Econometric analysis of realised volatility and its use in estimating stochastic volatility models

Ole E. Barndorff-Nielsen

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

oebn@mi.aau.dk

Neil Shephard

Nuffield College, Oxford OX1 1NF, UK

neil.shephard@nuf.ox.ac.uk

First submitted November 2000 (First circulated 26th October 2000).
This version November 2001.

Abstract

The availability of intra-day data on the prices of speculative assets means that we can use quadratic variation like measures of activity in financial markets, called realised volatility, to study the stochastic properties of returns. Here, under the assumption of a rather general stochastic volatility model, we derive the moments and the asymptotic distribution of the realised volatility error — the difference between realised volatility and the discretised integrated volatility (which we call actual volatility). These properties can be used to allow us to estimate the parameters of stochastic volatility models without the recourse to the use of simulation intensive methods.

Keywords: Kalman filter; Leverage; Lévy process; Power variation; Quadratic variation; Realised volatility; Stochastic volatility; Subordination; Superposition.

1 Introduction

1.1 Stochastic volatility

In the stochastic volatility (SV) model for log-prices of stocks and for log exchange rates a basic Brownian motion is generalised to allow the volatility term to vary over time. Then the log-price $y^*(t)$ follows the solution to the stochastic differential equation (SDE),

$$\mathrm{d}y^*(t) = \left\{\mu + \beta\sigma^2(t)\right\}\mathrm{d}t + \sigma(t)\mathrm{d}w(t), \tag{1}$$

where $\sigma^2(t)$, the *instantaneous* or *spot volatility*, is going to be assumed to (almost surely) have locally square integrable sample paths, while being stationary and stochastically independent of the standard Brownian motion $w(t)$. We will label μ the drift and β the risk premium. Over an interval of time of length $\Delta > 0$ returns are defined as

$$y_n = y^*(\Delta n) - y^*((n-1)\Delta), \qquad n = 1, 2, ..., \tag{2}$$

which implies that whatever the model for σ^2, it follows that

$$y_n|\sigma_n^2 \sim N(\mu\Delta + \beta\sigma_n^2, \sigma_n^2),$$

where

$$\sigma_n^2 = \sigma^{2*}(n\Delta) - \sigma^{2*}\{(n-1)\Delta\}, \quad \text{and} \quad \sigma^{2*}(t) = \int_0^t \sigma^2(u)\mathrm{d}u.$$

In econometrics $\sigma^{2*}(t)$ is called *integrated volatility*, while we call σ_n^2 *actual volatility*. Both definitions play a central role in the probabilistic analysis of SV models. Reviews of the literature on this topic are given in Taylor (1994), Shephard (1996) and Ghysels, Harvey, and Renault (1996), while statistical and probabilistic aspects are studied in detail in Barndorff-Nielsen and Shephard (2001a). One of the key results in this literature (Barndorff-Nielsen and Shephard (2001a)) is that if we write (when they exist) ξ, ω^2 and r, respectively, as the mean, variance and the autocorrelation function of the process $\sigma^2(t)$ then

$$\mathrm{E}\left(\sigma_n^2\right) = \xi\Delta, \quad \mathrm{Var}\left(\sigma_n^2\right) = 2\omega^2 r^{**}(\Delta) \qquad \text{and} \qquad \mathrm{Cov}\{\sigma_n^2, \sigma_{n+s}^2\} = \omega^2 \diamondsuit r^{**}(\Delta s), \tag{3}$$

where

$$\diamondsuit r^{**}(s) = r^{**}(s+\Delta) - 2r^{**}(s) + r^{**}(s-\Delta), \tag{4}$$

and

$$r^*(t) = \int_0^t r(u)\mathrm{d}u \quad \text{and} \quad r^{**}(t) = \int_0^t r^*(u)\mathrm{d}u. \tag{5}$$

That is the second order properties of $\sigma^2(t)$ completely determine the second order properties of actual volatility.

One of the most important aspects of SV models is that $\sigma^{2*}(t)$ can be exactly recovered using the entire path of $y^*(t)$. In particular, for the above SV model the *quadratic variation* is $\sigma^{2*}(t)$, i.e. we have

$$[y^*](t) = \text{p}-\lim_{q\to\infty} \sum \{y^*(t_{i+1}^q) - y^*(t_i^q)\}^2 = \sigma^{2*}(t) \tag{6}$$

for any sequence of partitions $t_0^q = 0 < t_1^q < ... < t_{m_r}^q = t$ with $\sup_i\{t_{i+1}^q - t_i^q\} \to 0$ for $q \to \infty$. Here $\text{p} - \lim$ denotes the probability limit of the sum. This is a powerful result for it does not depend upon the model for instantaneous volatility nor the drift terms in the SDE for log-prices given in (1). The quadratic variation estimation of integrated volatility has recently been highlighted, following the initial draft of Barndorff-Nielsen and Shephard (2001a) and the concurrent independent work of Andersen and Bollerslev (1998a), by Andersen, Bollerslev, Diebold, and Labys (2001) and Maheu and McCurdy (2001) in foreign exchange markets and Andersen, Bollerslev, Diebold, and Ebens (2001) and Areal and Taylor (2002) in equity markets. See also the contribution of Comte and Renault (1998).

In practice, although we often have a continuous record of quotes or transaction prices, at a very fine level the SV model is a poor fit to the data. This is due to market microstructure effects (e.g. discreteness of prices, bid/ask bounce, irregular trading etc. See Bai, Russell, and Tiao (2000)). As a result we should regard the above quadratic variation result as indicating that we can estimate actual volatility, for example over a day, reasonably accurately by sums of squared returns, say, using thirty minute periods but keeping in mind that taking returns over increasingly finer time periods will lead to the introduction of important biases. Hence the limit argument in quadratic variation is interesting but of limited direct practical use on its own. Suppose instead that we have a fixed M intra-day observations during each day, then the sum of squared intra-day changes over a day is

$$\{y\}_n = \sum_{j=1}^{M} \left\{ y^* \left((n-1)\Delta + \frac{\Delta j}{M} \right) - y^* \left((n-1)\Delta + \frac{\Delta (j-1)}{M} \right) \right\}^2, \tag{7}$$

which is an estimate of σ_n^2. It is a consistent estimate as $M \to \infty$, while it is unbiased when μ and β are zero. In econometrics $\{y\}_n$ has recently been labelled *realised volatility*, and we will follow that convention here. Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001) and Andersen, Bollerslev, Diebold, and Labys (2000) have empirically studied the properties of $\{y\}_n$ in foreign exchange and equity markets (earlier, less formal work on this topic includes Poterba and Summers (1986), Schwert (1989), Taylor and Xu (1997) and Christensen and Prabhala (1998)). In their econometric analysis they have regarded $\{y\}_n$ as a very accurate estimate of σ_n^2. Indeed they often regard the estimate as basically revealing the true value of actual volatility so that $y_n/\sqrt{\{y\}_n}$ is more or less Gaussian.

So far no measure of error has been obtained which indicates the difference between $\{y\}_n$ and σ_n^2. We will show that this difference is approximately mixed Gaussian, can be substantial and that more accurate estimates of σ_n^2 are readily available if we are prepared to use a model for $\sigma^2(t)$. Andreou and Ghysels (2001) have independently approximated the properties of realised volatility using the methods of Foster and Nelson (1996) in their study of rolling estimators of the spot volatility $\sigma^2(t)$.

In this paper we will discuss a simple way of formally bridging the gap between realised and actual volatility, providing a discussion of the properties of $\{y\}_n$ which has so far been lacking in the literature. Inevitably for finite M these properties will depend upon the dynamics of the instantaneous volatility as well as the drift term in the SDE for log-prices. This has to be the case, for the short-hand of ignoring the small sample effects of estimating σ_n^2 with the consistent $\{y\}_n$ is only valid for infeasibly large values of M.

In particular the contribution of our paper will be to allow us to:

- derive the asymptotic distribution of $\sqrt{M}\left(\{y\}_n - \sigma_n^2\right)$ for large M, showing that this does not depend upon μ and β;
- analyse the properties of realised volatility by assuming $\mu = \beta = 0$ as the corresponding error has been shown to be small;
- understand the exact second order properties of $\{y\}_n$ when $\mu = \beta = 0$;
- use the models for instantaneous volatility to provide *model based* estimates of actual volatility (rather than model free estimates which assume $M \to \infty$) using the series of realised volatility measurements when $\mu = \beta = 0$. These model based estimates can be based on past, current or historical sequences of realised volatilities; and
- estimate the parameters of SV models using simple and rather accurate statistical procedures when $\mu = \beta = 0$.

1.2 Empirical example

To illustrate some of the empirical features of realised volatility we have used the same return data as employed by Andersen, Bollerslev, Diebold, and Labys (2001), although the Appendix will describe the slightly different adjustments we have made to deal with some missing data. This United States Dollar/ German Deutsche Mark series covers the ten year period from 1st December 1986 until 30th November 1996. Every five minutes it records the most recent quote to appear on the Reuters screen. It has been kindly supplied to us by Olsen and Associates in

Zurich and preprocessed by Tim Bollerslev. It will be used throughout our paper to illustrate the results we will develop. In the top left graph in Figure 1 we have drawn the correlogram of the squared five minute returns over the ten years sample. It shows the well known very strong diurnal effect (the x-axis is drawn in days). This will be discussed in detail in Section 6 but for now will be ignored entirely. The graph on the top right of the Figure shows the correlogram of realised volatility, $\{y\}_n$, computed using $M = 288$ (i.e. based on five minute data) and again using the whole ten years of data. The graph starts out at around 0.6, decays very quickly for a number of days and then decays at a slower rate. The graphs on the bottom show a cumulative version of the squared five minute returns drawn on a small scale, while on the right the same cumulative function is drawn over a larger time scale. It is the daily increments of this process which makes up realised volatility.

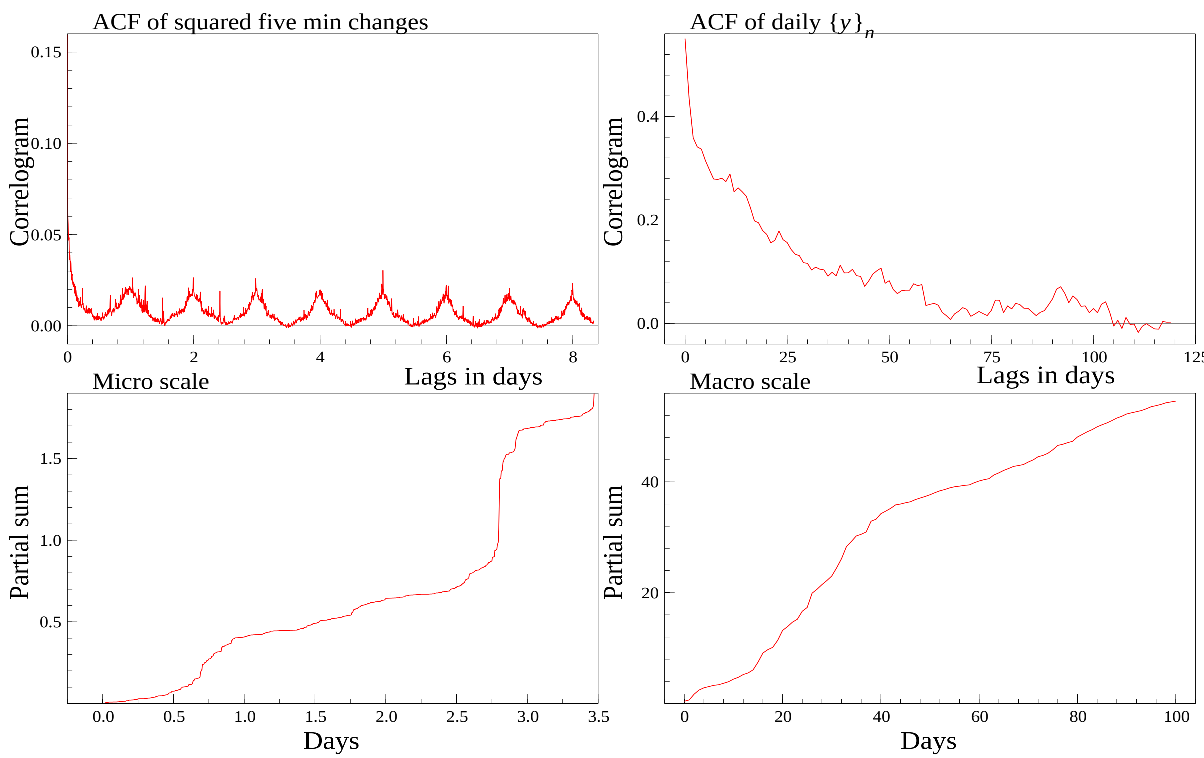

Figure 1: *Summary statistics for Olsen group's five minute changes data. Top left: ACF of five minute returns. Bottom left: cumulative sum of squared 5 minute changes over short interval, where the x-axis is marked off in days. Top right: ACF of realised volatility, where the x-axis is marked off in days. Bottom right: cumulative sum of squared five minute changes over long interval.*

