JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **17**: 457–477 (2002)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jae.691

ESTIMATING QUADRATIC VARIATION USING REALIZED VARIANCE

OLE E. BARNDORFF-NIELSEN[a] AND NEIL SHEPHARD[b]*

[a] *The Centre for Mathematical Physics and Stochastics (MaPhySto), University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark*

[b] *Nuffield College, University of Oxford, Oxford OX1 1NF, UK*

SUMMARY

This paper looks at some recent work on estimating quadratic variation using realized variance (RV)—that is, sums of M squared returns. This econometrics has been motivated by the advent of the common availability of high-frequency financial return data. When the underlying process is a semimartingale we recall the fundamental result that RV is a consistent (as $M \to \infty$) estimator of quadratic variation (QV). We express concern that without additional assumptions it seems difficult to give any measure of uncertainty of the RV in this context. The position dramatically changes when we work with a rather general SV model—which is a special case of the semimartingale model. Then QV is integrated variance and we can derive the asymptotic distribution of the RV and its rate of convergence. These results do not require us to specify a model for either the drift or volatility functions, although we have to impose some weak regularity assumptions. We illustrate the use of the limit theory on some exchange rate data and some stock data. We show that even with large values of M the RV is sometimes a quite noisy estimator of integrated variance.

1. INTRODUCTION

In this paper we ask two questions about *realized variance*[1] (RV), that is, the sum of M squared returns.

- What does RV estimate?
- How precise is RV?

The answer to the first question is straightforward and well known: it is a consistent estimator (as $M \to \infty$) of the corresponding quadratic variation (QV), for all semimartingales. We will see that this is potentially helpful for QV is quite often an econometrically revealing quantity in special cases of semimartingales models. Unfortunately, although RV is a consistent estimator of QV in general, we do not know anything about its precision or indeed even its rate of convergence as $M \to \infty$. This is potentially troublesome, for it does not allow us to deal with the issue that RV and QV are distinct. This obstacle can be overcome when we work within a stochastic volatility (SV)

* Correspondence to: Neil Shephard, Nuffield College, Oxford OX1 1NF, UK.
E-mail: neil.shephard@nuffield.oxford.ac.uk

Contract/grant sponsor: Danish Social Science Research Council.
Contract/grant sponsor: Danish National Research Foundation.
Contract/grant sponsor: ESRC; Contract/grant number: R00023839.

[1] Sums of squared returns are often called realized volatility in econometrics. We will discuss later in this introduction why we prefer to call this statistic realized variance.

Received 23 October 2001
Revised 3 April 2002

framework, which is an important special case of semimartingales with continuous sample paths. For such models we have been able to derive the rate of convergence and indeed the asymptotic distribution. We will illustrate this theory in the context of some high-frequency exchange rate data and daily stock data, showing in particular that RV can sometimes be a very noisy estimator of QV even when M is large.

In order to formalize some of these issues we begin with some definitions and notation. We write the log-price as $y^*(t)$, where t denotes time. Such a price series is usually synthesized from quote or transaction data. Over small time intervals the details of this construction greatly matters and has been studied extensively in the econometrics literature on market microstructure (e.g. see the interesting work of Andreou and Ghysels, 2002 and Bai, Russell and Tiao, 'Beyond Merton's utopia: effects of non-normality and dependence on the precision of variance estimates using high-frequency financial data', unpublished manuscript, 2000, in this context). For the moment we abstract from this issue. If we think of a fixed interval of time of length $\hbar > 0$, then the returns over the ith such interval are defined as

$$y_i = y^*(i\hbar) - y^*((i-1)\hbar), \quad i = 1, 2, \ldots$$

During this interval, we can also compute M intra-$\hbar$ returns. These are defined, for the i-th period, as

$$y_{j,i} = y^*\left((i-1)\hbar + \frac{\hbar j}{M}\right) - y^*\left((i-1)\hbar + \frac{\hbar(j-1)}{M}\right), \quad j = 1, 2, \ldots, M$$

Then many financial economists have measured variability during this period using *realized variance*, defined as

$$[y_M^*]_i = \sum_{j=1}^{M} y_{j,i}^2$$

This term is often called the *realized volatility* in econometrics, although we will keep back that name for

$$\sqrt{\sum_{j=1}^{M} y_{j,i}^2}$$

reflecting our use of volatility to mean standard deviations rather than variances[2]. Examples of the use of realized variances are given by, for example, Merton (1980), Poterba and Summers (1986), Schwert (1989, 1990), Richardson and Stock (1989), Taylor and Xu (1997) and Christensen and Prabhala (1998). An elegant survey of the literature on this topic, including a discussion of its economic importance, is given by Andersen, Bollerslev and Diebold (2002). See also the recent important contribution by Meddahi (2002).

From a formal econometric viewpoint we consider $[y_M^*]_i$ as an *estimator*, allowing us to study its finite sample behaviour for fixed M or its asymptotic properties as $M \to \infty$. Unfortunately, although we know RV converges to QV in probability, this result lacks a theory of measurement error which makes it hard to use this estimator. It would seem additional assumptions are needed. One set of assumptions is to say that y^* is Brownian motion with drift which is deformed by

[2] The use of volatility to denote standard deviations rather than variances is standard in financial economics. See, for example, the literature on volatility and variance swaps, which are derivatives written on realized volatility or variance, which includes Demeterfi *et al*. (1999), Howison, Rafailidis and Rasmussen ('A note on the pricing and hedging of Volatility derivates', unpublished manuscript, 2000) and Chriss and Morokoff (1999). We have chosen to follow this nomenclature rather than the one more familiar in econometrics.

a subordinator (that is, a process with non-negative, independent and stationary increments). We study QV in this context in Section 2 of this paper. Such models are frequently used in finance in order to derive derivative pricing formulas. An alternative is to assume a rather general stochastic volatility (SV) model. The latter framework is the mainstay of the discussion we give here.

The SV model we work with has a very flexible form. We assume

$$y^*(t) = \alpha(t) + \int_0^t \sigma(s)\mathrm{d}w(s), \quad t \geq 0$$

where α, the drift, and $\sigma > 0$, the spot volatility, obey some weak assumptions outlined in Section 3 and w in standard Brownian motion. In particular the spot volatility can have, for example, deterministic diurnal effects, jumps, long memory, no unconditional mean or be non-stationary. No knowledge of the form of the stochastic processes which govern α and σ are needed. SV models are a fundamental special type of semimartingale; in particular most semimartingales which possess continuous sample paths can be represented as SV models.

In these models, assuming σ and α are jointly independent from obey, returns w

$$y_i|\alpha_i, \sigma_i^{[2]} \sim N(\alpha_i, \sigma_i^{[2]}) \tag{1}$$

where

$$\alpha_i = \alpha\,(i\hbar) - \alpha\,((i-1)\hbar) \tag{2}$$

and

$$\sigma_i^{[2]} = \sigma^{2*}\,(i\hbar) - \sigma^{2*}\,((i-1)\hbar)\,, \quad \text{where} \quad \sigma^{2*}(t) = \int_0^t \sigma^2(s)\mathrm{d}s \tag{3}$$

We call $\sigma^2(t)$ the spot variance and $\sigma^{2*}(t)$ the integrated variance. Importantly, for all SV models σ^{2*} exactly equals QV

$$\sigma^{2*}(t) = [y^*](t) \tag{4}$$

and so

$$\sigma_i^{[2]} = [y^*]\,(i\hbar) - [y^*]\,((i-1)\hbar) = [y^*]_i$$

Thus QV reveals exactly the actual variance $\sigma_i^{[2]}$ in SV models. It does this without knowledge of the actual processes which govern α or σ.

