Variation, jumps and high frequency data in financial econometrics

Ole E. Barndorff-Nielsen
The T.N. Thiele Centre for Mathematics in Natural Science,
Department of Mathematical Sciences,
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
oebn@imf.au.dk

Neil Shephard
Nuffield College, Oxford OX1 1NF, UK
neil.shephard@nuf.ox.ac.uk

September 20, 2005

1 Introduction

We will review the econometrics of non-parametric estimation of the components of the variation of asset prices. This very active literature has been stimulated by the recent advent of complete records of transaction prices, quote data and order books. In our view the interaction of the new data sources with new econometric methodology is leading to a paradigm shift in one of the most important areas in econometrics: volatility measurement, modelling and forecasting.

We will describe this new paradigm which draws together econometrics with arbitrage free financial economics theory. Perhaps the two most influential papers in this area have been Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002), but many other papers have made important contributions. This work is likely to have deep impacts on the econometrics of asset allocation and risk management. One of the most challenging problems in this context is dealing with various forms of market frictions, which obscure the efficient price from the econometrician. Here we briefly discuss how econometricians have been attempting to overcome them.

In section 2 we will set out the basis of the econometrics of arbitrage-free price processes, focusing on the centrality of quadratic variation. In section 3 we will discuss central limit theorems for estimators of the QV process, while in section 4 the role of jumps in QV will be highlighted, with bipower and multipower variation being used to identify them and to test the hypothesis that there are no jumps in the price process. In section 5 we write about the econometrics of market frictions, while in section 6 we conclude.

2 Arbitrage-free, frictionless price processes

2.1 Semimartingales and quadratic variation

Given a complete record of transaction or quote prices it is natural to model prices in continuous time (e.g. Engle (2000)). This matches with the vast continuous time financial economic arbitrage-free theory based on a frictionless market. In this section and the next, we will discuss how to make inferences on the degree of variation in such frictionless worlds. Section 5 will extend this by characterising the types of frictions seen in practice and discuss strategies econometricians have been using to overcome these difficulties.

In its most general case the fundamental theory of asset prices says that a vector of log-prices at time t,

$$Y_t = \left(Y_t^1, ..., Y_t^p\right)',$$

must obey a *semimartingale* process (written $Y \in \mathcal{SM}$) on some filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P\right)$ in a frictionless market. The semimartingale is defined as being a process which can be written as

$$Y = A + M, \tag{1}$$

where A is a local finite variation process ($A \in \mathcal{FV}_{loc}$) and M is a local martingale ($M \in \mathcal{M}_{loc}$). Compact introductions to the economics and mathematics of semimartingales are given in Back (1991) and Protter (2004), respectively.

The Y process can exhibit jumps. It is tempting to decompose $Y = Y^{ct} + Y^d$, where Y^{ct} and Y^d are the purely continuous and discontinuous sample path components of Y. However, technically this definition is not clear as the jumps of the Y process can be so active that they cannot be summed up. Thus we will define

$$Y^{ct} = A^c + M^c,$$

where M^c is the continuous part of the local martingale component of Y and A^c is A minus the sum of the jumps of A[1]. Likewise, the continuous sample path subsets of $\mathcal{SM}$ and $\mathcal{M}$ will be denoted by $\mathcal{SM}^c$ and $\mathcal{M}^c$.

Crucial to semimartingales, and to the economics of financial risk, is the *quadratic variation* (QV) process of $(Y', X')' \in \mathcal{SM}$. This can be defined as

$$[Y, X]_t = \text{p}-\lim_{n\to\infty} \sum_{j=1}^{t_j \leq t} \left(Y_{t_j} - Y_{t_{j-1}}\right)\left(X_{t_j} - X_{t_{j-1}}\right)', \tag{2}$$

(e.g. Protter (2004, p. 66–77)) for any deterministic sequence[2] of partitions $0 = t_0 < t_1 < ... < t_n = T$ with $\sup_j\{t_{j+1} - t_j\} \to 0$ for $n \to \infty$. The convergence is also locally uniform in time. It can be shown that this probability limit exists for all semimartingales.

Throughout we employ the notation that

$$[Y]_t = [Y, Y]_t,$$

while we will sometimes refer to $\sqrt{[Y^l]_t}$ as the quadratic volatility (QVol) process for Y^l where $l = 1, 2, ..., p$. It is well known that[3]

$$[Y] = [Y^{ct}] + [Y^d], \quad \text{where} \quad [Y^d]_t = \sum_{0\leq u\leq t} \Delta Y_u \Delta Y_u' \tag{3}$$

where $\Delta Y_t = Y_t - Y_{t-}$ are the jumps in Y and noting that $[A^{ct}] = 0$. In the probability literature QV is usually defined in a different, but equivalent, manner (see, for example, Protter (2004, p. 66))

$$[Y]_t = Y_t Y_t' - 2\int_0^t Y_{u-} \mathrm{d}Y_u'. \tag{4}$$

[1]It is tempting to use the notation Y^c for Y^{ct}, but in the probability literature if $Y \in \mathcal{SM}$ then $Y^c = M^c$, so Y^c ignores A^c.

[2]The assumption that the times are deterministic can be relaxed to allow them to be any Riemann sequence of adapted subdivisions. This is discussed in, for example, Jacod and Shiryaev (2003, p. 51). Economically this is important for it means that we can also think of the limiting argument as the result of a joint process of Y and a counting process N whose arrival times are the t_j. So long as Y and N are adapted to at least their bivariate natural filtration the limiting argument holds as the intensity of N increases off to infinity with n.

[3]Although the sum of jumps of Y does not exist in general when $Y \in \mathcal{SM}$, the sum of outer products of the jumps always does exist. Hence $[Y^d]$ can be properly defined.

2.2 Brownian semimartingales

In economics the most familiar semimartingale is the *Brownian semimartingale* ($Y \in \mathcal{BSM}$)

$$Y_t = \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u, \tag{5}$$

where a is a vector of predictable drifts, σ is a matrix volatility process whose elements are càdlàg and W is a vector Brownian motion. The stochastic integral $(\sigma \bullet W)_t$, where $(f \bullet g)_t$ is generic notation for the process $\int_0^t f_u \mathrm{d}g_u$, is said to be a stochastic volatility process ($\sigma \bullet W \in \mathcal{SV}$) — e.g. the reviews in Ghysels, Harvey, and Renault (1996) and Shephard (2005). This vector process has elements which are $\mathcal{M}_{loc}^c$. Doob (1953) showed that all continuous local martingales with absolutely continuous quadratic variation can be written in the form of a SV process (see Karatzas and Shreve (1991, p. 170–172))[4]. The drift $\int_0^t a_u du$ has elements which are absolutely continuous — an assumption which looks ad hoc, however arbitrage freeness plus the SV model implies this property must hold (Karatzas and Shreve (1998, p. 3) and Andersen, Bollerslev, Diebold, and Labys (2003, p. 583)). Hence $Y \in \mathcal{BSM}$ is a rather canonical model in the finance theory of continuous sample path processes. Its use is bolstered by the facts that Ito calculus for continuous sample path processes is relatively simple.

If $Y \in \mathcal{BSM}$ then

$$[Y]_t = \int_0^t \Sigma_u \mathrm{d}u$$

the integrated covariance process, while

$$\mathrm{d}Y_t | \mathcal{F}_t \sim N\left(a_t \mathrm{d}t, \Sigma_{t-} \mathrm{d}t\right), \quad \text{where} \quad \Sigma_t = \sigma_t \sigma_t', \tag{6}$$

where $\mathcal{F}_t$ is the natural filtration – that is the information from the entire sample path of Y up to time t. Thus $a_t \mathrm{d}t$ and $\Sigma_t \mathrm{d}t$ have clear interpretations as the infinitesimal predictive mean and covariance of asset returns. This implies that $A_t = \int_0^t \mathrm{E}\left(\mathrm{d}Y_u | \mathcal{F}_u\right) \mathrm{d}u$ while, centrally to our interests,

$$\mathrm{d}[Y]_t = \mathrm{Cov}\left(\mathrm{d}Y_t | \mathcal{F}_t\right) \quad \text{and} \quad [Y]_t = \int_0^t \mathrm{Cov}\left(\mathrm{d}Y_u | \mathcal{F}_u\right) \mathrm{d}u.$$

[4]An example of a continuous local martingale which has no SV representation is a time-change Brownian motion where the time-change takes the form of the so-called "devil's staircase," which is continuous and non-decreasing but not absolutely continuous (see, for example, Munroe (1953, Section 27)). This relates to the work of, for example, Calvet and Fisher (2002) on multifractals.

Thus A and $[Y]$ are the integrated infinitesimal predictive mean and covariance of the asset prices, respectively.

2.3 Adding jump processes

There is no plausible economic theory which says that prices must follow continuous sample path processes. Indeed we will see later that statistically it is rather easy to reject this hypothesis even for price processes drawn from very thickly traded markets. In this paper we will add a finite activity jump process (this means there are a finite number of jumps in a fixed time interval) $J_t = \sum_{j=1}^{N_t} C_j$, adapted to the filtration generated by Y, to the Brownian semimartingale model. This yields

$$Y_t = \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u + \sum_{j=1}^{N_t} C_j. \tag{7}$$

Here N is a simple counting process and the C are the associated non-zero jumps (which we assume have a covariance) which happen at times $0 = \tau_0 < \tau_1 < \tau_2 < ...$. It is helpful to decompose J into $J = J^A + J^M$, where, assuming J has an absolutely continuous intensity, $J_t^A = \int_0^t c_u \mathrm{d}u$, and $c_t = \mathrm{E}(\mathrm{d}J_t|\mathcal{F}_t)$. Then J^M is the compensated jump process, so $J^M \in \mathcal{M}$, while $J^A \in \mathcal{FV}_{loc}^{ct}$. Thus Y has the decomposition as in (1), with $A_t = \int_0^t (a_u + c_u) \mathrm{d}u$ and

$$M_t = \int_0^t \sigma_u \mathrm{d}W_u + \sum_{j=1}^{N_t} C_j - \int_0^t c_u \mathrm{d}u.$$

It is easy to see that $[Y^d]_t = \sum_{j=1}^{N_t} C_j C_j'$ and so

$$[Y]_t = \int_0^t \Sigma_u \mathrm{d}u + \sum_{j=1}^{N_t} C_j C_j'.$$

Again we note that $\mathrm{E}(\mathrm{d}Y_t|\mathcal{F}_t) = (a_t + c_t)\mathrm{d}t$, but now,

$$\mathrm{Cov}(\sigma_t \mathrm{d}W_t, \mathrm{d}J_t|\mathcal{F}_t) = 0, \tag{8}$$

so

$$\mathrm{Cov}(\mathrm{d}Y_t|\mathcal{F}_t) = \Sigma_t \mathrm{d}t + \mathrm{Cov}(\mathrm{d}J_t|\mathcal{F}_t) \neq \mathrm{d}[Y]_t.$$

This means that the QV process aggregates the components of the variation of prices and so is not sufficient to learn the integrated covariance process $\int_0^t \Sigma_u \mathrm{d}u$.

To identify the components of the QV process we can use the bipower variation (BPV) process introduced by Barndorff-Nielsen and Shephard (2006). So long as it exists, the $p \times p$ matrix BPV process $\{Y\}$ has l, k-th element

$$\left\{Y^l, Y^k\right\} = \frac{1}{4}\left(\left\{Y^l + Y^k\right\} - \left\{Y^l - Y^k\right\}\right), \quad l, k, = 1, 2, ..., p, \tag{9}$$

where, so long as the limit exists and the convergence is locally uniform in t,[5]

$$\left\{Y^l\right\}_t = \text{p}-\lim_{\delta \downarrow 0} \sum_{j=1}^{\lfloor t/\delta \rfloor} \left|Y^l_{\delta(j-1)} - Y^l_{\delta(j-2)}\right| \left|Y^l_{\delta j} - Y^l_{\delta(j-1)}\right|. \tag{10}$$

Here $\lfloor x \rfloor$ is the floor function, which is the largest integer less than or equal to x. Combining the results in Barndorff-Nielsen and Shephard (2006) and Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) if Y is the form of (7) then, without any additional assumptions,

$$\mu_1^{-2}\left\{Y\right\}_t = \int_0^t \Sigma_u \mathrm{d}u,$$

where $\mu_r = \mathrm{E}\left|U\right|^r$, $U \sim N(0,1)$ and $r > 0$, which means that

$$[Y]_t - \mu_1^{-2}\left\{Y\right\}_t = \sum_{j=1}^{N_t} C_j C_j'.$$

At first sight the robustness of BPV looks rather magical, but it is a consequence of the fact that only a finite number of terms in the sum (10) are affected by jumps, while each return which does not have a jump goes to zero in probability. Therefore, since the probability of jumps in contiguous time intervals goes to zero as $\delta \downarrow 0$, those terms which do include jumps do not impact the probability limit. The extension of this result to the case where J is an infinite activity jump process is discussed in Section 4.4.

[5] In order to simplify some of the later results we consistently ignore end effects in variation statistics. This can be justified in two ways, either by (a) setting $Y_t = 0$ for $t < 0$, (b) letting Y start being a semimartingale at zero at time before $t = 0$. The latter seems realistic when dealing with markets open 24 hours a day, borrowing returns from small periods of the previous day. It means that there is a modest degree of wash over from one days variation statistics into the next day. There seems little econometric reasons why this should be a worry. Assumption (b) can also be used in equity markets when combined with some form of stochastic imputation, adding in artifical simulated returns for the missing period — see the related comments in Barndorff-Nielsen and Shephard (2002).

2.4 Forecasting

Suppose Y obeys (7) and introduce the generic notation

$$\begin{aligned} y_{t+s,t} &= Y_{t+s} - Y_t \\ &= a_{t+s,t} + m_{t+s,t}, \quad t, s > 0. \end{aligned}$$

So long as the covariance exists,

$$\begin{aligned} \operatorname{Cov}\left(y_{t+s,t}|\mathcal{F}_t\right) &= \operatorname{Cov}\left(a_{t+s,t}|\mathcal{F}_t\right) + \operatorname{Cov}\left(m_{t+s,t}|\mathcal{F}_t\right) \\ &\quad + \operatorname{Cov}\left(a_{t+s,t}, m_{t+s,t}|\mathcal{F}_t\right) + \operatorname{Cov}\left(m_{t+s,t}, a_{t+s,t}, |\mathcal{F}_t\right). \end{aligned}$$

Notice how complicated this expression is compared to the covariance in (6), which is due to the fact that s is not necessarily $\mathrm{d}t$ and so $a_{t+s,t}$ is no longer known given $\mathcal{F}_t$ — while $\int_t^{t+\mathrm{d}t} a_u \mathrm{d}u$ was. However, in all likelihood for small s, a makes a rather modest contribution to the predictive covariance of Y.

This suggests using the approximation that

$$\operatorname{Cov}\left(y_{t+s,t}|\mathcal{F}_t\right) \simeq \operatorname{Cov}\left(m_{t+s,t}|\mathcal{F}_t\right).$$

Now using (8) so

$$\operatorname{Cov}\left(m_{t+s,t}|\mathcal{F}_t\right) = \operatorname{E}\left([Y]_{t+s} - [Y]_t|\mathcal{F}_t\right) - \operatorname{E}\left\{\left(\int_t^{t+s} c_u \mathrm{d}u\right)\left(\int_t^{t+s} c_u \mathrm{d}u\right)' |\mathcal{F}_t\right\}.$$

Hence if c or s is small then we might approximate

$$\begin{aligned} \operatorname{Cov}\left(Y_{t+s} - Y_t|\mathcal{F}_t\right) &\simeq \operatorname{E}\left([Y]_{t+s} - [Y]_t|\mathcal{F}_t\right) \\ &= \operatorname{E}\left([\sigma \bullet W]_{t+s} - [\sigma \bullet W]_t|\mathcal{F}_t\right) + \operatorname{E}\left([J]_{t+s} - [J]_t|\mathcal{F}_t\right). \end{aligned}$$

Thus an interesting forecasting strategy for covariances is to forecast the increments of the QV process or its components. As the QV process and its components are themselves estimable, though with substantial possible error, this is feasible. This approach to forecasting has been advocated in a series of influential papers by Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001) and Andersen, Bollerslev, Diebold, and Labys (2003), while the important earlier paper by

Andersen and Bollerslev (1998a) was stimulating in the context of measuring the forecast performance of GARCH models. The use of forecasting using estimates of the increments of the components of QV was introduced by Andersen, Bollerslev, and Diebold (2003). We will return to it in section 3.9 when we have developed an asymptotic theory for estimating the QV process and its components.