1.3 Outline of the paper

The outline of the rest of the paper is as follows. In Section 2 we discuss the basic approach in the most straightforward setup where μ and β are zero, providing the second order properties of realised volatility. These can be used in estimating the value of actual volatility from a time series of realised volatilities. This is discussed in Section 3, which also contains a discussion of using the realised volatilities to provide estimates of continuous stochastic volatility models. Section 4 gives an empirical illustration of the methods developed in the previous two sections. Section 5 provides the asymptotic distribution of $\sqrt{M}\left(\{y\}_n - \sigma_n^2\right)$, covering the case where there is drift and a risk premium. This section shows that the effect on realised volatility of the drift and the risk premium is extremely small. Section 6 studies diurnal effects and leverage extensions. Section 7 concludes, while the Appendix contains a discussion of the dataset used in this paper together with a proof of Lemma 1 and Theorem 1.

2 Relating actual to realised volatility

2.1 Generic results

Actual volatility, σ_n^2, plays a crucial role in SV models. It can be estimated using realised volatility, $\{y\}_n$, given in (7). Here we discuss this in the simplest context where $\mu = \beta = 0$, delaying our discussion of the effect on $\{y\}_n$ of the drift and risk premium until Section 5. In that section we will show that the effect is minor and so the results we develop here will still be important in that wider case.

In SV models we can always make the decomposition

$$\{y\}_n = \sigma_n^2 + u_n, \qquad \text{where} \qquad u_n = \{y\}_n - \sigma_n^2. \tag{8}$$

Here we call u_n the *realised volatility error*, which has the property that $\mathrm{E}\left(u_n|\sigma_n^2\right) = 0$. Hence realised volatility is an unbiased estimator of actual volatility. We know that as $M \to \infty$ then $\{y\}_n \overset{a.s.}{\to} \sigma_n^2$, so it also consistent. However, the purpose of this section is to discuss the properties of $\{y\}_n$ for finite M. We can see that

$$\mathrm{E}\left(\{y\}_n\right) = \Delta\xi, \qquad \mathrm{Var}\left(\{y\}_n\right) = \mathrm{Var}(u_n) + \mathrm{Var}(\sigma_n^2), \qquad \mathrm{Cov}(\{y\}_n, \{y\}_{n+s}) = \mathrm{Cov}(\sigma_n^2, \sigma_{n+s}^2).$$

Further, writing

$$\sigma_{j,n}^2 = \sigma^{2*}\left((n-1)\Delta + \frac{\Delta j}{M}\right) - \sigma^{2*}\left((n-1)\Delta + \frac{\Delta(j-1)}{M}\right)$$

we have that

$$u_n \overset{\mathcal{L}}{=} \sum_{j=1}^{M} \sigma_{j,n}^2 \left(\varepsilon_{j,n}^2 - 1\right),$$

where $\varepsilon_{j,n} \overset{i.i.d.}{\sim} N(0,1)$ and independent of $\left\{\sigma^2_{j,n}\right\}$. It is clear that $\{u_n\}$ is a weak white noise sequence which is uncorrelated to the actual volatility series $\{\sigma^2_n\}$.

Now unconditionally,

$$\begin{aligned} \mathrm{Var}(u_n) &= 2M\mathrm{E}\left\{\left(\sigma^2_{1,n}\right)^2\right\} \\ &= 2M\left\{\mathrm{Var}\left(\sigma^2_{1,n}\right) + \mathrm{E}\left(\sigma^2_{1,n}\right)^2\right\}, \end{aligned} \tag{9}$$

for $\sigma^2_{1,n}$ has the same marginal distribution as each element of $\left\{\sigma^2_{j,n}\right\}$. In general we have from (3) that

$$\mathrm{E}\left(\sigma^2_{1,n}\right) = \Delta M^{-1}\xi, \quad \mathrm{Var}\left(\sigma^2_{1,n}\right) = 2\omega^2 r^{**}\left(\Delta M^{-1}\right). \tag{10}$$

Hence we can compute $\mathrm{Var}(u_n)$ for all SV models when $\mu = \beta = 0$. In turn, having established the second order properties of σ^2_n and u_n, we can immediately use the results in Whittle (1983) to provide best linear prediction and smoothing results for the unobserved actual volatilities σ^2_n from the time series of realised volatilities $\{y\}_n$. The only issues which remain on this front are computational. Otherwise this covers all covariance stationary models for $\sigma^2(t)$ — including long-memory processes.

One of the implications of the result given above is that

$$\begin{aligned} \mathrm{Cor}(\{y\}_n, \{y\}_{n+s}) &= \frac{\mathrm{Cov}(\sigma^2_n, \sigma^2_{n+s})}{\mathrm{Var}(u_n) + \mathrm{Var}(\sigma^2_n)} \\ &= \frac{\omega^2 \diamondsuit r^{**}(\Delta s)}{2M^{-1}\left\{2\omega^2 M^2 r^{**}\left(\Delta M^{-1}\right) + (\Delta\xi)^2\right\} + 2\omega^2 r^{**}(\Delta)}. \end{aligned}$$

Notice that

$$\mathrm{Cor}(y^2_n, y^2_{n+s}) = \frac{\omega^2 \diamondsuit r^{**}(\Delta s)}{2\left\{2\omega^2 r^{**}(\Delta) + (\Delta\xi)^2\right\} + 2\omega^2 r^{**}(\Delta)},$$

can be derived from this result, for $\{y\}_n = y^2_n$ when $M = 1$. Hence the decay rates in the autocorrelation function of $\{y\}_n$, σ^2_n and y^2_n are the same in general but the correlation varies considerably, being the highest for σ^2_n, followed by $\{y\}_n$ and lowest for y^2_n.

In practice we tend to use realised volatility measures with M being moderately large. Hence it is of interest to think of a central limit approximation to the distribution of u_n. This will depend upon the limit of $t^{-2}r^{**}(t)$ as $t \to 0$ from above. Now, by Taylor expansion

$$r^{**}(t) = r^{**}(0+) + tr^{*}(0+) + \frac{1}{2}t^2 r(0+) + o(t^2) = \frac{1}{2}t^2 r(0+) + o(t^2).$$

This means the limit of $t^{-2}r^{**}(t)$ is 1/2. A consequence of this is that

$$\lim_{M\to\infty} M^2 \mathrm{Var}\left(\sigma^2_{1,n}\right) = \Delta^2\omega^2, \tag{11}$$

implying that, as M goes to infinity,

$$\mathrm{Var}\left(\sqrt{M}u_n\right) = \mathrm{Var}\left\{\sqrt{M}\left(\{y\}_n - \sigma_n^2\right)\right\} \to 2\Delta^2\left(\omega^2 + \xi^2\right).$$

This is an important result. We have moved away from the standard consistency result of $\{y\}_n \xrightarrow{p} \sigma_n^2$ as $M \to \infty$ which follows from familiar quadratic variation results. Now we have the more refined measure of the uncertainty of this error term.

2.2 Simple example

Suppose the volatility process has the autocorrelation function $r(t) = \exp(-\lambda|t|)$. Here we recall two classes of processes which have this property. The first is the constant elasticity of variance (CEV) process which is the solution to the SDE

$$\mathrm{d}\sigma^2(t) = -\lambda\left\{\sigma^2(t) - \xi\right\}\mathrm{d}t + \omega\sigma(t)^\eta \mathrm{d}b(\lambda t), \qquad \eta \in [1,2],$$

where $b(t)$ is standard Brownian motion uncorrelated with $w(t)$. Of course the special case of $\eta = 1$ delivers the square root process, while when $\eta = 2$ we have Nelson's GARCH diffusion. These models have been heavily favoured by Meddahi and Renault (2002) in this context. The second process is the non-Gaussian Ornstein–Uhlenbeck, or OU type for short, process which is the solution to the SDE

$$\mathrm{d}\sigma^2(t) = -\lambda\sigma^2(t)\mathrm{d}t + \mathrm{d}z(\lambda t), \tag{12}$$

where $z(t)$ is a Lévy process with non-negative increments. These models have been developed in this context by Barndorff-Nielsen and Shephard (2001a). In Figure 2 we have drawn a curve to represent a simulated sample path of σ_n^2 from an OU process where $\sigma^2(t)$ has a $\Gamma(4,8)$ stationary distribution, $\lambda = -\log(0.99)$ and $\Delta = 1$, along with the associated realised volatility (depicted using crosses) computed using a variety of values of M. We see that as M increases the precision of realised volatility increases, while Figure 2.d shows that the variance of the realised volatility error increases with the volatility, a result we will come back to in Section 5 where the asymptotic distribution we develop for $\sqrt{M}u_n$ will reflect this feature.