The above theory means that $[y_M^*]_i$ consistently estimates $\sigma_i^{[2]}$, just using the theory of semimartingales. Barndorff-Nielsen and Shephard (2002) have shown that $[y_M^*]_i$ converges to $\sigma_i^{[2]}$ at rate $\sqrt{M}$ and have additionally derived the asymptotic distribution of the estimator:

$$\frac{\displaystyle\sum_{j=1}^{M} y_{j,i}^2 - \int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s}{\sqrt{\displaystyle\frac{2}{3}\sum_{j=1}^{M} y_{j,i}^4}} \xrightarrow{\mathcal{L}} N(0,1) \tag{5}$$

as $M \to \infty$, thus providing a measure of the precision of this estimator. Their preferred form of the result, due to its superior finite sample behaviour (see Barndorff-Nielsen and Shephard (2002a)), is that as $M \to \infty$ then

$$\frac{\log\left(\sum_{j=1}^{M} y_{j,i}^2\right) - \log\left(\sigma_i^{[2]}\right)}{\sqrt{\frac{2}{3}\frac{\sum_{j=1}^{M} y_{j,i}^4}{\left\{\sum_{j=1}^{M} y_{j,i}^2\right\}^2}}} \xrightarrow{\mathcal{L}} N(0,1) \tag{6}$$

This is a mixed Gaussian limit theory, that is, the denominator is itself random. Of course, this theory can be used to provide approximations for realized volatility as well as realized variance. The distribution of realized volatilies can also be approximated directly via (5) using the delta method which gives

$$\frac{\sqrt{\sum_{j=1}^{M} y_{j,i}^2} - \sqrt{\int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s}}{\sqrt{\frac{2}{12}\frac{\sum_{j=1}^{M} y_{j,i}^4}{\sum_{j=1}^{M} y_{j,i}^2}}} \xrightarrow{\mathcal{L}} N(0,1) \tag{7}$$

The log-based approximation (6) is likely to be preferred in practice when we construct confidence intervals for realized volatility.

To illustrate this result we have used the same return data employed by Andersen *et al.* (2001b) in their empirical study of the properties of realized variance, although we have made slightly different adjustments to deal with some missing data (in the context of this paper the effect of these differences are tiny, but were made here to be consistent with our other work on this dataset). Full details of this are given in Barndorff-Nielsen and Shephard (2002). The data was kindly supplied to us by the Olsen group in Zurich. This United States Dollar/German Deutsche Mark series covers the ten year period from 1 December 1986 until 30 November 1996. It records every five minutes the most recent quote to appear on the Reuters screen. Throughout we take $\hbar$ to represent a day and so have up to 288 five-minute returns to work with each day. This constrains our choice of M to taking the values 288, 144, 96, 72, 48, 36, 32, 24, 18, 16, 12, 9, 8, 6, 4, 3, 2, 1.

In Figure 1 we record, for a variety of values of M, RV and its 95% confidence intervals (based on (6)) for the first 9 days of the dataset. This is the first time such graphs have been produced. The result suggests that the confidence intervals do indeed narrow considerably with M. However, even with $M = 288$ the intervals are sometimes quite wide. The implication is that RV is a consistent but quite noisy estimator of $\sigma_i^{[2]}$, especially when volatility is high. We will return to this issue in more detail in Section 4.

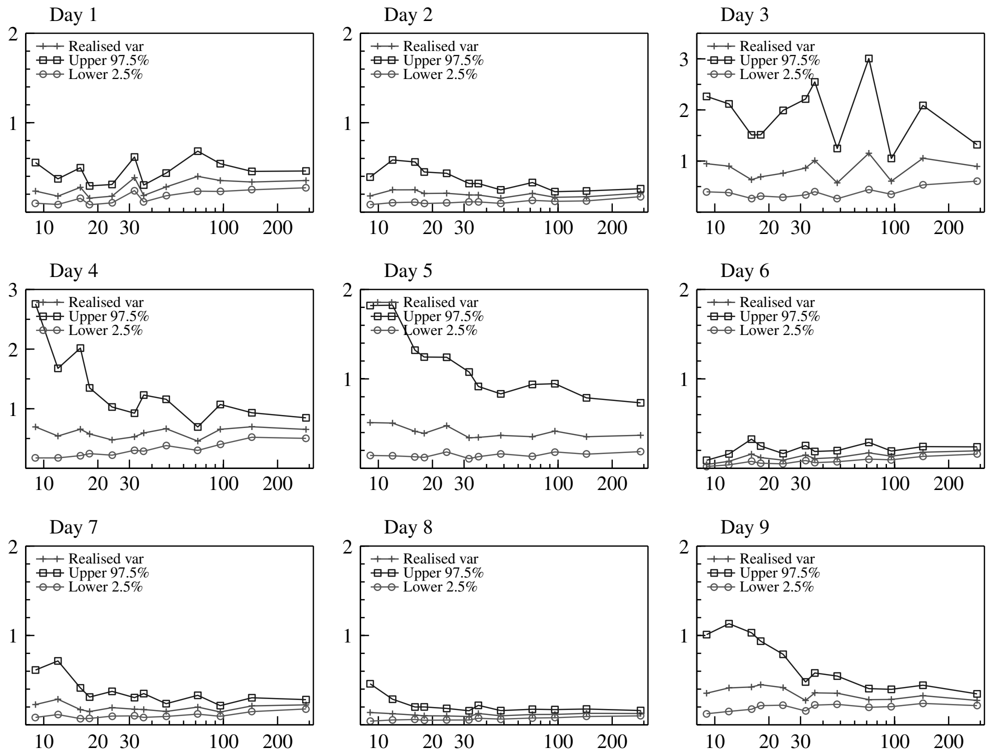

Figure 1. First 9 days. RV is plotted against M, with the smallest M being 8, the largest 288. Also plotted are the 95% intervals. Code: **se_realised.ox**

The structure of this paper is as follows. In Section 2 we discuss the definition of quadratic variation in the context of semimartingales. We will see that realized variance, by definition, converges in probability to QV and so RV is a consistent estimator of QV. We give some examples where QV does not reveal the conditional variance of returns in a subordinated Brownian motion model. Our conclusion is that more detailed assumptions are needed in order to provide a coherent analysis of RV. We also discuss the advantage of working with the conditional expectation of RV and the convergence of it to the conditional expectation of quadratic variation. This follows some recent work by Andersen *et al*. ('Modeling and forecasting realized volatility', unpublished manuscript, 2001). In Section 3 we move on to consider the properties of RV in the context of SV models. We develop our, rather robust, asymptotic theory for RV as $M \to \infty$. Section 4 discusses various aspects of the empirical implications of the theory of measurement error for RV. Section 5 provides a conclusion.

2. QUADRATIC VARIATION AND SEMIMARTINGALES

2.1. One Period is Enough

In applied work RV

$$[y_M^*]_i = \sum_{j=1}^{M} y_{j,i}^2$$

is usually computed for each $\hbar$ period (usually a day) separately. Hence from a theoretical viewpoint we only have to think about a single period, starting from time 0 until time t but working with M equally spaced high-frequency returns to calculate RV. This approach allows us to use a rather simpler notation.