2.5 Realised QV & BPV

The QV process can be estimated in many different ways. The most immediate is the realised QV estimator

$$[Y_\delta]_t = \sum_{j=1}^{\lfloor t/\delta \rfloor} \left(Y_{j\delta} - Y_{(j-1)\delta}\right)\left(Y_{j\delta} - Y_{(j-1)\delta}\right)',$$

where $\delta > 0$. This is the outer product of returns computed over a fixed interval of time of length δ. By construction, as $\delta \downarrow 0$, $[Y_\delta]_t \xrightarrow{p} [Y]_t$. Likewise

$$\left\{Y_\delta^l\right\}_t = \sum_{j=1}^{\lfloor t/\delta \rfloor} \left|Y_{\delta(j-1)}^l - Y_{\delta(j-2)}^l\right| \left|Y_{\delta j}^l - Y_{\delta(j-1)}^l\right|, \quad l = 1, 2, ..., p, \tag{11}$$

$\left\{Y_\delta^l, Y_\delta^k\right\} = \frac{1}{4}\left(\left\{Y_\delta^l + Y_\delta^k\right\} - \left\{Y_\delta^l - Y_\delta^k\right\}\right)$ and $\{Y_\delta\} \xrightarrow{p} \{Y\}$.

In practice, the presence of market frictions can potentially mean that this limiting argument is not really available as an accurate guide to the behaviour of these statistics for small δ. Such difficulties with limiting arguments, which are present in almost all areas of econometrics and statistics, do not invalidate the use of asymptotics, for it is used to provide predictions about finite sample behaviour. Probability limits are, of course, coarse and we will respond to this by refining our understanding by developing central limit theorems and hope they will make good predictions when δ is moderately small. For very small δ these asymptotic predictions become poor guides as frictions bite hard and this will be discussed in section 5.

In financial econometrics the focus is often on the increments of the QV and realised QV over set time intervals, like one day. Let us define the daily QV

$$V_i = [Y]_{hi} - [Y]_{h(i-1)}, \quad i = 1, 2, ...$$

while it is estimated by the realised daily QV

$$\widehat{V}_i = [Y_\delta]_{hi} - [Y_\delta]_{h(i-1)}, \quad i = 1, 2, ...$$

Clearly $\widehat{V}_i \xrightarrow{p} V_i$ as $\delta \downarrow 0$. The l-th diagonal element of $\widehat{V}_i$, written $\widehat{V}_i^{l,l}$ is called the realised variance[6] of asset l, while its square root is its realised volatility. The latter estimates $\sqrt{V_i^{l,l}}$, the daily QVol process of asset l. The l,k-th element of $\widehat{V}_i$, $\widehat{V}_i^{l,k}$, is called the realised covariance between assets l and k. Off these objects we can define standard dependence measures, like realised regression

$$\widehat{\beta}_i^{l,k} = \frac{\widehat{V}_i^{l,k}}{\widehat{V}_i^{k,k}} \xrightarrow{p} \beta_i^{l,k} = \frac{V_i^{l,k}}{V_i^{k,k}},$$

which estimates the QV regression and the realised correlation

$$\widehat{\rho}_i^{l,k} = \frac{\widehat{V}_i^{l,k}}{\sqrt{\widehat{V}_i^{l,l}\widehat{V}_i^{k,k}}} \xrightarrow{p} \rho_i^{l,k} = \frac{V_i^{l,k}}{\sqrt{V_i^{l,l}V_i^{k,k}}},$$

which estimates the QV correlation. Similar daily objects can be calculated off the realised BPV process

$$\widehat{B}_i = \mu_1^{-2}\left\{\{Y_\delta\}_{hi} - \{Y_\delta\}_{h(i-1)}\right\}, \quad i = 1, 2, ...$$

which estimates

$$B_i = \left[Y^{ct}\right]_{hi} - \left[Y^{ct}\right]_{h(i-1)} = \int_{h(i-1)}^{hi} \sigma_u^2 \mathrm{d}u, \quad i = 1, 2, ...$$

Realised volatility has a very long history in financial economics. It appears in, for example, Rosenberg (1972), Officer (1973), Merton (1980), French, Schwert, and Stambaugh (1987), Schwert (1989) and Schwert (1998), with Merton (1980) making the implicit connection with the case where $\delta \downarrow 0$ in the pure scaled Brownian motion plus drift case. Of course, in probability theory QV was discussed as early as Wiener (1924) and Lévy (1937) and appears as a crucial object in the development of the stochastic analysis of semimartingales which occurred in the second half of the last century. For more general financial processes a closer connection between realised QV and QV, and its use for econometric purposes, was made in a series of independent and concurrent papers by Comte and Renault (1998), Barndorff-Nielsen and Shephard (2001) and Andersen, Bollerslev, Diebold, and Labys (2001). The realised regressions and correlations were defined and studied in detail by Andersen, Bollerslev, Diebold, and Labys (2003) and Barndorff-Nielsen and Shephard (2004).

[6] Some authors call $\widehat{V}_i^{l,l}$ the realised volatility, but throughout this paper we follow the tradition in finance of using volatility to mean standard deviation type objects.

A major motivation for Barndorff-Nielsen and Shephard (2002) and Andersen, Bollerslev, Diebold, and Labys (2001) was the fact that volatility in financial markets is highly and unstably diurnal within a day, responding to regularly timed macroeconomic news announcements, social norms such as lunch times and sleeping or the opening of other markets. This makes estimating

$$\lim_{\varepsilon \downarrow 0} \left([Y]_{t+\varepsilon} - [Y]_t \right) / \varepsilon$$

extremely difficult. The very stimulating work of Genon-Catalot, Larédo, and Picard (1992), Foster and Nelson (1996), Mykland and Zhang (2002) and Mykland and Zhang (2005) tries to tackle this problem using a double asymptotics, as $\delta \downarrow 0$ and $\varepsilon \downarrow 0$. However, in the last five years many econometrics researchers have mostly focused on naturally diurnally robust quantities like the daily or weekly QV.

2.6 Derivatives based on realised QV and QVol

In the last ten years an over the counter market in realised QV and QVol has been rapidly developing. This has been stimulated by interests in hedging volatility risk — see Neuberger (1990), Carr and Madan (1998), Demeterfi, Derman, Kamal, and Zou (1999) and Carr and Lewis (2004). Examples of such options are where the payoffs are

$$\max \left([Y_\delta]_t - K_1, 0 \right), \quad \max \left(\sqrt{[Y_\delta]_t} - K_2, 0 \right). \tag{12}$$

Interesting δ is typically taken as a day. Such options approximate, potentially poorly,

$$\max \left([Y]_t - K_1, 0 \right), \quad \max \left(\sqrt{[Y]_t} - K_2, 0 \right). \tag{13}$$

The fair value of options of the type (13) has been studied by a number of authors, for various volatility models. For example, Brockhaus and Long (1999) employs the Heston (1993) SV model, Javaheri, Wilmott, and Haug (2002) GARCH diffusion, while Howison, Rafailidis, and Rasmussen (2004) study log-Gaussian OU processes. Carr, Geman, Madan, and Yor (2005) look at the same problem based upon pure jump processes. Carr and Lee (2003a) have studied how one might value such options based on replication without being specific about the volatility model. See also the overview of Branger and Schlag (2005).

The common feature of these papers is that the calculations are based on replacing (12) by (13). These authors do not take into account, to our knowledge, the potentially large difference between using $[Y_\delta]_t$ and $[Y]_t$.

2.7 Empirical illustrations: measurement

To illustrate some of the empirical features of realised daily QV, and particularly their precision as estimators of daily QV, we have used a series which records the log of the number of German Deutsche Mark a single US Dollar buys (written Y^1) and the log of the Japanese Yen/Dollar rate (written Y^2). It covers 1st December 1986 until 30th November 1996 and was kindly supplied to us by Olsen and Associates in Zurich (see Dacorogna, Gencay, Müller, Olsen, and Pictet (2001)), although we have made slightly different adjustments to deal with some missing data (described in detail in Barndorff-Nielsen and Shephard (2002)). Capturing time stamped indicative bid and ask quotes from a Reuters screen, they computed prices at each 5-minute period by linear interpolation by averaging the log bid and log ask for the two closest ticks.

Figure 1 provides some descriptive statistics for the exchange rates starting on 4th February, 1991. Figure 1(a) shows the first four active days of the dataset, displaying the bivariate 10 minute returns[7]. Figure 1(b) details the daily realised volatilities for the DM $\sqrt{\widehat{V}_i^1}$, together with 95% confidence intervals. These confidence intervals are based on the log-version of the limit theory for the realised variance we will develop in the next subsection. When the volatility is high, the confidence intervals tend to be very large as well. In Figure 1(c) we have drawn the realised covariance $\widehat{V}_i^{1,2}$ against i, together with the associated confidence intervals. These terms move rather violently through this period. The corresponding realised correlations $\widehat{\rho}_i^{1,2}$ are given in Figure 1(d). These are quite stable through time with only a single realised correlation standing out from the others in the sample. The correlations are not particularly precisely estimated, with the confidence intervals typically being around 0.2 wide.

Table 1 provides some additional daily summary statistics for 100 times the daily data (the scaling is introduced to make the Tables easier to read). It shows the means of the

[7] This time resolution was selected so that the results are not very sensitive to market frictions.

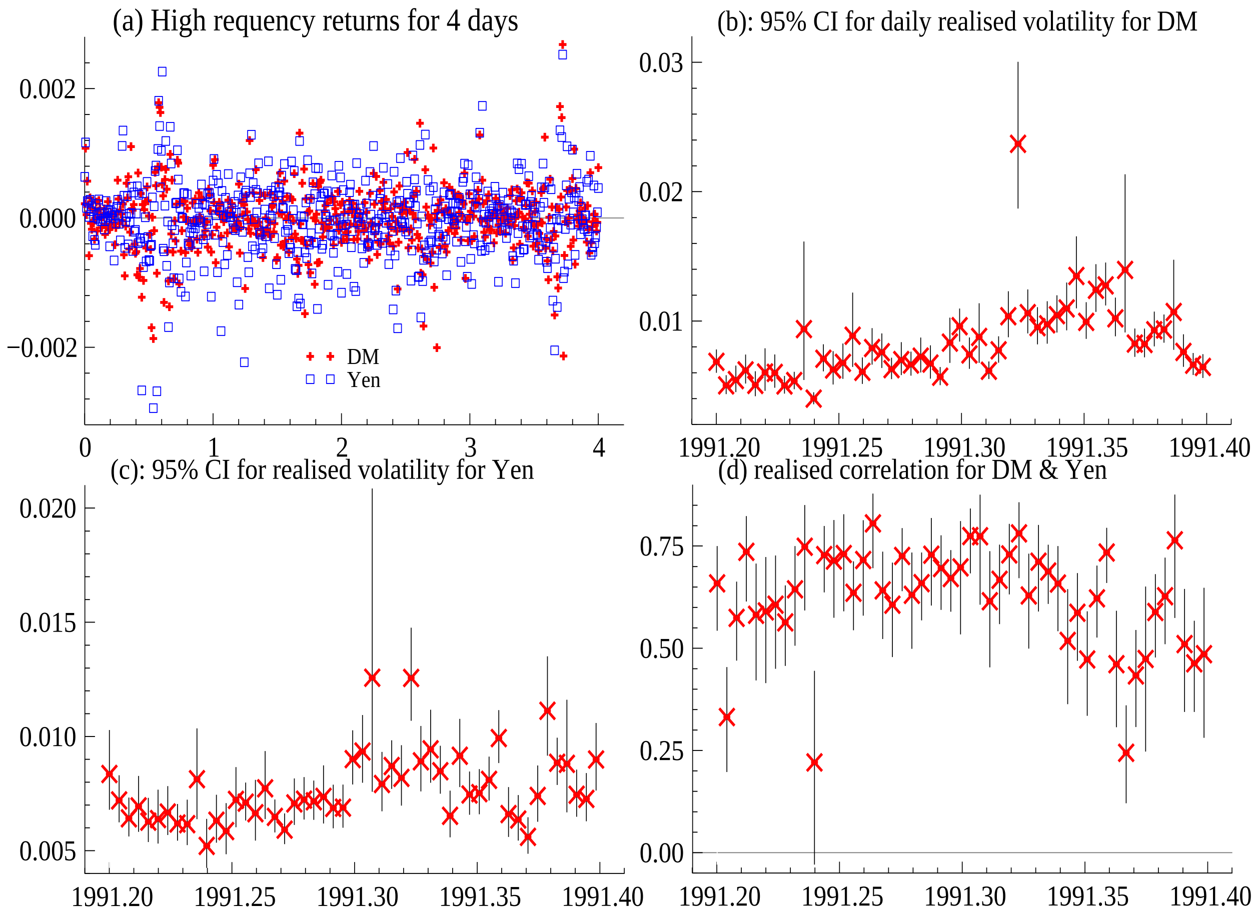

Figure 1: *DM and Yen against the Dollar. Data is 4th February 1991 onwards for 50 active trading days. (a) 10 minute returns on the two exchange rates for the first 4 days of the dataset. (b) Realised volatility for the DM series. This is marked with a cross, while the bars denote 95% confidence intervals. (c) Realised covariance. (d) Realised correlation.*

squared daily returns $\left(Y_i^1 - Y_{i-1}^1\right)^2$ and the estimated daily QVs $\widehat{V}_i^1$ are in line, but that the realised BPV $\widehat{B}_i^1$ is below them. The RV and BPV quantities are highly correlated, but the BPV has a smaller standard deviation. A GARCH(1,1) model is also fitted to the daily return data and its conditional, one-step ahead predicted variances h_i, computed. These have similar means and lower standard deviations, but h_i is less strongly correlated with squared returns than the realised measures.

2.8 Empirical illustration: time series behaviour

Figure 2 shows summaries of the time series behaviour of daily raw and realised DM quantities. They are computed using the whole run of 10 years of 10 minute return data.

Daily	Mean	Standard Dev/*Cor*			
QV: $\widehat{V}_i^1$	0.509	.50			
BPV: $\widehat{B}_i^1$	0.441	*.95*	.40		
GARCH: h_i	0.512	*.55*	*.57*	.22	
$\left(Y_i^1 - Y_{i-1}^1\right)^2$	0.504	*.54*	*.48*	*.39*	1.05

Table 1: *Daily statistics for 100 times DM/Dollar return series: estimated QV, BPV, conditional variance for GARCH and squared daily returns. Reported is the mean, standard deviation and correlations.*

Figure 2(a) shows the raw daily returns and 2(b) gives the corresponding correlogram of daily squared and absolute returns. As usual absolute returns are moderately more autocorrelated than squared returns, with the degree of autocorrelation in these plots being modest, while the memory lasts a large number of lags.

Figure 2(c) shows a time series plot of the daily realised volatilities $\sqrt{\widehat{V}_i^1}$ for the DM series, indicating bursts of high volatility and periods of rather tranquil activity. The correlogram for this series is given in Figure 2(d). This shows lagged one correlations of around one half and is around 0.25 at 10 lags. The correlogram then declines irregularly at larger lags. Figure 2(e) shows $\sqrt{\widehat{B}_i^1}$ using the lagged two bipower variation measure. This series does not display the peaks and troughs of the realised QVol statistics and its correlogram in Figure 2(d) is modestly higher with its first lag being around 0.56 compared to 0.47. The corresponding estimated jump QVol measure $\sqrt{\max\left(0, \widehat{V}_i^1 - \widehat{B}_i^1\right)}$ is displayed in Figure 2(f), while its correlogram is given in Figure 2(d), which shows a very small degree of autocorrelation.