For both CEV and OU models

$$r^{**}(t) = \lambda^{-2}\left\{e^{-\lambda|t|} - 1 + \lambda t\right\} \qquad \text{and} \qquad \Diamond r^{**}(\Delta s) = \lambda^{-2}(1 - e^{-\lambda\Delta})^2 e^{-\lambda\Delta(s-1)}, \quad s > 0.$$

This implies that

$$\mathrm{E}\left(\sigma_n^2\right) = \Delta\xi, \qquad \mathrm{Var}\left(\sigma_n^2\right) = \frac{2\omega^2}{\lambda^2}\left(e^{-\lambda\Delta} - 1 + \lambda\Delta\right),$$

and

$$\mathrm{Cor}\{\sigma_n^2, \sigma_{n+s}^2\} = de^{-\lambda\Delta(s-1)}, \qquad s = 1, 2, ..., \tag{13}$$

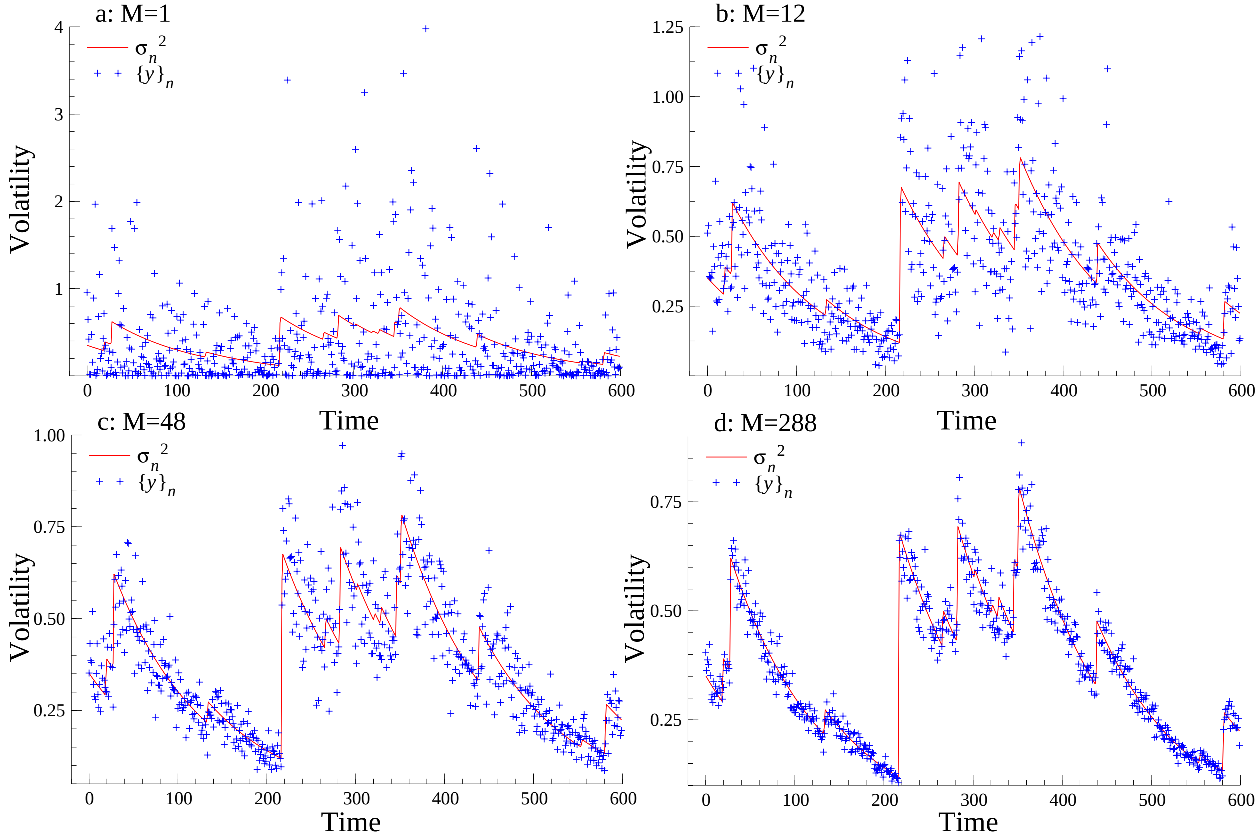

Figure 2: *Actual* σ_n^2 *and realised* $\{y\}_n$ *(with* M varying*) volatility based upon a* $\Gamma(4,8)$-OU process with $\lambda = -\log(0.98)$ and $\Delta = 1$. This implies $\xi = 0.5$ and $\xi\omega^{-2} = 8$.

where

$$d = \frac{(1-e^{-\lambda\Delta})^2}{2\left(e^{-\lambda\Delta} - 1 + \lambda\Delta\right)} \leq 1.$$

Finally

$$\begin{aligned} \operatorname{Var}(u_n) &= 2M\left\{\operatorname{Var}\left(\sigma_{1,n}^2\right) + \operatorname{E}\left(\sigma_{1,n}^2\right)^2\right\} \\ &= 2M\left\{2\omega^2\lambda^{-2}\left(e^{-\lambda\Delta/M} - 1 + \lambda\Delta M^{-1}\right) + \left(\Delta M^{-1}\right)^2\xi^2\right\}. \end{aligned} \quad (14)$$

Importantly the above analysis implies that actual volatility has the autocorrelation function of an autoregressive moving average (ARMA) model of order $(1,1)$. Its autoregressive root is $e^{-\lambda\Delta}$, which will be typically close to one unless Δ is very large, while the moving average root is also determined by $e^{-\lambda\Delta}$ but has to be found numerically. A graph of the moving average root against $e^{-\lambda\Delta}$ is given in the left hand side of Figure 3 and shows that for a wide range of the autoregressive root the moving average root is around 0.265. Likewise Figure 3 shows a plot of d against $e^{-\lambda\Delta}$ and indicates a rapid decline in this coefficient as the autoregressive root

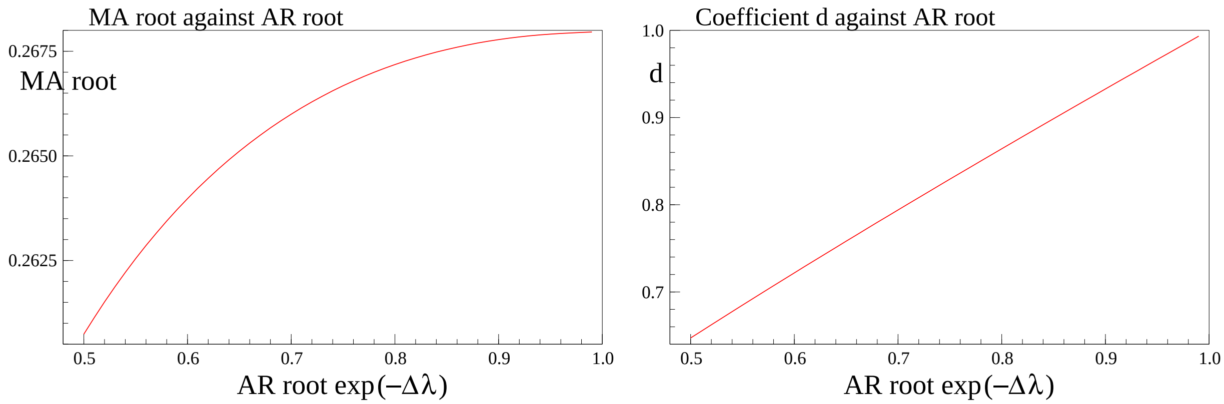

Figure 3: *Left graph shows plot of moving average root against autoregressive root* $e^{-\Delta\lambda}$ *for ARMA(1,1) representation. Right graph shows* d *in expression for* $Cor\{\sigma_n^2, \sigma_{n+s}^2\}$ *against autoregressive root* $e^{-\Delta\lambda}$.

falls. In particular, in financial econometrics the literature suggests volatility is quite persistent, which would imply that d should be close to one. Thus if t is recorded in days and Δ is set to one day, then empirically reasonable values of λ will imply that d should be close to one.

In turn the autocorrelation function for σ_n^2 implies that the squares of returns have autocorrelations of the form

$$\text{Cor}\{y_n^2, y_{n+s}^2\} = c' e^{-\lambda\Delta(s-1)}, \tag{15}$$

where

$$\frac{1}{3} \geq \frac{1}{3}d \geq c' = \frac{(1-e^{-\lambda\Delta})^2}{6\left\{e^{-\lambda\Delta} - 1 + \lambda\Delta\right\} + 2(\lambda\Delta)^2\left(\xi/\omega\right)^2} \geq 0.$$

This means y_n^2 also has a linear ARMA(1,1) representation. Further, it has the same autocorrelation function as the familiar generalised autoregressive conditional heteroskedastic (GARCH) model used extensively in econometrics (see, for example, Bollerslev, Engle, and Nelson (1994)). Finally, the autoregressive root of the ARMA representation is the same for y_n^2 as for σ_n^2, however the moving average root of the square changes is much larger in absolute value. The implication is that the correlograms for y_n^2 will be less clear than if we had observed the correlograms of the σ_n^2. This can be most easily seen by noting that for small λ,

$$c' \simeq \frac{1-\lambda\Delta}{3+2\left(\xi/\omega\right)^2},$$

which is much smaller than d which is approximately $1-\lambda\Delta$. For example if the $\xi = \omega$, then c' will be approximately 0.2 for daily data.

2.3 Extension of the example: superpositions

The OU/CEV volatility models are often too simple to accurately fit the types of dependence structures we observe in financial economics. This can be seen in the top right piece of Figure 1 which displays the autocorrelation function of realised volatility for the Olsen group's five minute data. This graph shows a relatively quick initial decline in the autocorrelation function, followed by a slower decay. This single observation is sufficient to dismiss the OU/CEV models.

One mathematically tractable way of improving the flexibility of the volatility model is to let the instantaneous volatility be the sum, or superposition, of independent OU or CEV processes. As the processes do not need to be identically distributed, this offers a great deal of flexibility while still being mathematically tractable. Superpositions of such processes also have potential for modelling long-range dependence and self-similarity in volatility. This is discussed in the OU case in Barndorff-Nielsen and Shephard (2001a) and in more depth by Barndorff-Nielsen (2001) who formalises the use of superpositions as a way of modelling long-range dependence. This follows earlier related work by Granger (1980), Barndorff-Nielsen, Jensen, and Sørensen (1990), Cox (1991), Ding and Granger (1996), Engle and Lee (1999), Shephard (1996, pp. 36–37), Andersen and Bollerslev (1997a), Barndorff-Nielsen (1998) and Comte and Renault (1998).

Consider volatility based on the sum of J independent OU or CEV processes

$$\sigma^2(t) = \sum_{i=1}^{J} \tau^{(i)}(t),$$

where the $\tau^{(i)}(t)$ process has the memory parameter λ_i. We assume

$$\mathrm{E}(\tau^{(i)}(t)) = w_i\xi, \qquad \mathrm{Var}(\tau^{(i)}(t)) = w_i\omega^2,$$

where

$$\{w_i \geq 0\} \quad \text{and} \quad \sum_{i=1}^{J} w_i = 1,$$

implying that $\mathrm{E}(\sigma^2(t)) = \xi$ and $\mathrm{Var}(\sigma^2(t)) = \omega^2$. The implication is that

$$\mathrm{Cov}(\sigma^2(t), \sigma^2(t+s)) = \sum_{i=1}^{J} \mathrm{Cov}\left(\tau^{(i)}(t), \tau^{(i)}(t+s)\right) = \omega^2 \sum_{i=1}^{J} w_i \exp\left(-\lambda_i\left|s\right|\right).$$

Hence the autocorrelation function of instantaneous volatility can have components which are a mix of quickly and slowly decaying components. For fixed J the statistical identification of this model (imposing a constraint like $\lambda_1 < ... < \lambda_J$) is a consequence of the form of the autocorrelation function and the uniqueness of the Laplace transformation.