2.2. Semimartingales

Most modern finance theory is based on semimartingales (see, for example, the excellent exposition in Shiryaev, 1999, pp. 294–313). In econometrics such processes are not so familiar, so we remind the reader of the definition. Suppose $y^*(t)$ is a stochastic process and that for ease of exposition we assume that $y^*(0) = 0$. Then $y^*(t)$ is said to be a *semimartingale* if it is decomposable as

$$y^*(t) = \alpha(t) + m(t), \quad \alpha(0) = m(0) = 0 \tag{8}$$

where $\alpha(t)$, a drift term, is a process with *locally bounded variation* paths (i.e. of bounded variation[3] on any finite subinterval of $[0, \infty)$) and $m(t)$ is a *local martingale*. For an excellent discussion of probabilistic aspects of this see Protter (1990). We will later additionally assume that $\alpha(t)$ is a *predictable* process[4], in which case $y^*(t)$ is said to be a *special semimartingale* (e.g. Protter, 1990, p. 107). Back (1991) discusses why constraining ourselves to live within the class of special semimartingales makes sense from an economic viewpoint. For this subset of semimartingales the canonical decomposition (8) is unique.

Let $y^*(t)$ be a general semimartingale. The *quadratic variation* (process) $[y^*]$ of y^* is defined by

$$[y^*](t) = y^{*2}(t) - 2\int_0^t y^*(s-)\mathrm{d}y^*(s) \tag{9}$$

Much more interestingly from an econometric view point it can be shown that

$$[y^*](t) = p\text{-}\lim \sum_{j=0}^{M-1} \left\{ y^*(s_{j+1}) - y^*(s_j) \right\}^2 \tag{10}$$

where $0 = s_0 < s_1 < \ldots < s_M = t$ and the limit is for the mesh size

$$\max_{1 \le j \le M} |s_j - s_{j-1}| \to 0 \quad as \quad M \to \infty$$

In general (e.g. Jacod and Shiryaev, 1987, p. 55)

$$[y^*](t) = [y^{*c}](t) + \sum_{0 \le s \le t} \left\{ \Delta y^*(s) \right\}^2$$

---

[3] If the real-valued function f on $[a, b]$ is such that

$$\sup_{\kappa} \sum |f(x_i) - f(x_{i-1})| < \infty$$

where the supremum is taken over all subdivisions κ of $[a, b]$ then the function is of bounded variation.

[4] The value at time t of a predictable process is known an instant before time t. Examples of predictable processes are deterministic trends and all càglàd processes (processes which are continuous from the left and have limits from the right).

where y^{*c} is the continuous martingale component of y^* and

$$\Delta y^*(s) = y^*(s) - y^*(s-)$$

is the jump at time s. In the context of special semimartingales this becomes

$$\begin{aligned}
[y^*](t) &= [y^{*c}](t) + \sum_{0 \le s \le t} \{\Delta m(s)\}^2 + \sum_{0 \le s \le t} \{\Delta \alpha(s)\}^2 + 2 \sum_{0 \le s \le t} \Delta m(s) \Delta \alpha(s) \\
&= [m](t) + \sum_{0 \le s \le t} \{\Delta \alpha(s)\}^2 + 2 \sum_{0 \le s \le t} \Delta m(s) \Delta \alpha(s)
\end{aligned}$$

the QV of m plus terms which are influenced by the jumps in α and m. If α is continuous then we obtain the important simplification

$$[y^*](t) = [m](t) \tag{11}$$

This holds even if there are jumps in m. The fact that in this case the QV for y^* equals the QV for m means that the QV is robust to smooth α processes. This is an important feature. A thorough discussion of the quadratic variation of semimartingales is given in Protter (1990). In the univariate case it is discussed in the econometrics literature by independent and concurrent work by Comte and Renault (1998), Barndorff-Nielsen and Shephard (2001) and Andersen and Bollerslev (1998). It was later developed and applied in some empirical work by Andersen *et al.* (2001b). See also Barndorff-Nielsen and Shephard ('Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics', unpublished manuscript, 2002) and Andersen *et al.* (2002) for an incisive survey of this area. Andersen *et al.* (2001a) discuss the use of the multivariate theory in the context of equity prices.

In a stimulating paper Andersen *et al.* (2001b) noted that (10) implies

Remark 1 *By definition RV converges in probability*[5] *to QV as $M \to \infty$ for all semimartingales.*

Of course this result occurs by construction. Before we discuss the meaning of this result we note it has at least two technical limitations (to our knowledge).

Remark 2 *The rate of convergence of RV to QV is unknown, as is its asymptotic distribution.*

Consistency is an important feature of an estimator, but it is possible such estimators can converge very slowly indeed to their limit. As we do not know the rate of convergence, all we can conclude is that when α is continuous

$$\sum_{j=0}^{M-1} \left\{ y^*(s_{j+1}) - y^*(s_j) \right\}^2 - [m](t) = o_p(1)$$

[5] It is important to note that the convergence is in probability, not almost surely or in mean square error.

2.3. Semimartingales by Subordination

It is clear that RV converges to QV for semimartingales. However, is QV useful? We saw in the introduction that QV is key in the context of stochastic volatility. But elsewhere? Here we look at a second special case of the semimartingale class. We let log-prices be a standard Brownian motion $b(t)$, with $b(0) = 0$, which is deformed by a random clock $\tau(t)$ which is assumed to be a subordinator (that is, a Lévy process with non-negative increments, i.e. the increments are independent and stationary, in addition to the non-negativity). We obtain

$$y^*(t) = b(\tau(t)) \tag{12}$$

which implies

$$y^*(t)|\tau(t) \sim N(0, \tau(t))$$

Influential special cases of this include the variance gamma (where τ is a gamma Lévy process), normal inverse Gaussian (where τ is an inverse Gaussian Lévy process) and generalized hyperbolic (where τ is a generalized inverse Gaussian Lévy process) processes discussed by Madan and Seneta (1990), Barndorff-Nielsen (1998) and Eberlein (2001) respectively. These types of models are frequently used to price European-style derivatives, allowing for non-Gaussianity in the risk-neutral process.

As the random time clock τ, which we call the chronometer, is a subordinator it is a pure (upward) jump process. Consequently y^* is also a pure jump process. Hence it is fundamentally different from the SV model which inherits continuous sample paths from its chronometer—integrated variance σ^{2*}. Like integrated variance, the subordinator determines the conditional variance of the returns. Hence one might hope that, just like the SV case, QV reveals τ. However, this is not true. The following is a simple example of this.

Example 1 *Suppose τ is a homogeneous Poisson process, then the subordinated Brownian motion takes on the form*

$$y^*(t) = \sum_{j=1}^{\tau(t)} z_j, \qquad z_j \sim NID(0, 1)$$

which has

$$[y^*](t) = \sum_{j=1}^{\tau(t)} z_j^2$$

Thus the QV is a $\chi^2_{\tau(t)}$ random variable, which is a potentially noisy version of $\tau(t)$. Of course the expectations of QV and τ are equal. This is a very important point and will be discussed in the next subsection at some length. There is an interesting and powerful probability literature on this topic where researchers calculate the posterior distribution of the subordinator τ given the QV—thus quantifying the measurement error of QV as an estimator of τ, see Carr *et al.* (2002) and Winkel (2001). These results are possible due to additional structure being assumed in that work. Such intricate results seem not to be possible for general semimartingales.

2.4. Expectations, QV and RV

The above observations are rather bleak. A sensible question to ask is if there are any useful properties of QV that RV can estimate? In an important contribution Andersen, *et al*. 2001c) argued that

- The conditional expectation of QV is economically interesting.
- The conditional expectation of RV is a good approximation to QV.