2.9 Empirical illustration: a more subtle example

2.9.1 Interpolation, last price, quotes and trades

So far we have not focused on the details of how we compute the prices used in these calculations. This is important if we wish to try to exploit information buried in returns recorded for very small values of δ, such as a handful of seconds. Our discussion will be based on data taken from the London Stock Exchange's electronic order book, called SETS, in January 2004. The market is open from 8am to 4.30pm, but we remove the first 15 minutes of each day following Engle and Russell (1998). Times are accurate up to one second. We will use three pieces of the database: transactions, best bid and best

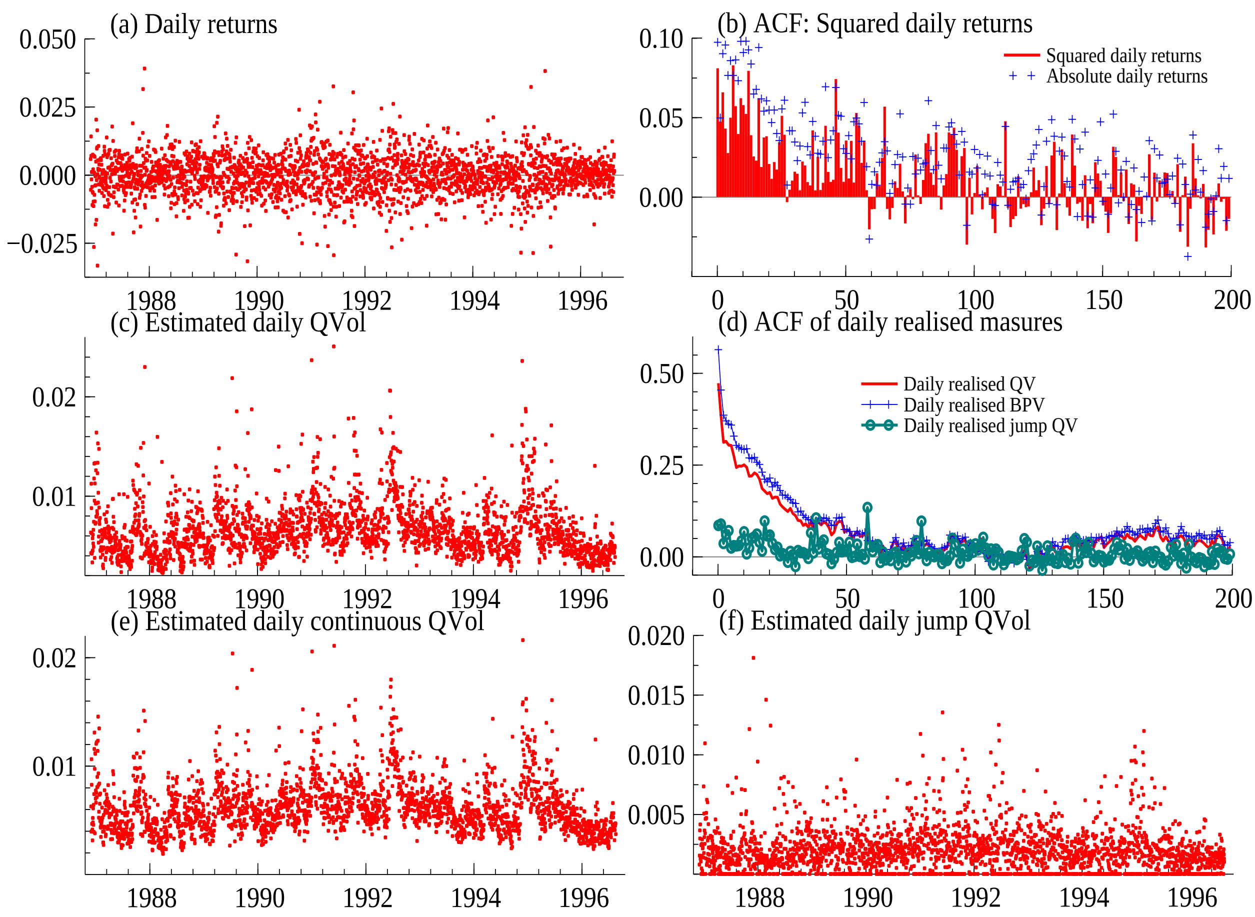

Figure 2: *All graphs use five minute changes data DM/Dollar. Top left: daily returns. Middle left: estimated daily QVol* $\sqrt{\widehat{V}_i}$, *bottom left: estimated daily continuous QVol* $\sqrt{\widehat{B}_i}$. *Bottom right: estimated continuous QVol* $\sqrt{\max\left(0, \widehat{V}_i - \widehat{B}_i\right)}$. *Top right: ACF of squared and absolute returns. X-axis is marked off in days. Middle right: ACF of various realised estimators.*

ask. Note the bid and ask are firm quotes, not indicative like the exchange rate data previous studied. We average the bid and ask to produce a mid-quote, which is taken to proxy the efficient price. We also give some results based on transaction prices. We will focus on four high value stocks: Vodafone (telecoms), BP (hydrocarbons), AstraZeneca (pharmaceuticals) and HSBC (banking).

The top row of Figure 3 shows the log of the mid-quotes, recorded every six seconds on the 2nd working day in January. The graphs indicate the striking discreteness of the price processes, which is particularly important for the Vodafone series. Table 2 gives the

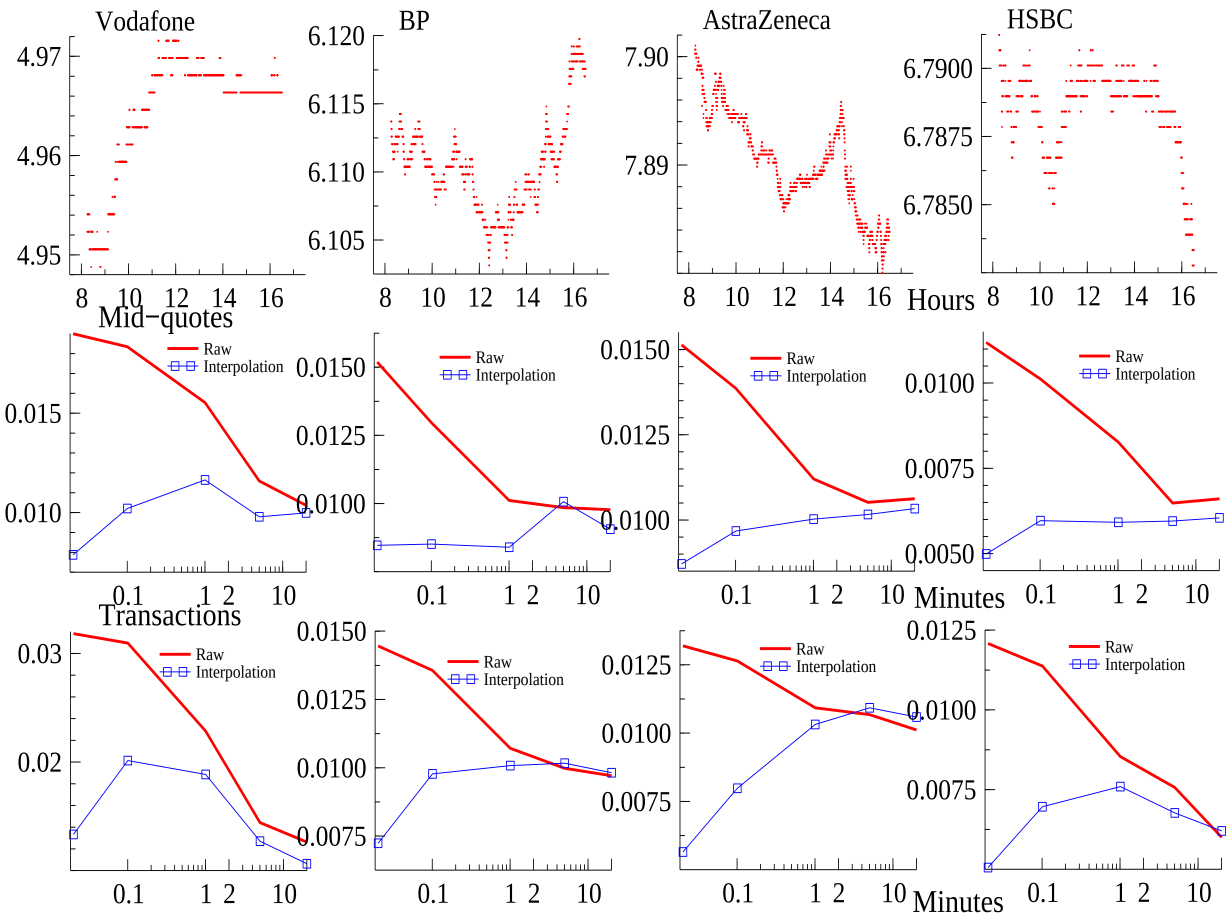

Figure 3: *LSE's electronic order book on the 2nd working day in January 2004. Top graphs: mid-quote log-price every 6 seconds, from 8.15am to 4.30pm. X-axis is in hours. Middle graphs: realised daily QVol computed using 0.015, 0.1, 1, 5 and 20 minute midpoint returns. X-axis is in minutes. Lower graphs: realised daily QVol computed using 0.1, 1, 5 and 20 minute transaction returns. Middle and lower graphs are computed using interpolation and the last tick method.*

tick size, the number of mid-point updates and transactions for each asset. It shows the usual result that as the tick size, as a percentage of the price increases, then the number of mid-quote price updates will tend to fall as larger tick sizes mean that there is a larger cost to impatience, that is jumping the queue in the order book by offering a better price than the best current and so updating the best quotes.

The middle row of Figure 3 shows the corresponding daily realised QVol, computed using 0.015, 0.1, 1, 5 and 20 minute intervals based on mid-quotes. These are related to the signature plots of Andersen, Bollerslev, Diebold, and Labys (2000). As the times of the mid-quotes fall irregularly in time, there is the question of how to approximate the

	Vodafone	BP	AstraZeneca	HSBC
		Daily volatility		
open-close	0.00968	0.00941	0.0143	0.00730
open-open	0.0159	0.0140	0.0140	0.00720
Correlation	0.861	0.851	0.912	0.731
Tick size	0.25	0.25	1.0	0.5
# of Mid-quotes per day	333	1,434	1,666	598
# of Transactions per day	3,018	2,995	2,233	2,264

Table 2: *Top part of table: Average daily volatility. Open is the mid-price at 8.15am, close is the mid-price at 4.30pm. Open-open looks at daily returns. Reported are the sample standard deviations of the returns over 20 days and sample correlation between the open-close and open-open daily returns. Bottom part of table: descriptive statistics about the size of the dataset.*

price at these time points. The Olsen method uses linear interpolation between the prices at the nearest observations before and after the correct time point. Another method is to use the last datapoint before the relevant time — the last tick or raw method (e.g. Wasserfallen and Zimmermann (1985)). Typically, the former leads to falls in realised QVol as δ falls, indeed in theory it converges to zero as $\delta \downarrow 0$ as its interpolated price process is of continuous bounded variation (Hansen and Lunde (2006)), while the latter increases modestly. The sensitivity to δ tends to be larger in cases where the tick size is large as a percentage of price and this is the case here. Overall we have the conclusion that the realised QVol does not change much when δ is 5 minutes or above and that it is more stable for interpolation than for last price. When we use smaller time intervals there are large dangers lurking. We will formally discuss the effect of market frictions in section 5.

The bottom row in Figure 3 shows the corresponding results for realised QVols computed using the transactions database. This ignores some very large over the counter trades. Realised QVol increases more strongly as δ falls when we use the last tick rather than mid-quote data. This is particularly the case for Vodafone, where bid/ask bounce has a large impact. Even the interpolation method has difficulties with transaction data. Overall, one gets the impression from this study that basing the analysis on mid-quote data is sound for the LSE data[8].

[8] A good alternative would be to carry out the entire analysis on either all the best bids or all the best

A fundamental difficulty with equity data is that the equity markets are only open for a fraction of the whole day and so it is quite possible that a large degree of their variation is at times when there is little data. This is certainly true for the U.K. equity markets which are closed during a high percentage of the time when U.S. markets are open. Table 2 gives daily volatility for open to close and open to open returns, as well as the correlation between the two return measures. It shows the open to close measures account for a high degree of the volatility in the prices, with high correlations between the two returns. The weakest relationship is for the Vodafone series, with the strongest for AstraZeneca. Hansen and Lunde (2005c) have studied how one can use high-frequency information to estimate the QV throughout the day, taking into account closed periods.

2.9.2 Epps effects

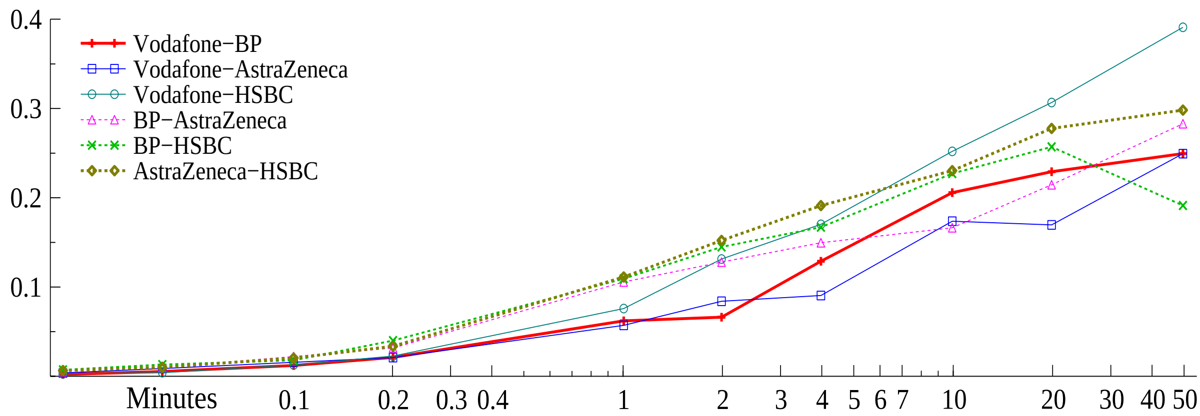

Figure 4: *LSE data during January 2004. Realised correlation computed daily, averaged over the month. Realised quantities are computed using data at the frequency on the x-axis.*

Market frictions affect the estimation of QVol, but if the asset is highly active, the tick size is small as a percentage of the price, δ is well above a minute and the mid-quote/interpolation method is used, then the effects are modest. The situation is much less rosy when we look at estimating quadratic covariations due to the so called Epps (1979) effect. This has been documented in very great detail by Sheppard (2005), who provides various theoretical explanations. We will come back to them in sections 3.8.3 and 5. For the moment it suffices to look at Figure 4 which shows the average daily realised

asks. This approach is used by Hansen and Lunde (2005c) and Large (2005).

correlation computed in January 2004 for the four stocks looked at above. Throughout prices are computed using mid-quotes and interpolation. The graph shows how this average varies with respect to δ. It trends downwards to zero as $\delta \downarrow 0$, with extremely low dependence measures for low values of δ. This is probably caused by the fact that asset prices tend not to simultaneously move due to non-synchronous trading and the differential rate at which information of different types is absorbed into individual stock prices.

3 Measurement error when $Y \in \mathcal{BSM}$

3.1 Infeasible asymptotics

Market frictions mean that it is not wise to use realised variation objects based on very small δ. This suggests refining our convergence in probability arguments to give a central limit theorem which may provide reasonable predictions about the behaviour of RV statistics for moderate values of δ, such as 5 or 10 minutes, where frictions are less likely to bite hard. Such CLTs will be the focus of attention in this section. At the end of the section, in addition, we will briefly discuss various alternative measures of variation, such as realised range, subsampling and kernel, which have recently been introduced to the literature. Finally we will also discuss how realised objects can contribute to the practical forecasting of volatility.

We will derive the central limit theorem for $[Y_\delta]_t$ which can then be discretised to produce the CLT for $\widehat{V}_i$. Univariate results will be presented, since this has less notational clutter. The results were developed in a series of papers by Jacod (1994), Jacod and Protter (1998), Barndorff-Nielsen and Shephard (2002) and Barndorff-Nielsen and Shephard (2004).

