Impact of jumps on returns and realised variances: econometric analysis of time-deformed Lévy processes

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, University of Oxford, Oxford OX1 1NF, U.K.
neil.shephard@nuf.ox.ac.uk

April 2003

Abstract

In order to assess the effect of jumps on realised variance calculations, we study some of the econometric properties of time-changed Lévy processes. We show that in general we can derive the second order properties of realised variances and use these to estimate the parameters of such models. Our analytic results give a first indication of the degrees of inconsistency of realised variance as an estimator of the time-change in the non-Brownian case. Further, our results suggest volatility is even more predictable than has been shown by the recent econometric work on realised variance.

Keywords: Kalman filter, Lévy process, Long-memory, Quasi-likelihood, Realised variance, Stochastic volatility, Time-change.

1 Introduction

1.1 Time-deformed Lévy processes

Here we study time-deformed Lévy processes. By doing this we can assess the impact of jumps on using realised variances which recent work has demonstrated has significantly improved our ability to forecast the volatility of financial markets. In particular this paper will derive the first four moments of returns, the second order properties of realised variances and the degree of inconsistency of the realised variance estimator of the time-change. These features will be used to estimate parameters of time-deformed Lévy processes using a quasi-likelihood constructed out of these moments, focusing on OU based models, superpositions, log-normal OU processes and long-memory models.

Time-deformed Lévy processes have recently been introduced into financial economics by Geman, Madan, and Yor (2003) and Carr and Wu (2003). To understand this class of processes we start with the standard setup in asset pricing models. We let log-prices y^* follow a semimartingale process

$$y^*(t) = \alpha^*(t) + m^*(t), \quad t \geq 0,$$

where α^* is a bounded variation process, m^* is a local martingale and t represents time. Throughout, for convenience of calculation and without loss of generality, we assume that $\alpha^*(0) = 0$ and $m^*(0) = 0$, which implies that $y^*(0) = 0$. Hence y^* can be thought of as the return process. For an excellent discussion of probabilistic aspects of semimartingales see Protter (1990), while its attraction from an economic viewpoint is discussed by Back (1991).

We then follow Carr, Geman, Madan, and Yor (2003), Geman, Madan, and Yor (2003) and Carr and Wu (2003) in basing the model on a time-deformed Lévy process. In particular we take

$$y^*(t) = \mu t + z(\tau^*(t)),$$

where z is a Lévy process (that is a process with independent and stationary increments i.e. a continuous time random walk) with the added condition that $\text{Var}(z(1)) < \infty$ and $z(0) = 0$. Textbook expositions on Lévy processes can be found in Bertoin (1996), Sato (1999) and Barndorff-Nielsen and Shephard (2003, Ch. 2). Here τ^* is a time-change (that is a process with non-decreasing paths) such that, for all t, $\tau^*(t) < \infty$. A review of relevant aspects of such time-change processes is given in Barndorff-Nielsen and Shephard (2003, Ch. 4). This structure means that $\alpha^*(t) = \mu t + \tau^*(t)\text{E}(z(1))$ while $m^*(t) = z(\tau^*(t)) - \tau^*(t)\text{E}(z(1))$. Finally, we will assume that the z and τ^* processes are independent. This rules out leverage type effects.

The four most well known examples of Lévy processes in financial economics are

(i) z is a Brownian motion, the workhorse of modern financial economics (e.g. Duffie (1996)). A sample path of a scaled Brownian motion is given in Figure 1(a).

(ii) z is a jump diffusion (Merton (1976)), that is the addition of Brownian motion and a compound Poisson process with Gaussian jumps.

(iii) z is a variance gamma (also called a normal gamma) process (Madan and Seneta (1990)).

(iv) z is a zero mean normal inverse Gaussian process (Barndorff-Nielsen (1998)). A sample path of a normal inverse Gaussian motion is given in Figure 1(b). It is designed to have the same variance, per unit of time, as the corresponding Brownian motion given in Figure 1(a).

Brownian motion has continuous sample paths, while all non-Brownian Lévy processes have jumps. Importantly example (ii) is fundamentally different from examples (iii) and (iv). (ii) is said to be of finite activity for it has a finite number of jumps in any finite period of time. (iii) and (iv) have the property that there are an infinite number of jumps in any finite period. Unfortunately non-Brownian motion Lévy processes are often equated with stable processes in the econometric literature (Mandelbrot (1963), Mandelbrot and Taylor (1967)). Such stable

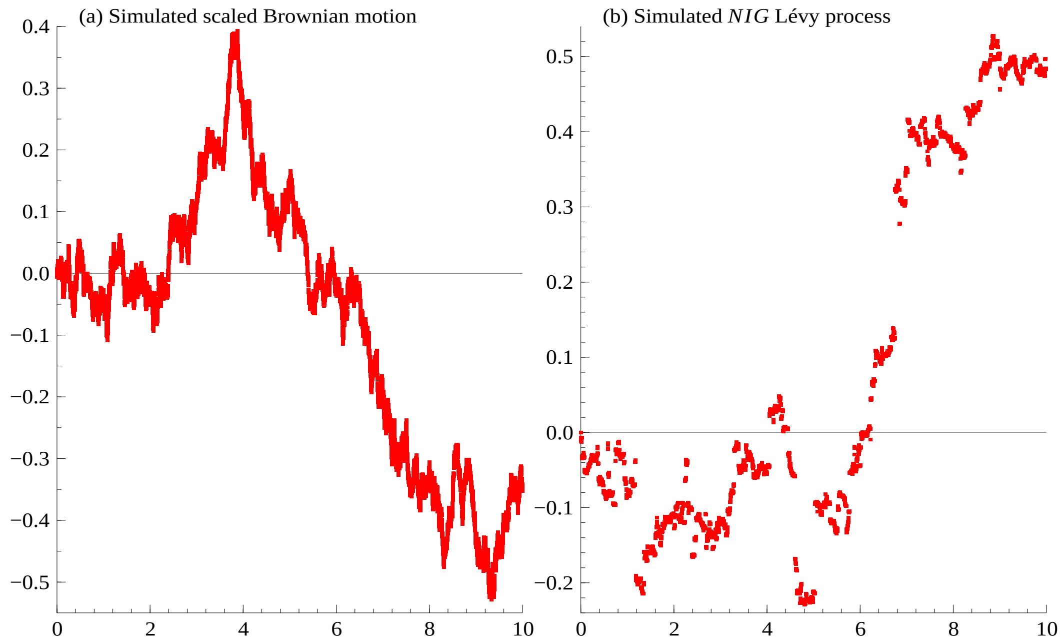

Figure 1: *(a) Sample path of* $\sqrt{0.02}$ *times standard Brownian motion. (b) Sample path of a NIG(0.2,0,0,10) Lévy process. Thus the increments of both processes have the same variance. Code:* `levy_graphs.ox`.

processes have a poor record of accurately modelling the log-prices of returns through time, e.g. they have infinite variances, while actual asset returns do not. Equating stable and Lévy processes is simply a technical misunderstanding, the class of Lévy processes is much wider than is commonly held in that literature.

The most well known example of a time-changed Lévy process is where z is Brownian motion, a model developed by Bochner (1949) and first used in economics by Clark (1973). Econometric research which followed these early papers include Stock (1988) and Ghysels and Jasiak (1994).

We additionally restrict our attention to time-changes of the form

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

where τ is a non-negative process. This means that τ^* has a continuous, but not necessarily differentiable, sample path. Under this assumption m^* has a continuous sample path with probability one iff z is Brownian motion. In the Brownian motion case this process is equivalent to a stochastic volatility process. This is studied in, for example, Taylor (1982), Hull and White (1987), Harvey, Ruiz, and Shephard (1994), Shephard (1996) and Ghysels, Harvey, and Renault (1996).

1.2 Returns and realised variance

Associated with the continuous time price process, is a sequence of discrete returns. In this paper we will use them to study the properties of time-deformed Lévy processes, both in theory and application. Here we establish a notation for returns and realised variances, which are a function of high frequency returns.

Consider a fixed interval of time of length $\hbar > 0$. For concreteness we typically refer to $\hbar$ as representing a day. Traditional daily returns are computed as

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

where i indexes the day. In the next Section we will calculate some of their properties and then later use them to make inference on the parameters indexing time-change models.

We will also focus on the case where we additionally have M intra-$\hbar$ high frequency observations during each $\hbar$ time period. The j-th intra-$\hbar$ return for the i-th period (e.g. if $\hbar$ is a day and $M = 1440$, then this is the return for the j-th minute on the i-th day) will be calculated as

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

High frequency returns allow us to compute

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

the *realised variance* (RV) for the i-th day. In econometrics the RV is used to proxy the variability of the i-th return. We will justify this in the context of time-changed Lévy processes by seeing that RV is an estimator of $\mathrm{Var}(y_i|\tau_i) = \tau_i \mathrm{Var}(z(1))$, where $\tau_i = \tau^*(i\hbar) - \tau^*((i-1)\hbar)$. That is it estimates the variance of y_i if we had known the path of τ^*.

RV has been studied in quite some detail, first by Comte and Renault (1998), Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2001). Later empirical and methodological work by Andersen, Bollerslev, Diebold, and Labys (2001) has been influential. A distribution theory for realised variance under time-changed Brownian motion was developed by Barndorff-Nielsen and Shephard (2002a), while some of this work was extended by Meddahi (2002a). See also Barndorff-Nielsen and Shephard (2002b) for a discussion of the multivariate case and Andersen, Bollerslev, and Diebold (2003) and Barndorff-Nielsen and Shephard (2003, Ch. 7) for surveys of this area. Bai, Russell, and Tiao (2000) provide some simulation evidence for the effects of jumps on realised variances.

The notation $[y_M^*]_i$ is designed to reflect the fact it is based on the y^* process using M intra-$\hbar$ observations and computed on the i-th day. The reason for the use of the square brackets will become clearer later when we recall the idea of quadratic variation.

The realised variance is related to, but different from, the standard empirical variance of returns

$$\frac{1}{M}\sum_{j=1}^{M} y_{j,i}^{2}-\left(\frac{1}{M}\sum_{j=1}^{M} y_{j,i}\right)^{2}=\frac{1}{M}\sum_{j=1}^{M} y_{j,i}^{2}-\frac{1}{M^{2}}y_{i}^{2}.$$

In high frequency finance this quantity does not make sense for it will converge in probability to zero as $M \to \infty$. This is because the individual high frequency returns will shrink in size as M increases. The realised variance is roughly M times the empirical variance of returns, the difference being that realised variance ignores the $\frac{1}{M}y_i^2$ term as it is stochastically of smaller order than $\sum_{j=1}^{M} y_{j,i}^2$.

1.3 Structure of the paper

The outline of this paper is as follows. In Section 2 we derive various cumulants of the returns from time-deformed Lévy processes. In Section 3 we extend this to the RV case, giving us a first analytic handle on the inconsistency of realised variance as an estimator of the time-change hidden in the price process. In Section 4 we use these properties to derive rather simple and computationally tractable quasi-likelihood estimators of the parameters which index time-deformation models. We illustrate these results in Section 5 based on squared daily observations estimating various short memory and long memory Lévy and Brownian motion based SV models. In Section 6 we carry out the same exercise but using realised variances rather than squared data. In Section 7 we calculate best linear filters of the time-change. We also use them to derive news impact functions. Section 8 concludes.