The linearity of the superposition of OU processes means that actual volatility has the form $\sigma_n^2 = \sum_{i=1}^{J} \tau_n^{(i)}$ where

$$\tau_n^{(i)} = \tau^{(i)*}(n\Delta) - \tau^{(i)*}\left\{(n-1)\Delta\right\}, \quad \text{and} \quad \tau^{(i)*}(t) = \int_0^t \tau^{(i)}(u)\mathrm{d}u.$$

The key feature is that each $\tau_n^{(i)}$ has an ARMA$(1,1)$ representation of the type discussed above. As the autocovariance function of a sum of independent components is the sum of the autocovariances of the terms in the sum, we can compute the autocorrelation function of σ_n^2 without any new work. Computationally it is helpful to realise that the sum of uncorrelated ARMA$(1,1)$ processes can be fed into a linear state space representation when combined with (8). The only new issue is computing

$$\mathrm{Var}(u_t) = 2M\left\{\mathrm{Var}\left(\sigma_{1,n}^2\right) + \mathrm{E}\left(\sigma_{1,n}^2\right)^2\right\}.$$

Clearly $\mathrm{E}\left(\sigma_{1,n}^2\right) = \xi\Delta M^{-1}$ while

$$\begin{aligned}
\mathrm{Var}\left(\sigma_{1,n}^2\right) &= \sum_{i=1}^{J}\mathrm{Var}\left(\tau_{1,n}^{(i)}\right) = 2\omega^2\sum_{i=1}^{J} w_i r_i^{**}\left(\Delta M^{-1}\right) \\
&= 2\omega^2\sum_{i=1}^{J}\frac{w_i}{\lambda_i^2}\left\{e^{-\lambda_i\Delta M^{-1}} - 1 + \lambda_i\Delta M^{-1}\right\} \\
&= 2\omega^2\sum_{i=1}^{J}\frac{w_i}{2\lambda_i^2}\left(\lambda_i\Delta M^{-1}\right)^2 + o(M^{-2}).
\end{aligned}$$

Importantly, for large M this expression simplifies and so we again obtain

$$\mathrm{Var}\left(\sqrt{M}u_n\right) \to 2\Delta^2\left(\omega^2+\xi^2\right), \qquad \text{as} \qquad M\to\infty.$$

3 Efficiency gains: model based estimators of volatility

3.1 State space representation

If $\sigma^2(t)$ is OU or CEV then σ_n^2 has an ARMA$(1,1)$ representation and so it is computationally convenient to place (8) into a linear state space representation (see, for example, Harvey (1989, Ch. 3) and Hamilton (1994, Ch. 13)). In particular we write $\alpha_{1n} = (\sigma_n^2 - \Delta\xi)$ and $u_n = \sigma_u v_{1n}$, then the state space is explicitly

$$\begin{aligned}
\{y\}_n &= \Delta\xi + (1 \quad 0)\,\alpha_n + \sigma_u v_{1n}, \\
\alpha_{n+1} &= \begin{pmatrix}\phi & 1\\ 0 & 0\end{pmatrix}\alpha_n + \begin{pmatrix}\sigma_\sigma \\ \sigma_\sigma\theta\end{pmatrix}v_{2n},
\end{aligned} \qquad (16)$$

where v_n is a zero mean, white noise sequence with an identity covariance matrix. The parameters ϕ, θ and σ_σ^2 represent the autoregressive root, the moving average root and the variance of the innovation to this process, while σ_u^2 is found from (9) and (10). Software for handling linear state space models is available in Koopman, Shephard, and Doornik (1999). Having constructed this representation we can use a Kalman filter to unbiasedly and efficiently (in a linear sense) estimate σ_n^2 by prediction (that is the estimate of σ_n^2, using $\{y\}_1, ..., \{y\}_{n-1}$) and smoothing (that is the estimate of σ_n^2, using $\{y\}_1, ..., \{y\}_T$ where T is the sample size). Biproducts of the

Kalman filter are the mean square errors of these *model based* (that is they depend upon the assumption that σ_n^2 has an ARMA$(1,1)$ representation) estimators.

M	$\xi = 0.5$, $\xi\omega^{-2} = 8$			$\xi = 0.5$, $\xi\omega^{-2} = 4$			$\xi = 0.5$, $\xi\omega^{-2} = 2$		
$e^{-\Delta\lambda} = 0.99$	Smooth	Predict	$\{y\}_n$	Smooth	Predict	$\{y\}_n$	Smooth	Predict	$\{y\}_n$
1	.0134	.0226	.624	.0209	.0369	.749	.0342	.0625	.998
12	.00383	.00792	.0520	.00586	.0126	.0624	.00945	.0211	.0833
48	.00183	.00430	.0130	.00276	.00692	.0156	.00440	.0116	.0208
288	.000660	.00206	.00217	.000967	.00343	.00260	.00149	.00600	.00347
$e^{-\Delta\lambda} = 0.9$	Smooth	Predict	$\{y\}_n$	Smooth	Predict	$\{y\}_n$	Smooth	Predict	$\{y\}_n$
1	.0345	.0456	.620	.0569	.0820	.741	.0954	.148	.982
12	.0109	.0233	.0520	.0164	.0396	.0624	.0259	.0697	.0832
48	.00488	.0150	.0130	.00707	.0260	.0156	.0108	.0467	.0208
288	.00144	.00966	.00217	.00195	.0178	.00260	.00280	.0338	.00347

Table 1: *Exact mean square error (steady state) of the estimators of actual volatility. The first two estimators are model based (smoother and predictor) and the third is model free (realised volatility $\{y\}_n$). These measures are calculated for different values of $\omega^2 = \mathrm{Var}(\sigma^2(t))$ and λ, keeping $\xi = \mathrm{E}(\sigma^2(t))$ fixed at 0.5.*

Table 1 reports the mean square error of the model based predictor and smoother of actual volatility, as well as the corresponding result for the model free raw realised volatility (the mean square errors of the model based estimators will be above the figures quoted towards the very start and end of the sample, for we have quoted steady state quantities). The results in the left hand block of the Table corresponds to the model which was simulated in Figure 2, while the other blocks vary the ratio of ξ to ω^2. The exercise is repeated for two values of λ.

The main conclusion from the results in Table 1 is that model based approaches can potentially lead to very significant reductions in mean square error, with the reductions being highest for persistent (low value of λ) volatility processes with high values of $\xi\omega^{-2}$. Even for moderately large values of M the model based predictor can be more accurate than realised volatility, sometimes by a considerable amount. This is an important result from a forecasting viewpoint. However, when there is not much persistence and M is very large, this result is reversed and realised volatility can be moderately more accurate. The smoother is always substantially more accurate than realised volatility, even when M is very large and there is not much memory in volatility. This suggests that model based methods may be particularly helpful in estimating historical records of actual volatility. Finally, we should place a number of caveats on these conclusions. The above results represent a somewhat favourable setup for the model based approach. In the above calculations we have assumed knowledge of the second order properties of volatility while in practice we will have to build such a model and then estimate it, inducing additional biases that we have not reported on.

3.2 Estimating parameters: a numerical illustration

Estimating the parameters of continuous time stochastic volatility models is known to be difficult due to our inability to compute the appropriate likelihood function. This has prompted the development of a sizable collection of methods to deal with this problem. A very incomplete list of references include Gourieroux, Monfort, and Renault (1993), Gallant and Long (1997), Kim, Shephard, and Chib (1998), Elerian, Chib, and Shephard (2001) and Sørensen (2000). Here we study a simple approach based on the realised volatilities. The closest paper to ours in this respect is a recent one by Bollerslev and Zhou (2001) who use a method of moment approach based on assuming that the actual volatility process $\{\sigma_n^2\}$ is observed via the quadratic variation estimator. That is they assume there is no realised volatility error.

Table 2 shows the result of a small simulation experiment which investigates the effectiveness of the quasi-likelihood estimation methods based on the time series of realised volatility. The quasi-likelihood is constructed using the output of the Kalman filter. It is suboptimal for it does not exploit the non-Gaussian nature of the volatility dynamics, but it provides a consistent and asymptotically normal set of estimators. This follows from the fact that the Kalman filter builds the Gaussian quasi-likelihood function for the ARMA representation of the process, where the noise in the representation is both white and strong mixing (strong mixing follows from Sørensen (2000) and Genon-Catalot, Jeantheau, and Larédo (2000) who show that if volatility is strong mixing then squared returns are strong mixing). This means we can immediately apply the asymptotic theory results of Francq and Zakoïan (2000) in this context so long as $\sigma^2(t)$ is strong mixing. Further the estimation takes only around 5 seconds on a modern notebook computer.

M	$\lambda = 0.01$	$\xi = 0.5$	$\omega^2 = \xi/8 = .0625$	$\lambda = 0.01$	$\xi = 0.5$	$\omega^2 = \xi/4 = .125$
1	.00897, 1.76	.318, .659	.00751, .152	.00750, .400	.272, .752	.0172, .225
12	.00891, .0409	.341, .669	.0130, .0759	.00789, .0367	.265, .751	.0197, .168
48	.00920, .0348	.339, .672	.0134, .0715	.00920, .0320	.266, .727	.0199, .149
288	.00928, .0336	.334, .674	.0130, .0755	.00906, .0299	.269, .731	.0207, .152

M	$\lambda = 0.1$	$\xi = 0.5$	$\omega^2 = \xi/8 = .0625$	$\lambda = 0.1$	$\xi = 0.5$	$\omega^2 = \xi/4 = .125$
1	.0451, 1.57	.400, .573	.0271, .151	.0505, .312	.374, .599	.0548, .226
12	.0725, .165	.420, .572	.0383, .0847	.0713, .158	.397, .593	.0717, .170
48	.0748, .152	.421, .566	.0397, .0829	.0754, .148	.398, .592	.0763, .163
288	.0792, .141	.425, .572	.0410, .0788	.0755, .136	.403, .619	.0774, .176

Table 2: *Monte Carlo estimates of the 0.1 and 0.9 quantiles of the maximum quasi-likelihood estimator of SV model with OU volatility. Volatility model has $\sigma^2(t) \sim \Gamma(\nu, a)$ with 500 daily observations, which implies $\xi = \nu a^{-1}$* and *$\omega^2 = \nu a^{-2}$. The true value of ξ is always fixed at 0.5, while ω^2* and *λ* vary. *M denotes the number of intra-day observations used. 1,000 replications are used in the study.*

The setup of the simulation study uses 500 daily observations where the volatility is an OU process with a gamma marginal distribution. The Table varies the value of M. When $M = 1$ this corresponds to using the classical approach of squared daily returns. When M is higher we are using intra-day data. The results suggest that the intra-day data allows us to estimate the parameters much more efficiently. Indeed when M is large the estimators have very little bias and turn out to be quite close to being jointly Gaussian. The experiment also suggests that when λ is larger, which corresponds to the process having less memory, then the estimates of ξ and ω^2 are sharper. Taken together the results are quite encouraging for they are based on only two years of data but suggest we can construct quite precise estimates of these models with this.

4 Empirical illustration

To illustrate some of these results we have fitted a set of superposition based OU/CEV SV models to the realised volatility time series constructed from the five minute exchange rate return data discussed in the introduction to this paper. Here we use the quasi-likelihood method to estimate the parameters of the model — ξ, ω^2, $\lambda_1, ..., \lambda_J$ and $w_1, ..., w_J$. We do this for a variety of values of M, starting with $M = 6$, which corresponds to working with four hour returns. The resulting parameter estimates are given in Table 3. For the moment we will focus on this case.

M	J	ξ	ω^2	λ_1	λ_2	λ_3	w_1	w_2	Quasi-L	BP
6	3	0.4783	0.376	0.0370	1.61	246	0.212	0.180	-113,258	11.2
6	2	0.4785	0.310	0.0383	3.76	—	0.262	—	-113,261	11.3
6	1	0.4907	0.358	1.37	—	—	—	—	-117,397	302
18	3	0.460	0.373	0.0145	0.0587	3.27	0.0560	0.190	-101,864	26.4
18	2	0.460	0.533	0.0448	4.17	—	0.170	—	-101,876	26.5
18	1	0.465	0.497	1.83	—	—	—	—	-107,076	443
144	3	0.508	4.79	0.0331	0.973	268	0.0183	0.0180	-68,377	15.3
144	2	0.509	0.461	0.0429	3.74	—	0.212	—	-68,586	23.3
144	1	0.513	0.374	1.44	—	—	—	—	-76,953	765

Table 3: *Fit of the superposition of J volatility processes for a SV model based on realised volatility computed using $M = 6$, $M = 18$ and $M = 144$. We do not record w_J as this is 1 minus the sum of the other weights. Estimation method: quasi-likelihood using output from a Kalman filter. BP denotes Box–Pierce statistic, based on 20 lags, which is a test of serial dependence in the scaled residuals.*

The fitted parameters suggests a dramatic shift in the fitted model as we go from $J = 1$ to $J = 2$ or 3. The more flexible models allow for a factor which has quite a large degree of memory, as well as a more rapidly decaying component or two. A simple measure of fit of the model is the Box–Pierce statistic, which shows a large jump from a massive 302 when $J = 1$, down to an acceptable number for a superposition model.