Theorem 1 *Suppose that $Y \in \mathcal{BSM}$ is one-dimensional and that (for all $t < \infty$) $\int_0^t a_u^2 \mathrm{d}u < \infty$, then as $\delta \downarrow 0$ so*

$$\delta^{-1/2}\left([Y_\delta]_t - [Y]_t\right) \to \sqrt{2}\int_0^t \sigma_u^2 \mathrm{d}B_u, \tag{14}$$

where B is a Brownian motion which is independent from Y and the convergence is in law stable as a process.

Proof. By Ito's lemma for continuous semimartingales

$$Y^2 = [Y] + 2Y \bullet Y,$$

then

$$\left(Y_{j\delta} - Y_{(j-1)\delta}\right)^2 = [Y]_{\delta j} - [Y]_{\delta(j-1)} + 2\int_{\delta(j-1)}^{\delta j} (Y_u - Y_{(j-1)\delta})\mathrm{d}Y_u.$$

This implies that

$$\begin{aligned} \delta^{-1/2}\left([Y_\delta]_t - [Y]_t\right) &= 2\delta^{-1/2}\sum_{j=1}^{\lfloor t/\delta \rfloor}\int_{\delta(j-1)}^{\delta j}(Y_u - Y_{(j-1)\delta})\mathrm{d}Y_u \\ &= 2\delta^{-1/2}\int_0^{\delta\lfloor t/\delta \rfloor}(Y_u - Y_{\delta\lfloor u/\delta \rfloor})\mathrm{d}Y_u. \end{aligned}$$

Jacod and Protter (1998, Theorem 5.5) show that for Y satisfying the conditions in Theorem 1 then[9]

$$\delta^{-1/2}\int_0^t (Y_u - Y_{\delta\lfloor u/\delta \rfloor})\mathrm{d}Y_u \to \frac{1}{\sqrt{2}}\int_0^t \sigma_u^2 \mathrm{d}B_u,$$

where $B \perp\!\!\!\perp Y$ and the convergence is in law stable as a process. This implies

$$\delta^{-1/2}\left([Y_\delta] - [Y]\right) \to \sqrt{2}\left(\sigma^2 \bullet B\right).$$

□

The most important point of this Theorem is that $B \perp\!\!\!\perp Y$. The appearance of the additional Brownian motion B is striking. This means that Theorem 1 implies, for a single t,

$$\delta^{-1/2}\left([Y_\delta]_t - [Y]_t\right) \xrightarrow{L} MN\left(0, 2\int_0^t \sigma_u^4 \mathrm{d}u\right), \quad (15)$$

where MN denotes a mixed Gaussian distribution. This result implies in particular that, for $i \neq j$,

$$\delta^{-1/2}\begin{pmatrix} \widehat{V}_i - V_i \\ \widehat{V}_j - V_j \end{pmatrix} \xrightarrow{L} MN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 2\begin{pmatrix} \int_{h(i-1)}^{hi}\sigma_u^4 \mathrm{d}u & 0 \\ 0 & \int_{h(j-1)}^{hj}\sigma_u^4 \mathrm{d}u \end{pmatrix}\right),$$

[9]At an intuitive level, if we ignore the drift then

$$\int_{\delta(j-1)}^{\delta j}(Y_u - Y_{(j-1)\delta})\mathrm{d}Y_u \simeq \sigma_{\delta(j-1)}^2 \int_{\delta(j-1)}^{\delta j}(W_u - W_{(j-1)\delta})\mathrm{d}W_u,$$

which is a martingale difference sequence in j with zero mean and conditional variance of $\frac{1}{2}\sigma_{\delta(j-1)}^2$. Applying a triangular martingale CLT one would expect this result, although formalising it requires a considerable number of additional steps.

so $\widehat{V}_i - V_i$ are asymptotically uncorrelated, so long as $\text{Var}\left(\widehat{V}_i - V_i\right) < \infty$, through time.

Barndorff-Nielsen and Shephard (2002) showed that Theorem 1 can be used in practice as the *integrated quarticity* $\int_0^t \sigma_u^4 \mathrm{d}u$ can be consistently estimated using $(1/3)\{Y_\delta\}_t^{[4]}$ where

$$\{Y_\delta\}_t^{[4]} = \delta^{-1} \sum_{j=1}^{\lfloor t/\delta \rfloor} \left(Y_{j\delta} - Y_{(j-1)\delta}\right)^4. \tag{16}$$

In particular then

$$\frac{\delta^{-1/2}\left([Y_\delta]_t - [Y]_t\right)}{\sqrt{\frac{2}{3}\{Y_\delta\}_t^{[4]}}} \xrightarrow{L} N(0,1). \tag{17}$$

This is a nonparametric result as it does not require us to specify the form of a or σ.

The multivariate version of (14) has that as $\delta \downarrow 0$ so

$$\delta^{-1/2}\left([Y_\delta]_{(kl)} - [Y]_{(kl)}\right) \to \frac{1}{\sqrt{2}} \sum_{b=1}^{q} \sum_{c=1}^{q} \left\{\left(\sigma_{(kb)}\sigma_{(cl)} + \sigma_{(lb)}\sigma_{(ck)}\right) \bullet B_{(bc)}\right\}, \quad k, l = 1, ..., q, \tag{18}$$

where B is a $q \times q$ matrix of independent Brownian motions, independent of Y and the convergence is in law stable as a process. In the mixed normal version of this result, the asymptotic covariance is a $q \times q \times q \times q$ array with elements

$$\left\{\int_0^t \{\Sigma_{(kk')u}\Sigma_{(ll')u} + \Sigma_{(kl')u}\Sigma_{(lk')u} + \Sigma_{(kl)u}\Sigma_{(k'l')u}\}\mathrm{d}u\right\}_{k,k',l,l'=1,...,q}. \tag{19}$$

Barndorff-Nielsen and Shephard (2004) showed how to use high frequency data to estimate this array of processes. We refer the reader to that paper, and also Mykland and Zhang (2005), for details.

3.2 Finite sample performance & the bootstrap

Our analysis of $[Y_\delta]_t - [Y]_t$ has been asymptotic as $\delta \downarrow 0$. Of course it is crucial to know if this analysis is informative for the kind of moderate values of δ we see in practice. A number of authors have studied the finite sample behaviour of the feasible limit theory given in (17) and a log-version, derived using the delta-rule

$$\frac{\delta^{-1/2}\left(\log[Y_\delta]_t - \log[Y]_t\right)}{\sqrt{\frac{2}{3}\frac{\{Y_\delta\}_t^{[4]}}{\left([Y_\delta]_t\right)^2}}} \xrightarrow{L} N(0,1). \tag{20}$$

We refer readers to Barndorff-Nielsen and Shephard (2005a), Meddahi (2002), Goncalves and Meddahi (2004), and Nielsen and Frederiksen (2005). The overall conclusion is that (17) is quite poorly sized, but that (20) performs pretty well. The asymptotic theory is challenged in cases where there are components in volatility which are very quickly mean reverting. In the multivariate case, Barndorff-Nielsen and Shephard (2004) studied the finite sample behaviour of realised regression and correlation statistics. They suggest various transformations which improve the finite sample behaviour of these statistics, including the use of the Fisher transformation for the realised correlation.

Goncalves and Meddahi (2004) have studied how one might try to bootstrap the realised daily QV estimator. Their overall conclusions are that the usual Edgeworth expansions, which justify the order improvement associated with the bootstrap, are not reliable guides to the finite sample behaviour of the statistics. However, it is possible to design bootstraps which provide very significant improvements over the limiting theory in (17). This seems an interesting avenue to follow up, particularly in the multivariate case.

3.3 Irregularly spaced data

Mykland and Zhang (2005) have recently generalised (14) to cover the case where prices are recorded at irregular time intervals. See also the related work of Barndorff-Nielsen and Shephard (2005b). Mykland and Zhang (2005) define a random sequence of times, independent of Y,[10] over the interval $t \in [0, T]$,

$$\mathcal{G}_n = \{0 = t_0 < t_1 < ... < t_n = T\},$$

then continue to have $\delta = T/n$, and define the estimated QV process

$$[Y_{\mathcal{G}_n}]_t = \sum_{j=1}^{t_j \leq t} \left(Y_{t_j} - Y_{t_{j-1}}\right)^2 \xrightarrow{p} [Y]_t.$$

[10]It is tempting to think of the t_j as the time of the j-th trade or quote. However, it is well know that the process generating the times of trades and price movements in tick time are not statistically independent (e.g. Engle and Russell (2005) and Rydberg and Shephard (2000)). This would seem to rule out the direct application of the methods we use here in tick time, suggesting care is needed in that case.

They show that as $n \to \infty$ so[11]

$$\begin{aligned} \delta^{-1/2}\left([Y_{\mathcal{G}_n}]_t - [Y]_t\right) &= 2\delta^{-1/2}\sum_{j=1}^{t_j \le t}\int_{t_{j-1}}^{t_j}(Y_u - Y_{t_{j-1}})\mathrm{d}Y_u \\ &\xrightarrow{L} MN\left(0, 2\int_0^t \left(\frac{\partial H_u^{\mathcal{G}}}{\partial u}\right)\sigma_u^4 \mathrm{d}u\right), \end{aligned}$$

where

$$H_t^{\mathcal{G}} = \lim_{n\to\infty} H_t^{\mathcal{G}_n}, \quad \text{where} \quad H_t^{\mathcal{G}_n} = \delta^{-1}\sum_{j=0}^{t_j \le t}(t_j - t_{j-1})^2,$$

and we have assumed that σ follows a diffusion and $H^{\mathcal{G}}$, which is a bit like a QV process but is scaled by δ^{-1}, is differentiable with respect to time. The $H^{\mathcal{G}}$ function is non-decreasing and runs quickly when the sampling is slower than normal. For regularly space data, $t_j = \delta j$ and so $H_t^{\mathcal{G}} = t$, which reproduces (15).

It is clear that

$$[Y_{\mathcal{G}_n}]_t^{[4]} = \delta^{-1}\sum_{j=1}^{t_j \le t}\left(Y_{t_j} - Y_{t_{j-1}}\right)^4 \xrightarrow{p} 3\int_0^t \left(\frac{\partial H_u^{\mathcal{G}}}{\partial u}\right)\sigma_u^4 \mathrm{d}u,$$

which implies the feasible distributional result in (17) and (20) also holds for irregularly spaced data, which was one of the results in Barndorff-Nielsen and Shephard (2005b).

3.4 Multiple grids

Zhang (2004) extended the above analysis to the simultaneous use of multiple grids — allowing the same $[Y]$ to be estimated using a variety of realised QV type objects based on slightly different spacing between observations. In our exposition we will work with $\mathcal{G}_n(i) = \{0 = t_0^i < t_1^i < ... < t_n^i = T\}$ for $i = 0, 1, 2, ..., I$ and $\delta_i = T/n_i$. Then define the i-th estimated QV process $[Y_{\mathcal{G}_{n(i)}}]_t = \sum_{j=1}^{t_j^i \le t}\left(Y_{t_j^i} - Y_{t_{j-1}^i}\right)^2$. Additionally we need a new cross term for the covariation between the time scales. The appropriate term is

$$H_t^{\mathcal{G}(i)\cup\mathcal{G}(k)} = \lim_{n\to\infty} H_t^{\mathcal{G}_n(i)\cup\mathcal{G}_n(k)}, \quad \text{where} \quad H_t^{\mathcal{G}_n(i)\cup\mathcal{G}_n(k)} = (\delta_i\delta_k)^{-1/2}\sum_{j=1}^{t_j^{i,k} \le t}\left(t_j^{i,k} - t_{j-1}^{i,k}\right)^2,$$

[11]At an intuitive level, if we ignore the drift then

$$\int_{t_{j-1}}^{t_j}(Y_u - Y_{t_{j-1}})\mathrm{d}Y_u \simeq \sigma_{t_{j-1}}^2 \int_{t_{j-1}}^{t_j}(W_u - W_{t_{j-1}})\mathrm{d}W_u,$$

which is a martingale difference sequence in j with zero mean and conditional variance of $\frac{1}{2}\sigma_{t_{j-1}}^2 (t_j - t_{j-1})$. Although this suggests the stated result, formalising it requires a considerable number of additional steps.

where $t_j^{i,k}$ comes from

$$\mathcal{G}_n(i) \cup \mathcal{G}_n(k) = \left\{ 0 = t_0^{i,k} < t_1^{i,k} < ... < t_{2n}^{i,k} = T \right\}, \quad i,k = 0,1,2,...,I.$$

Clearly, for all i,

$$\delta_i^{-1/2}\left([Y_{\mathcal{G}_n(i)}]_t - [Y]_t\right) = 2\delta_i^{-1/2} \sum_{j=1}^{t_j^i \leq t} \int_{t_{j-1}^i}^{t_j^i} (Y_u - Y_{t_{j-1}^i}) \mathrm{d}Y_u$$

so the scaled (by $\delta_i^{-1/2}$ and $\delta_k^{-1/2}$, respectively) asymptotic covariance matrix of $[Y_{\mathcal{G}_n(i)}]_t$ and $[Y_{\mathcal{G}_n(k)}]_t$ is

$$2 \begin{pmatrix} \int_0^t \left(\frac{\partial H_u^{\mathcal{G}(i)}}{\partial u} \right) \sigma_u^4 \mathrm{d}u & \bullet \\ \int_0^t \left(\frac{\partial H_u^{\mathcal{G}(i)\cup\mathcal{G}(k)}}{\partial u} \right) \sigma_u^4 \mathrm{d}u & \int_0^t \left(\frac{\partial H_u^{\mathcal{G}(k)}}{\partial u} \right) \sigma_u^4 \mathrm{d}u \end{pmatrix}.$$

Example 1 *Let $t_j^0 = \delta(j+\varepsilon)$, $t_j^1 = \delta(j+\eta)$ where $|\varepsilon - \eta| \in [0,1]$ are temporal offsets, then $H_t^{\mathcal{G}(0)} = H_t^{\mathcal{G}(1)} = t$,*

$$H_t^{\mathcal{G}(0)\cup\mathcal{G}(1)} = t\left((\eta-\varepsilon)^2 + (1 - |\eta - \varepsilon|)^2\right).$$

Thus

$$\delta^{-1/2} \begin{pmatrix} [Y_{\mathcal{G}_n(0)}]_t - [Y]_t \\ [Y_{\mathcal{G}_n(1)}]_t - [Y]_t \end{pmatrix} \xrightarrow{L} MN\left(0, 2\begin{pmatrix} 1 & \bullet \\ (\eta-\varepsilon)^2 + (1-|\eta-\varepsilon|)^2 & 1 \end{pmatrix} \int_0^t \sigma_u^4 \mathrm{d}u\right)$$

The correlation between the two measures is minimised at $1/2$ by setting $|\eta - \varepsilon| = 1/2$.

Example 1 extends naturally to when $t_j^k = \delta\left(j + \frac{k}{K+1}\right)$, $k = 0,1,2,...,K$, which allows many equally spaced realised QV like estimators to be defined based on returns measured over δ periods. The scaled asymptotic covariance of $[Y_{\mathcal{G}_n(i)}]_t$ and $[Y_{\mathcal{G}_n(k)}]_t$ is

$$2\left\{ \left(\frac{k-i}{K+1}\right)^2 + \left(1 - \left|\frac{k-i}{K+1}\right|\right)^2 \right\} \int_0^t \sigma_u^4 \mathrm{d}u.$$

If $K=1$ or $K=2$ then the correlation between the estimates is $1/2$ and $5/9$, respectively. As the sampling points become more dense the correlation quickly escalates which means that each new realised QV estimator brings out less and less additional information.