2 Some cumulants of returns

2.1 Background material on cumulants

Here we will be interested in various moments of

$$y^{*}(t)=\mu t+z(\tau^{*}(t)).$$

In order to do this it will turn out to be useful to have some knowledge of cumulants. To start, recall that cumulants are derived via the cumulant function, which is $\log \mathrm{E}e^{i\theta X}$ for some arbitrary random variable X. Then the j-th cumulant is defined as (assuming it exists)

$$\kappa_j=\left.\frac{\partial^j \log \mathrm{E}\left(e^{i\theta X}\right)}{\partial\theta^j}\right|_{\theta=0}, \quad j=1,2,\dots .$$

These cumulants are related to the more familiar uncentred moments. Recall (e.g. Barndorff-Nielsen and Cox (1989, p. 7)) that if we write

$$\mu_j'=\mathrm{E}\left(X^j\right), \quad j=1,2,\dots ,$$

then

$$\begin{array}{rcl} \mu_1' & = & \kappa_1 \\ \mu_2' & = & \kappa_2 + \kappa_1^2 \\ \mu_3' & = & \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3 \\ \mu_4' & = & \kappa_4 + 4\kappa_1\kappa_3 + 3\kappa_2^2 + 6\kappa_1^2\kappa_2 + \kappa_1^4. \end{array} \tag{3}$$

Often we will be interested in the cumulants and moments of processes as a function of t. In this case it is useful to include this explicitly in terms of the notation. Thus we will sometimes write the cumulants and moments of some process $z(t)$ as $\kappa_j(t)$ and $\mu_j'(t)$, respectively.

Example 1 *The leading case is where $z(t) = \beta t + \sigma w(t)$, with w being standard Brownian motion. Then for this process*

$$\kappa_1(t) = \beta t, \quad \kappa_2(t) = \sigma^2 t \quad \textit{and} \quad \kappa_j(t) = 0 \textit{ for } j > 2.$$

This implies that the uncentred moments are

$$\begin{array}{rcl} \mu_1'(t) & = & \beta t, \\ \mu_2'(t) & = & \sigma^2 t + \beta^2 t^2, \\ \mu_3'(t) & = & 3\beta\sigma^2 t^2 + \beta^3 t^3, \\ \mu_4'(t) & = & 3\sigma^4 t^2 + 6\beta^2\sigma^2 t^3 + \beta^4 t^4. \end{array}$$

It is interesting to think of t as being small, for this will be relevant when we study RVs. Then β effects only the higher order terms in $\mu_2'(t)$ and $\mu_4'(t)$.

□

2.2 Cumulants of Lévy processes

A characterising feature of Lévy processes is that the cumulants of the process $z(t)$ at time t, again written as $\kappa_j(t)$, have a simple relationship to the cumulants of the process at time 1, written as κ_j. Thus

$$\kappa_j(t) = t\kappa_j.$$

This follows from the fact that

$$\mathrm{E}\left\{e^{i\varsigma z(t)}\right\} = \left\{\mathrm{E}e^{i\varsigma z(1)}\right\}^t,$$

the fundamental property of continuous time processes with independent and stationary increments, which implies that the cumulant function has the feature that

$$\log \mathrm{E}\left\{e^{i\varsigma z(t)}\right\} = t \log \mathrm{E}e^{i\varsigma z(1)}.$$

Example 2 *Suppose z is a symmetric Lévy process. Then for this process $\kappa_3 = 0$ and*

$$\begin{aligned}
\mu_1'(t) &= \kappa_1 t, \\
\mu_2'(t) &= \kappa_2 t + \kappa_1^2 t^2, \\
\mu_3'(t) &= 3\kappa_1\kappa_2 t^2 + \kappa_1^3 t^3, \\
\mu_4'(t) &= \kappa_4 t + 3\kappa_2^2 t^2 + 6\kappa_1^2\kappa_2 t^3 + \kappa_1^4 t^4.
\end{aligned}$$

This example contains Example 1 as a special case by setting $\kappa_1 = \beta$, $\kappa_2 = \sigma^2$ and $\kappa_4 = 0$. The interesting change in the results is the impact on $\mu_4'(t)$, which now has a term involving just t. For small t this term will be dominant, which will have important implications for our analysis of RVs.

2.3 Conditional cumulants and moments of log-prices

We recall the model structure for log-prices is

$$y^*(t) = \mu t + z(\tau^*(t)),$$

where we have assumed that z is independent from τ^*. We will initially focus on quantities connected to the cumulants of $y^*(t)|\tau^*(t)$ which are, when they exist,

$$\tau^*(t)\kappa_j. \tag{4}$$

We can use this result and the relationship between cumulants and uncentred moments (3) to derive the following results in terms of uncentred conditional moments. Throughout we set $\mu = 0$, although the corresponding results for $\mu \neq 0$ can be backed out straightforwardly.

Proposition 1 *If $\kappa_4 < \infty$, then*

$$\mathrm{E}(y^*(t)|\tau^*(t)) = \kappa_1\tau^*(t),$$

$$\mathrm{E}\left\{y^*(t)^2|\tau^*(t)\right\} = \kappa_2\tau^*(t) + \kappa_1^2\tau^*(t)^2, \tag{5}$$

$$\mathrm{E}\left\{y^*(t)^3|\tau^*(t)\right\} = \kappa_3\tau^*(t) + 3\kappa_1\kappa_2\tau^*(t)^2 + \kappa_1^3\tau^*(t)^3, \tag{6}$$

$$\mathrm{E}\left\{y^*(t)^4|\tau^*(t)\right\} = \kappa_4\tau^*(t) + \left(4\kappa_1\kappa_3 + 3\kappa_2^2\right)\tau^*(t)^2 + 6\kappa_1^2\kappa_2\tau^*(t)^3 + \kappa_1^4\tau^*(t)^4.$$

Further,

$$\mathrm{Var}(y^*(t)|\tau^*(t)) = \kappa_2\tau^*(t), \tag{7}$$

$$\mathrm{Cov}(y^*(t)^2, y^*(t)|\tau^*(t)) = \kappa_3\tau^*(t) + 2\kappa_1\kappa_2\tau^*(t)^2, \tag{8}$$

$$\mathrm{Var}(y^*(t)^2|\tau^*(t)) = \kappa_4\tau^*(t) + 2\left(2\kappa_1\kappa_3 + \kappa_2^2\right)\tau^*(t)^2 + 4\kappa_1^2\kappa_2\tau^*(t)^3. \tag{9}$$

$$\mathrm{E}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}^2|\tau^*(t)\right] = \kappa_4\tau^*(t) + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\tau^*(t)^2 + 4\kappa_1^2\kappa_2\tau^*(t)^3 + \kappa_1^4\tau^*(t)^4.$$

□

Proof. The first four results follow from (4) together with (3). (7)-(9) follow from the use of these results with the standard formulae that, generically for some X and Y,

$$\operatorname{Var}(Y|X) = \mathrm{E}\left(Y^2|X\right) - \mathrm{E}\left(Y|X\right)^2, \ \operatorname{Var}\left(Y^2|X\right) = \mathrm{E}\left(Y^4|X\right) - \mathrm{E}\left(Y^2|X\right)^2,$$

and

$$\operatorname{Cov}\left(Y^2, Y|X\right) = \mathrm{E}\left(Y^3|X\right) - \mathrm{E}\left(Y^2|X\right)\mathrm{E}\left(Y|X\right).$$

Finally,

$$\begin{aligned}
\mathrm{E}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}^2 |\tau^*(t)\right] &= \mathrm{E}\left[\left\{y^*(t)^4 - 2y^*(t)^2\kappa_2\tau^*(t) + \kappa_2^2\tau^*(t)^2\right\}|\tau^*(t)\right] \\
&= \mathrm{E}\left\{y^*(t)^4|\tau^*(t)\right\} - 2\kappa_2\tau^*(t)\mathrm{E}\left\{y^*(t)^2|\tau^*(t)\right\} + \kappa_2^2\tau^*(t)^2 \\
&= \mathrm{E}\left\{y^*(t)^4|\tau^*(t)\right\} + 3\kappa_2^2\tau^*(t)^2 - 2\kappa_1^2\kappa_2\tau^*(t)^3 \\
&= \kappa_4\tau^*(t) + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\tau^*(t)^2 + 4\kappa_1^2\kappa_2\tau^*(t)^3 + \kappa_1^4\tau^*(t)^4,
\end{aligned}$$

as stated

□

2.4 Unconditional moments of log-prices

In order to derive unconditional moments of $y^*(t)$ we need to remind ourselves of the second order properties of τ^* when τ is covariance stationary.

Remark 1 *Suppose τ is a second-order or covariance stationary process with ξ, ω^2 and r being, respectively, the mean, variance and the autocorrelation function of the process τ. Then Barndorff-Nielsen and Shephard (2001) showed that*

$$\mathrm{E}\left\{\tau^*(t)\right\} = \xi t, \quad \textit{and} \quad \operatorname{Var}\{\tau^*(t)\} = 2\omega^2 r^{**}(t) \tag{10}$$

where

$$r^*(t) = \int_0^t r(u)\mathrm{d}u \quad \textit{and} \quad r^{**}(t) = \int_0^t r^*(u)\mathrm{d}u. \tag{11}$$

□

When this result is combined with Proposition 1 we have the following result.

Proposition 2 *If $\kappa_4 < \infty$ and τ is stationary and such that $\mathrm{E}\left\{\tau(t)^4\right\} < \infty$, then*

$$\mathrm{E}(y^*(t)) = \kappa_1 t\xi,$$

$$\mathrm{E}\left\{y^*(t)^2\right\} = \kappa_2 t\xi + \kappa_1^2\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\},$$

$$\mathrm{E}\left\{y^*(t)^3\right\} = \kappa_3 t\xi + 3\kappa_1\kappa_2\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\} + \kappa_1^3\mathrm{E}\left\{\tau^*(t)^3\right\},$$

$$\mathrm{E}\left\{y^*(t)^4\right\} = \kappa_4 t\xi + \left(4\kappa_1\kappa_3 + 3\kappa_2^2\right)\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\} + 6\kappa_1^2\kappa_2\mathrm{E}\left\{\tau^*(t)^3\right\} + \kappa_1^4\mathrm{E}\left\{\tau^*(t)^4\right\}.$$

Consequently

$$\mathrm{Var}(y^*(t)) = \kappa_2 t\xi + 2\kappa_1^2\omega^2 r^{**}(t),$$

$$\mathrm{Cov}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}, \kappa_2\tau^*(t)\right] = \kappa_1^2\kappa_2\mathrm{E}\left\{\tau^*(t)^3\right\},$$

$$\begin{aligned}\mathrm{E}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}^2\right] &= \kappa_4 t\xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\} \\ &\quad + 4\kappa_1^2\kappa_2\mathrm{E}\left\{\tau^*(t)^3\right\} + \kappa_1^4\mathrm{E}\left\{\tau^*(t)^4\right\}.\end{aligned}$$

□

Proof. This follows by the application of iterative expectations and (10).

□

Example 3 *The results in Proposition 2 were derived in Barndorff-Nielsen and Shephard (2001) in the special case of* $\kappa_1 = \kappa_3 = \kappa_4 = 0$. *They found that*

$$\mathrm{E}(y^*(t)) = 0, \quad \mathrm{E}\left\{y^*(t)^2\right\} = \kappa_2 t\xi, \quad \mathrm{E}\left\{y^*(t)^3\right\} = 0,$$

and, most importantly from our viewpoint,

$$\mathrm{Cov}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}, \kappa_2\tau^*(t)\right] = 0,$$

$$\mathrm{E}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}^2\right] = 2\kappa_2^2\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\}.$$

□

When $\kappa_1 \neq 0$, the third and fourth moments of the log-prices require us to calculate the third and fourth moments of τ^*, which are non-standard. The following Proposition is helpful in thinking about these extra terms when t is small, which will be important when we work with RVs in the next Section.