Their argument proceeds in the following manner. First, they note that, using the first $\hbar$-time period to simplify the notation and definition (9),

$$\begin{aligned}\operatorname{Var}(y^*(\hbar)|\mathcal{F}_0) &= \operatorname{Var}(m(\hbar)|\mathcal{F}_0) + \operatorname{Var}(\alpha(\hbar)|\mathcal{F}_0) + 2\operatorname{Cov}(\alpha(\hbar), m(\hbar)|\mathcal{F}_0) \qquad (13)\\ &= \operatorname{E}([m](\hbar)|\mathcal{F}_0) + \operatorname{Var}(\alpha(\hbar)|\mathcal{F}_0) + 2\operatorname{Cov}(\alpha(\hbar), m(\hbar)|\mathcal{F}_0) \qquad (14)\end{aligned}$$

where $\mathcal{F}_0$ is the natural filtration at time 0. Then they point out that in many models $\operatorname{Cov}(\alpha(\hbar), m(\hbar)|\mathcal{F}_0) = 0$ while it is commonly the case that $\operatorname{Var}(\alpha(\hbar)|\mathcal{F}_0)$ is small unless $\hbar$ is large. If we assume both terms are exactly zero[6] then if α is continuous we have

$$\begin{aligned}\operatorname{Var}(y^*(\hbar)|\mathcal{F}_0) &= \operatorname{E}([m](\hbar)|\mathcal{F}_0)\\ &= \operatorname{E}\left([y^*](\hbar)|\mathcal{F}_0\right)\end{aligned}$$

This is important for it says that the conditional variance of future returns is the conditional expectation of QV, which in turn is an object which can be consistently estimated by RV. Hence it is tempting to conclude that the conditional variance of returns is asymptotically conditionally unbiasedly estimated by RV, that is,

$$\operatorname{Var}(y^*(\hbar)|\mathcal{F}_0) - \operatorname{E}\left([y_M^*]_1|\mathcal{F}_0\right) = o(1) \qquad (15)$$

as a function of M. This is feasible for it would allow us to build an empirical time series model for the conditional mean of RV, but then claim it gives a valid approximation to the conditional variance of returns. This empirical approach has been pioneered in financial econometrics by Andersen *et al*. (2001c). However, convergence in probability of RV to QV does not imply convergence in the means of the two objects, although we believe that this convergence property is likely to hold under very weak additional conditions on α and m. The simplest such condition is when α is zero. In this case

$$\operatorname{E}\{[y^*](\hbar)|\mathcal{F}_0\} = \operatorname{E}\{[y_M^*]_1|\mathcal{F}_0\}$$

exactly, by the properties of local martingales.

We can formalize a more general discussion of the convergence of conditional expectations in the following manner. Recall $[x_M]_i$ is our notation for the RV in the ith period for an arbitrary

[6] This holds exactly if α is a deterministic function of time.

special semimartingale x, then by the canonical decomposition (8) of a special semimartingale, we have that

$$[y_M^*]_i = [\alpha_M]_i + 2[\alpha_M, m_M]_i + [m_M]_i, \qquad \text{with} \quad [\alpha_M, m_M]_i = \sum_{j=1}^{M} \alpha_{j,i} m_{j,i}$$

where we are using the general notation for a process x:

$$x_{j,i} = x\left((i-1)\hbar + \hbar j M^{-1}\right) - x\left((i-1)\hbar + \hbar(j-1)M^{-1}\right)$$

The implication is that

$$\begin{aligned} \mathrm{E}\{[y_M^*]_1|\mathcal{F}_0\} &= \mathrm{E}\{[\alpha_M]_1|\mathcal{F}_0\} + 2\mathrm{E}\{[\alpha_M, m_M]_1|\mathcal{F}_0\} + \mathrm{E}\{[m_M]_1|\mathcal{F}_0\} \\ &= \mathrm{E}\{[\alpha_M]_1|\mathcal{F}_0\} + 2\mathrm{E}\{[\alpha_M, m_M]_1|\mathcal{F}_0\} + \mathrm{E}\{[m]_1|\mathcal{F}_0\} \end{aligned}$$

Then (15) follows under the following five sufficient conditions:

- α is continuous.
- $\mathrm{Cov}(\alpha(\hbar), m(\hbar)|\mathcal{F}_0)$ is zero.
- $\mathrm{Var}(\alpha(\hbar)|\mathcal{F}_0)$ is zero.
- $2\mathrm{E}\{[\alpha_M, m_M]_1|\mathcal{F}_0\} \to 0 \quad \text{as} \quad M \to \infty$ (16)
- $\mathrm{E}\{[\alpha_M]_1|\mathcal{F}_0\} \to 0 \quad \text{as} \quad M \to \infty$ (17)

We have already discussed the first three of these points. On the latter two, we know that both $[\alpha_M, m_M]_1$ and $[\alpha_M]_1$ converge in probability to zero and so we would expect (16) and (17) to hold under very weak conditions. In particular if α is continuous and deterministic then all five conditions always hold. We will return to checking the conditions (16) and (17) at the end of the next section.

3. SV MODELS AND REALIZED VARIANCE

3.1. SV Models and Integrated Variance

In response to the above difficulties of dealing with general semimartingales, we advocate making some additional assumptions. In particular we specialize the semimartingale assumption down to a stochastic volatility model.

In the stochastic volatility model for log-prices a basic Brownian motion is generalized to allow the volatility term to vary over time. See, for example, Barndorff-Nielsen and Shephard (2001) and Ghysels, Harvey and Renault (1996) on some of the literature on this topic. Here we use a rather flexible model:

$$y^*(t) = \alpha(t) + \int_0^t w(s)\mathrm{d}w(s) \qquad t \geq 0 \tag{18}$$

where $\sigma > 0$ and α are assumed to be stochastically independent of the standard Brownian motion w. We call σ the *spot volatility*, σ^2 the *spot variance* and α the *mean* process. By allowing the spot volatility to be random and serially dependent, this model will imply returns that exhibit volatility clustering and have unconditional distributions which are fat-tailed. This allows it to be used in finance and econometrics as the basis for option pricing models which overcome some of the major failings in the Black–Scholes option pricing approach. Leading references in this regard include Hull and White (1987), Heston (1993) and Renault (1997). Importantly,

$$y^*(t)|\alpha(t), \sigma^{2*}(t) \sim N(\alpha(t), \sigma^{2*}(t))$$

where

$$\sigma^{2*}(t) = \int_0^t \sigma^2(s)\mathrm{d}s$$

is called the integrated variance. Hence for an econometrician it is the object to be estimated.

Throughout we will maintain the following assumptions on the volatility and mean processes:

(1) σ^2 and α pathwise of local bounded variation on $[0, \infty)$.
(2) For every $t > 0$, α has the property

$$\overline{\lim_{\delta\downarrow 0}} \max_{1\le j\le M} \delta^{-1}|\alpha(j\delta) - \alpha((j-1)\delta)| < \infty \tag{19}$$

where M denotes a positive integer and $\delta = t/M$. This condition is implied by Lipschitz continuity and itself implies continuity[7] of α.

These regularity conditions are quite mild. Of some special interest are cases where α is of the form

$$\alpha(t) = \int_0^t g(\sigma(s))\mathrm{d}s$$

for g a smooth function. Condition 2 holds in general for such models.

Note that the assumptions allow the spot volatility to have, for example, deterministic diurnal effects, jumps, long memory, no unconditional mean or to be non-stationary. Also the conditions imply that σ^2 and α are bounded Riemann integrable functions, while y^* is a semimartingale with a continuous local martingale component $\int_0^t \sigma(s)\mathrm{d}w(s)$. In particular this SV model is a special case of the semimartingale model we discussed in the previous section. This implies the well known result that:

Remark 3 *For the SV model*

$$[y^*](t) = \left[\int_0^t \sigma(s)\mathrm{d}w(s)\right](t) = \sigma^{2*}(t)$$

This result appeared in concurrent and independent work by Andersen and Bollerslev (1998) (assuming $\alpha(t) = 0$) and Barndorff-Nielsen and Shephard (2001). Both of these papers were presented at the 1997 Olsen conference. If we combine this result with Remark 1 it immediately implies that RV consistently estimates $\sigma^{2*}(t)$, which was first noted explicitly (assuming $\alpha(t) = 0$) in Andersen and Bollerslev (1998).