3.5 Subsampling

The multiple grid allows us to create a pooled grid estimator of QV — which is a special case of subsampling a statistic based on a random field, see for example the review of Politis, Romano, and Wolf (1999, Ch. 5). A simple example of this is

$$[Y_{\mathcal{G}_n^+(K)}]_t = \frac{1}{K+1}\sum_{i=0}^{K}[Y_{\mathcal{G}_n(i)}]_t, \tag{21}$$

which was mentioned in this context by Müller (1993) and Zhou (1996, p. 48). Clearly $[Y_{\mathcal{G}_n^+(K)}]_t \xrightarrow{p} [Y]_t$ as $\delta \downarrow 0$, while the properties of this estimator were first studied when $Y \in \mathcal{BSM}$ by Zhang, Mykland, and Aït-Sahalia (2005). Zhang (2004) also studies the properties of unequally weighted pooled estimators, while additional insights are provided by Aït-Sahalia, Mykland, and Zhang (2005).

Example 2 *Let* $t_j^k = \delta\left(j + \frac{k}{K+1}\right)$, $k = 0, 1, 2, ..., K$. *Then, for fixed* K *as* $\delta \downarrow 0$ *so*

$$\delta^{-1/2}\begin{pmatrix} [Y_{\mathcal{G}_n(0)}]_t - [Y]_t \\ [Y_{\mathcal{G}_n(1)}]_t - [Y]_t \end{pmatrix}$$

$$\xrightarrow{L} MN\left(0, \frac{2}{(K+1)^2}\sum_{i=0}^{K}\sum_{k=0}^{K}\left\{\left(\frac{k-i}{K+1}\right)^2 + \left(1 - \left|\frac{k-i}{K+1}\right|\right)^2\right\}\int_0^t \sigma_u^4 \mathrm{d}u\right)$$

This subsampler is based on a sample size $K+1$ *times the usual one but returns are still recorded over intervals of length* δ. *When* $K = 1$ *then the constant in front of integrated quarticity is* 1.5 *while when* $K = 2$ *it drops to* 1.4074. *The next terms in the sequence are* 1.3750, 1.3600, 1.3519 *and* 1.3469 *while it asymptotes to* 1.333, *a result due to Zhang, Mykland, and Aït-Sahalia (2005). Hence the gain from using the entire sample path of* Y *via multiple grids is modest and almost all the available gains occur by the time* K *reaches* 2. *However, we will see later that this subsampler has virtues when there are market frictions.*

3.6 Serial covariances

Suppose we define the notation $\mathcal{G}_\delta(\varepsilon, \eta) = \{\delta(\varepsilon + \eta), \delta(2\varepsilon + \eta), ...\}$, then the above theory implies that

$$\begin{pmatrix} \delta^{-1/2}\left([Y_{\mathcal{G}_n(2,0)}]_t - \int_0^t \sigma_u^2 \mathrm{d}u\right) \\ \delta^{-1/2}\left([Y_{\mathcal{G}_n(2,-1)}]_t - \int_0^t \sigma_u^2 \mathrm{d}u\right) \\ \delta^{-1/2}\left([Y_{\mathcal{G}_n(1,0)}]_t - \int_0^t \sigma_u^2 \mathrm{d}u\right) \end{pmatrix} \xrightarrow{L} MN\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 2 \end{pmatrix}\int_0^t \sigma_u^4 \mathrm{d}u\right).$$

Define the realised serial covariance as

$$\widehat{\gamma}_s(Y_\delta, X_\delta) = \sum_{j=1}^{\lfloor t/\delta \rfloor} \left(Y_{\delta j} - Y_{\delta(j-1)}\right)\left(X_{\delta(j-s)} - X_{\delta(j-s-1)}\right), \quad s = 0, 1, 2, ..., S,$$

and say $\widehat{\gamma}_{-s}(Y, X) = \widehat{\gamma}_s(Y, X)$ while $\widehat{\gamma}_s(Y_\delta) = \widehat{\gamma}_s(Y_\delta, Y_\delta)$. Derivatives on such objects have recently been studied by Carr and Lee (2003b). We have that

$$2\widehat{\gamma}_1(Y_\delta) = [Y_{\mathcal{G}_n(2,0)}]_t + [Y_{\mathcal{G}_n(2,-1)}]_t - 2[Y_{\mathcal{G}_n(1,0)}]_t + o_p(\delta^{1/2}).$$

Note that $\widehat{\gamma}_0(Y_\delta) = [Y_{\mathcal{G}_n(1,0)}]_t$. Then, clearly

$$\delta^{-1/2}\begin{pmatrix} \widehat{\gamma}_0(Y_\delta) - \int_0^t \sigma_u^2 \mathrm{d}u \\ \widehat{\gamma}_1(Y_\delta) \\ \vdots \\ \widehat{\gamma}_S(Y_\delta) \end{pmatrix} \xrightarrow{L} MN\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \int_0^t \sigma_u^4 \mathrm{d}u \right), \tag{22}$$

see Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004, Theorem 2). Consequently

$$\delta^{-1/2}\begin{pmatrix} \widehat{\gamma}_1(Y_\delta)/\widehat{\gamma}_0(Y_\delta) \\ \vdots \\ \widehat{\gamma}_S(Y_\delta)/\widehat{\gamma}_0(Y_\delta) \end{pmatrix} \xrightarrow{L} MN\left(0, I \frac{\int_0^t \sigma_u^4 \mathrm{d}u}{\left(\int_0^t \sigma_u^2 \mathrm{d}u\right)^2}\right),$$

which differs from the result of Bartlett (1946), inflating the usual standard errors as well as making inference multivariate mixed Gaussian. There is some shared characteristics with the familiar Eicker (1967) robust standard errors but the details are, of course, rather different.

3.7 Kernels

Following Bartlett (1950) and Eicker (1967), long run estimates of variances are often computed using kernels. We will see this idea may be helpful when there are market frictions and so we take some time discussing this here. It was introduced in this context by Zhou (1996) and Hansen and Lunde (2006), while a thorough discussion was given by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004, Theorem 2). A kernel takes on the form of

$$RV_w(Y) = w_0[Y_\delta] + 2\sum_{i=1}^{q} w_i \widehat{\gamma}_i(Y_\delta), \tag{23}$$

where the weights w_i are non-stochastic. It is clear from (22) that if the estimator is based on δ/K returns, so that it is compatible with (21), then

$$\left\{\frac{\delta}{K}\left(w_0^2 + 2\sum_{i=1}^{q} w_i^2\right)\right\}^{-1/2}\left(RV_w(Y_{\frac{\delta}{K}}) - w_0 \int_0^t \sigma_u^2 \mathrm{d}u\right) \xrightarrow{L} MN\left(0, 2\int_0^t \sigma_u^4 \mathrm{d}u\right). \tag{24}$$

In order for this method to be consistent for integrated variance as $q \to \infty$ we need that $w_0 = 1 + o(1)$ and $\sum_{i=1}^{q} w_i^2 / K = O(1)$ as a function of q.

Example 3 *The Bartlett kernel puts $w_0 = 1$ and $w_i = (q+1-i)/(q+1)$. When $q = 1$ then $w_1 = 1/2$ and the constant in front of integrated quarticity is 3, while when $q = 2$ then $w_1 = 2/3$, $w_2 = 1/3$ and the constant becomes $4 + 2/9$. For moderately large q this is well approximated by $4(q+1)/3$. This means that we need $q/K \to 0$ for this method to be consistent. This result appears in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004, Theorem 2).*

3.8 Other measures

3.8.1 Realised range

Suppose $Y = \sigma W$, a scaled Brownian motion, then

$$\mathrm{E}\left(\sup_{0 \le s \le t} Y_s^2\right) = \varphi_2 \sigma^2 t, \quad \text{where} \quad \varphi_r = \mathrm{E}\left(\sup_{0 \le s \le 1} |W_s|^r\right),$$

noting that $\varphi_2 = 4 \log 2$ and $\varphi_4 = 9\zeta(3)$, where ζ is the Riemann function. This observation led Parkinson (1980) to provide a simple estimator of σ^2 based on the highs and lows of asset prices. See also the work of Rogers and Satchell (1991), Alizadeh, Brandt, and Diebold (2002), Ghysels, Santa-Clara, and Valkanov (2004) and Brandt and Diebold (2004). One reason for the interest in ranges is the belief that they are quite informative and somewhat robust to market frictions. The problem with this analysis is that it does not extend readily when $Y \in \mathcal{BSM}$.

In independent work, Christensen and Podolskij (2005) and Martens and van Dijk (2005) have studied the realised range process. Christensen and Podolskij (2005) define the process as

$$\backslash Y \backslash_t = \mathrm{p-}\lim_{\delta \downarrow 0} \sum_{j=1}^{\lfloor t/\delta \rfloor} \sup_{s \in [(j-1)\delta, j\delta]} (Y_s - Y_{(j-1)\delta})^2 , \qquad (25)$$

which is estimated by the obvious realised version, written $\backslash Y_\delta \backslash_t$. Christensen and Podolskij (2005) have proved that if $Y \in \mathcal{BSM}$, then $\varphi_2^{-1} \backslash Y \backslash_t = \int_0^t \sigma_u^2 \mathrm{d}u$. Christensen and Podolskij (2005) also shows that under rather weak conditions

$$\delta^{-1/2}\left(\varphi_2^{-1} \backslash Y_\delta \backslash_t - [Y]_t\right) \xrightarrow{L} MN\left(0, \frac{\varphi_4 - \varphi_2^2}{\varphi_2^2} \int_0^t \sigma_u^4 \mathrm{d}u\right),$$

where $\varphi' = (\varphi_4 - \varphi_2^2)/\varphi_2^2 \simeq 0.4$. This shows that it is around five time as efficient as the usual realised QV estimator. Christensen and Podolskij (2005) suggest estimating integrated quarticity using

$$\delta^{-1}\varphi_4^{-1}\sum_{j=1}^{\lfloor t/\delta \rfloor}\sup_{s\in[(j-1)\delta,j\delta]}\left(Y_s - Y_{(j-1)\delta}\right)^4 ,$$

which means this limit theorem is feasible. Martens and van Dijk (2005) have also studied the properties of $\backslash Y_\delta \backslash_t$ using simulation and empirical work.

As far as we know no results are known about estimating $[Y]$ using ranges when there are jumps in Y, although it is relatively easy to see that a bipower type estimator could be defined using contiguous ranges which would robustly estimate $[Y^{ct}]$.

3.8.2 Discrete sine transformation

Curci and Corsi (2003) have argued that before computing realised QV we should prefilter the data using a discrete sine transformation to the returns in order to reduce the impact of market frictions. This is efficient when the data X is a Gaussian random walk Y plus independent Gaussian noise ε model, where we think of the noise as market frictions. The Curci and Corsi (2003) method is equivalent to calculating the realised QV process on the smoother $\mathrm{E}(Y|X;\theta)$, where θ are the estimated parameters indexing the Gaussian model. This type of approach was also advocated in Zhou (1996, p. 112).

3.8.3 Fourier and overlapping approaches

Motivated by the problem of irregularly spaced data, where the spacing is independent of Y, Malliavin and Mancino (2002) showed that if $Y \in \mathcal{BSM}$ then

$$\left[Y_J^l, Y_J^k\right]_{2\pi} = \pi^2\left[\frac{1}{J}\sum_{j=1}^{J}\left(a_j^l a_j^k + b_j^l b_j^k\right)\right] \xrightarrow{p} \left[Y^l, Y^k\right]_{2\pi}, \qquad (26)$$

as $J \to \infty$, where the Fourier coefficients of Y are

$$a_j^l = \frac{1}{\pi}\int_0^{2\pi}\cos(ju)\mathrm{d}Y_u^l, \quad b_j^l = \frac{1}{\pi}\int_0^{2\pi}\sin(ju)\mathrm{d}Y_u^l.$$

The Fourier coefficients are computed by, for example, integration by parts

$$\begin{aligned} a_j^l &= \frac{1}{\pi}\left(Y_{2\pi}^l - Y_0^l\right) + \frac{j}{\pi}\int_0^{2\pi} \sin(ju) Y_u^l \mathrm{d}u \\ &\simeq \frac{1}{\pi}\left(Y_{2\pi}^l - Y_0^l\right) + \frac{1}{\pi}\sum_{i=0}^{n-1} \left\{\cos\left(jt_i\right) - \cos\left(jt_{i+1}\right)\right\} Y_{t_i}^l, \\ b_j^l &\simeq \frac{1}{\pi}\sum_{i=0}^{n-1} \left\{\sin\left(jt_i\right) - \sin\left(jt_{i+1}\right)\right\} Y_{t_i}^l. \end{aligned}$$

This means that, in principle, one can use all the available data for all the series, even though prices for different assets appear at different points in time. Indeed each series has its Fourier coefficients computed separately, only performing the multivariate aspect of the analysis at step (26). A similar type of analysis could be based on wavelets, see Hog and Lunde (2003).

The performance of this Fourier estimator of QV is discussed by, for example, Barucci and Reno (2002b), Barucci and Reno (2002a), Kanatani (2004b), Precup and Iori (2005), Nielsen and Frederiksen (2005) and Kanatani (2004a) who carry out some extensive simulation and empirical studies of the procedure. Reno (2003) has used a multivariate version of this method to study the Epps effects, while Mancino and Reno (2005) use it to look at dynamic principle components. Kanatani (2004a, p. 22) has shown that in the univariate case the finite J Fourier estimator can be written as a kernel estimator (23). For regularly spaced data he derived the weight function, noting that as J increases, so each of these weights declined and so for fixed δ so $[Y_J]_{2\pi} \to [Y_\delta]_{2\pi}$. An important missing component in this analysis is any CLT for this estimator.

A related approach has been advocated by Corsi (2005, Ch. 5), Martens (2003) and Hayashi and Yoshida (2005, Definition 3.1). They study the estimator

$$\}Y^l, Y^m\{_t = \sum_{i=1}^{t_i \leq t} \sum_{j=1}^{t_j \leq t} \left(Y_{t_i}^l - Y_{t_{i-1}}^l\right)\left(Y_{t_j}^m - Y_{t_{j-1}}^m\right) I\left\{(t_{i-1}, t_i) \cap (t_{j-1}, t_j) \neq \varnothing\right\}. \tag{27}$$

This multiplies returns together whenever time intervals of the returns have any component which are overlapping. This artificially includes terms with components which are approximately uncorrelated (inflating the variance of the estimator), but it does not exclude any terms and so does not miss any of the contributions to quadratic covariation. They show under various assumptions that as the times of observations become denser over the interval from time 0 to time t, this estimator converges to the desired quadratic covariation quantity.

3.8.4 Generalised bipower variation

The realised bipower variation process suggests studying generic statistics of the form introduced by Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) and Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006)

$$Y_\delta(g,h)_t = \delta \sum_{j=1}^{\lfloor t/\delta \rfloor} g\left(\delta^{-1/2}\left(Y_{\delta(j-1)} - Y_{\delta(j-2)}\right)\right) h\left(\delta^{-1/2}\left(Y_{\delta j} - Y_{\delta(j-1)}\right)\right), \tag{28}$$

where the multivariate $Y \in \mathcal{BSM}$ and g, h are conformable matrices with elements which are continuous with at most polynomial growth in their arguments. Both QV and multivariate BPV can be cast in this form by the appropriate choice of g, h. Some of the choices of g, h will deliver statistics which will be robust to jumps.

Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) have shown that as $\delta \downarrow 0$ the probability limit of this process is always the generalised BPV process

$$\int_0^t \rho_{\sigma_u}(g)\rho_{\sigma_u}(h)\mathrm{d}u,$$

where the convergence is locally uniform, $\rho_\sigma(g) = \mathrm{E}g(X)$ and $X \sim N(0, \sigma\sigma')$. They also provide a central limit theorem for the generalised power variation estimator.