Proposition 3 *If* $\alpha_j = \mathrm{E}\{|\tau(t) - \xi|^j\}$ *is finite for some natural number* j *then*

$$\mathrm{E}\{(\tau^*(t) - \xi t)^j\} = O(t^j).$$

In fact,

$$|\mathrm{E}\{(\tau^*(t) - \xi t)^j\}| \leq \alpha_j t^j.$$

□

Proof.

$$\begin{aligned}
|\mathrm{E}\{(\tau^*(t)-\xi t)^j\}| &\leq \mathrm{E}\left|\int_0^t \dots \int_0^t (\tau(s_1)-\xi)\dots(\tau(s_j)-\xi)\mathrm{d}s_1\dots\mathrm{d}s_j\right| \\
&\leq \int_0^t \dots \int_0^t \mathrm{E}\left\{|\tau(s_1)-\xi)\dots(\tau(s_j)-\xi)|\right\}\mathrm{d}s_1\dots\mathrm{d}s_j.
\end{aligned}$$

Further, we can use the result that for arbitrary random variables $x_1, ..., x_j$ we have

$$\mathrm{E}\{|x_1\cdots x_j|\} \leq (\mathrm{E}\{|x_1|^j\}\cdots\mathrm{E}\{|x_j|^j\})^{1/j},$$

which is a consequence of Jensen's inequality. This implies that

$$\begin{aligned}
|\mathrm{E}\{(\tau^*(t)-\xi t)^j\}| &\leq \int_0^t \dots \int_0^t \alpha_j^{1/j}\dots\alpha_j^{1/j}\mathrm{d}s_1\dots\mathrm{d}s_j \\
&= \alpha_j \int_0^t \dots \int_0^t \mathrm{d}s_1\dots\mathrm{d}s_j \\
&= \alpha_j t^j.
\end{aligned}$$

□

If we combine Propositions 2 and 3, then we have immediately the following result.

Proposition 4 *If $\kappa_4 < \infty$ and τ is stationary and $\mathrm{E}\tau^4(t) < \infty$, then for $t \downarrow 0$,*

$$\mathrm{E}\left\{y^*(t)^3\right\} = \kappa_3 t\xi + 3\kappa_1\kappa_2\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\} + O(t^3),$$

$$\mathrm{E}\left\{y^*(t)^4\right\} = \kappa_4 t\xi + \left(3\kappa_2^2 + 2\kappa_1\kappa_3\right)\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\} + O(t^3),$$

while

$$\mathrm{Cov}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}, \kappa_2\tau^*(t)\right] = O(t^3),$$

$$\mathrm{E}\left[\left\{y^*(t)^2 - \kappa_2\tau^*(t)\right\}^2\right] = \kappa_4 t\xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 r^{**}(t) + (t\xi)^2\right\} + O(t^3).$$

□

This result means that $y^*(t)^2 - \kappa_2\tau^*(t)$ is, to a higher order approximation, uncorrelated to $\kappa_2\tau^*(t)$ and that we can characterise the variability of $y^*(t)^2 - \kappa_2\tau^*(t)$.

2.5 Second order properties of squared returns

Now let us look at the econometric properties of a sequence of returns over an interval of length $\hbar > 0$,

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

It will be convenient for us to also define the associated *actual time-changes*

$$\tau_i = \tau^*(\hbar i) - \tau^*(\hbar(i-1)), \quad i = 1, 2, \dots, n.$$

If τ is stationary, then y_i is stationary and so has the same marginal distribution as

$$y_1 = y^*(\hbar).$$

Hence Proposition 2 directly computes

$$\mathrm{E}\left(y_i\right), \quad \mathrm{E}\left(y_i^2\right), \quad \mathrm{E}\left(y_i^3\right), \quad \mathrm{E}\left(y_i^4\right), \quad \mathrm{Var}\left(y_i^2\right), \tag{12}$$

and

$$\mathrm{Var}\left\{\left(y_i^2 - \tau_i\right)\right\}, \quad \mathrm{Cov}\left[\left(y_i^2 - \kappa_2\tau_i\right), \kappa_2\tau_i\right]$$

by just setting $t = \hbar$.

The only new issue is to give a discussion of the dynamics of y_i and y_i^2. Our analysis will be based on a result in Barndorff-Nielsen and Shephard (2001) which showed that, for $s \neq 0$,

$$\mathrm{Cov}(\tau_i, \tau_{i+s}) = \omega^2 \diamondsuit r^{**}(\hbar s) \tag{13}$$

where

$$\diamondsuit r^{**}(\hbar s) = r^{**}((s+1)\,\hbar) - 2r^{**}(\hbar s) + r^{**}((s-1)\,\hbar). \tag{14}$$

This can be combined with an extension of Proposition 3 which is that as $\hbar \downarrow 0$ then

$$\mathrm{Cov}\left(\tau_i, \tau_{i+s}\right) = O(\hbar^2), \quad \mathrm{Cov}\left(\tau_i, \tau_{i+s}^2\right) = O(\hbar^3), \quad \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}^2\right) = O(\hbar^4).$$

The above results imply the following.

Proposition 5 *If τ is covariance stationary and $\kappa_2 < \infty$, then for $s \neq 0$*

$$\mathrm{Cov}\left(y_i, y_{i+s}\right) = \kappa_1^2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}\right), \quad \mathrm{Cov}\left(y_i, y_{i+s}^2\right) = \kappa_1\kappa_2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}\right) + \kappa_1^3 \mathrm{Cov}\left(\tau_i, \tau_{i+s}^2\right),$$

$$\begin{aligned} \mathrm{Cov}\left(y_i^2, y_{i+s}^2\right) &= \kappa_2^2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}\right) + \kappa_1^2\kappa_2 \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}\right) \\ &\quad + \kappa_2\kappa_1^2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}^2\right) + \kappa_1^4 \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}^2\right), \end{aligned}$$

and

$$\mathrm{Cov}\left\{\left(y_i^2 - \kappa_2\tau_i\right), \left(y_{i+s}^2 - \kappa_2\tau_{i+s}\right)\right\} = \kappa_1^4 \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}^2\right).$$

Further, for $\hbar \downarrow 0$ then

$$\mathrm{Cov}\left(y_i, y_{i+s}\right) = \kappa_1^2\omega^2 \diamondsuit r^{**}(\hbar s),$$

$$\mathrm{Cov}\left(y_i, y_{i+s}^2\right) = \kappa_1\kappa_2\omega^2 \diamondsuit r^{**}(\hbar s) + O(\hbar^3),$$

$$\mathrm{Cov}\left(y_i^2, y_{i+s}^2\right) = \kappa_2^2\omega^2 \diamondsuit r^{**}(\hbar s) + O(\hbar^3),$$

$$\mathrm{Cov}\left\{\left(y_i^2 - \kappa_2\tau_i\right), \left(y_{i+s}^2 - \kappa_2\tau_{i+s}\right)\right\} = O(\hbar^4).$$

$\square$

Proof. First

$$\begin{aligned} \mathrm{Cov}\left(y_i, y_{i+s}\right) &= \mathrm{E}\left\{\mathrm{E}(y_i y_{i+s}|\tau_i, \tau_{i+s})\right\} - \mathrm{E}\left(y_i\right)^2 \\ &= \kappa_1^2 \mathrm{E}\left(\tau_i \tau_{i+s}\right) - \kappa_1^2 \mathrm{E}\left(\tau_i\right)^2 \\ &= \kappa_1^2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}\right). \end{aligned}$$

Second

$$\begin{aligned} \mathrm{Cov}\left(y_i, y_{i+s}^2\right) &= \mathrm{E}\left\{\mathrm{E}(y_i y_{i+s}^2|\tau_i, \tau_{i+s})\right\} - \mathrm{E}\left(y_i\right)\mathrm{E}\left(y_i^2\right) \\ &= \kappa_1\kappa_2 \mathrm{E}\left(\tau_i \tau_{i+s}\right) + \kappa_1^3 \mathrm{E}\left(\tau_i \tau_{i+s}^2\right) - \kappa_1 \mathrm{E}\left(\tau_i\right)\left\{\kappa_2 \mathrm{E}\left(\tau_i\right) + \kappa_1^2 \mathrm{E}\left(\tau_i^2\right)\right\} \\ &= \kappa_1\kappa_2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}\right) + \kappa_1^3 \mathrm{Cov}\left(\tau_i, \tau_{i+s}^2\right). \end{aligned}$$

Third

$$\begin{aligned} \mathrm{Cov}\left(y_i, y_{i+s}^2\right) &= \mathrm{E}\left\{\mathrm{E}(y_i^2 y_{i+s}^2|\tau_i, \tau_{i+s})\right\} - \mathrm{E}\left(y_i^2\right)^2 \\ &= \mathrm{E}\left\{\left(\kappa_2\tau_i + \kappa_1^2\tau_i^2\right)\left(\kappa_2\tau_{i+s} + \kappa_1^2\tau_{i+s}^2\right)\right\} - \left\{\kappa_2 \mathrm{E}\left(\tau_i\right) + \kappa_1^2 \mathrm{E}\left(\tau_i^2\right)\right\}^2 \\ &= \kappa_2^2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}\right) + \kappa_1^2\kappa_2 \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}\right) \\ &\quad + \kappa_2\kappa_1^2 \mathrm{Cov}\left(\tau_i, \tau_{i+s}^2\right) + \kappa_1^4 \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}^2\right). \end{aligned}$$

Finally

$$\begin{aligned} &\mathrm{Cov}\left\{\left(y_i^2 - \kappa_2\tau_i\right), \left(y_{i+s}^2 - \kappa_2\tau_{i+s}\right)\right\} \\ =\ &\mathrm{E}\left\{\mathrm{E}\left(y_i^2 - \kappa_2\tau_i\right)\left(y_{i+s}^2 - \kappa_2\tau_{i+s}\right)|\tau_i, \tau_{i+s}\right\} - \left\{\mathrm{E}\left(y_{i+s}^2 - \kappa_2\tau_{i+s}\right)\right\}^2 \\ =\ &\kappa_1^4 \mathrm{Cov}\left(\tau_i^2, \tau_{i+s}^2\right). \end{aligned}$$

□

3 Cumulants of realised variance

3.1 Setting the scene: the theory of quadratic variation

It is well known that if y^* is a semimartingale then the probability limit of $[y_M^*]_i$ as $M \to \infty$ is defined by quadratic variation. Recall the *quadratic variation* (QV) process is well defined (e.g. Jacod and Shiryaev (1987, p. 55)) for any semimartingale y^* and can be written as

$$[y^*](t) = \mathrm{p}\!-\!\lim_{M\to\infty} \sum_{j=0}^{M-1} \{y^*(t_{j+1}) - y^*(t_j)\}^2, \tag{15}$$

for any sequence of partitions $t_0 = 0 < t_1 < ... < t_M = t$ with $\sup_j\{t_{j+1} - t_j\} \to 0$ for $M \to \infty$. Here $\mathrm{p}-\lim$ denotes the probability limit of the sum. Thus QV can be thought of as the sum of squares of returns computed over infinitesimal time intervals calculated during the period from time 0 up to time t. Geman, Madan, and Yor (2002) has studied the joint law of $\tau^*(t)$ and $[y^*](t)$ in the case where $z = w$ and τ^* is a subordinator, that is a Lévy process with non-negative increments. Winkel (2002) extends some of this work, in particular to more general Lévy processes.

The definition of QV immediately implies that for all semimartingales as $M \to \infty$

$$[y_M^*]_i \xrightarrow{p} [y^*](\hbar i) - [y^*](\hbar\,(i-1)) = [y^*]_i.$$

Importantly, if α^* is continuous then generically this simplifies to

$$[y_M^*]_i \xrightarrow{p} [m^*]_i = [m^*](\hbar i) - [m^*](\hbar\,(i-1)).$$

This is very well known (e.g. Barndorff-Nielsen and Shephard (2002b)).