Again this result is attractive for it does not depend upon the particular structure of the mean and volatility process. Unlike the general semimartingale case, the RV is converging to the object we wish to have. The only problem that remains is that we do not know the rate of convergence nor

[7] Continuity of α implies α is a predictable process—hence this assumption is a restriction on the class of special semimartingales we can analyse.

the asymptotic distribution. As we have seen from the previous section, the general semimartingale theory is silent on this issue.

3.2. Asymptotic Distribution of Realized Variance

In some recent work Barndorff-Nielsen and Shephard (2002) derive the following asymptotic approximation to the distribution of realised variance:

Remark 4 *For $M \to \infty$*

$$\frac{\sum_{j=1}^{M} y_{j,i}^2 - \int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s}{\sqrt{\frac{2}{3}\sum_{j=1}^{M} y_{j,i}^4}} \xrightarrow{\mathcal{L}} N(0,1) \tag{20}$$

This holds under conditions (1) and (2).

This is a considerable strengthening of the above consistency result. Now we have a measure of error. In line with the above approach, it has the advantage that it is model free for the denominator does not require any knowledge of α or σ. An improved understanding of this result can be gained by noting the following result on fourth-order power variation, which is called realized quarticity, due to Barndorff-Nielsen and Shephard (2002b):

Remark 5 *Under conditions (1) and (2) for $M \to \infty$*

$$\frac{M}{3\hbar}\sum_{j=1}^{M} y_{j,i}^4 \xrightarrow{p} \int_{\hbar(i-1)}^{\hbar i} \sigma^4(s)\mathrm{d}s$$

This result allows us to recast (20) into a theoretically informative, although no longer statistically feasible, form

$$\frac{\sqrt{M}\left(\sum_{j=1}^{M} y_{j,i}^2 - \int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s\right)}{\sqrt{2\hbar\int_{\hbar(i-1)}^{\hbar i} \sigma^4(s)\mathrm{d}s}} \xrightarrow{\mathcal{L}} N(0,1)$$

This shows five things:

Remark 6 *(i) RV is converging to integrated variance at rate $\sqrt{M}$, (ii) the asymptotic distribution is mixed Gaussian and so*

$$\sum_{j=1}^{M} y_{j,i}^2 - \int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s$$

will be marginally (much) heavier tailed than Gaussian, (iii) typically the variability of $\sum_{j=1}^{M} y_{j,i}^2 - \int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s$ *will be higher if the level of the volatility is higher, (iv) if the fourth moment of returns does not exist then the unconditional variance of* $\sum_{j=1}^{M} y_{j,i}^2 - \int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s$ *will not exist. This last result echoes an earlier Monte Carlo study by Bai et al. (unpublished manuscript, 2000) who noted the very poor mean square error performance of realised variance in the case where the fourth moment is close to being not bounded. (v) the fourth moment not existing does not invalidate the asymptotic result (20) for RV. It still holds in this case.*

Importantly the rate of convergence and asymptotic distribution is not impacted by non-stationarity in the volatility or other types of irregularities in the drift.

In a recent paper Barndorff-Nielsen and Shephard (2003a) showed that if we transform the limit theory to the log scale then

Remark 7 *For* $M \to \infty$

$$\frac{\log\left(\sum_{j=1}^{M} y_{j,i}^2\right) - \log\left(\int_{\hbar(i-1)}^{\hbar i} \sigma^2(s)\mathrm{d}s\right)}{\sqrt{\frac{2}{3}\frac{\sum_{j=1}^{M} y_{j,i}^4}{\left\{\sum_{j=1}^{M} y_{j,i}^2\right\}^2}}} \xrightarrow{\mathcal{L}} N(0,1) \tag{21}$$

Their Monte Carlo experiments showed that this limit law (21) had better finite sample performance than the raw result (20).

3.3. Expectations, QV and RV Revisited

We saw in the previous section that it is not possible to generally assert that

$$\mathrm{Var}(y^*(\hbar)|\mathcal{F}_0) - \mathrm{E}\left([y_M^*]_1|\mathcal{F}_0\right) = o(1)$$

for all semimartingales. Under assumptions (1) and (2) in Section 3.1, $\mathrm{E}\{[\alpha_M, m_M]_1|\mathcal{F}_0\} = 0$ while α is continuous. If we ignore the contributions of $\mathrm{Cov}(\alpha(\hbar), m(\hbar)|\mathcal{F}_0)$ and $\mathrm{Var}(\alpha(\hbar)|\mathcal{F}_0)$ as they are likely to be insignificant in applications, the only issue is to show that

$$\mathrm{E}\{[\alpha_M]_1|\mathcal{F}_0\} \to 0 \qquad \text{as} \qquad M \to \infty \tag{22}$$

In this subsection we consider, for $\delta = \hbar/M$, $\mathrm{E}\{[\alpha_M]_1|\mathcal{F}_0\}$ where $[\alpha_M]_1 = \sum_{j=1}^{M} \alpha_j^2$ and $\alpha_j = \alpha(j\delta) - \alpha((j-1)\delta)$. We will give a condition under which

$$\mathrm{E}\{[\alpha_M]|\mathcal{F}_0\} = O(\delta^{\gamma})$$

for some $\gamma > 0$. For this it is sufficient that

$$\mathrm{E}\{[\alpha(t+\delta) - \alpha(t)]^2 | \mathcal{F}_t\} = O(\delta^{1+\gamma}) \tag{23}$$

for a deterministic $O(\delta^{1+\gamma})$, uniformly in t. In fact, under this condition,

$$\begin{aligned} \mathrm{E}\{[\alpha_M] | \mathcal{F}_0\} &= \sum_{j=1}^{M} \mathrm{E}\left(\alpha_j^2 | \mathcal{F}_0\right) \\ &= \sum_{j=1}^{M} \mathrm{E}\left\{\mathrm{E}\left(\alpha_j^2 | \mathcal{F}_{(j-1)\delta}\right) | \mathcal{F}_0\right\} \\ &= \sum_{j=1}^{M} O(\delta^{1+\gamma}) = O(\delta^{\gamma}) \end{aligned}$$

We now give an example where we can show (23) usually holds but under another set of conditions (22) does not hold.

