volatility. The mathematics becomes rather cumbersome, so the derivation of the variance of the model in (19) is not given here.

4 Practical simulations and results

The same method used to model the influence of microstructure effects as described in Hasbrouck (1993, 1996) and as used in Corsi and Cursi (2003) is used here. Seventy-two thousand (*i.e.* (250 days)·(24 hours)·(60 minutes)/5) returns of length 5 minutes were simulated a 100 times, with different characteristics in each simulation. We compared the microstructure realised volatility (micr.RV) with the realised volatility (RV), the 30 minutes realised volatility (every 6th return is to be taken) and the DST measurement. In Tables 1–5, the mean, the mean squared error (MSE), and the variance (Var) of the four measurements are given. In determining the DST measurement, m is taken as 40, and the bias is subtracted. The DST measurement is therefore unbiased.

	Mean	MSE	Var
micr.RV	29.5350	0.0175	0.0173
DST	28.5715	0.9107	0.0186
RV	29.5233	0.0044	0.0044
RV.30min	29.5356	0.0335	0.0334

Table 1: *Simulation 1. Constant instantaneous volatility and no autocorrelation. Annualised volatility: 29.5161%.*

In simulation 1 the returns are independent with constant instantaneous volatility. From Table 1 it is clear that all estimators perform well. The respective means are all close to

the annualised volatility of 29.5161. The mean squared errors and variances are small, with the realised volatility measurement outperforming the others in this regard.

A non-zero signal to noise ratio was used in the simulations that follow. The signal to noise ratio is defined as $\sqrt{\sigma_1}/\sigma$, where σ is the tick-by-tick (5 minutes) volatility and σ_1 is the tick-by-tick (5 minutes) first-order auto-covariance.

In simulations 2 and 3 the returns have first order autocorrelation with constant instantaneous volatility. We notice in Table 3 that if the noise to signal ratio and the autocorrelation are very high, the DST measurement performs best, but this is in an unrealistic environment. As the noise to signal ratio decreases, the microstructure realised volatility quickly becomes the best estimator. Even at a very high noise to signal ratio and with an autocorrelation of 0.92 and -0.32 respectively, the microstructure realised volatility easily out-performs the other measurements, as may be seen in Table 2.

	mean	MSE	Var
micr.RV	29.5372	0.0713	0.0716
DST	28.2334	1.8505	0.2072
RV	47.9422	369.4281	30.2076
RV.30min	47.9396	368.8906	29.7636

Table 2: *Simulation 2. Constant instantaneous volatility and $p(1) = -0.32$ with noise to signal ratio 0.92. Annualised volatility: 29.5161%.*

	mean	MSE	Var
micr.RV	29.9199	5.6545	5.5470
DST	29.1315	0.5399	0.3959
RV	155.2504	16 408.39	605.3382
RV.30min	155.1981	16 392.86	602.9326

Table 3: *Simulation 3. Constant instantaneous volatility and $p(1) = -0.48$ with noise to signal ratio 3.5. Annualised volatility: 29.5161%.*

In simulation 4 a changing instantaneous volatility was assumed. In this simulation a high annualised volatility, seldom seen in real life, was assumed in order to illustrate how the other measurements break down under non-constant instantaneous volatility and under microstructure effects, as my be seen in Table 4.

In simulation 5 non-zero second order autocorrelation of returns was assumed. It is clear from Table 5 that in the presence of second order autocorrelation the microstructure realised volatility again performs the best. The DST measurement is fairly robust against model misspecifications.

The microstructure realised volatility is the only measurement that performs well in all the simulations. It is also the only estimator that is unbiased and which has a small mean square error. In real life return data, where volatility changes and where microstructures are present, the microstructure realised volatility is the only measurement that still gives

	mean	MSE	Var
micr.RV	100.3488	0.5413	0.6693
DST	90.8785	97.9993	9.0542
RV	119.0969	996.8801	651.9170
RV.30min	119.1650	992.2786	644.6082

Table 4: *Simulation 4. Autocorrelation and changing instantaneous volatility. Annualised volatility: 100.3243%.*

	mean	MSE	Var
micr.RV	29.5086	0.0326	0.0326
DST	28.2386	1.7895	0.1596
RV	46.3981	304.2548	19.5824
RV.30min	47.9396	304.5106	19.7060

Table 5: *Simulation 5. Constant instantaneous volatility and $p(1) = -0.36$, $p(2) = 0.124$ with noise to signal ratio 0.92. Annualised volatility: 29.5161%.*

satisfactory results.