Example 4 *Suppose α^* and τ^* have continuous sample paths and $m^*(t) = \int_0^t \tau^{1/2}(u)\mathrm{d}w(u) = w(\tau^*(t))$, a SV process, then $y^*(t)$ has continuous sample paths and $[y^*](t) = [m^*](t)$. Further, it can be shown that $[m^*](t) = \tau^*(t)$, which implies that*

$$[y_M^*]_i \xrightarrow{p} \tau^*(i\hbar) - \tau^*((i-1)\,\hbar) = \tau_i.$$

Hence RV consistently estimates the time change of the Brownian motion. Under some additional regularity assumptions Barndorff-Nielsen and Shephard (2002a) and Barndorff-Nielsen and Shephard (2002b) have additionally proved that

$$\frac{\sqrt{\frac{M}{\hbar}}\left([y_M^*]_i - \tau_i\right)}{\sqrt{2\int_{(i-1)\hbar}^{i\hbar} \tau^2(u)\mathrm{d}u}} \xrightarrow{d} N(0,1), \tag{16}$$

as $M \to \infty$. $\square$

In the more general case of a time-changed Lévy process, $m^* = z(\tau^*)$, we have the more challenging situation that

$$[m^*](t) \neq \kappa_2 \tau^*(t), \quad \text{so} \quad [y_M^*]_i \xrightarrow{p} [y^*]_i \neq \kappa_2 \tau_i.$$

3.2 Basic results on realised variance

3.2.1 Mean and variance of $[y_M^*]_i - \kappa_2\tau_i$

It is clear that an econometrician interested in forecasting the variability of returns should focus on forecasting $\kappa_2\tau_i$, which we estimate by $[y_M^*]_i$. Here we study the properties of the RV error $[y_M^*]_i - \kappa_2\tau_i$.

Proposition 6 *Suppose τ is covariance stationary and $\kappa_4 < \infty$, then*

$$\begin{aligned}
\mathrm{E}\left([y_M^*]_i - \kappa_2\tau_i\right) &= \kappa_1^2\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\} \\
\mathrm{Var}\left([y_M^*]_i - \kappa_2\tau_i\right) &= \kappa_4\hbar\xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\} + O(M^{-2}),
\end{aligned}$$

$$\mathrm{Cov}\left\{\left([y_M^*]_i - \kappa_2\tau_i\right), \kappa_2\tau_i\right\} = O(M^{-2})$$

and

$$\mathrm{Cov}\left\{\left([y_M^*]_i - \kappa_2\tau_i\right), \left([y_M^*]_{i+s} - \kappa_2\tau_{i+s}\right)\right\} = O(M^{-3}).$$

□

Proof. It is useful to write

$$\tau_{j,i} = \tau^*\left((i-1)\hbar + \frac{\hbar j}{M}\right) - \tau^*\left((i-1)\hbar + \frac{\hbar(j-1)}{M}\right), \quad j = 1, ..., M.$$

Then we can decompose

$$[y_M^*]_i - \kappa_2\tau_i = \sum_{j=1}^{M}\left(y_{j,i}^2 - \tau_{j,i}\right).$$

Now Propositions 5, 4 and 1 tells us that

$$\begin{aligned}
\mathrm{E}(y_{j,i}^2 - \tau_{j,i}) &= \kappa_1^2\left\{2\omega^2 r^{**}\left(\tfrac{\hbar}{M}\right) + \left(\tfrac{\hbar}{M}\xi\right)^2\right\} \\
\mathrm{Var}(y_{j,i}^2 - \tau_{j,i}) &= \begin{array}{l}\kappa_4\left(\tfrac{\hbar}{M}\right)\xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 r^{**}\left(\tfrac{\hbar}{M}\right) + \left(\tfrac{\hbar}{M}\xi\right)^2\right\} \\ +4\kappa_1^2\kappa_2\mathrm{E}\left\{\tau^*\left(\tfrac{\hbar}{M}\right)^3\right\} + \kappa_1^4\mathrm{E}\left\{\tau^*\left(\tfrac{\hbar}{M}\right)^4\right\},\end{array} \\
\mathrm{Cov}(y_{j,i}^2 - \tau_{j,i}, y_{j+s,i}^2 - \tau_{j+s,i}) &= O(M^{-4}).
\end{aligned}$$

This implies the desired result.

□

In the $\kappa_1 = \kappa_4 = 0$ and $\kappa_2 = 1$ case this reproduces the Barndorff-Nielsen and Shephard (2002a) result, who also showed that $r^{**}(t)$ is $O(t^2)$ as $t \downarrow 0$ and so $Mr^{**}\left(\hbar M^{-1}\right) = O(M^{-1})$ as $M \to \infty$.

The above means our Proposition 6 tells us the following.

- The bias of $[y_M^*]_i$ as an estimator of $\kappa_2\tau_i$. We see this is $O(M^{-1})$.
- The variability of $[y_M^*]_i - \kappa_2\tau_i$, the RV error. When $\kappa_4 > 0$ this involves a term which does not disappear as $M \to \infty$. This is important as this captures the inconsistency of $[y_M^*]_i$ as an estimator of $\kappa_2\tau_i$ in the non-Brownian time-change model. The next order term is analytically calculable and is important even if $\kappa_4 = 0$, while the other terms are of $O(M^{-2})$. Typically, all terms of $O(1)$ and $O(M^{-1})$ dominate the bias in terms of the mean square error for the RV as an estimator of the time-change. Hence the bias has little impact in practice.

- The $\text{Cov}\left\{([y_M^*]_i - \kappa_2\tau_i), \kappa_2\tau_i\right\}$ tells us that the time-change and RV error are, to a high order, uncorrelated.

- The $\text{Cov}\left\{([y_M^*]_i - \kappa_2\tau_i), ([y_M^*]_{i+s} - \kappa_2\tau_{i+s})\right\}$ informs us that to a very high order, the RV errors are uncorrelated.

Overall, the movement from Brownian motion to the Lévy time-change model has really only impacted the variability of the RV error. This is the most important point we make in this paper.

3.2.2 Second order properties of $[y_M^*]_i$

To carry out inference on these types of models it is helpful to know the second order properties of $[y_M^*]_i$.

Proposition 7 *If $\kappa_4 < \infty$ and τ is covariance stationary, then*

$$\text{E}\left([y_M^*]_i\right) = \kappa_2\hbar\xi + \kappa_1^2\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\},$$

$$\text{Var}\left([y_M^*]_i\right) = 2\omega^2\kappa_2^2 r^{**}(\hbar) + \kappa_4\hbar\xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\} + O(M^{-2}),$$

while, for $s \neq 0$,

$$\begin{aligned}\text{Cov}\left([y_M^*]_i, [y_M^*]_{i+s}\right) &= \kappa_2^2\text{Cov}\left(\tau_i, \tau_{i+s}\right) + O(M^{-2}) \\ &= \kappa_2^2\omega^2\diamondsuit r^{**}(\hbar s) + O(M^{-2}).\end{aligned}$$

□

Proof. We have that

$$[y_M^*]_i = \kappa_2\tau_i + \left([y_M^*]_i - \kappa_2\tau_i\right).$$

Then (10) and Proposition 6 imply that

$$\text{E}\left([y_M^*]_i\right) = \kappa_2\hbar\xi + \kappa_1^2\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\}.$$

Likewise

$$\begin{aligned}\text{Var}\left([y_M^*]_i\right) &= \text{Var}\left(\kappa_2\tau_i\right) + \text{Var}\left([y_M^*]_i - \kappa_2\tau_i\right) + 2\text{Cov}\left\{([y_M^*]_i - \kappa_2\tau_i), \kappa_2\tau_i\right\} \\ &= 2\omega^2\kappa_2^2 r^{**}(\hbar) + \kappa_4\hbar\xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\} + O(M^{-2}).\end{aligned}$$

Finally,

$$\text{Cov}\left([y_M^*]_i, [y_M^*]_{i+s}\right) = \kappa_2^2\text{Cov}\left(\tau_i, \tau_{i+s}\right) + O(M^{-2}).$$

□

From now on in the paper we will ignore $O(M^{-2})$ terms in the variability terms and $O(M^{-1})$ in the bias. This leads to the approximations

$$\mathrm{E}\left([y_M^*]_i\right) \simeq \kappa_2 \hbar \xi, \quad \mathrm{Cov}\left\{\left([y_M^*]_i - \kappa_2 \tau_i\right), \kappa_2 \tau_i\right\} \simeq 0, \quad \mathrm{Cov}\left\{\left([y_M^*]_i - \kappa_2 \tau_i\right), \kappa_2 \tau_{i+s}\right\} \simeq 0,$$

$$\mathrm{Var}\left([y_M^*]_i\right) \simeq 2\omega^2 \kappa_2^2 r^{**}(\hbar) + \kappa_4 \hbar \xi + \left(4\kappa_1\kappa_3 + 2\kappa_2^2\right)\left\{2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + M^{-1}\hbar^2\xi^2\right\},$$

$$\mathrm{Cov}\left([y_M^*]_i, [y_M^*]_{i+s}\right) \simeq \kappa_2^2 \omega^2 \diamondsuit r^{**}(\hbar s),$$

4 Some implications for autocorrelation and quarticity

4.1 Autocorrelation function of realised variances

Notice that $\mathrm{E}\left([y_M^*]_i\right)$ and $\mathrm{Cov}\left([y_M^*]_i, [y_M^*]_{i+s}\right)$ are unaffected by the jumps in the process. Taken together these results mean that

$$\mathrm{Cor}\left([y_M^*]_i, [y_M^*]_{i+s}\right) = \frac{\kappa_2^2 \omega^2 \diamondsuit r^{**}(\hbar s)}{2\omega^2 \kappa_2^2 r^{**}(\hbar) + \kappa_4 \hbar \xi + 4\kappa_2^2 \omega^2 M r^{**}\left(\frac{\hbar}{M}\right) + 2M^{-1}\kappa_2^2\left(\hbar\xi\right)^2}.$$

Thus the autocorrelation function of $[y_M^*]_i$ is monotonically decreasing in κ_4. Further, as $M \to \infty$ and assuming $\kappa_2 = 1$,

$$\begin{aligned}
\mathrm{Cor}\left([y_M^*]_i, [y_M^*]_{i+s}\right) &\to \frac{\omega^2 \diamondsuit r^{**}(\hbar s)}{2\omega^2 r^{**}(\hbar) + \kappa_4 \hbar \xi} \\
&= \mathrm{Cor}\left([y^*]_i, [y^*]_{i+s}\right) \\
&\leq \; = \frac{\diamondsuit r^{**}(\hbar s)}{2 r^{**}(\hbar)} \\
&= \mathrm{Cor}\left(\tau_i, \tau_{i+s}\right).
\end{aligned}$$

Of course the equality, in the inequality, is obtained only in the Brownian case. Otherwise the autocorrelation amongst the $[y_M^*]_i$ will systematically underestimate the autocorrelation in the τ_i and so the predictability in the volatility process. This resonates with the modern methodological literature on volatility forecasting by Andersen and Bollerslev (1998), Andersen, Bollerslev, and Meddahi (2002a) and Andersen, Bollerslev, and Meddahi (2002b) which has shown that the volatility of financial markets is much more predictable than is widely believed in the academic literature.

An important alternative asymptotics is to allow $\kappa_4 \to \infty$, while fixing $\kappa_2 = 1$. In this case

$$\mathrm{Cor}\left([y_M^*]_i, [y_M^*]_{i+s}\right) \underset{\kappa_4 \to \infty}{\to} 0,$$

as the variance of the squares of the returns becomes infinity. This holds whatever the value of M and however dependent is the volatility process.

The general result as $M \to \infty$ implies that

$$\begin{aligned}
\mathrm{Cor}\left([y_M^*]_i, \kappa_2\tau_i\right) &= \frac{\mathrm{Var}\left(\kappa_2\tau_i\right)}{\sqrt{\mathrm{Var}\left(\kappa_2\tau_i\right)\left\{\mathrm{Var}\left(\kappa_2\tau_i\right)+\mathrm{Var}(u_{M:i})\right\}}} \\
&= \frac{2\omega^2\kappa_2^2 r^{**}(\hbar)}{\sqrt{2\omega^2\kappa_2^2 r^{**}(\hbar)\left\{2\omega^2\kappa_2^2 r^{**}(\hbar)+\kappa_4\hbar\xi+4\kappa_2^2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right)+2M^{-1}\kappa_2^2\left(\hbar\xi\right)^2\right\}}} \\
&\to \frac{r^{**}(\hbar)}{\sqrt{r^{**}(\hbar)\left\{r^{**}(\hbar)+\frac{\kappa_4}{2\omega^2\kappa_2^2}\hbar\xi\right\}}},
\end{aligned}$$

as $M \to \infty$. Of course this is 1 iff $\kappa_4 = 0$. Again this correlation can be driven to be arbitrarily close to zero by allowing κ_4 to become large.