An example of the above framework which we have not covered yet is achieved by selecting $h(y) = \left|y^l\right|^r$ for $r > 0$ and $g(y) = 1$, then (28) becomes

$$\delta^{1-r/2} \sum_{j=1}^{\lfloor nt \rfloor} \left|Y^l_{\delta(j-1)} - Y^l_{\delta(j-2)}\right|^r, \tag{29}$$

which is called the realised r-th order power variation. When r is an integer it has been studied from a probabilistic viewpoint by Jacod (1994) while Barndorff-Nielsen and Shephard (2003) look at the econometrics of the case where $r > 0$. The increments of these types of high frequency volatility measures have been informally used in the financial econometrics literature for some time when $r = 1$, but until recently without a strong understanding of their properties. Examples of their use include Schwert (1990), Andersen and Bollerslev (1998b) and Andersen and Bollerslev (1997), while they have also been abstractly discussed by Shiryaev (1999, pp. 349–350) and Maheswaran and Sims (1993). Following the work by Barndorff-Nielsen and Shephard (2003), Ghysels, Santa-Clara, and Valkanov (2004) and Forsberg and Ghysels (2004) have successfully used realised power variation as an input into volatility forecasting competitions.

It is unclear how the greater flexibility over the choice of g, h will help econometricians in the future to learn about new features of volatility and jumps, perhaps robustly to market frictions. It would also be attractive if one could generalise (28) to allow g and h to be functions of the path of the prices, not just returns.

3.9 Non-parametric forecasting

3.9.1 Background

We saw in section 2.4 that if s is small then

$$\mathrm{Cov}\left(Y_{t+s} - Y_t|\mathcal{F}_t\right) \simeq \mathrm{E}\left([Y]_{t+s} - [Y]_t|\mathcal{F}_t\right).$$

This suggests:

1. estimating components of the increments of QV;

2. projecting these terms forward using a time series model.

This separates out the task of historical measurement of past volatility (step 1) from the problem of forecasting (step 2).

Suppose we wish to make a sequence of one-step or multi-step ahead predictions of $V_i = [Y]_{hi} - [Y]_{h(i-1)}$ using their proxies $\widehat{V}_i = [Y_\delta]_{hi} - [Y_\delta]_{h(i-1)}$, raw returns $y_i = Y_{hi} - Y_{h(i-1)}$ (to try to deal with leverage effects) and components $\widehat{B}_i = \{Y_\delta\}_{hi} - \{Y_\delta\}_{h(i-1)}$, where $i = 1, 2, ..., T$. For simplicity of exposition we set $h = 1$. This setup exploits the high frequency information set, but is somewhat robust to the presence of complicated intraday effects. Clearly if $Y \in \mathcal{BSM}$ then the CLT for realised QV implies that as $\delta \downarrow 0$, so long as the moments exist,

$$\mathrm{E}\left(V_i|\mathcal{F}_{i-1}\right) \simeq \mathrm{E}\left(\widehat{V}_i|\mathcal{F}_{i-1}\right) + o(\delta^{1/2}).$$

It is compelling to choose to use the coarser information set, so

$$\begin{aligned} & \operatorname{Cov}\left(Y_i - Y_{i-1}|\widehat{V}_{i-1}, \widehat{V}_{i-2}, ..., \widehat{V}_1, \widehat{B}_{i-1}, \widehat{B}_{i-2}, ..., \widehat{B}_1, y_{i-1}, ..., y_1\right) \\ \simeq & \operatorname{E}\left(V_i|\widehat{V}_{i-1}, \widehat{V}_{i-2}, ..., \widehat{V}_1, \widehat{B}_{i-1}, \widehat{B}_{i-2}, ..., \widehat{B}_1, y_{i-1}, ..., y_1\right) \\ \simeq & \operatorname{E}\left(\widehat{V}_i|\widehat{V}_{i-1}, \widehat{V}_{i-2}, ..., \widehat{V}_1, \widehat{B}_{i-1}, \widehat{B}_{i-2}, ..., \widehat{B}_1, y_{i-1}, ..., y_1\right). \end{aligned}$$

Forecasting can be carried out using structural or reduced form models. The simplest reduced form approach is to forecast $\widehat{V}_i$ using the past history $\widehat{V}_{i-1}, \widehat{V}_{i-2}, ..., \widehat{V}_1$, $y_{i-1}, y_{i-2}, ..., y_1$ and $\widehat{B}_{i-1}, \widehat{B}_{i-2}, ..., \widehat{B}_1$ based on standard forecasting methods such as autoregressions. The earliest modelling of this type that we know of was carried out by Rosenberg (1972) who regressed $\widehat{V}_i$ on $\widehat{V}_{i-1}$ to show, for the first time in the academic literature, that volatility was partially forecastable.

This approach to forecasting is convenient but potentially inefficient for it fails to use all the available high frequency data. In particular, for example, if $Y \in \mathcal{SV}$ then accurately modelled high frequency data may allow us to accurately estimate the spot covariance $\Sigma_{(i-1)h}$, which would be a more informative indicator than $\widehat{V}_{i-1}$. However, the results in Andersen, Bollerslev, and Meddahi (2004) are reassuring on that front. They indicate that if $Y \in \mathcal{SV}$ there is only a small loss in efficiency by forgoing $\Sigma_{(i-1)h}$ and using $\widehat{V}_{i-1}$ instead. Further, Ghysels, Santa-Clara, and Valkanov (2004) and Forsberg and Ghysels (2004) have forcefully argued that by additionally conditioning on low power variation statistics (29) very significant forecast gains can be achieved.

3.9.2 Illustration

In this subsection we will briefly illustrate some of these suggestions in the univariate case. Much more sophisticated studies are given in, for example, Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001), Andersen, Bollerslev, Diebold, and Labys (2003), Bollerslev, Kretschmer, Pigorsch, and Tauchen (2005) and Andersen, Bollerslev, and Meddahi (2004), who look at various functional forms, differing asset types and more involved dynamics. Ghysels, Santa-Clara, and Valkanov (2004) suggest an alternative method, using high frequency data but exploiting more sophisticated dynamics through so-called MIDAS regressions.

Const	Realised QV terms				Realised BPV terms				Summary measures	
Const	$\widehat{V}_{i-1}$	$\widehat{V}_{i-5}$	$\widehat{V}_{i-20}$	$\widehat{V}_{i-40}$	$\widehat{B}_{i-1}$	$\widehat{B}_{i-5}$	$\widehat{B}_{i-20}$	$\widehat{B}_{i-40}$	$\log L$	$Port_{49}$
0.503									-1751.42	4660
(0.010)										
0.170	0.413	0.153	0.061	0.030					-1393.41	199
(0.016)	(0.018)	(0.018)	(0.018)	(0.017)						
0.139	-0.137	-0.076	-0.017	0.116	0.713	0.270	0.091	-0.110	-1336.81	108
(0.017)	(0.059)	(0.059)	(0.058)	(0.058)	(0.075)	(0.074)	(0.074)	(0.073)		
0.139					0.551	0.180	0.071	0.027	-1342.03	122
(0.017)					(0.023)	(0.023)	(0.022)	(0.021)		

Table 3: *Prediction for 100 times returns on the DM/Dollar series. Dynamic regression, predicting future daily RV $\widehat{V}_i$ using lagged values and lagged values of estimated realised BPV terms $\widehat{B}_i$. Software used was PcGive. Subscripts denote the lag length in this table. Everything is computed using 10 minute returns. Figures in brackets are asymptotic standard errors. $Port_{49}$ denotes the Box-Ljung portmantau statistic computed with 49 lags, while log-L denotes the Gaussian likelihood.*

Table 3 gives a simple example of this approach for 100 times the returns on the DM/Dollar series. It shows the result of regressing $\widehat{V}_i$ on a constant, and simple lagged versions of $\widehat{V}_i$ and $\widehat{B}_i$. We dropped a priori the use of y_i as regressors for this exchange rate, where leverage effects are usually not thought to be important. The unusual spacing, using 1, 5, 20 and 40 lags, mimics the approach used by Corsi (2003) and Andersen, Bollerslev, and Diebold (2003). The results are quite striking. None of the models have satisfactory Box-Ljung portmanteau tests (this can be fixed by including a moving average error term in the model), but the inclusion of lagged information is massively significant. The lagged realised volatilities seem to do a reasonable job at soaking up the dependence in the data, but the effect of bipower variation is more important. This is in line with the results in Andersen, Bollerslev, and Diebold (2003) who first noted this effect. See also the work of Forsberg and Ghysels (2004) on the effect of inclusion of other power variation statistics in forecasting.

Table 4 shows some rather more sophisticated results. Here we model returns directly using a GARCH type model, but also include lagged explanatory variables in the conditional variance. This is in the spirit of the work of Engle and Gallo (2005). The results above the line show the homoskedastic fit and the improvement resulting from the standard GARCH(1,1) model. Below the line we include a variety of realised variables as explanatory variables; including longer lags of realised variables does not improve the fit.

	Realised terms		Standard GARCH terms		
Const	$\widehat{V}_{i-1}$	$\widehat{B}_{i-1}$	$(Y_{i-1}-Y_{i-2})^2$	h_{i-1}	$\log L$
0.504					-2636.59
(0.021)					
0.008			0.053	0.930	-2552.10
(0.003)			(0.010)	(0.013)	
0.017	-0.115	0.253	0.019	0.842	-2533.89
(0.009)	(0.039)	(0.076)	(0.019)	(0.052)	
0.011	0.085		0.015	0.876	-2537.49
(0.008)	(0.042)		(0.017)	(0.049)	
0.014		0.120	0.013	0.853	-2535.10
(0.009)		(0.058)	(0.019)	(0.055)	
0.019	-0.104	0.282		0.822	-2534.89
(0.010)	(0.074)	(0.116)		(0.062)	

Table 4: *Prediction for 100 times returns $Y_i - Y_{i-1}$ on the DM/Dollar series. GARCH type model of the conditional variance h_i of daily returns, using lagged squared returns $(Y_{i-1} - Y_{i-2})^2$, realised QV $\widehat{V}_{i-1}$, realised BPV $\widehat{B}_{i-1}$ and lagged conditional variance h_{i-1}. Throughout a Gaussian quasi-likelihood is used. Robust standard errors are reported. Carried out using PcGive.*

The best combination has a large coefficient on realised BPV and a negative coefficient on realised QV. This means when there is evidence for a jump then the impact of realised volatility is tempered, while when there is no sign of jump the realised variables are seen with full force. What is interesting from these results is that the realised effects are very much more important than the lagged daily returns. In effect the realised quantities have basically tested out the traditional GARCH model.

Overall this tiny empirical study confirms the results in the literature about the predictability of realised volatility. However, we have also seen that it is quite easy to outperform a simple autoregressive model for RV. We can see how useful bipower variation is and that taken together the realised quantities do provide a coherent way of empirically forecasting future volatility.

3.10 Parametric inference and forecasting

Throughout we have emphasised the non-parametric nature of the analysis. This is helpful due to the strong and complicated diurnal patterns we see in volatility. These effects tend also to be unstable through time and so are difficult to model parametrically. A literature which mostly avoids this problem is that on estimating parametric SV models from low frequency data. Much of this is reviewed in Shephard (2005, Ch. 1). Examples include the use of Markov chain Monte Carlo methods (e.g. Kim, Shephard, and Chib (1998)) and efficient method of moments (e.g. Chernov, Gallant, Ghysels, and Tauchen (2003)). Both approaches are computationally intensive and intricate to code. Simpler method of moment procedures (e.g. Andersen and Sørensen (1996)) have the difficulty that they are sensitive to the choice of moments and can be rather inefficient.

Recently various researchers have used the time series of realised daily QV to estimate parametric SV models. These models ignore the intraday effects and so are theoretically misspecified. Typically the researchers use various simple types of method of moments estimators, relying on the great increase in information available from realised statistics to overcome the inefficiency caused by the use of relatively crude statistical methods. The first papers to do this were Barndorff-Nielsen and Shephard (2002) and Bollerslev and Zhou (2002), who studied the first two dynamic moments of the time series $\widehat{V}_1$, $\widehat{V}_2$, ..., $\widehat{V}_T$ implied by various common volatility models and used these to estimate the parameters embedded within the SV models. More sophisticated approaches have been developed by Corradi and Distaso (2004) and Phillips and Yu (2005). Barndorff-Nielsen and Shephard (2002) also studied the use of these second order properties of the realised quantities to estimate V_1, V_2, ..., V_T from the time series of $\widehat{V}_1$, $\widehat{V}_2$, ..., $\widehat{V}_T$ using the Kalman filter. This exploited the asymptotic theory for the measurement error (15). See also the work of Meddahi (2002), Andersen, Bollerslev, and Meddahi (2004) and Andersen, Bollerslev, and Meddahi (2005).

3.11 Forecast evaluation

One of the main early uses of realised volatility was to provide a instrument for measuring the success for various volatility forecasting methods. Andersen and Bollerslev (1998a) studied the correlation between V_i or $\widehat{V}_i$ and h_i, the conditional variance from a GARCH model based on daily returns from time 1 up to time $i-1$. They used these results to argue that GARCH models were more successful than had been previously understood in the empirical finance literature. Hansen and Lunde (2005b) study a similar type of problem, but look at a wider class of forecasting models and carry out formal testing of the superiority of one modelling approach over another.

Hansen and Lunde (2005a) and Patton (2005) have focused on the delicate implications of the use of different loss functions to discriminate between competing forecasting models, where the object of the forecasting is $\operatorname{Cov}(Y_i - Y_{i-1}|\mathcal{F}_{i-1})$. They use $\widehat{V}_i$ to proxy this unobserved covariance. See also the related work of Koopman, Jungbacker, and Hol (2005).

4 Addition of jumps

4.1 Bipower variation

In this short section we will review some material which non-parametrically identifies the contribution of jumps to the variation of asset prices. A focus will be on using this method for testing for jumps from discrete data. We will also discuss some work by Cecilia Mancini which provides an alternative to BPV for splitting up QV into its continuous and discontinuous components.

Recall $\mu_1^{-2}\{Y\}_t = \int_0^t \Sigma_u \mathrm{d}u$ when Y is a $\mathcal{BSM}$ plus jump process given in (7). The BPV process is consistently estimated by the $p \times p$ matrix realised BPV process $\{Y_\delta\}$, defined in (10). This means that we can, in theory, consistently estimate $[Y^{ct}]$ and $[Y^d]$ by $\mu_1^{-2}\{Y_\delta\}$ and $[Y_\delta] - \mu_1^{-2}\{Y_\delta\}$, respectively.

One potential use of $\{Y_\delta\}$ is to test for the hypothesis that a set of data is consistent with a null hypothesis of continuous sample paths. We can do this by asking if $[Y_\delta]_t - \mu_1^{-2}\{Y_\delta\}_t$ is statistically significantly bigger than zero — an approach introduced by Barndorff-Nielsen and Shephard (2006). This demands a distribution theory for realised BPV objects, calculated under the null that $Y \in \mathcal{BSM}$ with $\sigma > 0$.

Building on the earlier CLT of Barndorff-Nielsen and Shephard (2006), Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) have established a CLT which covers this situation when $Y \in \mathcal{BSM}$. We will only present the univariate result, which has that as $\delta \downarrow 0$ so

$$\delta^{-1/2}\left(\{Y_\delta\}_t - \{Y\}_t\right) \to \mu_1^2 \sqrt{(2+\vartheta)} \int_0^t \sigma_u^2 \mathrm{d}B_u, \tag{30}$$

where $B \perp\!\!\!\perp Y$, the convergence is in law stable as a process and

$$\vartheta = \left(\pi^2/4\right) + \pi - 5 \simeq 0.6090.$$

This result, unlike Theorem 1, has some quite technical conditions associated with it in order to control the degree to which the volatility process can jump; however we will not discuss those issues here. Extending the result to cover the joint distribution of the estimators of the QV and the BPV processes, they showed that

$$\delta^{-1/2}\left(\begin{array}{c}\mu_1^{-2}\{Y_\delta\}_t-\mu_1^{-2}\{Y\}_t\\ [Y_\delta]_t-[Y]_t\end{array}\right)\xrightarrow{L} MN\left(\left(\begin{array}{c}0\\0\end{array}\right),\left(\begin{array}{cc}(2+\vartheta) & 2\\ 2 & 2\end{array}\right)\int_0^t\sigma_u^4\mathrm{d}u\right),$$

a Hausman (1978) type result as the estimator of the QV process is, of course, fully asymptotically efficient when $Y\in\mathcal{BSM}$. Consequently

$$\frac{\delta^{-1/2}\left([Y_\delta]_t-\mu_1^{-2}\{Y_\delta\}_t\right)}{\sqrt{\vartheta\int_0^t\sigma_u^4\mathrm{d}u}}\xrightarrow{L} N(0,1), \qquad (31)$$

which can be used as the basis of a test of the null of no jumps.