To provide a more detailed assessment of the fit of the model we have drawn a series of graphs in Figure 4. Except where explicitly noted we have computed the graphs using the $J = 3$ fitted model, although there would be very little difference if we had taken $J = 2$. Figure 4(a) draws the computed realised volatility $\{y\}_n$, together with the corresponding smoothed estimate of actual volatility using the fitted SV model. We can see that realised volatility is much more jagged than the smoothed quantity. In Figure 4(b) we have drawn a kernel based estimate of the

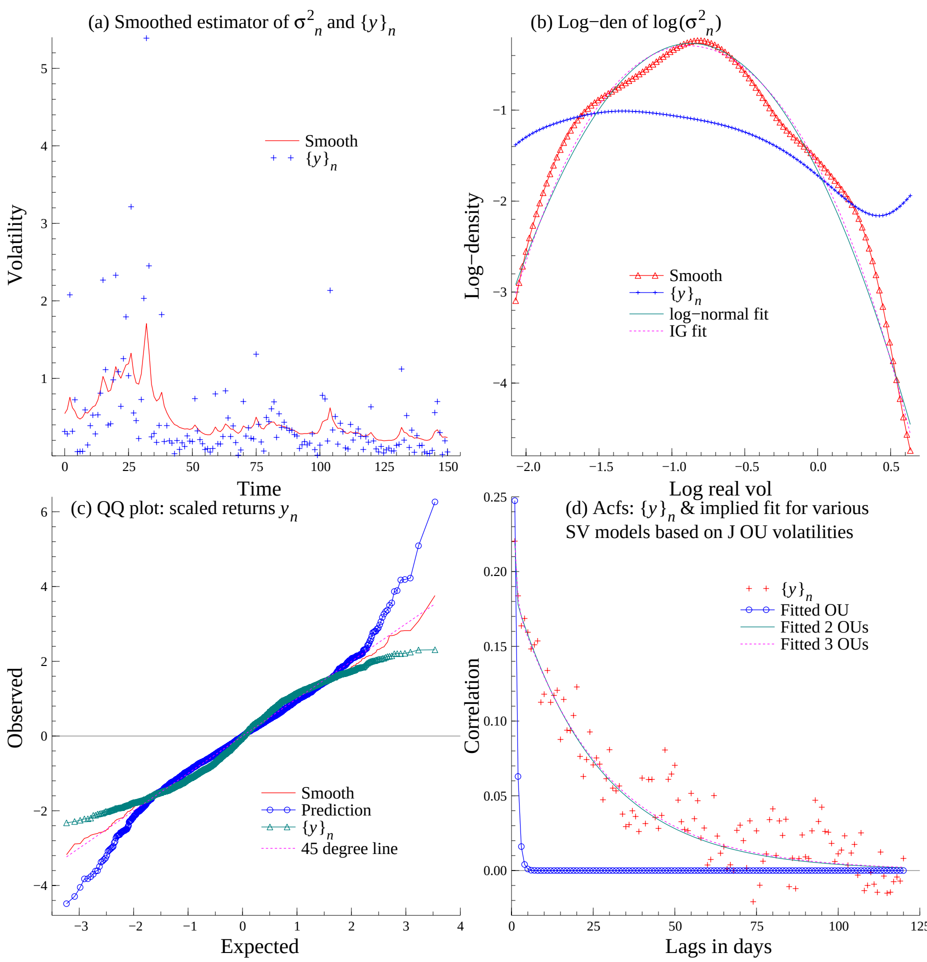

Figure 4: *Results from SV model using $M = 6$ (four hour returns). (a) first 150 observations of $\{y\}_n$ & smoothed estimates of σ_n^2. (b) kernel based estimates of the density of log($\{y\}_n$) and log-smoothed log(σ_n^2). Also the log-normal and inverse Gaussian fits (they appear on top of one another). (c) QQ plot for y_n scaled by estimated σ_n, using $\sqrt{\{y\}_n}$, predicted and smoothed volatility. (d) Acf of $\{y\}_n$ and the fit of SV model for various values of J.*

log-density of the log of realised volatility. The bandwidths were taken to be $1.06\widehat{\sigma}T^{-1/5}$, where T is the sample size and $\widehat{\sigma}$ is the empirical standard deviation of the log of realised volatility (this is an optimal choice against a mean square error loss for Gaussian data, e.g. Silverman

(1986)) while we have chosen the range of the display to match the upper and lower 0.05% of the data — so trimming very little of the data. Andersen, Bollerslev, Diebold, and Labys (2001) have suggested that the marginal distribution of realised volatility is closely approximated by a log-normal distribution when M is high, and that this would support a model for actual volatility which is log-normal. Such models go back to Clark (1973) and Taylor (1986). However, when we draw the corresponding fitted log-normal log-density, choosing the parameters by using maximum likelihood based upon the estimated smoothed realised volatilities as data, we see that the fit is poor. The same holds for the inverse Gaussian log-density. This is also drawn on the figure, but is so close to the fit of the log-normal that it is extremely hard to tell the difference between the two curves. Inverse Gaussian models for volatility were suggested by Barndorff-Nielsen and Shephard (2001a). The rejection of the log-normal and inverse Gaussian marginal distributions for realised volatility itself seems conclusive here. However, when we carry out the same action on the smoothed realised volatilities this rejection no longer holds, implying that realised volatility error really matters here. The kernel based estimate of the log-density of the log smoothed estimates is very much in line with the log-normal or inverse Gaussian hypothesis. This seems to extend the observations of Andersen, Bollerslev, Diebold, and Labys (2001) in at least two directions: (i) our model based estimated actual volatility is fitted well not just by the log-normal distribution, but equally well by the inverse Gaussian, (ii) by using a model-based smoother the above stylised fact can be deduced using quite a low value of M. Of course we have yet to see if these results continue to hold as M increases.

Figure 4(c) draws a QQ plot of returns y_n divided by a number of estimates of σ_n. If the SV model holds correctly and there is no measurement error then these variables should be Gaussian and the QQ plot should appear on a 45° line. The Figure indicates that when we scale returns by realised volatility the returns are highly non-Gaussian, while when we use the smoothed estimate then the model seems to fit extremely well. If we replace the smoothed estimate by the predictor of actual volatility, then we see that the fit is as poor as the plot based on the realised volatility. Overall, Figure 4(c) again confirms the fit of the model, while suggests when $M = 6$ the difference between realised and smoothed volatility is important.

Figure 4(d) shows the corresponding autocorrelation function for the realised volatility series together with the corresponding empirical correlogram. We see from this figure that when $J = 1$ we are entirely unable to fit the data, as its' autocorrelation function starts at around 0.6 and then decays to zero in a couple of days. A superposition of two processes is much better, picking up the longer-range dependence in the data. The superposition of two and three processes give very similar fits, indeed in the graph they are hardly distinguishable.

We next ask how these results vary as M increases. We reanalyse the situation when $M = 144$, which corresponds to working with ten minute returns. Table 3 contains the estimated parameters for this problem. They suggest that moving to a superposition of three processes has an important impact on the fit of the model. Again the fitted models indicate volatility has elements which have a substantial memory, while other components are much more transitory. An important feature of this table is the jump in the value of the estimated ω^2 when we move to having $J = 3$. This is caused by the third component which has a very high value of λ, which does not overly change the variance of actual volatility.

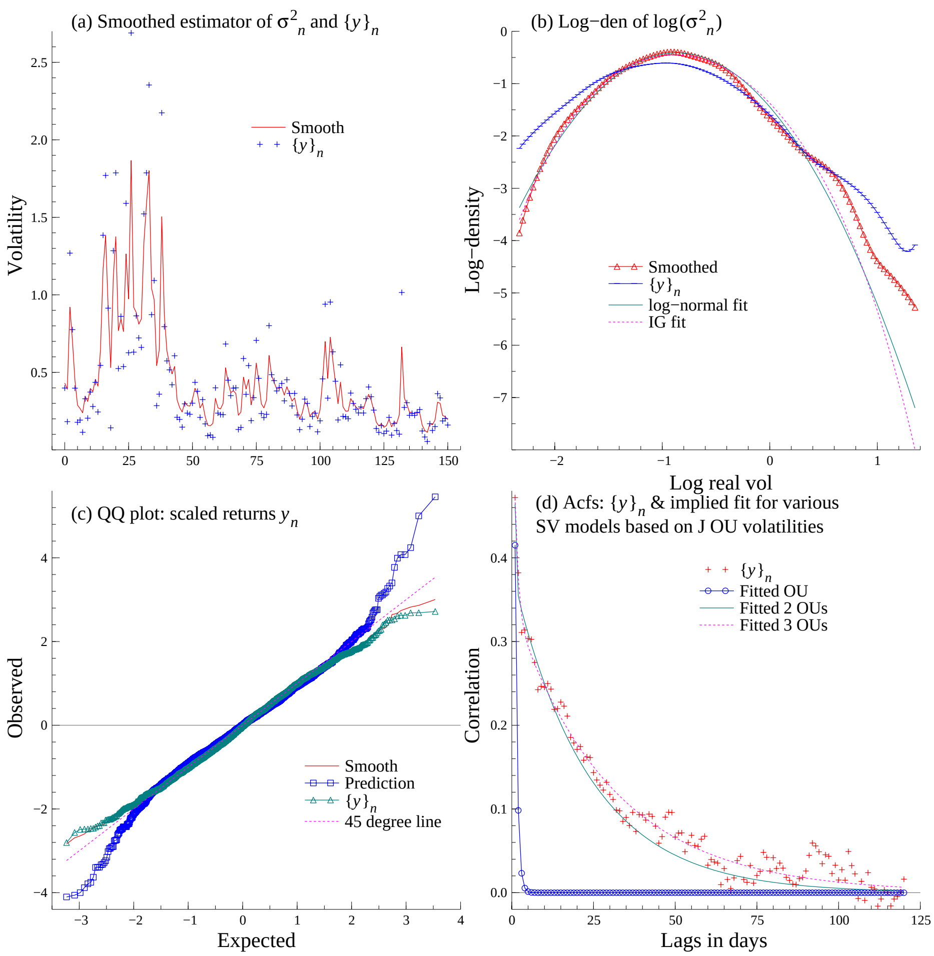

Figure 5: *Results from the fit of the SV model using $M = 144$ (10 minute returns). (a) gives $\{y\}_n$ together with the smoothed estimator of σ_n^2, (b) estimates of the log-density of log($\{y\}_n$) and smoothed estimates of log(σ_n^2). Also log-normal and inverse Gaussian fits. (c) QQ plot for standardised returns. (d) Autocorrelation function of $\{y\}_n$ and the fit of SV model with various superpositions of J processes.*

The fit of the model can also be seen from Figure 5. This broadly shows the same results as Figure 4 except for the following. Realised volatility is now less jagged, and so the smoothed estimator of actual volatility and realised volatility are much more in line. The plots of the estimated log-densities show that realised and smoothed volatilities are again closer, with both being quite well fitted by the log-normal and inverse Gaussian distributions. The smoothed estimators are still more closely approximated than the realised however. The QQ-plots for realised and smoothed volatility are roughly similar, while the one for prediction is still not satisfactory. This indicates that the uncertainty of predicting volatility one day ahead is substantial. Finally, the autocorrelation functions show an improvement in fit as we go from $J = 2$ to $J = 3$ in the SV model.

We finish this section by briefly repeating this exercise with an intermediate value of M, taking $M = 18$, which corresponds to working with returns calculated over 80 minute periods. The results are given in Table 3. They, and the corresponding plots not reproduced here, are very much in line with the previous graphs with the smoothed estimates of actual volatility performing well, although the QQ plot is not as good as it was when we used four hour data.