4.1.1 Meddahi regression

In the Brownian case Meddahi (2002a) has argued that we should replace $[y_M^*]_i$ by the regression estimator

$$\widehat{\kappa_2\tau_i} = (1-\beta_M)\,\mathrm{E}\left([y_M^*]_i\right) + \beta_M[y_M^*]_i, \quad \text{where} \quad \beta_M = \frac{\mathrm{Var}\left(\kappa_2\tau_i\right)}{\mathrm{Var}\left([y_M^*]_i\right)} \in [0,1].$$

Of course, as $M \to \infty$,

$$\begin{aligned}
\beta_M &= \frac{\mathrm{Var}\left(\kappa_2\tau_i\right)}{\mathrm{Var}\left([y_M^*]_i\right)} \\
&= \frac{2\omega^2\kappa_2^2 r^{**}(\hbar)}{2\omega^2\kappa_2^2 r^{**}(\hbar)+\kappa_4\hbar\xi+4\kappa_2^2\omega^2 M r^{**}\left(\frac{\hbar}{M}\right)+2M^{-1}\kappa_2^2\left(\hbar\xi\right)^2} \\
&\to \frac{r^{**}(\hbar)}{r^{**}(\hbar)+\frac{\kappa_4}{2\omega^2\kappa_2^2}\hbar\xi} \\
&= \beta \leq 1.
\end{aligned}$$

Now

$$\widehat{\kappa_2\tau_i} \underset{M\to\infty}{\overset{p}{\to}} (1-\beta)\,\kappa_2\hbar\xi + \beta[y^*]_i.$$

Again, inevitably, this is an inconsistent estimator of $\kappa_2\tau_i$ for $z \neq w$. Andersen, Bollerslev, and Meddahi (2002a) and Andersen, Bollerslev, and Meddahi (2002b) have studied in detail other properties of these estimators in the Brownian motion time-deformation case.

4.2 Quarticity

Barndorff-Nielsen and Shephard (2002a) introduced the idea of quarticity $\sum_{j=1}^{M} y_{j,i}^4$ as a measure of the variability of $[y_M^*]_i - \kappa_2\tau_i$. In the Brownian time-change model they showed that, under suitable regularity conditions,

$$\frac{M}{3\hbar}\sum_{j=1}^{M} y_{j,i}^4 \to \int_{(i-1)\hbar}^{i\hbar} \tau^2(u)\mathrm{d}u,$$

which can be used to convert the infeasible limit theory (16) into the feasible version

$$\frac{\left(\sum_{j=1}^{M} y_{j,i}^2 - \tau_i\right)}{\sqrt{\frac{2}{3}\sum_{j=1}^{M} y_{j,i}^4}} \xrightarrow{d} N(0,1).$$

An alternative feasible limit theory has recently been introduced by Barndorff-Nielsen and Shephard (2002b). They showed that the dynamic quarticities

$$\frac{M}{\hbar}\sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2 \xrightarrow{p} \int_{(i-1)\hbar}^{i\hbar} \tau^2(u)\mathrm{d}u, \quad \text{and} \quad \frac{M}{\hbar}\left(\sum_{j=1}^{M} y_{j,i}^4 - \sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2\right) \xrightarrow{p} 2\int_{(i-1)\hbar}^{i\hbar} \tau^2(u)\mathrm{d}u$$

which implies the alternative asymptotic relations

$$\frac{\left(\sum_{j=1}^{M} y_{j,i}^2 - \tau_i\right)}{\sqrt{2\sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2}} \xrightarrow{d} N(0,1), \quad \text{or} \quad \frac{\left(\sum_{j=1}^{M} y_{j,i}^2 - \tau_i\right)}{\sqrt{\sum_{j=1}^{M} y_{j,i}^4 - \sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2}} \xrightarrow{d} N(0,1).$$

It is $\sum_{j=1}^{M} y_{j,i}^4 - \sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2$ which will turn out in a moment to prove the most helpful in the context of robustifying this result to Lévy jumps[1].

This implies that if $\kappa_4 < \infty$ and τ is stationary and $\mathrm{E}\tau^4(t) < \infty$, then as $M \to \infty$

$$\mathrm{E}\left\{y_{i,j}^4\right\} = \kappa_4\left(\frac{\hbar}{M}\right)\xi + \left(3\kappa_2^2 + 4\kappa_1\kappa_3\right)\left\{2\omega^2 r^{**}\left(\frac{\hbar}{M}\right) + \left(\xi\frac{\hbar}{M}\right)^2\right\} + O(M^{-3}),$$

while

$$\mathrm{E}\left(y_{j,i}^2 y_{j+1,i}^2\right) = \kappa_2^2\left\{\omega^2 \diamondsuit r^{**}\left(\frac{\hbar}{M}\right) + \left(\xi\frac{\hbar}{M}\right)^2\right\} + O(M^{-3}).$$

The important point is that the expectation of $y_{j,i}^2 y_{j+1,i}^2$ is unaffected by κ_1, κ_3 and, most importantly, κ_4.

It will be helpful to remind ourselves that Barndorff-Nielsen and Shephard (2002a) calculated the moments of $\sum_{j=1}^{M} y_{j,i}^4$, under covariance stationarity of τ and Brownian motion. Now

$$3M^2\left\{2\omega^2 r^{**}\left(\frac{\hbar}{M}\right) + \left(\xi\frac{\hbar}{M}\right)^2\right\} \to 3\hbar^2\left(\omega^2 + \xi^2\right),$$

as $M \to \infty$, using the fact proved by Barndorff-Nielsen and Shephard (2002a) that $2M^2 r^{**}\left(\frac{\hbar}{M}\right) \to \hbar^2$. Likewise

$$\mathrm{E}\left(M\sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2\right) \to \hbar^2\left(\omega^2 + \xi^2\right).$$

[1] The non-negativity follows from the result that for all a, b then $a^2 - 2ab + b^2 \geq 0$. Hence $\sum_{j=1}^{M} y_{j,i}^4 - \sum_{j=1}^{M-1} y_{j,i}^2 y_{j+1,i}^2$ is zero iff $y_{j,i} = 0$ for all j.

Thus

$$\sum_{j=1}^{M} \mathrm{E}\left(y_{j,i}^{4}\right) - \sum_{j=1}^{M-1} \mathrm{E}\left(y_{j,i}^{2} y_{j+1,i}^{2}\right) = \kappa_4 \hbar \xi + \left(2\kappa_2^2 + 4\kappa_1\kappa_3\right) \frac{\hbar^2}{M} \left(\omega^2 + \xi^2\right) + o(M^{-1}).$$

Hence $\sum_{j=1}^{M} y_{j,i}^{4} - \sum_{j=1}^{M-1} y_{j,i}^{2} y_{j+1,i}^{2}$ is a model free, although noisy, higher order unbiased estimator of the variance of the RV error so long as either $\kappa_1 = 0$ or $\kappa_3 = 0$.

5 Quasi-likelihood estimation

5.1 Based on low-frequency squared returns

The theory we have been developing show that if $\kappa_1 = 0$ and $\kappa_4 < \infty$ and τ is second order stationary then we can compute the unconditional mean and covariance of $\underline{y} = \left(y_1^2, \ldots, y_n^2\right)'$ as a function of the parameters of the model. We write the parameter vector as θ. Here n denotes the sample size.

We can define a Gaussian quasi-likelihood function

$$\log L_Q(\theta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log \left|\mathrm{Cov}(\underline{y})\right| - \frac{1}{2} \left\{\underline{y} - \mathrm{E}(\underline{y})\right\}' \left\{\mathrm{Cov}(\underline{y})\right\}^{-1} \left\{\underline{y} - \mathrm{E}(\underline{y})\right\},$$

which allows us to find

$$\widehat{\theta}_Q = \arg\max_{\theta} \log L_Q(\theta).$$

Clearly $\widehat{\theta}_Q$ is suboptimal as $\underline{y}$ is not Gaussian, however it should be consistent and its asymptotic distribution theory can be computed using the general theory of method of moments estimation using the fact that the score has zero expectation under the second order properties of the model. In particular, if we write

$$\mathcal{J} = \lim_{n\to\infty} \frac{1}{n} \mathrm{Cov}\left(\frac{\partial \log L_Q(\theta)}{\partial \theta}\right) \quad \text{and} \quad \mathcal{I} = -\lim_{n\to\infty} \frac{1}{n} \mathrm{E}\left\{\frac{\partial^2 \log L_Q(\theta)}{\partial \theta \partial \theta'}\right\},$$

then

$$\sqrt{n}\left(\widehat{\theta}_Q - \theta\right) \xrightarrow{d} N(0, \mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}).$$

Typically, in theory, we have to estimate $\mathcal{J}$ using spectral matrix methods (e.g. Newey and West (1987)). However, extensive empirical work we have carried out suggests that the serial dependence in the elements of the quasi-score vector is very small indeed.

In practice, when it comes to the estimation of the parameters which index these models, it makes sense to constrain $\kappa_2 = 1$ and allow κ_4 to be freely determined. This type of constraint has to be imposed when the model is estimated using just the second order properties of y_i^2 or $[y_M^*]_i$. Otherwise the model will be unidentified.

Usually n is quite large in financial economics when $\hbar = 1$ represents one day and so computing $\log_Q$ is onerous due to the need to compute the inverse of the $n \times n$ matrix $\mathrm{Cov}(\underline{y})$, which is typically an $O\left(n^3\right)$ operation. However, stationarity of τ means that $\underline{y}$ is itself stationary and so Cov($\underline{y}$) must be Toeplitz. The Durbin algorithm[2] (e.g. Golub and Van Loan (1989, p. 187) and Doornik and Ooms (2003, p. 29)) can be used to compute $\log L_Q$ in $O\left(n^2\right)$. This works with the Choleski decomposition of the inverse of Cov($\underline{y}$). We write Cov($\underline{y}$) as LDL', where L is lower triangular and D is diagonal, so

$$\log L_Q(\theta) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|D| - \frac{1}{2}e'e, \quad \text{where} \quad e = L^{-1}D^{-1/2}\left\{\underline{y} - \mathrm{E}(\underline{y})\right\}.$$

Note that L^{-1} is also lower triangular. Importantly the i-th diagonal elements of D are the variances of the best linear, unbiased one-step ahead forecasts of y_i^2,

$$\widehat{y_{i|i-1}^2} = P_L\left(y_i^2|y_1^2, ..., y_{i-1}^2\right)$$

while the i-th element of e has the associated one-step ahead forecast errors

$$e_i = y_i^2 - \widehat{y_{i|i-1}^2}.$$

We should note that Engle (2002) has recently used simulation to approximate these best linear estimators in the context of the Brownian-SV models.