Example 2 *Suppose*

$$m(t) = \int_0^t \tau^{1/2}(u) \mathrm{d}w(u)$$

where τ and w are independent and

$$\alpha(t) = \tau^*(t) = \int_0^t \tau(u) \mathrm{d}wu$$

with τ a stationary process satisfying

$$\mathrm{d}\tau(t) = -\lambda\tau(t)\mathrm{d}t + \mathrm{d}z(\lambda t)$$

where z is a non-negative Lévy process (a process with independent, stationary and non-negative increments). Then α is a locally bounded variation process and

$$Cov(m(t), \alpha(t)) = 0 \qquad and \qquad E\{[\alpha_M, m_M]_1 | \mathcal{F}_0\} = 0$$

The non-Gaussian OU processes τ were highlighted in Barndorff-Nielsen and Shephard (2001) as analytically tractable models for the spot volatility. Then a feature of the mean process is that

$$\alpha(t) = \lambda^{-1}\{z(\lambda t) - \tau(t) + \tau(0)\}$$

which implies, writing $\varepsilon(t;\lambda) = \lambda^{-1}\left(1 - e^{-\lambda t}\right)$, that

$$\begin{aligned} \alpha(t) &= \int_0^t \varepsilon(t-s;\lambda)\mathrm{d}z(\lambda s) + \varepsilon(t;\lambda)\tau(0) \\ &\leq \varepsilon(t;\lambda)\{z(\lambda t) + \tau(0)\} \\ &\leq t\{z(\lambda t) + \tau(0)\} \end{aligned}$$

Hence we have that

$$\xi = E\{\tau(t)\} = E\{z(1)\} \qquad \textit{and} \qquad \omega^2 = Var\{\tau(t)\} = Var\{z(1)\}/2$$

exist and that

$$E\{\alpha(\delta)^2|\tau(0)\} \leq \delta^2\{\lambda\delta\omega + \lambda^2\delta^2\xi^2 + 2\lambda\delta\xi\tau(0) + \tau(0)^2\}$$

and hence (23) holds with $\gamma = 1$. *Of course when the variance of* $\tau(t)$ *does not exist, then the result (22) fails as* $E([y^*]_1|\mathcal{F}_0)$ *does not exist.* □

4. EMPIRICAL EXAMPLES

4.1. A Time Series of Daily RVs

Figure 2 shows the daily time series of the realized variance for $M = 144$, which corresponds to utilizing 10-minute returns, using the foreign exchange data discussed in Section 1. Here we report the first 50 days of the series.

The 95% daily confidence intervals for RV are based on the accurate log-based asymptotic result given in equation (6). We can see the important widening and closing of the 95% confidence intervals, with the intervals seemingly being very large when the volatility is high. In summary,

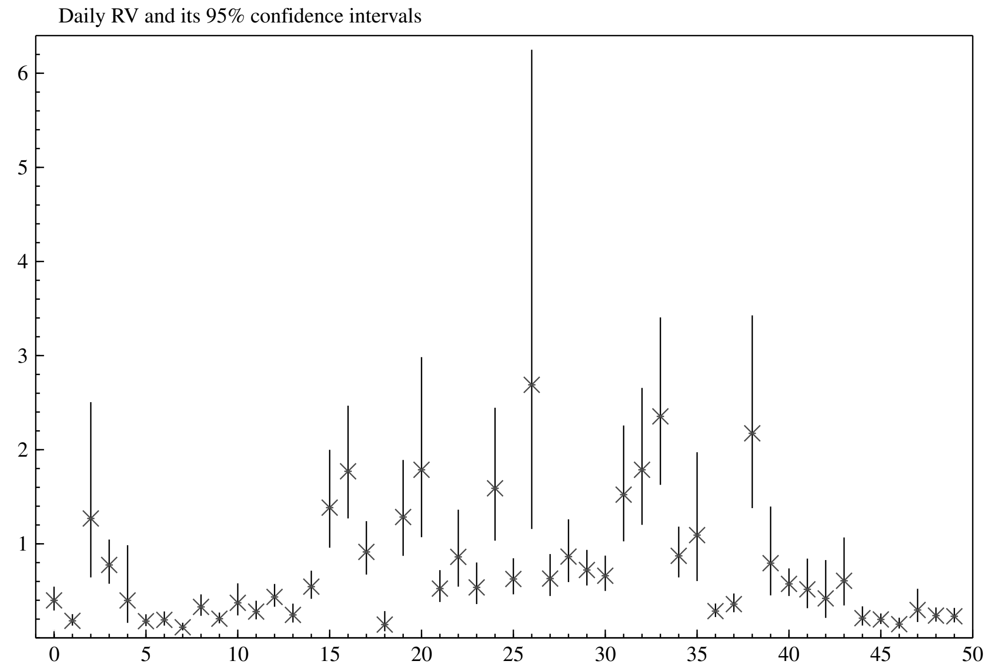

Figure 2. Daily RV, $[y_M^*]_i$, drawn against n for the first 50 days of the sample. Also drawn as vertical bars are the 95% intervals based on the log transformation. Throughout $M = 144$. Code: **se_realised.ox**

Remark 8 *When volatility is low RV is quite accurate, but in periods of high volatility the measurement error can be very large indeed.*

Further, in our applied work the confidence intervals are typically wider than the level of volatility itself.

4.2. Market Microstructure Biases

There are substantial efficiency gains to be made in estimating $\int \sigma^2(u)\mathrm{d}u$ not by just low-frequency squared returns but by computing $[y_M^*]_i$ with a high value of M (see, for example, Andersen and Bollerslev, 1998). However, a number of econometricians are worried that this will lead to biases due to market microstructure effects. In particular Andreou and Ghysels (2001) and Bai *et al.* (unpublished manuscript, 2000) have argued that irregular spaced trading and price discreteness can impact realized variance and that these effects become more dangerous when M is large. Can the above theory throw any light on the above arguments? One approach to thinking about this issue is to plot

$$\frac{1}{N}\sum_{i=1}^{N}[y_M^*]_i$$

against M. This is called a volatility signature plot by Andersen *et al.* (2000). Now for large values of M and N and assuming that the volatility between days is mixing we have that

$$\left(\frac{1}{N}\sum_{i=1}^{N}[y_M^*]_i - \frac{1}{N}\int_0^{N\hbar}\sigma^2(u)\mathrm{d}u\right) \quad \dot{\sim} \quad N\left(0, \frac{2}{3}\frac{1}{N}\sum_{i=1}^{N}[y_M^*]_i^{[4]}\right) \tag{24}$$

where

$$[y_M^*]_i^{[4]} = \sum_{j=1}^{M} y_{j,i}^4$$

which allows us to compute confidence intervals for the sums of realized volatilities. These are given in Figure 3 for the 2448 days in the exchange rate data discussed above. We have not drawn the graph for very small values of M as the asymptotics would be totally unreliable.

The approximation given in (24) suggests that if the continuous time SV model was literally true then the average RV should not change significantly as we alter M. The figures indicate that there is an upward movement in the average RV for large values of M. The size of the move is around 10% of the level of volatility. The standard errors suggest than this is significant, however there are quite a large number of assumptions made in their computation and so we are reluctant to put a great deal of weight on this issue. However, the figure is indicative of problems which arise when we take M to be above around 50.

4.3. A Time Series of Annual Realized Volatilities

One use of the asymptotics for realized variances and volatilities is to compute confidence intervals for low-frequency data such as annual measures of volatility. Here the high-frequency data would