5 Asymptotic distribution of realised volatility error

5.1 The theory

In Section 2 we derived the mean and variance of the realised volatility error for a continuous time SV model when $\mu = \beta = 0$. Although it is possible to derive the corresponding result when $\mu \neq 0$ but $\beta = 0$, adapting to the risk premium case seems difficult. Instead we take an asymptotic route. In this section we obtain a limit theory for

$$\sqrt{M}\left(\{y\}_n - \sigma_n^2\right),$$

which covers the case of a drift and risk premium.

Theorem 1 *For the SV model in (1) suppose the volatility process σ^2 is of locally bounded variation (i.e. with probability* 1 *the paths of σ^2 are of bounded variation on any compact subinterval of $[0,\infty)$). Then, for any positive Δ and $M \to \infty$*

$$\frac{\{y\}_n - \sigma_n^2}{\sqrt{2\sum_{j=1}^{M}\left(\sigma_{j,n}^2\right)^2}} \xrightarrow{\mathcal{L}} N(0,1), \tag{17}$$

where

$$\sigma_{j,n}^2 = \sigma^{2*}\left((n-1)\Delta + \frac{\Delta j}{M}\right) - \sigma^{2*}\left((n-1)\Delta + \frac{\Delta(j-1)}{M}\right).$$

Furthermore,

$$\Delta^{-1} M \sum_{j=1}^{M}\left(\sigma_{j,n}^2\right)^2 \overset{a.s.}{\to} \sigma_n^{[4]} \tag{18}$$

where

$$\sigma_n^{[4]} = \int_{(n-1)\Delta}^{n\Delta} \sigma^4(s)\mathrm{d}s$$

In particular, then, the limiting law of $\sqrt{M}\left(\{y\}_n - \sigma_n^2\right)$ *is a normal variance mixture.* □

Proof. *Given in the Appendix.*

This theorem implies

$$\sqrt{M}\left(\{y\}_n - \sigma_n^2\right)|\sigma_n^{[4]} \xrightarrow{\mathcal{L}} N\left(0, 2\Delta\sigma_n^{[4]}\right), \tag{19}$$

which has the important implication that we can strengthen the usual quadratic variation result that the drift and risk premium has no impact on the limit of $\{y\}_n$ to the result that the asymptotic distribution of $\sqrt{M}\left(\{y\}_n - \sigma_n^2\right)$ does not depend upon μ or β. Thus the effect on realised volatility of drift and risk premium is of only *third order*, which suggests it maybe safe to ignore it in many cases.

The above theorem also implies that we would expect the variance of the realised volatility error to depend positively on the level of volatility. We have already seen an example of this in the simulated data in Figure 2. Further, the marginal distribution of the realised volatility error should be thicker than normal due to the normal variance mixture (19) averaged over the random $\sigma_n^{[4]}$. We call $\sigma_n^{[4]}$ *actual quarticity*, while the associated $\sigma^4(t)$ is the *spot quarticity.*

We should note that

$$\sum_{j=1}^{M}\left(\sigma_{j,n}^2\right)^2$$

is the realised volatility of $\sigma^{2*}(t)$. Of course the limit, as $M \to \infty$, of this realised volatility is zero, however the above theorem shows that the scaled

$$M\sum_{j=1}^{M}\left(\sigma_{j,n}^2\right)^2$$

has a stochastic limit. This is a special case of a more general lemma we prove in the Appendix on what we call *power variation.*

Lemma 1. *Assume that* $\tau(t)$ *is of locally bounded variation. Then, for* $M \to \infty$ *and* r a positive integer*,*

$$\Delta^{-r+1}M^{r-1}\sum_{j=1}^{M}\tau_j^r \xrightarrow{a.s.} \tau^{r*}(\Delta) \tag{20}$$

where

$$\tau_j = \tau^*(jM^{-1}\Delta) - \tau^*((j-1)M^{-1}\Delta)$$

and

$$\tau^{r*}(t) = \int_0^t \tau^r(s)\mathrm{d}s.$$

□

Proof. *Given in the Appendix.*

We intend to report fully on the implications of this result, and various possible extensions, elsewhere. One of these extensions is that, again writing

$$y_{j,n} = y^*((n-1)\Delta + j\Delta M^{-1}) - y^*((n-1)\Delta + (j-1)\,\Delta M^{-1}),$$

it can be shown that the *realised quarticity*

$$\frac{1}{3}M\Delta^{-1}\sum_{j=1}^{M} y_{j,n}^4 \overset{a.s.}{\to} \sigma_n^{[4]}.$$

which is also the limit for $M \to \infty$ of $M\Delta^{-1}\sum_{j=1}^{M}(\sigma_{j,n}^2)^2$. Consequently, the former — known — sum can be used instead of the latter — unknown — sum in the denominator on the left hand side of the key limiting result (17). In particular

$$\frac{\{y\}_n - \sigma_n^2}{\sqrt{\frac{2}{3}\sum_{j=1}^{M} y_{j,n}^4}} \overset{\mathcal{L}}{\to} N(0,1).$$

Following the first draft of this paper we have used Monte Carlo methods to study the finite sample behaviour of this asymptotic approximation. Results are reported in Barndorff-Nielsen and Shephard (2001b). These experiments suggest we need quite large values of M for the result to be reliable, however a better performance is obtained by transforming the approximation on to the log scale. Then the approximation becomes

$$\frac{\log\{y\}_n - \log\sigma_n^2}{\sqrt{\frac{2}{3}\frac{\sum_{j=1}^{M} y_{j,n}^4}{\left(\sum_{j=1}^{M} y_{j,n}^2\right)^2}}} \overset{\mathcal{L}}{\to} N(0,1).$$

This seems to be quite accurate even for moderate values of M. Following the developments of this paper, our further work on power variation has recently been reported in Barndorff-Nielsen and Shephard (2001c).

Finally we note that the theorem requires that τ is of locally bounded variation. In the OU case this is easily checked for we know that

$$\tau(t) = \exp(-\lambda t)\tau(0) + \int_0^t \exp\{-\lambda(t-s)\}\,\mathrm{d}z(\lambda s).$$

5.2 Application

Suppose our interest is in estimating μ and β, knowing that

$$y_n|\sigma_n^2 \sim N\left(\mu + \beta\sigma_n^2, \sigma_n^2\right).$$

A naive approach would be to regress returns on a constant and the sequence of feasible realised volatilities to produce a simple regression based estimator. Such an estimator will be both biased and inconsistent due to an *errors in variables* effect of mismeasuring actual volatility by using realised volatility (see Hendry (1995, Ch. 12) for a discussion of this in a historical context). The bias is determined by the variance of $(\{y\}_n - \sigma_n^2)$, which we have seen is $O(M^{-1})$ even in the presence of drift and risk premium. A smaller bias would result if we use a model based estimator of actual volatility, instead of the simpler realised volatility. We saw in Section 2 that this can substantially reduce the variance, and this will carry over to the bias reduction of the regression based estimator.

An alternative strategy is to employ an instrumental variable approach. This requires us to find an estimator of σ_n^2 which does not rely on data at time n, but is correlated with σ_n^2. A model-free candidate for this task is

$$\widehat{\sigma_n^2} = \frac{1}{2}\left(\{y\}_{n-1} + \{y\}_{n+1}\right),$$

the average of contiguous realised volatilities. If actual volatility is temporally dependent, the theory developed in Section 2 shows that $\widehat{\sigma_n^2}$ will be correlated with σ_n^2 and so is a valid instrument. Of course, within the context of a model, the best instrument will be the jackknife estimator, which is the best linear estimator of σ_n^2 given

$$\{y\}_1, \{y\}_2, ..., \{y\}_{n-1}, \{y\}_{n+1}, ..., \{y\}_T.$$

This is readily computed for models which can be placed within a linear state space form. A final approach to dealing with this issue is to append an extra measurement equation to the linear state space form (16) and estimate μ and β at the same time as other parameters in a fully specified model.

6 Extensions

6.1 Diurnal affects and actual volatility

An important aspect of the realised volatility series is that it is not very sensitive to the substantial and complicated intra-day diurnal pattern in volatility found in many empirical studies (e.g. Andersen and Bollerslev (1997b) and Andersen and Bollerslev (1998b)) as well as being clear from the top left of Figure 1. To understand this it is helpful to think of the spot volatility as the sum of a (potentially unknown) deterministic diurnal component, $\sigma_\psi^2\{\mathrm{mod}(t,\Delta)\}$ where Δ represents a day, plus a stochastic process, $\sigma_\lambda^2(t)$, then we have

$$\sigma^2(t) = \sigma_\psi^2\{\mathrm{mod}(t,\Delta)\} + \sigma_\lambda^2(t).$$

Hence in this model the spot volatility has a repeating intra-day (i.e. diurnal) component, but does not have a day of the week or monthly seasonal. As a result

$$\sigma_n^2 = c + \sigma_{n,\lambda}^2, \qquad \text{where} \qquad c = \int_0^{\Delta} \sigma_\psi^2 \left\{ \mathrm{mod}(u, \Delta) \right\} \mathrm{d}u$$

and

$$\sigma_{n,\lambda}^2 = \sigma_\lambda^{2*}(n) - \sigma_\lambda^{2*}\left\{(n-1)\right\}, \quad \text{and} \quad \sigma_\lambda^{2*}(t) = \int_0^t \sigma_\lambda^2(u)\mathrm{d}u.$$

In this structure the dynamics of realised volatility is unaffected by the presence of a diurnal effect. Of course, in practice this additive structure should be regarded as holding only approximately, in which case the diurnal effect may not be completely ignorable. However, in this paper we will neglect this deficiency.

6.2 Leverage

The above analysis has not included a leverage term in the model. This can be added in a number of ways. We follow Barndorff-Nielsen and Shephard (2001a) in parameterising the effect as

$$\mathrm{d}y^*(t) = \left\{ \mu + \beta \sigma^2(t) \right\} \mathrm{d}t + \sigma(t) \mathrm{d}w(t) + \rho \mathrm{d}\overline{z}(\lambda t),$$

where we assume $\overline{z}(t) = z(t) - \mathrm{E}\left\{z(t)\right\}$ and $z(t)$ is a Lévy process potentially correlated with $\sigma^2(t)$. The corresponding quadratic variation for this process is

$$[y^*](t) = \sigma^{2*}(t) + \rho^2 [\overline{z}](\lambda t),$$

while the realised volatility error

$$u_n = \left(\left\{ y^0 \right\}_n - \sigma_n^2 \right) + \rho^2 \left(\{\overline{z}\}_n - [\overline{z}]_n \right) + 2\rho c_n,$$

where

$$y^0(t) = \int_0^t \sigma(t) \mathrm{d}w(t), \quad c_n = \sum_{j=1}^{M} \overline{z}_{j,n} \varepsilon_{j,n} \sigma_{j,n} \quad \text{and} \quad \left\{ y^0 \right\}_n - \sigma_n^2 = \sum_{j=1}^{M} \sigma_{j,n}^2 \left(\varepsilon_{j,n}^2 - 1 \right)$$

using the generic realised volatility notation developed in Section 2 of the paper. Here the three terms which make up u_n are zero meaned and uncorrelated when we assume $\mu = \beta = 0$. The only task is to calculate the variances of each of the terms.

The new terms are straightforward to study once we have the following Lemma which relates a Lévy process to its quadratic variation and realised volatility.