Great computational gains can be made in the special case where τ_i can be represented as

$$\tau_i = \xi + x\varpi_i, \quad \varpi_{i+1} = \mathcal{T}\varpi_i + v_i, \tag{17}$$

where x is a selection matrix, $\mathcal{T}$ is a fixed matrix which may be indexed by some parameters, the "state vector" ϖ_i is a fixed, finite dimension and v_i is a zero mean, weak white noise process. Then we can compute $\log L_Q$ in $O(n)$ computations using the Kalman filter (e.g. Harvey (1993, Ch. 4) and Durbin and Koopman (2001, Ch. 4.2) for textbook expositions and Koopman, Shephard, and Doornik (1999) for computational tools for carrying out the calculations). This follows from the fact that

$$y_i^2 = \kappa_2\tau_i + u_i,$$

where the properties of u_i were given in Proposition 2, can be combined with (17) to put y_i^2 into a linear state space representation. In particular if we write the best linear predictor of ϖ_i using $y_1^2, y_2^2, \ldots, y_{i-1}^2$ as $a_{i|i-1}$ and the associate mean square error matrix as $P_{i|i-1}$, then

$$\widehat{\tau}_{i|i-1} = \xi + x_i a_{i|i-1}, \quad f_i = xP_{i|i-1}x' + \mathrm{Var}(u_i),$$

and so[3]

$$\log_Q(\theta) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{i=1}^n \log f_i - \frac{1}{2}\sum_{i=1}^n f_i^{-1}\left(y_i^2 - \kappa_2\widehat{\tau}_{i|i-1}\right)^2. \tag{18}$$

[2]Most matrix languages have functions which carry out the Durbin algorithm, e.g. in Ox it is `pacf`.

[3]Of course, this relates back to the output from Durbin's algorithm with $\widehat{y_{i|i-1}^2} = \kappa_2\widehat{\tau}_{i|i-1}$ and f_i being the diagonal elements of the D matrix.

Informal checks suggest that when $n = 3,000$ and the dimension of the state space is two then the Kalman filter computes $\log L_Q$ around 20 times faster than when we use the Durbin algorithm, although we should note that the Kalman filter's computational load is quadratic in the dimension of the state and so can become slower than the Durbin algorithm for very large dimensional state vectors. When n is under $1,000$ the difference between the two algorithms is not very substantial. Overall, in both the cases of the Durbin algorithm or the Kalman filter the calculation of the likelihood is, using a modern PC, very fast.

Being able to write τ_i into (17) is restrictive. We will give a number of examples where this is possible, however typically models with long-memory features will not be able to be written in this way. In this case we are forced back to the somewhat slower method of Durbin.

Before we move on we note there is a very large and stimulating literature on alternative ways of estimating SV models. Some of this literature is reviewed in Shephard (1996) and Ghysels, Harvey, and Renault (1996). Broadly this literature splits into (i) simulation based Bayesian analysis via Markov chain Monte Carlo (e.g. Kim, Shephard, and Chib (1998)), (ii) simulation based indirect inference (e.g. Gourieroux, Monfort, and Renault (1993), Gallant and Tauchen (1996)), (iii) generalised method of moments (e.g. Andersen and Sorensen (1996)), (iv) simulation based maximum likelihood using important sampling (e.g. Durham and Gallant (2002)).

5.2 Inference based on realised variances

Precisely the same approach can be used when we replace squared returns with RVs. Now we can argue exactly when $\kappa_1 = \kappa_3 = 0$ or asymptotically approximately for large M when this does not hold. In either case we can also compute the asymptotic approximations to the unconditional mean and covariance of $\underline{[y_M^*]} = ([y_M^*]_1, \ldots, [y_M^*]_n)'$ as a function of the parameters of the model.

We can define a Gaussian realised quasi-likelihood function

$$\begin{aligned} \log L_{RQ}(\theta) \quad &= \quad -\frac{n}{2}\log 2\pi - \frac{1}{2}\log\left|\mathrm{Cov}\left(\underline{[y_M^*]}\right)\right| \\ &\quad -\frac{1}{2}\left\{\underline{[y_M^*]} - \mathrm{E}\left(\underline{[y_M^*]}\right)\right\}'\left\{\mathrm{Cov}\left(\underline{[y_M^*]}\right)\right\}^{-1}\left\{\underline{[y_M^*]} - \mathrm{E}\left(\underline{[y_M^*]}\right)\right\}, \end{aligned}$$

which allows us to define

$$\widetilde{\theta}_{RQ} = \arg\max_{\theta} \log L_{RQ}(\theta).$$

Again $\widetilde{\theta}_{RQ}$ is suboptimal as $\underline{[y_M^*]}$ is not Gaussian, however it is again consistent and its asymptotic distribution theory can be computed using the general theory of method of moments estimation. Further, informally, one would expect it to be much more efficient than inference based on

$\log L_Q(\theta)$ as the added noise $[y_M^*]_i - \kappa_2 \tau_i$ is much smaller when M is large. Again the quasi-likelihood function can be computed using the Durbin algorithm as $[y_M^*]_i$ is stationary in $O(n^2)$ while in the special case where $[y_M^*]_i$ can be placed into a linear state space form then $\log L_{RQ}(\theta)$ can be computed in $O(n)$.

Previous related work on estimating parameters of SV models includes Bollerslev and Zhou (2002) and Barndorff-Nielsen and Shephard (2002a). The former paper looked at using a method of moments procedure on some special cases of Brownian time-change models, while Barndorff-Nielsen and Shephard (2002a) used the above quasi-likelihood in the case where the model can be handled by the Kalman filter.

6 Illustration based on daily returns

6.1 Single factor model

We start with the simplest model for τ where

$$r(s) = \mathrm{Cor}(\tau(t), \tau(t+s)) = \exp(-\lambda s).$$

This is the autocorrelation function of an OU process, suggested in this context by Barndorff-Nielsen and Shephard (2001) and also for the CIR variance process (e.g. Heston (1993)). Then $r^{**}(s) = \lambda^{-2}(e^{-\lambda s} - 1 + \lambda s)$, which is enough to analytically characterise the autocovariance function of y_i^2 and so compute the quasi-likelihood function directly using Durbin's method.

In order to use the Kalman filter in this context we need to perform more analytic calculations. A straightforward manipulation from (13) implies that

$$\mathrm{Cov}(\tau_i, \tau_{i+s}) = \omega^2 \lambda^{-2} (1 - e^{-\lambda \hbar})^2 e^{-\lambda \hbar (s-1)},$$

which is the autocovariance function of an ARMA$(1,1)$ process and hence can be placed into a linear state space form with a two dimensional state. We write this representation as

$$(\tau_i - \hbar \xi) = \phi \left(\tau_{i-1} - \hbar \xi\right) + e_i + \theta e_{i-1}.$$

The autoregressive root is $\phi = e^{-\lambda \hbar}$. Barndorff-Nielsen and Shephard (2002a) noted this but then worked out θ and $\mathrm{Var}(e_i)$ using numerical methods. We now see that the moving average root and variance of e_i can be found analytically using the following lemma. This result is based on having analytically available $\mathrm{Var}(\tau_i)$ and $\mathrm{Cor}(\tau_i, \tau_{i-1})$, which is the case in this class of models.

Lemma 1 *(after a related result by Meddahi (2002b) on the process for y_i^2). Straightforwardly* $\mathrm{Var}(e_i)\left(1 + \theta^2 + 2\theta\phi\right) = \mathrm{Var}(\tau_i)\left(1 - \phi^2\right)$. *Now write $c_i = \tau_i - \phi \tau_{i-1}$, then $c_i = e_i + \theta e_{i-1}$ so* *that*

$$\operatorname{Var}(c_i) = \left(1+\phi^2\right)\operatorname{Var}(\tau_i) - 2\phi\operatorname{Cov}(\tau_i, \tau_{i-1})$$

and

$$\begin{aligned}
\operatorname{Cov}(c_i, c_{i-1}) &= \left(1+\phi^2\right)\operatorname{Cov}(\tau_i, \tau_{i-1}) - \phi\operatorname{Var}(\tau_i) - \phi\operatorname{Cov}(\tau_i, \tau_{i-2}) \\
&= \left(1+\phi^2\right)\operatorname{Cov}(\tau_i, \tau_{i-1}) - \phi\operatorname{Var}(\tau_i) - \phi^2\operatorname{Cov}(\tau_i, \tau_{i-1}) \\
&= \operatorname{Cov}(\tau_i, \tau_{i-1}) - \phi\operatorname{Var}(\tau_i) \\
&= \operatorname{Var}(\tau_i)\left\{\operatorname{Cor}(\tau_i, \tau_{i-1}) - \phi\right\}.
\end{aligned}$$

Note that $\operatorname{Cor}(\tau_i, \tau_{i-1}) \geq \phi$ *as* $e^{\lambda\hbar} - e^{-\lambda\hbar} \geq 2\lambda\hbar$. *Write*

$$\rho_1 = \frac{\operatorname{Cov}(c_i, c_{i-1})}{\operatorname{Var}(c_i)} \in \left[0, \frac{1}{2}\right], \quad \textit{then} \quad \theta = \frac{1-\sqrt{1-4\rho_1^2}}{2\rho_1} \in [0,1].$$

Remark 2 *Note that*

$$\rho_1 = \frac{\left\{\operatorname{Cor}(\tau_i, \tau_{i-1}) - \phi\right\}}{(1+\phi^2) - 2\phi\operatorname{Cor}(\tau_i, \tau_{i-1})},$$

which only depends on λ *and* $\hbar$. *Hence this is also true for* θ. *Numerical experiments suggest that for a wide set of parameter values* θ *is usually around* 0.25 *when* $\hbar = 1$.

For this model we can place it into a linear state space form with $\tau_i = \xi + (1\ 0)\varpi_i$ and

$$\varpi_{i+1} = \begin{pmatrix} (\tau_{i+1} - \xi\hbar) \\ \theta e_{i+1} \end{pmatrix} = \begin{pmatrix} \phi & 1 \\ 0 & 0 \end{pmatrix}\varpi_i + \begin{pmatrix} 1 \\ \theta \end{pmatrix} e_i,$$

which allows a very fast quasi-likelihood evaluation when $n = 3,000$. For these models the estimation is carried out in just a couple of seconds.

6.2 Empirical results

To illustrate these methods we will study the value of various exchange rates against the US Dollar recorded daily by Datastream from 26 July 1985 to 28th July 2000. Throughout, for simplicity of exposition, we report 100 times the returns.

Table 1 shows the maximum quasi-likelihood estimates of this simple model, with and without Lévy effects. The Table suggests that Lévy based time-deformation models dominate Brownian SV models in a consistent manner using this criteria. When we impose Brownian motion we set κ_4 to be zero. Throughout when κ_4 is allowed to be freely determined on the positive half-line, it is estimated to be strongly positive and leads to an increase in the quasi-likelihood. Further, this Table shows that the presence of the Lévy process allows the estimated value of ξ to be in-line with the empirical version of $\operatorname{E}\left(y_i^2\right)$. Hence the model will exhibit so-called variance tracking — the estimated model will have returns with the same unconditional variance as that

Currency Implied moments	QML estimators $\xi = \mathrm{E}\left(y_i^2\right)$	ω^2	κ_4	λ	$\log L_Q$	Implied $\mathrm{Var}\left(y_i^2\right)$
Can	.110	.00633		.0322	730.4	.0434
$\widehat{\mathrm{E}\left(y_i^2\right)} = .0890$	.0887	.00495	.135	.0400	734.0	.0424
$\widehat{\mathrm{Var}\left(y_i^2\right)} = .0424$						
DM	.616	.110		.0196	-5626.3	1.09
$\widehat{\mathrm{E}\left(y_i^2\right)} = .464$	.464	.0722	.893	.0370	-5619.0	1.06
$\widehat{\mathrm{Var}\left(y_i^2\right)} = 1.06$						
FF	.576	.172		.0343	-5729.1	1.17
$\widehat{\mathrm{E}\left(y_i^2\right)} = .449$	.448	.129	.803	.0448	-5724.2	1.14
$\widehat{\mathrm{Var}\left(y_i^2\right)} = 1.14$						
SF	.724	.128		.0160	-6172.8	1.43
$\widehat{\mathrm{E}\left(y_i^2\right)} = .562$	.563	.0867	.904	.0286	-6167.3	1.40
$\widehat{\mathrm{Var}\left(y_i^2\right)} = 1.40$						
Yen	.560	.430		.245	-6623.4	1.82
$\widehat{\mathrm{E}\left(y_i^2\right)} = .530$	.529	.310	.684	.147	-6621.2	1.81
$\widehat{\mathrm{Var}\left(y_i^2\right)} = 1.81$						
UK	.535	.104		.00613	-5145.7	.886
$\widehat{\mathrm{E}\left(y_i^2\right)} = .388$	.397	.0700	.820	.0122	-5143.5	.852
$\widehat{\mathrm{Var}\left(y_i^2\right)} = .849$						

Table 1: Quasi-likelihood fit for the Brownian SV and tim deformed Lévy models using a single factor variance process. Code: quasi_track.ox

observed in the data. Further, the Lévy based model brings the implied variance of the squared returns in-line with the empirical variance. We will call this kurtosis tracking. Another interesting feature is that the estimated value of λ is moderated by the presence of the Lévy process, with initial high values being made to fall and low values increasing. Hence the estimated Lévy based model is more consistent across the series than the Brownian special case.