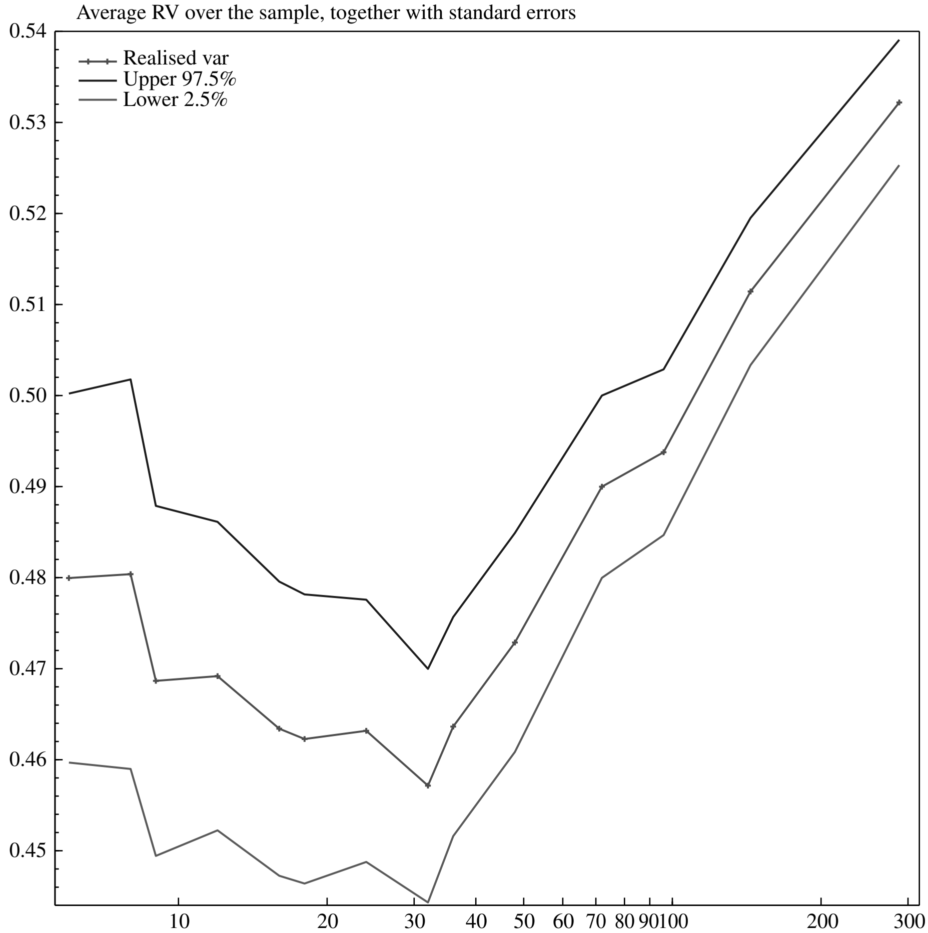

Figure 3. Average (across days) value of the RV drawn against M. Code: **se_realised.ox**

be daily observations and our goal in this subsection is to work with realized volatilities:

$$\sqrt{\sum_{j=1}^{M} y_{j,i}^2}$$

that is, the square root of realized variances. Such historical time series are very common in financial economics. See, for example, the work of Schwert (1989, 1990, 1998) who discusses realized volatilities for a wide variety of financial assets over long time periods.

In this subsection we take a long series on the closing prices on the Dow Jones Industrial Average, starting on 26 May 1896 and going up to 31 December 2001. This is taken from the Dow Jones website and so is in the public domain. This is a narrower index than some of the more widely used series discussed in the literature. In particular the series constructed by Schwert (1990) has many more advantages. However, the Dow Jones index has the virtue that it can be downloaded free of charge which is a requirement of this journal.

This series has a small number of recording breaks, which we have ignored as they make no substantial difference to our analysis. The series has the interesting feature that in the early part of it the markets were open six days a week, while in more recent years this has been reduced to five. Of course this makes no difference to the implementation of our theory.

There is a very substantial break from 30 July 1914 until 31 December 1914. This was caused by the start of World War I, with Germany declaring war on Russia on 1 August 1914. This creates some important difficulties for the index was at 71.42 when it closed, while it reopened at 54 after Christmas in 1914. If we ignore this break, it will imply a very high level of volatility for 1914 due to the massive movement in the index. To construct our data series we have followed the approach of Barndorff-Nielsen and Shephard (2002) who suggest stochastically interpolating prices during breaks. They argue for the use of a Brownian bridge added to a straight-line trend between prices, carrying out the computations on the log-price. The result is shown in Figure 4.

The Brownian bridge we used in this analysis links the closing price in August to the opening price at the end of December in a random way. There is only a single parameter in the linking, the variance of the Brownian motion. This is chosen *a priori* as 0.04/110 per day, which gives a standard deviation of yearly price movements of around 0.33. This is historically moderately high, reflecting the uncertainty of the period. The results we give below are not very sensitive to this choice for we will see that 1914 is not a particularly volatile year in this dataset.

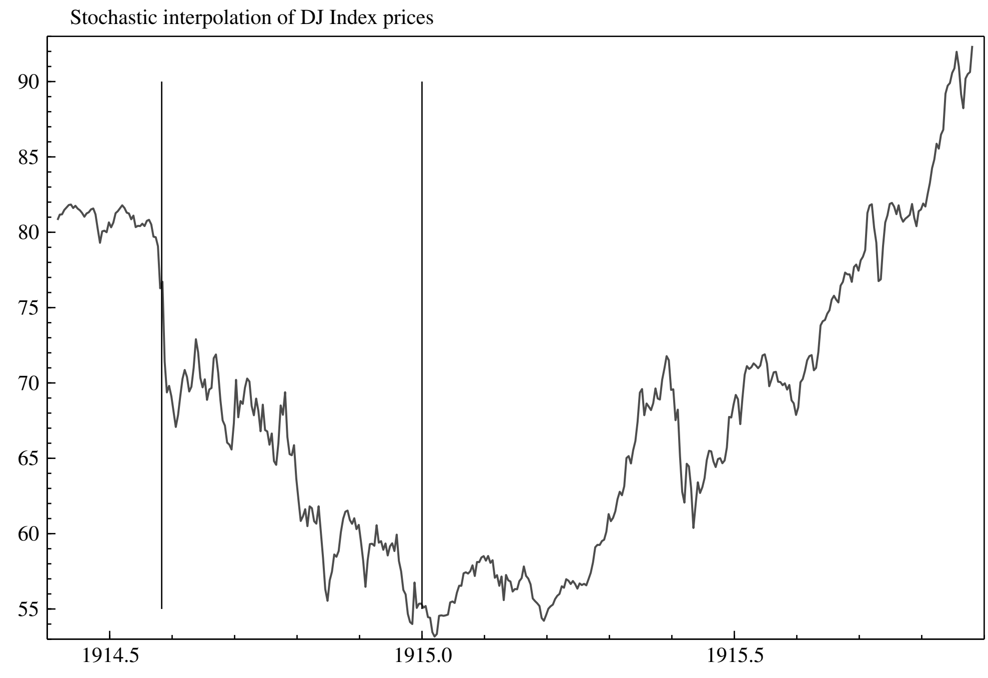

Figure 4. Between the vertical lines prices are interpolated using a Brownian bridge (on the log scale). Code: **schwert.ox**

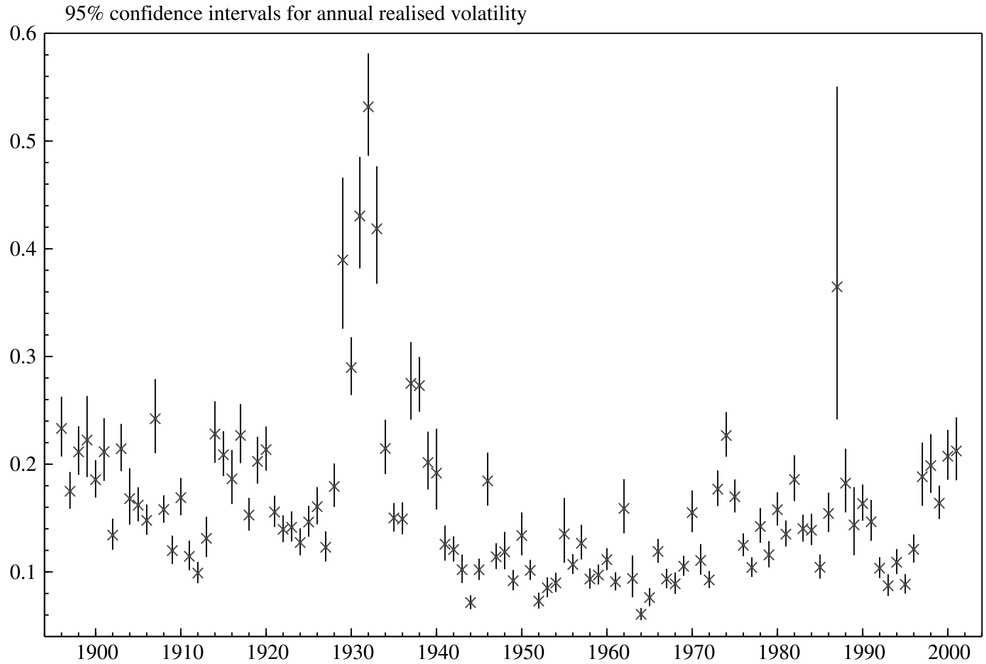

Figure 5. Annual realized volatility $\sqrt{\sum_{j=1}^{M} y_{j,n}^2}$ for the Dow Jones Industrial Average (marked with crosses) together with 95% confidence intervals. Code: **schwert.ox**