5 Conclusions

In the last few years many attempts have been made to define a volatility measurement that is able to deal with all the characteristics present in returns data in practice. The two problematic characteristics of volatility of high frequency returns are non-zero autocorrelations of lag greater than or equal to one, and changes in the instantaneous volatility. While some volatility measurements capture one of the two characteristics quite well, no previous measurement has been able to deal with both. To the best knowledge of the authors the microstructure realised volatility proposed in this paper is the first volatility measurement to successfully do so.

We have demonstrated the derivation of a microstructure realised volatility measurement that can deal with first order autocorrelation and have mentioned that by adding the correct terms, the model can easily be extended to deal with any autocorrelation lag. In the simulations reported in this study, the superiority of this measurement over previous measurements has been shown under realistic situations. Ultimately, a volatility measurement must lead to improved volatility forecasts to be considered superior to other measurements. Studies that examine how well the raw measurement proposed in this paper performs to other measurements (under various volatility models) will deliver the final verdict on its effectiveness.

References

[1] ANDERSON TG, BORRERSLEV T, DIEBOLD FX & LABYS P, 2001a, *The distribution of exchange rate volatility*, Journal of the American Statistical Association, **96**, pp. 42–55.

[2] ANDERSON TG, BORRERSLEV T, DIEBOLD FX & LABYS P, 2001b, *Modelling and forecasting realised volatility*, NBER Working Paper 8160. The National Bureau of Economic Research, Cambridge (MA).

[3] BARNDORFF-NIELSEN OE & SHEPARD N, 2001, *Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics*, Journal of the Royal Statistical Society, **63**, pp. 167–241.

[4] BARNDORFF-NIELSEN OE & SHEPARD N, 2002a, *Econometric analysis of realized volatility and its use in estimating stochastic volatility models*, Journal of the Royal Statistical Society, **64**, pp. 253–280.

[5] BARNDORFF-NIELSEN OE & SHEPARD N, 2002b, *Estimating quadratic variation using realized volatility*, Journal of Applied Econometrics, **(17)**, pp. 457–477.

[6] BARNDORFF-NIELSEN OE & SHEPARD N, 2002c, *How accurate is the asymptotic approximation to the distribution of realized volatility?*, pp. 306–331, in Andrews DWF & Stock JH (Eds), *Identification and Inference for Econometric models: Essays in honor of Thomas Rothenberg*, Cambridge University Press, Cambridge.

[7] BARNDORFF-NIELSEN OE & SHEPARD N, 2003, *Power and Bipower variation with stochastic volatility and jumps*, Manuscript, Oxford University, Oxford.

[8] COHEN K, COHEN S, COHEN M, COHEN R, COHEN S & COHEN DW, 1981, *Transaction costs, order placement strategy, and existence of the bid-ask spread*, Journal of Political Economy, **89**, pp. 287–305.

[9] COMTE F & RENAULT E, 1998, *Long memory in continuous time stochastic volatility models*, Mathematical Finance, **8**, pp. 291–323.

[10] CORSI F & CURSI G, 2003, *A Discrete Sine Transform Approach for Realized Volatility Measurement*, Manuscript, University of Pisa, Pisa.

[11] GLOSTEN L, 1994, *Is the electronic open limit-order book inevitable?*, Journal of Finance, **49**, pp. 1127–1161.

[12] GLOTTLIEB G & KALAY A, 1985, *Implication of the discreteness of observed stock prices*, Journal of Finance, **40**, pp. 135–153.

[13] HASBROUCK J, 1993, *Assessing the quality of a security market: a new approach to transaction-cost measurement*, Review of Financial Studies, **6(1)**, pp. 191–212.

[14] HASBROUCK J, 1996, *Modelling market microstructure time series*, pp. 647–692, in Maddala GS and Rao CR (Eds.), Statistical Methods in Finance, *Handbook of Statistics*, **14**, Amsterdam.

[15] ROLL R, 1984, *A simple implicit measure of the effective bid-ask spread in an efficient market*, Journal of Finance, **39**, pp. 1127–1139.