Lemma 2 *Let t be fixed and let $z(t)$ be a Lévy process with finite fourth cumulant. Then defining*

$$\left\{ z \right\}(t) = \sum_{j=1}^{M} \left\{ z(jtM^{-1}) - z((j-1)\, tM^{-1}) \right\}^2,$$

we have that

$$\mathrm{E}\left\{\begin{array}{c} z(t) \\ \{z\}(t) \\ [z](t) \end{array}\right\} = t\left(\begin{array}{c} \kappa_1 \\ \kappa_2 \\ \kappa_2 \end{array}\right), \quad \mathrm{Cov}\left\{\begin{array}{c} z(t) \\ \{z\}(t) \\ [z](t) \end{array}\right\} = t\left(\begin{array}{ccc} \kappa_2 & \kappa_3 & \kappa_3 \\ \kappa_3 & \kappa_4 + 3\kappa_2^2 tM^{-1} & \kappa_4 \\ \kappa_3 & \kappa_4 & \kappa_4 \end{array}\right),$$

where κ_r denotes the r-th cumulant of $z(1)$. An implication is that $\{z\}(t) - [z](t)$ has zero mean, while

$$\mathrm{Var}\left\langle \{z\}(t) - [z](t) \right\rangle = 3\kappa_2^2 t^2 M^{-1}.$$

□

Proof. *Most of the results follow immediately recognising that the r-th cumulant of $z(t)$ is $t\kappa_r$. This is a consequence of the Lévy-Khintchine representation. The only piece of this result which is not trivial is*

$$\begin{aligned} \mathrm{Var}(\{z\}(t)) &= M\mu_4\left[z(tM^{-1})\right] \\ &= M\left\{\kappa_4\left[z(tM^{-1})\right] + 3\kappa_2\left[z(tM^{-1})\right]^2\right\} \\ &= M\left(tM^{-1}\kappa_4 + 3t^2M^{-2}\kappa_2^2\right). \end{aligned}$$

Here $\mu_4[\bullet]$ and $\kappa_4[\bullet]$ denotes the fourth centred moment and cumulant of $\bullet$ respectively. □

We achieve our desired result immediately for

$$M\mathrm{Cov}\left\{\begin{array}{l} \{\overline{z}\}_n - [\overline{z}]_n \\ \{y^0\}_n - \sigma_n^2 \\ c_n \end{array}\right\} \to \Delta^2\left(\begin{array}{ccc} 3\kappa_2^2 & 0 & 0 \\ 0 & 2\left(\omega^2 + \xi^2\right) & 0 \\ 0 & 0 & \kappa_2\xi \end{array}\right),$$

as $M \to \infty$. Repeating the pattern we had before, no new issues arise when we allow for a drift or a risk premium for their effect will be small compared to the other terms. Of course, the central limit theory we developed in Section 5 of this paper will apply to $\sqrt{M}\left(\{y^0\}_n - \sigma_n^2\right)$ not to $\sqrt{M}\left(\{y\}_n - [y]_n\right)$.

Trivially the above analysis also deals with the situation of a model which is a SV process plus jumps, where the volatility is not correlated with the jumps.

7 Conclusion

In this paper we have studied the statistical properties of realised volatility in the context of SV models. Our results are entirely general, providing both a central limit theory approximation as well as an exact second order analysis. These results can be used, in conjunction with a model for the dynamics of volatility, to produce a more accurate estimate of actual volatility. Further, a simple quasi-likelihood results which could be used to perform computationally quite simple estimation. Potentially they allow us to exploit the availability of high frequency data in financial economics, giving us relatively simple and efficient ways of estimating these stochastic processes.

Finally, in our empirical work we have taken Δ to be one day. This choice is entirely ad hoc. Another possibility is to simultaneously look at several different Δ values. This may have virtue as a way of checking the fit of the model, as well as allowing potentially more efficient estimation. However, we have yet to explore this issue. To do so it would be convenient to have a functional limit theorem for $\left(\{y\}_n - \sigma_n^2\right)$, which we are currently studying.

8 Acknowledgments

This paper is a revised version of "Econometric analysis of realised volatility and its use in estimating Lévy based non-Gaussian OU type stochastic volatility models." Barndorff-Nielsen's work is supported by CAF, which is funded by the Danish Social Science Research Council, and by MaPhySto, which is funded by the Danish National Research Foundation. Neil Shephard's research is supported by the UK's ESRC through the grant R00023839. All the calculations made in this paper are based on software written by the second author using the `Ox` language of Doornik (2001) combined with the state space software library described by Koopman, Shephard, and Doornik (1999). We thank Michel M. Dacorogna 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. The comments of Tim Bollerslev, Frank Gerhard, Nour Meddahi, Enrique Sentana and the two excellent referees on an initial draft were particularly helpful.

9 Appendix

This appendix has three subsections. First we discuss some of the aspects of the data we use in this paper. Second we give a proof of Lemma 1, while in the third subsection we prove Theorem 1.

9.1 Data appendix

The Olsen group have kindly made available to us a dataset which records every five minutes the most recent quote to appear on the Reuters screen from 1st December 1986 until 30th November 1996. When prices are missing they have interpolated them. Details of this processing are given in Dacorogna, Gencay, Muller, Olsen, and Pictet (2001). The same dataset was analysed by Andersen, Bollerslev, Diebold, and Labys (2001). We follow the extensive work of Torben Andersen and Tim Bollerslev on this dataset, who remove much of the times when the market is basically closed. This includes almost all of the weekend, while they have taken out most US holidays. The result is what we will regard as a single time series of length 705,313 observations. Although many of the breaks in the series have been removed, sometimes there are sequences

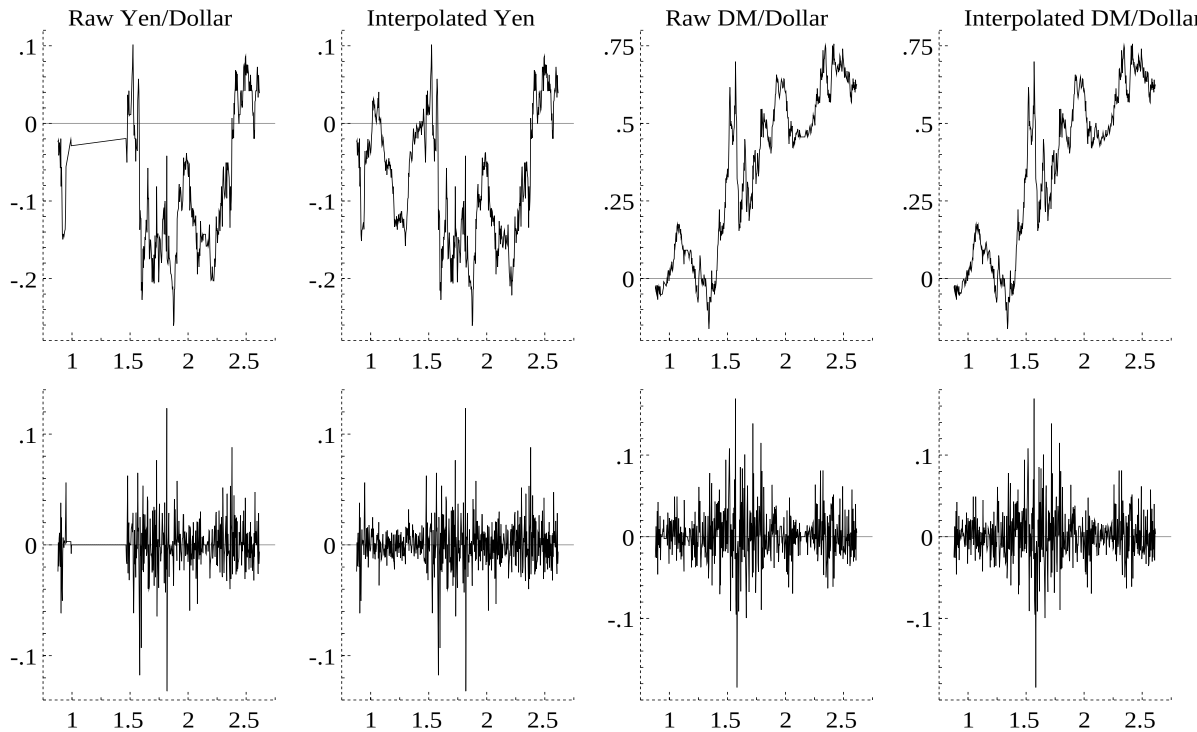

Figure 6: *Top line of graphs are the raw and interpolated data using a Brownian bridge interpolator. Bottom line of graphs is the corresponding returns. The x-axes are marked off in days.*

of very small price changes caused by, for example, unmodelled non-US holidays or data feed breakdowns. We deal with this by adding a Brownian bridge simulation to sequences of data where at each time point the absolute change in a five minute period is below 0.01%. That is, when this happens, we interpolate prices stochastically, adding a Brownian bridge with a standard deviation of 0.01 for each time period. By using a bridge process we are not affecting the long run trajectory of prices, while the impact on realised volatility is very small indeed. We have used this stochastic method here in order to be consistent with our other work on this topic where this effect is important. It is illustrated in Figure 6, which shows the first 500 observations in the Dollar/DM series we have used in this paper and another series which is for the Yen/Dollar. Later stretches of the data have fewer breaks in them, but this graph illustrates the effects of our intervention. Clearly our approach is ad hoc. However, a proper statistical modelling of these breaks is very complicated due to their many causes and the fact that our dataset is enormous.

9.2 Proof of Lemma 1

Recall for the process τ we use the notation

$$\tau^*(t) = \int_0^t \tau(s)\mathrm{d}s \quad \text{and} \quad \tau_j = \tau^*(jM^{-1}\Delta) - \tau^*((j-1)M^{-1}\Delta).$$

Proof of Lemma 1. By the definition of τ_j, for every j there exists a c_j such that

$$\inf_{(j-1)M^{-1}\Delta \leq s \leq jM^{-1}\Delta} \tau(s) \leq c_j \leq \sup_{(j-1)M^{-1}\Delta \leq s \leq jM^{-1}\Delta} \tau(s)$$

and

$$\tau_j = c_j \Delta M^{-1}. \tag{21}$$

The local bounded variation of τ implies that τ^r is locally bounded and Riemann integrable. Consequently

$$\Delta^{-r+1}M^{r-1}\sum_{j=1}^{M}\tau_j^r = \sum_{j=1}^{M} c_j^r \Delta M^{-1} \to \int_0^\Delta \tau^r(s)\mathrm{d}s = \tau^{r*}(\Delta).$$

□

The fact that τ^r is Riemann integrable is perhaps not immediately obvious. However, we recall that a bounded function f is Riemann integrable on an interval $[0,t]$ if and only if the set of discontinuity points of f has Lebesgue measure 0 (see Hobson (1927, pp. 465–466), Munroe (1953, p. 174, Theorem 24.4) or Lebesgue (1902)). In our case the latter property follows immediately from the bounded variation of τ (any function of bounded variation is the difference between an increasing and a decreasing function and any monotone function has at most countably many discontinuities).