Figure 2 focuses on the Canadian Dollar case. Figure 3 will deal with the other rates later. Figure 2 shows the correlogram of the squared data together with the implied fitted autocorrelation function for the Brownian based process and that of the Lévy OU time deformation process (we will return to the line denoted Lévy-OU$_2$ later). We can see that the autocorrelation function for the Brownian based model is higher than that for the Lévy-OU process. The Lévy based model has slightly less memory in it than the Brownian based case and we can just about discern this from the graph.

Remark 3 *If we approximate the setup in Lemma 1 by setting $\theta = 0$, then the linear represen-*

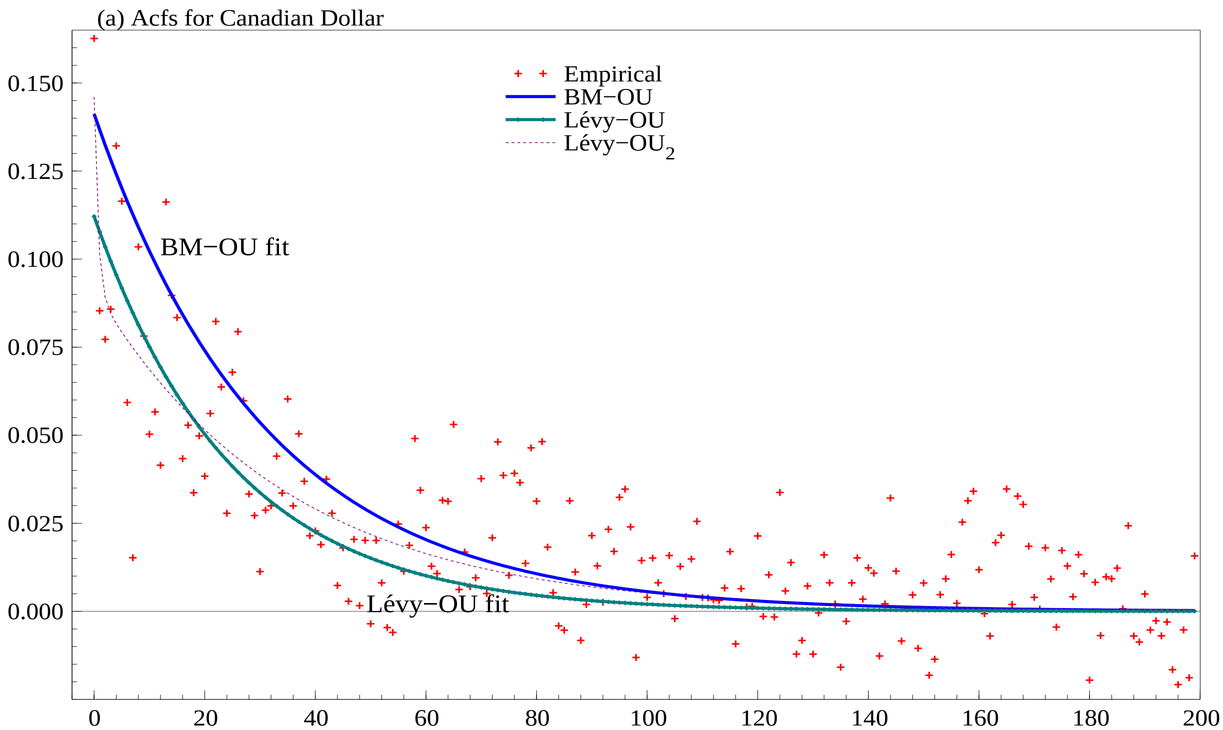

Figure 2: *Empirical correlogram for squared returns and the implied fit for the autocorrelation function for different SV models fitted by QML for the Canadian Dollar. Drawn are the implied fits from a BM-OU, Lévy-OU and Lévy-OU$_2$. On the graph the BM-OU symbol is used to indictate which curve is the one corresponding to the fitted BM-OU process. Likewise for the Lévy-OU process. Code:* `quasi.ox`.

tation is

$$y_i^2 = \xi\hbar + \varpi_i + u_i, \quad \varpi_{i+1} = (\tau_{i+1} - \xi\hbar) = \phi\varpi_i + e_i.$$

Writing $a_i = \mathrm{P}_L\left(\varpi_i | y_1^2, \ldots, y_{i-1}^2\right)$, *the Kalman filter implies that*

$$a_{i+1} = \phi a_i + k_i\left(y_i^2 - \xi\hbar - a_i\right), \quad k_i = \phi\frac{p_i}{p_i + \mathrm{Var}(u_i)} \in [0, \phi),$$

where $p_i = \mathrm{E}\left((\varpi_i - a_i)^2 | y_1^2, \ldots, y_{i-1}^2\right)$ *with*

$$p_{i+1} = \phi^2 p_i - \phi k_i p_i + \mathrm{Var}(v_i) \geq 0.$$

Rearranging the equation for a_{i+1} *and writing* $\widehat{\tau}_{i+1|i} = \xi\hbar + a_{i+1}$ *then*

$$\begin{aligned} \widehat{\tau}_{i+1|i} &= \xi\hbar + \phi\left(\tau_{i|i} - \xi\hbar\right) + k_i\left(y_i^2 - \tau_{i|i}\right) \\ &= (1-\phi)\xi\hbar + (\phi - k_i)\,\widehat{\tau}_{i|i-1} + k_i y_i^2 \\ &\geq 0 \end{aligned}$$

if $\widehat{\tau}_{i|i-1} \geq 0$. So long as $\text{Var}(v_i) > 0$ then p_i converges from above to a steady state, which implies that for moderate i the recursion for $\widehat{\tau}_{i+1|i}$ takes the form of a GARCH$(1,1)$ recursion. This was first discussed in Barndorff-Nielsen and Shephard (2001).

6.3 J factor model

When we move to a superposition model the results change somewhat. We assume that

$$r(s) = \sum_{j=1}^{J} w_j \exp(-\lambda_j s) \quad \text{where} \quad \sum_{j=1}^{J} w_j = 1 \quad \text{and} \quad w_j \geq 0. \tag{19}$$

Related work on building models for the spot variance out of Markov component models includes Shephard (1996), Engle and Lee (1999), Gallant, Hsu, and Tauchen (1999), Alizadeh, Brandt, and Diebold (2002), Barndorff-Nielsen and Shephard (2001), Barndorff-Nielsen and Shephard (2002a) and Chernov, Gallant, Ghysels, and Tauchen (2002).

Again the resulting process is analytically tractable with

$$r^{**}(s) = \sum_{j=1}^{J} w_j \lambda_j^{-2}(e^{-\lambda_j s} - 1 + \lambda_j s),$$

so the $\log L_Q$ can easily be calculated via the Durbin algorithm as terms such as

$$\begin{aligned} \text{Cov}(\tau^*(t)) &= 2\omega^2 r^{**}(t) \\ &= 2\omega^2 \sum_{j=1}^{J} w_j \lambda_j^{-2}(e^{-\lambda_j t} - 1 + \lambda_j t), \end{aligned}$$

are straightforward to calculate. Further, we can write τ_i as the sum of J uncorrelated ARMA$(1,1)$ processes. The parameters of each of these ARMA$(1,1)$ components can be calculated independently and analytically from the one factor results given in Lemma 1. An attractive feature of this setup is that we can again write this into a linear state space form with

$$\tau_i = \hbar\xi = (1\ 0\ 1\ 0 \ldots 1\ 0)\,\varpi_i,$$

and

$$\varpi_{i+1} = \begin{pmatrix} \phi_1 & 1 & 0 & 0 & & 0 & 0 \\ 0 & 0 & & & & 0 & 0 \\ 0 & 0 & \phi_2 & 1 & & 0 & 0 \\ 0 & 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \phi_J & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \varpi_i + \begin{pmatrix} 1 & 0 & & 0 \\ \theta_1 & 0 & & 0 \\ 0 & 1 & & 0 \\ 0 & \theta_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \theta_J \end{pmatrix} \begin{pmatrix} e_{1,i} \\ e_{2,i} \\ \\ e_{J,i} \end{pmatrix}.$$

Hence the Kalman filter can be used to rapidly compute $\log L_Q$.

In this subsection we apply this in the case of $J = 2$ to the exchange rate data discussed above. See Table 2. In the cases of the Canadian Dollar and Japanese Yen when κ_4 is estimated

	QML estimators					
	ξ	ω^2	κ_4	λ	w	$\log L_Q$
Can	.0838	.0124		$(1.47, .0286)$	$(.680, .319)$	741.1
DM	.625	.102		$(.0478, .000622)$	$(.638, .361)$	-5620.2
	.468	.0736	.861	$(.0489, .00405)$	$(.847, .152)$	-5618.4
FF	.593	.157		$(.0540, .00109)$	$(.766, .233)$	-5725.3
	.452	.130	.774	$(.0533, .00421)$	$(.898, .101)$	-5723.7
SF	.729	.125		$(.0301, .000176)$	$(.677, .322)$	-6168.9
	.563	.0867	.904	.0286		-6167.3
Yen	.502	.514		$(.694, .0143)$	$(.798, .201)$	-6603.7
UK	.528	.108		$(.0558, .00164)$	$(.347, .652)$	-5141.8
	.410	.0813	.674	$(.0634, .00361)$	$(.583, .416)$	-5141.1

Table 2: Quasi-likelihood fit for the Brownian and Lévy based SV models using a two factor variance process. Code: quasi_track.ox

it came out to be zero and so we do not repeat the row in the Table. The Table shows that the move to the superposition model removes some or all of the importance of the Lévy term. Instead, empirically, we have that sometimes it is enough to have an additional component of volatility which is quite fast decaying.

An interesting feature of the QML fit is the implied moments of the estimated model. The corresponding autocorrelation function for the squared process is given in Figure 3, which shows that even these rather simple models seem to match rather well the correlogram of the squared process.