4.2 Multipower variation

The "standard" estimator of integrated quarticity, given in (16), is not robust to jumps. One way of overcoming this problem is to use a multipower variation (MPV) measure — introduced by Barndorff-Nielsen and Shephard (2006). This is defined as

$$\{Y\}_t^{[\mathbf{r}]}=\mathrm{p}-\lim_{\delta\downarrow 0}\delta^{(1-r_+/2)}\sum_{j=1}^{\lfloor t/\delta\rfloor}\left\{\prod_{i=1}^{I}\left|Y_{\delta(j-i)}-Y_{\delta(j-1-i)}\right|^{r_i}\right\},$$

where $r_i>0$, $r=(r_1,r_2,...,r_I)'$ for all i and $r_+=\sum_{i=1}^{I}r_i$. The usual BPV process is the special case $\{Y\}_t=\{Y\}_t^{[1,1]}$.

If Y obeys (7) and $r_i<2$ then

$$\{Y\}_t^{[\mathbf{r}]}=\left(\prod_{i=1}^{I}\mu_{r_i}\right)\int_0^t\sigma_u^{r_+}\mathrm{d}u,$$

This process is approximated by the estimated MPV process

$$\{Y_\delta\}_t^{[\mathbf{r}]}=\delta^{(1-r_+/2)}\sum_{j=1}^{\lfloor t/\delta\rfloor}\left\{\prod_{i=1}^{I}\left|Y_{\delta(j-i)}-Y_{\delta(j-1-i)}\right|^{r_i}\right\}.$$

In particular the scaled realised tri and quadpower variation,

$$\mu_1^{-4}\{Y_\delta\}_t^{[1,1,1,1]}\quad\text{and}\quad\mu_{4/3}^{-3}\{Y_\delta\}_t^{[4/3,4/3,4/3]},$$

respectively, both estimate $\int_0^t \sigma_u^4 du$ consistently in the presence of jumps. Hence either of these objects can be used to replace the integrated quarticity in (31), so producing a non-parametric test for the presence of jumps in the interval $[0, t]$. The test is conditionally consistent, meaning if there is a jump, it will detected and has asymptotically the correct size. Extensive small sample studies are reported in Huang and Tauchen (2005), who favour ratio versions of the statistic like

$$\frac{\delta^{-1/2}\left(\frac{\mu_1^{-2}\{Y_\delta\}_t}{[Y_\delta]_t} - 1\right)}{\sqrt{\vartheta \frac{\{Y_\delta\}_t^{[1,1,1,1]}}{\left(\{Y_\delta\}_t\right)^2}}} \xrightarrow{L} N(0,1),$$

which has pretty reasonable finite sample properties. They also show that this test tends to under reject the null of no jumps in the presence of some forms of market frictions.

It is clearly possible to carry out jump testing on separate days or weeks. Such tests are asymptotically independent over these non-overlapping periods under the null hypothesis.

To illustrate this methodology we will apply the jump test to the DM/Dollar rate, asking if the hypothesis of a continuous sample path is consistent with the data we have. Our focus will mostly be on Friday January 15th 1988, although we will also give results for neighbouring days to provide some context. In Figure 5 we plot 100 times the change during the week of the discretised Y_δ, so a one unit uptick represents a 1% change, for a variety of values of $n = 1/\delta$, as well as giving the ratio jump statistics $\widehat{B}_i/\widehat{V}_i$ with their corresponding 99% critical values.

In Figure 5 there is a large uptick in the D-mark against the Dollar, with a movement of nearly two percent in a five minute period. This occurred on the Friday and was a response to the news of a large fall in the U.S. balance of payment deficit, which led to a large strengthening of the Dollar. The data for January 15th had a large $\widehat{V}_i$ but a much smaller $\widehat{B}_i$. Hence the statistics are attributing a large component of $\widehat{V}_i$ to the jump, with the adjusted ratio statistic being larger than the corresponding 99% critical value. When δ is large the statistic is on the borderline of being significant, while the situation becomes much clearer as δ becomes small. This illustration is typical of results presented in Barndorff-Nielsen and Shephard (2006) which showed that many of the large jumps in this exchange rate correspond to macroeconomic news announcements. This is consistent with

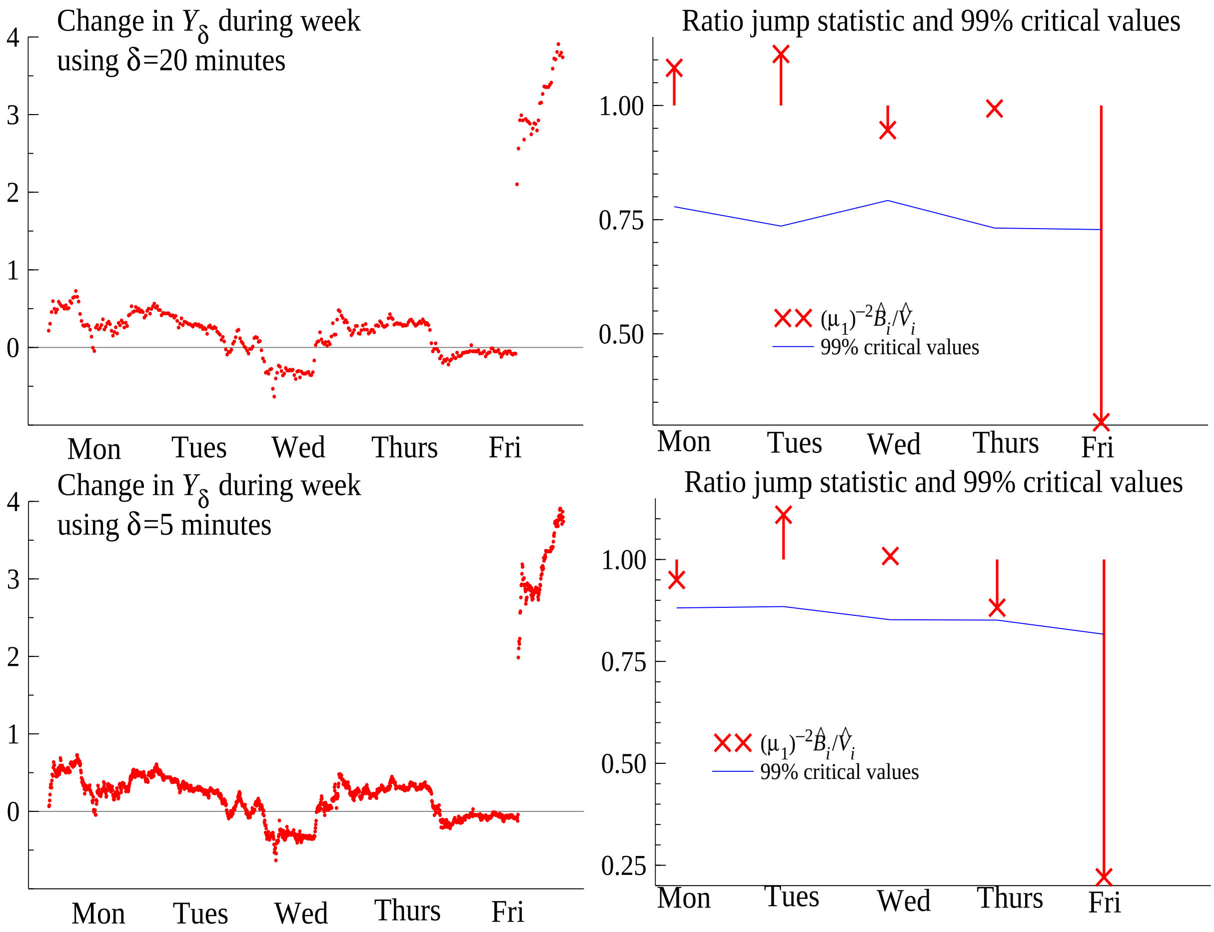

Figure 5: *Left hand side: change in* Y_δ *during a week, centred at* 0 *on Monday 11th January and running until Friday of that week. Drawn every 20 and 5 minutes. An up tick of 1 indicates strengthening of the Dollar by 1%. Right hand side shows an index plot of* $\widehat{B}_i/\widehat{V}_i$*, which should be around 1 if there are no jumps. Test is one sided, with criticial values also drawn as a line.*

the recent economics literature documenting significant intraday announcement effects, e.g. Andersen, Bollerslev, Diebold, and Vega (2003).

4.3 Grids

It is clear that the martingale based CLT for irregularly spaced data for the estimator of the QV process can be extended to cover the BPV case. We define

$$\{Y_{\mathcal{G}_n}\}_t = \sum_{j=1}^{t_j \leq t} \left|Y_{t_{j-1}} - Y_{t_{j-2}}\right| \left|Y_{t_j} - Y_{t_{j-1}}\right| \xrightarrow{p} \{Y\}_t.$$

Using the same notation as before, we would expect the following result to hold, due to the fact that $H^{\mathcal{G}}$ is assumed to be continuous,

$$\delta^{-1/2}\begin{pmatrix} \mu_1^{-2}\{Y_{\mathcal{G}_n}\}_t - \mu_1^{-2}\{Y\}_t \\ [Y_\delta]_t - [Y]_t \end{pmatrix} \xrightarrow{L} MN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} (2+\vartheta) & 2 \\ 2 & 2 \end{pmatrix} \int_0^t \left(\frac{\partial H_u^{\mathcal{G}}}{\partial u}\right)\sigma_u^4 \mathrm{d}u\right).$$

The integrated moderated quarticity can be estimated using $\mu_1^{-4}\{Y_\delta\}_t^{[1,1,1,1]}$, or a grid version, which again implies that the usual feasible CLT continues to hold for irregularly spaced data. This is the expected result from the analysis of power variation provided by Barndorff-Nielsen and Shephard (2005b).

Potentially there are modest efficiency gains to be had by computing the estimators of BPV on multiple grids and then averaging them. The extension along these lines is straightforward and will not be detailed here.

4.4 Infinite activity jumps

The probability limit of realised BPV is robust to finite activity jumps. A natural question to ask is: (i) is the CLT also robust to jumps, (ii) is the probability limit also unaffected by infinite activity jumps, that is jump processes with an infinite number of jumps in any finite period of time. Both issues are studied by Barndorff-Nielsen, Shephard, and Winkel (2004) in the case where the jumps are of Lévy type, while Woerner (2004) looks at the probability limit for more general jump processes.

Barndorff-Nielsen, Shephard, and Winkel (2004) find that the CLT for BPV is affected by finite activity jumps, but this is not true of tripower and high order measures of variation. The reason for the robustness of tripower results is quite technical and we will not discuss it here. However, it potentially means that inference under the assumption of jumps can be carried out using tripower variation, which seems an exciting possibility. Both Barndorff-Nielsen, Shephard, and Winkel (2004) and Woerner (2004) give results which prove that the probability limit of realised BPV is unaffected by some types of infinite activity jump processes. More work is needed on this topic to make these result definitive. It is somewhat related to the parametric study of Aït-Sahalia (2004). He shows that maximum likelihood estimation can disentangle a homoskedastic diffusive component from a purely discontinuous infinite activity Lévy component of prices. Outside the likelihood framework, the paper also studies the optimal combinations of moment functions for the generalized method of moment estimation of homoskedastic jump-diffusions. Further insights can be found by looking at likelihood inference for Lévy processes, which is studied by Aït-Sahalia and Jacod (2005a) and Aït-Sahalia and Jacod (2005b).

4.5 Testing the null of no continuous component

In some stimulating recent papers, Carr, Geman, Madan, and Yor (2003) and Carr and Wu (2004), have argued that it is attractive to build SV models out of pure jump processes, with no Brownian aspect. It is clearly important to be able to test this hypothesis, seeing if pure discreteness is consistent with observed prices.

Barndorff-Nielsen, Shephard, and Winkel (2004) showed that

$$\delta^{-1/2}\left(\{Y_\delta\}_t^{[2/3,2/3,2/3]} - [Y^{ct}]_t\right)$$

has a mixed Gaussian limit and is robust to jumps. But this result is only valid if $\sigma > 0$, which rules out its use for testing for pure discreteness. However, we can artificially add a scaled Brownian motion, $U = \sigma B$, to the observed price process and then test if

$$\delta^{-1/2}\left(\{Y_\delta + U_\delta\}_t^{[2/3,2/3,2/3]} - \sigma^2 t\right)$$

is statistically significantly greater than zero. In principle this would be a consistent non-parametric test of the maintained hypothesis of Peter Carr and his coauthors.

4.6 Alternative methods for identifying jumps

Mancini (2001), Mancini (2004) and Mancini (2003) has developed robust estimators of $[Y^{ct}]$ in the presence of finite activity jumps. Her approach is to use truncation

$$\sum_{j=1}^{\lfloor t/\delta \rfloor} \left(Y_{j\delta} - Y_{(j-1)\delta}\right)^2 I(\left|Y_{j\delta} - Y_{(j-1)\delta}\right| < r_\delta), \qquad (32)$$

where $I(.)$ is an indicator function. The crucial function r_δ has to have the property that $\sqrt{\delta \log \delta^{-1}} r_\delta^{-1} \downarrow 0$. It is motivated by the modulus of continuity of Brownian motion paths that almost surely

$$\lim_{\delta \downarrow 0} \sup_{\substack{0 \le s,t \le T \\ |t-s|<\delta}} \frac{|W_s - W_t|}{\sqrt{2\delta \log \delta^{-1}}} = 1.$$

This is an elegant theory, which works when $Y \in \mathcal{BSM}$. It is not prescriptive about the tuning function r_δ, which is an advantage and a drawback. Given the threshold in (32) is universal, this method will throw out more returns as jumps during a high volatility period than during a low volatility period.

Aït-Sahalia and Jacod (2005b, Section 7 onwards) provides additional insights into these types of truncation estimators in the case where Y is scaled Brownian motion plus a homogeneous pure jump process. They develop a two-step procedure, which automatically selects the level of truncation. Their analysis is broader still, providing additional insights into a range of power variation type objects.

5 Mitigating market frictions

The semimartingale model of the frictionless, arbitrage free market is a fiction. When we use high frequency data to perform inference on either transaction or quote data then various market frictions can become important. O'Hara (1995), Engle (2000), Hasbrouck (2003) and Engle and Russell (2005) review the detailed modelling of these effects. Inevitably such modelling is quite complicated.

With the exception of subsection 2.9, we have so far mostly ignored frictions by thinking of δ as being only moderately small. This is ad hoc and it is wise to try to more formally identify the impact of frictions. In this context the first econometric work was carried out by Fang (1996) and Andersen, Bollerslev, Diebold, and Labys (2000) who used so-called signature plots to assess the degree of bias caused by frictions using a variety of values of δ. The signature plots we draw show the square root of the time series average of estimators of V_i computed over many days, plotting this against δ. If the log-price process was a pure martingale then we would expect the plot to have roughly horizontal lines.

Hansen and Lunde (2006) have reviewed the literature on the effect of market frictions on realised QV statistics. Their broad conclusions are that for thickly traded stocks: (i) for returns measured over 10 to 20 minute returns using mid-quotes the central limit theories based on no noise give good approximations to the reality, (ii) for returns measure over 1 to 10 minutes, noise becomes important but it is empirically realistic to view the noise as independent of Y, (iii) for higher frequency data the situation is much more complicated.

Econometricians have recently started to try to use higher frequency data to estimate $[Y]$, taking into account the effect of market frictions. All the work we have seen assumes independence of Y with the frictions. Important approaches are (a) subsampling by Zhou (1996), Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2004) and Aït-Sahalia, Mykland, and Zhang (2005), (b) point process by Large (2005), (c) kernels by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004). It is unclear how this rapidly evolving literature will settle in the next few years. Particularly important contributions need to be made in the multivariate case where the effects of market frictions are most readily felt.