9.3 Proof of Theorem 1

We first recall some definitions. Consider the SV model

$$y^*(t) = \mu t + \beta\tau^*(t) + \int_0^t \tau^{1/2}(s)\mathrm{d}w(t),$$

with τ positive, stationary and independent of w (we have switched our notation for the volatility as it simplifies our later derivation). Now writing u and $\{y\}$ for u_1 and $\{y\}_1$ we have

$$u = \{y\} - \tau^*(\Delta) = \sum_{j=1}^{M} y_j^2 - \tau^*(\Delta)$$

where

$$y_j = y^*(jM^{-1}\Delta) - y^*((j-1)M^{-1}\Delta)$$

Conditionally on $\tau_1, ..., \tau_M$, the increments $y_1, ..., y_M$ are independent, and $y_j \stackrel{\mathcal{L}}{=} N(\mu M^{-1}\Delta + \beta\tau_j, \tau_j)$. Thus, conditionally, y_j^2 is noncentral χ^2 with cumulant function

$$\mathrm{C}\{\zeta \ddagger y_j^2|\tau_j\} = -\frac{1}{2}\log(1-2i\tau_j\zeta) + i\nu_j\zeta(1-2i\tau_j\zeta)^{-1}$$

where

$$\nu_j = (\mu M^{-1}\Delta + \beta\tau_j)^2 \tag{22}$$

Consequently

$$\mathrm{C}\{\zeta \ddagger u|\tau_1, ..., \tau_M\} = -\sum_{j=1}^{M}\left\{\tfrac{1}{2}\log(1-2i\tau_j\zeta) - i\nu_j\zeta(1-2i\tau_j\zeta)^{-1} + i\tau_j\zeta\right\}$$

By Taylor's formula with remainder (cf., for instance, Barndorff-Nielsen and Cox (1989, formula 6.122)) we find, provided

$$2|\zeta| \max_{1\le j\le M} \tau_j < 1$$

that

$$\tfrac{1}{2}\log(1-2i\tau_j\zeta) - i\nu_j\zeta(1-2i\tau_j\zeta)^{-1} + i\zeta\tau_j = \zeta^2\{\tau_j^2 Q_{0j}(\zeta) + 2\nu_j\tau_j Q_{1j}(\zeta)\} - i\nu_j\zeta,$$

where

$$Q_{0j}(\zeta) = 2\int_0^1 \frac{1-s}{(1-2i\tau_j\zeta s)^2}\mathrm{d}s$$

and

$$Q_{1j}(\zeta) = 2\int_0^1 \frac{1-s}{(1-2i\tau_j\zeta s)^3}\mathrm{d}s.$$

Hence

$$\mathrm{C}\{\zeta \ddagger u|\tau_1, ..., \tau_M\} = i\zeta\sum_{j=1}^{M}\nu_j - \zeta^2\sum_{j=1}^{M}\left\{\tau_j^2 Q_{0j}(\zeta) + 2\nu_j\tau_j Q_{1j}(\zeta)\right\}. \tag{23}$$

Proof of Theorem 1. Note first that (18) follows from Lemma 1. Next, rewrite (23) as

$$\begin{aligned}\mathrm{C}\{\zeta \ddagger u|\tau_1, ..., \tau_M\} &= i\zeta\sum_{j=1}^{M}\nu_j - \zeta^2\sum_{j=1}^{M}(\tau_j^2 + 2\nu_j\tau_j) \\ &\quad -\zeta^2\sum_{j=1}^{M}\left[\tau_j^2\{Q_{0j}(\zeta)-1\} + 2\nu_j\tau_j\{Q_{1j}(\zeta)-1\}\right] \\ &= \frac{1}{2}\zeta^2 2\sum_{j=1}^{M}\tau_j + R(\zeta),\end{aligned}$$

where

$$R(\zeta) = i\zeta\sum_{j=1}^{M}\nu_j - 2\zeta^2\sum_{j=1}^{M}\nu_j\tau_j - \zeta^2\sum_{j=1}^{M}\left[\tau_j^2\{Q_{0j}(\zeta)-1\} + 2\nu_j\tau_j\{Q_{1j}(\zeta)-1\}\right].$$

Thus, to verify (17) we must show that

$$\sum_{j=1}^{M} \nu_j \Big/ \sqrt{\textstyle\sum_{j=1}^M \tau_j^2} \to 0, \qquad \sum_{j=1}^{M} \nu_j \tau_j \Big/ \textstyle\sum_{j=1}^M \tau_j^2 \to 0,$$

$$\sum_{j=1}^{M} \tau_j^2 \left\{ Q_{0j}\left(\zeta/\sqrt{2\textstyle\sum_{j=1}^M \tau_j^2}\right) - 1 \right\} \Big/ \textstyle\sum_{j=1}^M \tau_j^2 \to 0,$$

and

$$\sum_{j=1}^{M} \nu_j \tau_j \left\{ Q_{1j}\left(\zeta/\sqrt{2\textstyle\sum_{j=1}^M \tau_j^2}\right) - 1 \right\} \Big/ \textstyle\sum_{j=1}^M \tau_j^2 \to 0$$

or, equivalently, by (20), that

$$\sqrt{M}\sum_{j=1}^{M} \nu_j \to 0, \qquad M\sum_{j=1}^{M} \nu_j \tau_j \to 0,$$

$$M\sum_{j=1}^{M} \tau_j^2 \left\{ Q_{0j}\left(\zeta/\sqrt{2\textstyle\sum_{j=1}^M \tau_j^2}\right) - 1 \right\} \to 0, \tag{24}$$

and

$$M\sum_{j=1}^{M} \nu_j \tau_j \left\{ Q_{1j}\left(\zeta/\sqrt{2\textstyle\sum_{j=1}^M \tau_j^2}\right) - 1 \right\} \to 0. \tag{25}$$

We have

$$\sqrt{M}\sum_{j=1}^{M} \nu_j = M^{-1/2}\left(\Delta^2\mu^2 + 2\mu\Delta\beta\sum_{j=1}^{M}\tau_j + \beta^2 M \sum_{j=1}^{M}\tau_j^2\right),$$

which tends to zero on account of (20). Furthermore, also by (20) we find

$$M\sum_{j=1}^{M} \nu_j\tau_j = M^{-1}\mu^2\Delta^2\sum_{j=1}^{M}\tau_j + 2\mu\Delta\beta\sum_{j=1}^{M}\tau_j^2 + \beta^2 M\sum_{j=1}^{M}\tau_j^3 \to 0.$$

Finally, to show (24)-(25) we first note that by (21), the local boundedness of τ and (20),

$$\tau_j \Big/ \sqrt{\textstyle\sum_{j=1}^M \tau_j^2} = \sqrt{M}\tau_j \Big/ \sqrt{M\textstyle\sum_{j=1}^M \tau_j^2} = M^{-1/2}\Delta c_j \Big/ \sqrt{M\textstyle\sum_{j=1}^M \tau_j^2} = O(M^{-1/2})$$

uniformly in j. Hence

$$Q_{0j}\left(\zeta/\sqrt{2\textstyle\sum_{j=1}^M \tau_j^2}\right) - 1 \to 0 \tag{26}$$

and

$$Q_{1j}\left(\zeta/\sqrt{2\textstyle\sum_{j=1}^M \tau_j^2}\right) - 1 \to 0 \tag{27}$$

uniformly in j. Moreover, again using (20), we have

$$M\sum_{j=1}^{M}(\tau_j^2 + \nu_j\tau_j) = (1+\Delta^2\mu^2)M\sum_{j=1}^{M}\tau_j^2 + 2\Delta\mu\beta M\sum_{j=1}^{M}\tau_j^3 + \beta^2 M\sum_{j=1}^{M}\tau_j^4 = O(1)$$

and (24)-(25) follows from this and (26)-(27). □

References

Andersen, T. G. and T. Bollerslev (1997a). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance 52*, 975–1005.

Andersen, T. G. and T. Bollerslev (1997b). 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, F. X. Diebold, and H. Ebens (2001). The distribution of realised stock return volatility. *Journal of Financial Economics 61*, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000). Exchange rate returns standardised by realised volatility are (nearly) Gaussian. *Multinational Finance Journal 4*, 159–179.

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.

Andreou, E. and E. Ghysels (2001). Rolling-sampling volatility estimators: some new theoretical, simulation and empirical results. *Journal of Business and Economic Statistics 19*. Forthcoming.

Areal, N. M. P. C. and S. J. Taylor (2002). The realised volatility of FTSE-100 futures prices. *Journal of Futures Markets 22*. Forthcoming.

Bai, X., J. R. Russell, and G. C. Tiao (2000). Beyond Merton's utopia: effects of non-normality and dependence on the precision of variance estimates using high-frequency financial data. Unpublished paper: Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics 2*, 41–68.

Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein-Uhlenbeck type processes. *Theory of Probability and Its Applications 45*, 175–194.

Barndorff-Nielsen, O. E. and D. R. Cox (1989). *Asymptotic Techniques for use in Statistics*. London: Chapman & Hall.

Barndorff-Nielsen, O. E., J. L. Jensen, and M. Sørensen (1990). Parametric modelling of turbulence. *Phil. Trans. Roy. Soc. Lond. 332*, 439–455.

Barndorff-Nielsen, O. E. and N. Shephard (2001b). How accurate is the asymptotic approximation to the distribution of realised volatility? In D. W. F. Andrews, J. L. Powell, P. A. Ruud, and J. H. Stock (Eds.), *Identification and Inference for Econometric Models.* Festschrift for Thomas J. Rothenberg. Forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (2001a). 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 (2001c). Realised power variation and stochastic volatility. Unpublished paper: Nuffield College, Oxford.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. In R. F. Engle and D. McFadden (Eds.), *The Handbook of Econometrics, Volume 4*, pp. 2959–3038. Amsterdam: North-Holland.

Bollerslev, T. and H. Zhou (2001). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*. Forthcoming.

Christensen, B. J. and N. R. Prabhala (1998). The relation between implied and realized volatility. *Journal of Financial Economics 37*, 125–150.

Clark, P. K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica 41*, 135–156.

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

Cox, D. R. (1991). Long-range dependence, non-linearity and time irreversibility. *Journal of Time Series Analysis 12*, 329–335.

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

Ding, Z. and C. W. J. Granger (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics 73*, 185–215.

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

Elerian, O., S. Chib, and N. Shephard (2001). Likelihood inference for discretely observed non-linear diffusions. *Econometrica 69*, 959–993.

Engle, R. F. and G. G. J. Lee (1999). A permanent and transitory component model of stock return volatility. In R. F. Engle and H. White (Eds.), *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, Chapter 20, pp. 475–497. Oxford: Oxford University Press.

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

Francq, C. and J.-M. Zakoïan (2000). Covariance matrix estimation for estimators of mixing weak ARMA models. *Journal of Statistical Planning and Inference 83*, 369–94.

Gallant, A. R. and J. R. Long (1997). Estimating stochastic differential equations efficiently by minimum chi-square. *Biometrika 84*, 125–41.

Genon-Catalot, V., T. Jeantheau, and C. Larédo (2000). Stochastic volatility as hidden Markov models and statistical applications. *Bernoulli 6*, 1051–1079.

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.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*, S85–S118.

Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics 14*, 227–238.

Hamilton, J. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hobson, E. W. (1927). *The Theory of Functions of a Real Variable and the Theory of Fourier's Series* (3 ed.). Cambridge: Cambridge University Press.

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., N. Shephard, and J. A. Doornik (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal 2*, 107–166.

Lebesgue, H. (1902). Integrale, longuer, aire. *Annali di Mathematica pura ed applicata 7*, 231–359.

Maheu, J. M. and T. H. McCurdy (2001). Nonlinear features of realised FX volatility. *Review of Economics and Statistics 83*. Forthcoming.

Meddahi, N. and E. Renault (2002). Temporal aggregation of volatility models. *Journal of Econometrics*. Forthcoming.

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

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

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

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In D. R. Cox, D. V. Hinkley, and O. E. Barndorff-Nielsen (Eds.), *Time Series Models in Econometrics, Finance and Other Fields*, pp. 1–67. London: Chapman & Hall.

Silverman, B. W. (1986). *Density Estimation for Statistical and Data Analysis*. London: Chapman & Hall.

Sørensen, M. (2000). Prediction based estimating equations. *Econometrics Journal 3*, 123–147.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Chichester: John Wiley.

Taylor, S. J. (1994). Modelling stochastic volatility. *Mathematical Finance 4*, 183–204.

Taylor, S. J. and X. Xu (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance 4*, 317–340.

Whittle, P. (1983). *Prediction and Regulation* (2 ed.). Oxford: Blackwell. 1st edition, 1963.