The realized volatilities and their 95% confidence intervals are given in Figure 5. The confidence intervals use the log-based limit theory given in (6). The results again reflect the tendency for the intervals to be wide when the level of volatility is high. However, the results are more varied in this case than in the high-frequency analysis we gave for the exchange rate data. In particular the volatility spike in 1987 is poorly measured for it is caused by high levels of price movements over a very short time interval. There is not enough data in the daily observations to pin down precisely the level of volatility in this case. In the 1930s, on the other hand, the high level of movements was sustained over a long time interval and so we produce quite a precise estimate of the level of volatility.

5. CONCLUSION

This paper has reviewed some recent work on the properties of RV. For the first time we have provided a measure of precision for empirical examples of RV based on a rather flexible SV model. We have seen that the confidence intervals for $\sigma_i^{[2]}$ are typically quite wide even when our estimates are based on 10-minute return data. This overturns a widely held view that integrated variance can be measured without much error. Our analysis shows that the error can be very large and is likely to be so when the volatility in the market is high. This will mean that when econometricians use realized variance they need to be very careful of the measurement error. We hope that the asymptotic analysis we have provided will help them in dealing with this task.

ACKNOWLEDGMENTS

Ole E. Barndorff-Nielsen's work is supported by CAF (**www.caf.dk**), which is funded by the Danish Social Science Research Council, and by MaPhySto (**www.maphysto.dk**), funded by the Danish National Research Foundation. Neil Shephard's research is supported by the UK's ESRC through the grant 'Econometrics of trade-by-trade price dynamics', which is coded R00023839. All the calculations made in this paper are based on software written by the authors using the Ox language of Doornik (2001). We owe a large debt to Michel M. Dacorogna at the Olsen Group for allowing us to use Olsen's high-frequency exchange rate data in our study and Tim Bollerslev for supplying us with a semi-cleaned version of this data. We thank Torben Andersen, Tim Bollerslev, Frank Diebold, Eric Ghysels and Nour Meddahi for both encouragement and stimulating conversations on this topic. Finally, the comments of the referees on the initial draft of the paper helped us to improve our exposition.

REFERENCES

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

Andersen TG, Bollerslev T, Diebold FX, Labys P. 2000. Great realizations. *Risk* **13**: 105–108.

Andersen TG, Bollerslev T, Diebold FX. 2002. Parametric and nonparametric measurement of volatility. In *Handbook of Financial Econometrics*, Ait-Sahalia Y, Hansen LP (eds). North-Holland: Amsterdam; (forthcoming).

Andersen TG, Bollerslev T, Diebold FX, Ebens H. 2001a. The distribution of realised stock return volatility. *Journal of Financial Economics* **61**: 43–76.

Andersen TG, Bollerslev T, Diebold FX, Labys P. 2001b. The distribution of exchange rate volatility. *Journal of the American Statistical Association* **96**: 42–55.

Andersen TG, Bollerslev T, Diebold FX, Labys P. 2001c. Modeling and forecasting realized volatility. Unpublished paper: Department of Economics, Duke University.

Andreou E, Ghysels E. 2002. Rolling-sampling volatility estimators: some new theoretical, simulation and empirical results. *Journal of Business and Economic Statistics* **20**: 363–376.

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

Barndorff-Nielsen OE. 1998. Processes of normal inverse Gaussian type. *Finance and Stochastics* **2**: 41–68.

Barndorff-Nielsen OE, Shephard N. 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 OE, Shephard N. 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 OE, Shephard N. 2002a. How accurate is the asymptotic approximation to the distribution of realized variance? In *Identification and Inference for Econometric Models. A Festschrift in Honour of Thomas J. Rothenberg*, Andrews D, Powell J, Ruud PA, Stock JH (eds). Econometric Society Manograph, Cambridge University Press: Cambridge, (forthcoming).

Barndorff-Nielsen OE, Shephard N. 2002b. Realized power variation and stochastic volatility. Bernoulli, (forthcoming).

Carr *et al*. 2002. Stochastic volatility for Lévy processes. *Mathematical Finance*, (forthcoming).

Chriss N, Morokoff W. 1999. Volatility and variance swaps. *Risk* **12**: 55–59.

Christensen BJ, Prabhala NR. 1998. The relation between implied and realized volatility. *Journal of Financial Economics* **37**: 125–150.

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

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

Doornik JA. 2001. *Ox: Object Oriented Matrix Programming, 3.0*. Timberlake Consultants Press: London.

Eberlein E. 2001. Application of generalized hyperbolic Lévy motion to finance. In *Lévy Processes–Theory and Applications*, Barndorff-Nielsen OE, Mikosch T, Resnick S (eds). Birkhäuser: Boston, MA; 319–337.

Ghysels E, Harvey AC, Renault E. 1996. Stochastic volatility. In *Statistical Methods in Finance*, Rao CR, Maddala GS (eds). North-Holland: Amsterdam; 119–191.

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

Hull J, White A. 1987. The pricing of options on assets with stochastic volatilities. *Journal of Finance* **42**: 281–300.

Jacod J, Shiryaev AN. 1987. *Limit Theorems for Stochastic Processes*. Springer-Verlag: Berlin.

Madan DB, Seneta E. 1990. The VG model for share market returns. *Journal of Business* **63**: 511–524.

Meddahi N. 2002. A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics* **17**: 479–508.

Merton RC. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* **8**: 323–361.

Poterba J, Summers L. 1986. The persistence of volatility and stock market fluctuations. *American Economic Review* **76**: 1124–1141.

Protter P. 1990. *Stochastic Integration and Differential Equations: A New Approach*. Springer-Verlag: New York.

Renault E. 1997. Econometric models of option pricing errors. In *Advances in Economics and Econometrics: Theory and Applications*, Kreps DM, Wallis KF (eds). Cambridge University Press: Cambridge; 223–278.

Richardson MP, Stock JH. 1989. Drawing inferences from statistics based on multi-year asset returns. *Journal of Financial Economics* **25**: 323–348.

Schwert GW. 1989. Why does stock market volatility change over time? *Journal of Finance* **44**: 1115–1153.

Schwert GW. 1990. Indexes of U.S. stock prices from 1802 to 1987. *Journal of Business* **63**: 399–426.

Schwert GW. 1998. Stock market volatility: Ten years after the crash. *Brookings–Wharton Papers on Financial Services* **1**: 65–114.

Shiryaev AN. 1999. *Essentials of Stochastic Finance: Facts, Models and Theory*. World Scientific: Singapore.

Taylor SJ, Xu X. 1997. The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance* **4**: 317–340.

Winkel M. 2001. The recovery problem for time-changed Lévy processes. MaPhySto Research Report 2001-37, Department of Mathematical Sciences, Aarhus University.