6 Conclusions

This paper has reviewed the literature on the measurement and forecasting of uncertainty through quadratic variation type objects. The econometrics of this has focused on realised objects, estimating QV and its components. Such an approach has been shown to provide a leap forward in our understanding of time varying volatility and jumps, which are crucial in asset allocation, derivative pricing and risk assessment. A drawback with these types of methods is the potential for market frictions to complicate the analysis. Recent research has been trying to address this issue and has introduced various innovative methods. There is still much work to be carried through in that area.

7 Acknowledgements

We are grateful to Tim Bollerslev, Eric Ghysels, Peter Hansen, Jean Jacod, Dmitry Kulikov, Jeremy Large, Asger Lunde, Andrew Patton, Mark Podolskij, Kevin Sheppard and Jun Yu for comments on an earlier draft. Talks based on this paper were given in 2005 as the Hiemstra Lecture at the 13th Annual conference of the Society of Non-linear Dynamics and Econometrics in London, the keynote address at the 3rd Nordic Econometric Meeting in Helsinki and as a Special Invited Lecture at the 25th European Meeting of Statisticians in Oslo. Ole Barndorff-Nielsen's work is supported by CAF (`www.caf.dk`), which is funded by the Danish Social Science Research Council. Neil Shephard's research is supported by the UK's ESRC through the grant "High frequency financial econometrics based upon power variation." The calculations made for this paper were carried out using PcGive of Doornik and Hendry (2005) and software written by the authors using the Ox language of Doornik (2001).

References

Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics 74*, 487–528.

Aït-Sahalia, Y. and J. Jacod (2005a). Fisher's information for discretely sampled Lévy processes. Unpublished paper: Department of Economics, Princeton University.

Aït-Sahalia, Y. and J. Jacod (2005b). Volatility estimators for discretely sampled Lévy processes. Unpublished paper: Department of Economics, Princeton University.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). Ultra high frequency volatility estimation with dependent microstructure noise. Unpublished paper: Department of Economics, Princeton University.

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 (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*, 115–158.

Andersen, T. G. and T. Bollerslev (1998a). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andersen, T. G. and T. Bollerslev (1998b). Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance 53*, 219–265.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2003). Some like it smooth, and some like it rough: untangling continuous and jump components in measuring, modeling and forecasting asset return volatility. Unpublished paper: Economics Dept, Duke University.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000). Great realizations. *Risk 13*, 105–108.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55. Correction published in 2003, volume 98, page 501.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review 93*, 38–62.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2004). Analytic evaluation of volatility forecasts. *International Economic Review 45*, 1079–1110.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2005). Correcting the errors: A note on volatility forecast evaluation based on high-frequency data and realized volatilities. *Econometrica 73*, 279–296.

Andersen, T. G. and B. Sørensen (1996). GMM estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Business and Economic Statistics 14*, 328–352.

Back, K. (1991). Asset pricing for general processes. *Journal of Mathematical Economics 20*, 371–395.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2005). A central limit theorem for realised power and bipower variations of continuous semimartingales. In Y. Kabanov and R. Lipster (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*. Springer. Forthcoming.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for realised bipower variation in econometrics. *Econometric Theory*. Forthcoming.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2004). Regular and modified kernel-based estimators of integrated variance: the case with independent noise. Unpublished paper: Nuffield College, Oxford.

Barndorff-Nielsen, O. E. and N. Shephard (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B 63*, 167–241.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2003). Realised power variation and stochastic volatility. *Bernoulli 9*, 243–265. Correction published in pages 1109–1111.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica 72*, 885–925.

Barndorff-Nielsen, O. E. and N. Shephard (2005a). How accurate is the asymptotic approximation to the distribution of realised volatility? In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg*, pp. 306–331. Cambridge: Cambridge University Press.

Barndorff-Nielsen, O. E. and N. Shephard (2005b). Power variation and time change. *Theory of Probability and Its Applications*. Forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*. Forthcoming.

Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel (2004). Limit theorems for multipower variation in the presence of jumps in financial econometrics. Unpublished paper: Nuffield College, Oxford.

Bartlett, M. S. (1946). On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, Supplement 8*, 27–41.

Bartlett, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika 37*, 1–16.

Barucci, E. and R. Reno (2002a). On measuring volatility and the GARCH forecasting performance. *Journal of International Financial Markets, Institutions and Money 12*, 182–200.

Barucci, E. and R. Reno (2002b). On measuring volatility of diffusion processes with high frequency data. *Economic Letters 74*, 371–378.

Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen (2005). The dynamics of bipower variation, realized volatility and returns. Unpublished paper: Department of Economics, Duke University.

Bollerslev, T. and H. Zhou (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics 109*, 33–65.

Brandt, M. W. and F. X. Diebold (2004). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business*. Forthcoming.

Branger, N. and C. Schlag (2005). An economic motivation for variance contracts. Unpublished paper: Faculty of Economics and Business Administration, Goethe University.

Brockhaus, O. and D. Long (1999). Volatility swaps made simple. *Risk 2*, 92–95.

Calvet, L. and A. Fisher (2002). Multifractality in asset returns: theory and evidence. *Review of Economics and Statistics 84*, 381–406.

Carr, P., H. Geman, D. B. Madan, and M. Yor (2003). Stochastic volatility for Lévy processes. *Mathematical Finance 13*, 345–382.

Carr, P., H. Geman, D. B. Madan, and M. Yor (2005). Pricing options on realized variance. *Finance and Stochastics*. Forthcoming.

Carr, P. and R. Lee (2003a). Robust replication of volatility derivatives. Unpublished paper: Courant Institute, NYU.

Carr, P. and R. Lee (2003b). Trading autocorrelation. Unpublished paper: Courant Institute, NYU.

Carr, P. and K. Lewis (2004). Corridor variance swaps. *Risk*, 67–72.

Carr, P. and D. B. Madan (1998). Towards a theory of volatility trading. In R. Jarrow (Ed.), *Volatility*, pp. 417–427. Risk Publications.

Carr, P. and L. Wu (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics 71*, 113–141.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative models of stock price dynamics. *Journal of Econometrics 116*, 225–257.

Christensen, K. and M. Podolskij (2005). Asymptotic theory for range-based estimation of integrated volatility of a continuous semi-martingale. Unpublished paper: Aarhus School of Business.

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*, 291–323.

Corradi, V. and W. Distaso (2004). Specification tests for daily integrated volatility, in the presence of possible jumps. Unpublished paper: Queen Mary College, London.

Corsi, F. (2003). A simple long memory model of realized volatility. Unpublished paper: University of Southern Switzerland.

Corsi, F. (2005). Measuring and modelling realized volatility: from tick-by-tick to long memory. Ph.D. thesis, Department of Economics, University of Lugano.

Curci, G. and F. Corsi (2003). A discrete sine transformation approach to realized volatility measurement. Unpublished paper.

Dacorogna, M. M., R. Gencay, U. A. Müller, R. B. Olsen, and O. V. Pictet (2001). *An Introduction to High-Frequency Finance*. San Diego: Academic Press.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou (1999). A guide to volatility and variance swaps. *Journal of Derivatives 6*, 9–32.

Doob, J. L. (1953). *Stochastic Processes*. New York: John Wiley and Sons.

Doornik, J. A. (2001). *Ox: Object Oriented Matrix Programming, 3.0*. London: Timberlake Consultants Press.

Doornik, J. A. and D. F. Hendry (2005). *PC Give, Version 10.4*. London: Timberlake Consultants Press.

Eicker, F. (1967). Limit theorems for regressions with unequal and dependent errors. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1*, pp. 59–82. Berkeley: University of California.

Engle, R. F. (2000). The econometrics of ultra-high frequency data. *Econometrica 68*, 1–22.

Engle, R. F. and J. P. Gallo (2005). A multiple indicator model for volatility using intra daily data. *Journal of Econometrics*. Forthcoming.

Engle, R. F. and J. R. Russell (1998). Forecasting transaction rates: the autoregressive conditional duration model. *Econometrica 66*, 1127–1162.

Engle, R. F. and J. R. Russell (2005). Analysis of high frequency data. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland. Forthcoming.

Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association 74*, 291–296.

Fang, Y. (1996). Volatility modeling and estimation of high-frequency data with Gaussian noise. Unpublished Ph.D. thesis, Sloan School of Management, MIT.

Forsberg, L. and E. Ghysels (2004). Why do absolute returns predict volatility so well. Unpublished paper: Economics Department, UNC, Chapel Hill.

Foster, D. P. and D. B. Nelson (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica 64*, 139–174.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*, 3–29.

Genon-Catalot, V., C. Larédo, and D. Picard (1992). Non-parametric estimation of the diffusion coefficient by wavelet methods. *Scandinavian Journal of Statistics 19*, 317–335.

Ghysels, E., A. C. Harvey, and E. Renault (1996). Stochastic volatility. In C. R. Rao and G. S. Maddala (Eds.), *Statistical Methods in Finance*, pp. 119–191. Amsterdam: North-Holland.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*. Forthcoming.

Goncalves, S. and N. Meddahi (2004). Bootstrapping realized volatility. Unpublished paper, CIRANO, Montreal.

Hansen, P. R. and A. Lunde (2005a). Consistent ranking of volatility models. *Journal of Econometrics*. Forthcoming.

Hansen, P. R. and A. Lunde (2005b). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*. Forthcoming.

Hansen, P. R. and A. Lunde (2005c). A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics 3*, 525–554.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise (with discussion). *Journal of Business and Economic Statistics 24*. Forthcoming.

Hasbrouck, J. (2003). *Empirical Market Microstructure: Economic and Statistical Perspectives on the Dynamics of Trade in Security Markets*. Unpublished manuscript, Department of Finance, Stern Business School, NYU.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*, 1251–1271.

Hayashi, T. and N. Yoshida (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli 11*, 359–379.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility, with applications to bond and currency options. *Review of Financial Studies 6*, 327–343.

Hog, E. and A. Lunde (2003). Wavelet estimation of integrated volatility. Unpublished paper: Aarhus School of Business.

Howison, S. D., A. Rafailidis, and H. O. Rasmussen (2004). On the pricing and hedging of volatility derivatives. *Applied Mathematical Finance 11*, 317–346.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variation. *Journal of Financial Econometrics 3*, 456–499.

Jacod, J. (1994). Limit of random measures associated with the increments of a Brownian semimartingale. Preprint number 120, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.

Jacod, J. and P. Protter (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability 26*, 267–307.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2 ed.). Springer-Verlag: Berlin.

Javaheri, A., P. Wilmott, and E. Haug (2002). GARCH and volatility swaps. Unpublished paper: available at www.wilmott.com.

Kanatani, T. (2004a). High frequency data and realized volatility. Ph.D. thesis, Graduate School of Economics, Kyoto University.

Kanatani, T. (2004b). Integrated volatility measuring from unevenly sampled observations. *Economics Bulletin 3*, 1–8.

Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus* (2 ed.), Volume 113 of *Graduate Texts in Mathematics*. Berlin: Springer–Verlag.

Karatzas, I. and S. E. Shreve (1998). *Methods of Mathematical Finance*. New York: Springer–Verlag.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies 65*, 361–393.

Koopman, S. J., B. Jungbacker, and E. Hol (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance 12*, 445–475.

Large, J. (2005). Estimating quadratic variation when quoted prices jump by a constant increment. Unpublished paper: Nuffield College, Oxford.

Lévy, P. (1937). *Théories de L'Addition Aléatories*. Paris: Gauthier-Villars.

Maheswaran, S. and C. A. Sims (1993). Empirical implications of arbitrage-free asset markets. In P. C. B. Phillips (Ed.), *Models, Methods and Applications of Econometrics*, pp. 301–316. Basil Blackwell.

Malliavin, P. and M. E. Mancino (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics 6*, 49–61.

Mancini, C. (2001). Does our favourite index jump or not. Dipartimento di Matematica per le Decisioni, Universita di Firenze.

Mancini, C. (2003). Statistics of a Poisson-Gaussian process. Dipartimento di Matematica per le Decisioni, Universita di Firenze.

Mancini, C. (2004). Estimation of the characteristics of jump of a general Poisson-diffusion process. *Scandinavian Actuarial Journal 1*, 42–52.

Mancino, M. and R. Reno (2005). Dynamic principal component analysis of multivariate volatilites via fourier analysis. *Applied Mathematical Finance 12*, 187–199.

Martens, M. (2003). Estimating unbiased and precise realized covariances. Unpublished paper.

Martens, M. and D. van Dijk (2005). Measuring volatility with the realized range. Unpublished paper: Econometric Institute, Erasmus University, Rotterdam.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatilities. *Journal of Applied Econometrics 17*, 479–508.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*, 323–361.

Müller, U. A. (1993). Statistics of variables observed over overlapping intervals. Unpublished paper, Olsen and Associates, Zurich.

Munroe, M. E. (1953). *Introduction to Measure and Integration*. Cambridge, MA: Addison-Wesley Publishing Company, Inc.

Mykland, P. A. and L. Zhang (2002). Inference for volatility-type objects and implications for hedging. Unpublished paper: Department of Statistics, University of Chicago.

Mykland, P. A. and L. Zhang (2005). ANOVA for diffusions. *Annals of Statistics 33*. Forthcoming.

Neuberger, A. (1990). Volatility trading. Unpublished paper: London Business School.

Nielsen, M. O. and P. H. Frederiksen (2005). Finite sample accuracy of integrated volatility estimators. Unpublished paper, Department of Economics, Cornell University.

Officer, R. R. (1973). The variability of the market factor of the New York stock exchange. *Journal of Business 46*, 434–453.

O'Hara, M. (1995). *Market Microstructure Theory*. Oxford: Blackwell Publishers.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*, 61–66.

Patton, A. (2005). Volatility forecast evaluation and comparison using imperfect volatility proxies. Unpublished paper: Department of Accounting and Finance, LSE.

Phillips, P. C. B. and J. Yu (2005). A two-stage realized volatility approach to the estimation for diffusion processes from discrete observations. Unpublished paper: Cowles Foundation for Research in Economics, Yale University.

Politis, D., J. P. Romano, and M. Wolf (1999). *Subsampling*. New York: Springer.

Precup, O. V. and G. Iori (2005). Cross-correlation measures in high-frequency domain. Unpublished paper: Department of Economics, City University.

Protter, P. (2004). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.

Reno, R. (2003). A closer look at the epps effect. *International Journal of Theoretical and Applied Finance 6*, 87–102.

Rogers, L. C. G. and S. E. Satchell (1991). Estimating variance from high, low, open and close prices. *Annals of Applied Probability 1*, 504–512.

Rosenberg, B. (1972). The behaviour of random variables with nonstationary variance and the distribution of security prices. Working paper 11, Graduate School of Business Administration, University of California, Berkeley. Reprinted in Shephard (2005).

Rydberg, T. H. and N. Shephard (2000). A modelling framework for the prices and times of trades made on the NYSE. In W. J. Fitzgerald, R. L. Smith, A. T. Walden, and P. C. Young (Eds.), *Nonlinear and Nonstationary Signal Processing*, pp. 217–246. Cambridge: Isaac Newton Institute and Cambridge University Press.

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance 44*, 1115–1153.

Schwert, G. W. (1990). Indexes of U.S. stock prices from 1802 to 1987. *Journal of Business 63*, 399–426.

Schwert, G. W. (1998). Stock market volatility: Ten years after the crash. *Brookings-Wharton Papers on Financial Services 1*, 65–114.

Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.

Sheppard, K. (2005). Measuring realized covariance. Unpublished paper: Department of Economics, University of Oxford.

Shiryaev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models and Theory*. Singapore: World Scientific.

Wasserfallen, W. and H. Zimmermann (1985). The behaviour of intraday exchange rates. *Journal of Banking and Finance 9*, 55–72.

Wiener, N. (1924). The quadratic variation of a function and its Fourier coefficients. *Journal of Mathematical Physics 3*, 72–94.

Woerner, J. (2004). Power and multipower variation: inference for high frequency data. Unpublished paper.

Zhang, L. (2004). Efficient estimation of stochastic volatility using noisy observations: a multiscale approach. Unpublished paper: Department of Statistics, Carnegie Mellon University.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*. Forthcoming.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics 14*, 45–52.