6.4 Log-normal OU process

An important class of models is where we put $\log \tau$ as a $\mathrm{N}(\xi_{\log}, \omega^2_{\log})$-OU process. This appeared in the work of, for example, Hull and White (1987) while discrete time versions of this model was pioneered by Taylor (1982) in the SV literature. We call this the log-normal OU process, denoted LNOU. In that case

$$r(s) = \frac{\exp\left(\omega^2_{\log} e^{-\lambda|s|}\right) - 1}{1 - e^{-\omega^2_{\log}}},$$

while

$$\xi = e^{\xi_{\log} + \omega^2_{\log}/2} \quad \text{and} \quad \omega^2 = e^{2\xi_{\log} + \omega^2_{\log}} \left(e^{\omega^2_{\log}} - 1\right).$$

This model can be thought of as being a sup-OU process, an infinite dimensional superposition of OU process, for we can write

$$r(s) = \mathrm{E}_X \left(e^{-\lambda s X}\right) = \sum_{j=1}^{\infty} \Pr(X = j) e^{-\lambda s j},$$

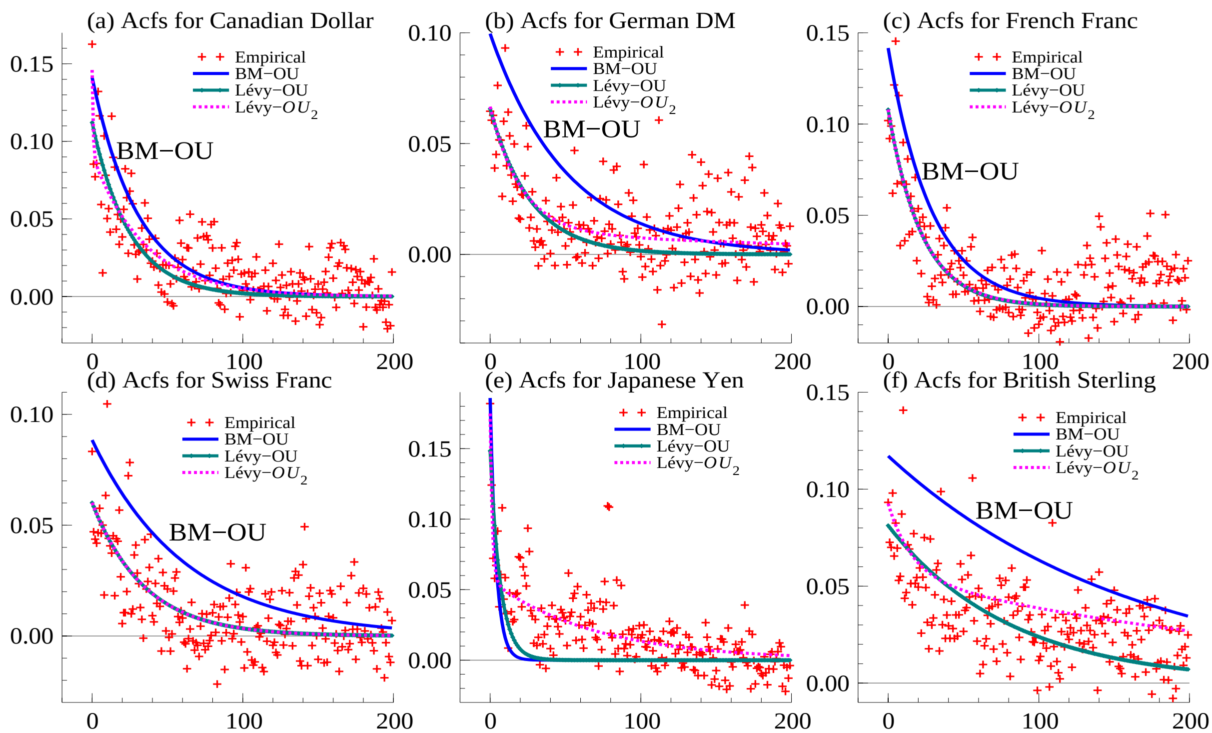

Figure 3: *Implied fit for the autocorrelation function for squared returns for different SV models fitted by QML using 6 different currencies. Drawn are the implied fit from an OU or OU_2 and the empirical correlogram. Code:* `quasi.ox`.

where

$$X \sim Po_{>0}(\omega^2_{\log}),$$

a truncated Poisson variable with the atom at 0 being knocked out and the probability function being renormalised so it sums to one over the strictly positive integers. This representation is helpful for we can use software for superposition models to compute the quasi-likelihood function. A similar, but rather differently motivated, type of approach is used by Meddahi (2001) in his work on eigenfunction process approximation to other processes. In practice we tend to place a large upper bound (e.g. 10) on the support of the truncated Poisson distribution. In this case the Durbin algorithm is highly competitive in terms of computational speed compared to the Kalman filter.

We have estimated this class of models in the BM-LNOU SV and Lévy-LNOU SV cases. The results are given in Table 3. We can see that the BM-LNOU SV processes has a slightly better fit in terms of the $\log L_Q$ than the BM-OU SV process, although the improvement is very modest indeed. The implied $\mathrm{E}(y_i^2)$ and $\mathrm{Var}(y_i^2)$ are very similar. We have not drawn the implied autocorrelation function of the squared returns of the two models as they are so close it is hard to tell them apart graphically.

Table 3 shows that when we move to the Lévy case the models become even more similar. It is not possible to give any general preference for one model over the other.

	QML estimators					Implied parameters		
	$\xi_{\log}$	$\omega^2_{\log}$	κ_4	λ	$\log L_Q$	$\xi = \mathrm{E}(y_i^2)$	ω^2	$\mathrm{Var}(y_i^2)$
Can	-2.40	.416		.0279	730.8	.110	.00634	.0433
	-2.66	.493	.133	.0348	734.4	.0887	.00502	.0424
DM	-.608	.253		.0158	-5625.9	.617	.109	1.09
	-.912	.290	.890	.0308	-5619.0	.464	.0727	1.06
FF	-.758	.416		.0283	-5728.9	.576	.172	1.17
	-1.05	.498	.796	.0375	-5724.3	.448	.130	1.14
SF	-.430	.218		.0131	-6172.5	.725	.128	1.43
	-.695	.243	.898	.0243	-6167.4	.563	.0877	1.40
Yen	-1.01	.866		.205	-6622.1	.560	.432	1.82
	-1.01	.756	.658	.126	-6620.2	.529	.316	1.81
UK	-.772	.303		.00521	-5145.4	.537	.102	.884
	-1.10	.366	.816	.0104	-5143.2	.398	.0701	.851

Table 3: Quasi-likelihood estimation of various LNOU SV processes. Fitted are both the Brownian motion and Lévy based models. Code: quasi_track.ox

6.5 A long memory model

Suppose we parameterise r

$$r(s) = \int_0^\infty e^{-s\lambda}\pi(\mathrm{d}\lambda),$$

where π is a probability measure on λ. This autocorrelation function arises as the limit of a superposition process, with the number of components potentially going off to infinity. A rigourous theory for this is provided by Barndorff-Nielsen (2001), whose work is related to earlier papers on building long-memory models by the addition of short memory processes by Mandelbrot (1971) and subsequently by Granger (1980) and Granger and Joyeau (1980). We do not know of any papers which estimate genuine long-memory continuous time time-deformation models without employing some form of discretisation. Papers which employ discretisation include Comte and Renault (1998), Gallant, Hsu, and Tauchen (1999) and Meddahi (2001). See also the earlier work on long memory SV models by Harvey (1998).

In the present setting, letting $\varepsilon(t;\lambda) = \lambda^{-1}(1-e^{-\lambda t})$ and $\varepsilon^*(t;\lambda) = \lambda^{-2}\left(e^{-\lambda t} - 1 + \lambda t\right)$, we have that

$$r^*(t) = \int_0^\infty \varepsilon(t;\lambda)\pi(\mathrm{d}\lambda) \quad \text{and} \quad r^{**}(t) = \int_0^\infty \varepsilon^*(t;\lambda)\pi(\mathrm{d}\lambda).$$

Example 5 *In the special case where π corresponds to a $\Gamma(2H,\alpha)$. Recall a $\Gamma(\nu,\alpha)$ density is*

$$f(x) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1}\exp(-\alpha x), \quad x > 0,$$

so then

$$r(u) = \left(1 + \frac{u}{\alpha}\right)^{-2H}.$$

Thus we find that for $H < 1/2$ (which deliver long-memory models)

$$r^*(t) = (1-2H)^{-1}\alpha\left\{\left(1+\frac{t}{\alpha}\right)^{1-2H} - 1\right\}$$

and

$$r^{**}(t) = (1-2H)^{-1}\alpha^2\left[(2-2H)^{-1}\left\{\left(1+\frac{t}{\alpha}\right)^{2-2H} - 1\right\} - \frac{t}{\alpha}\right].$$

An attractive feature of this model is that it can be added to a short memory model, like an OU process, to deliver an autocorrelation function which has flexibility over both the shorter and longer runs.

Table 4 shows the results from fitting Brownian and Lévy time-changed processes. These are quite encouraging. If we focus first on the Brownian motion case, then the four parameter model seems to have a better fit than simple three parameter OU based models (whose fit is given in Table 1), although the fit is quite similar and slightly worse than a five parameter superposition model with two components (given in Table 2). An interesting feature is that the estimated κ_4 is quite similar between the long-memory model and the OU_2 process.

	QML estimators						
	ξ	ω^2	κ_4	α	H	$\log L_Q$	Implied $\text{Var} y_i^2$
Can	.0765	.0177	0	.0456	.143	736.3	.0435
DM	.625	.101		8.03	.223	-5621.7	1.08
	.468	.0790	.834	16.5	.471	-5619.6	1.06
FF	.587	.162		15.3	.421	-5726.9	1.16
	.454	.145	.697	14.4	.457	-5725.5	1.15
SF	.727	.125		11.6	.251	-6170.4	1.42
	.566	.0996	.833	19.2	.453	-6168.7	1.40
Yen	.469	.644	0	.308	.274	-6606.8	1.81
UK	.526	.110		9.00	.124	-5141.9	.882
	.405	.0815	.699	17.1	.249	-5141.5	.855

Table 4: Estimates of a pure gamma long memory variance model using squared daily returns from a variety of exchange rates. Code: quasi_track.ox

The estimates of the long-memory parameters H increase as we move from the Brownian model to the Lévy model. Half the series have estimates of H which are larger than 0.4 in the Lévy cases, which suggests the long-memory models provides a parsimonious and reasonable description of the dynamics.

We can mix the long memory variance model with a short memory OU. We do this by writing that

$$r(u) = w\left(1+\frac{u}{\alpha}\right)^{-2H} + (1-w)\,e^{-\lambda u}, \quad w \in [0,1]\,, \quad \lambda > 0.$$

Clearly, with this structure

$$\begin{aligned} r^{**}(u) &= w(1-2H)^{-1}\alpha^2\left[(2-2H)^{-1}\left\{(1+\frac{u}{\alpha})^{2-2H}-1\right\}-\frac{u}{\alpha}\right] \\ &\quad +(1-w)\,\lambda^{-2}\left(e^{-\lambda u}-1+\lambda u\right), \end{aligned}$$

which is straightforward to calculate and so we can use Durbin's method to compute the quasi-likelihood. The resulting estimates are given in Table 5 in the Brownian and Lévy process cases.

In this Table in the Brownian motion case the estimated values of H and α do not vary very much across the series, with the long-memory component having only a minor part of the variation of the series. The only exception to this is the UK Sterling, which again has quite a small value of H. Throughout λ is estimated to be quite large, which means this component is modelling very short term fluctuations in the spot variance. Hence its main use is in modelling the fat tails of the series.

	QML estimators								
	ξ	ω^2	κ_4	α	λ	H	w	$\log L_Q$	Implied $\mathrm{Var} y_i^2$
Can	.0776	.0138	0	19.0	1.69	.453	.340	740.27	.0426
DM	.469	3.04		16.5	44.5	.463	.0258	-5619.6	1.06
	.468	.0836	.831	16.1	17.8	.470	.946	-5619.6	1.06
FF	.456	1.88		14.4	32.2	.461	.0766	-5725.6	1.15
	.454	.151	.696	14.7	32.1	.463	.955	-5725.5	1.15
SF	.563	.472	0	26.1	3.17	.462	.194	-6167.5	1.41
Yen	.481	.541	0	26.8	.800	.458	.245	-6604.2	1.81
UK	.405	.486	0	17.2	7.34	.235	.167	-5141.3	.856

Table 5: Estimation of a mixed OU and gamma long memory variance model for squared daily returns using a variety of exchange rates. Code: quasi_track.ox

When we add the flexibility of a Lévy process the quasi-likelihood function hardly moves. In cases where κ_4 is estimated to be non-zero, the impact of the short-memory component is small with w being estimated to be over 0.9. Hence, for these series the long-memory component dominates. Overall, for daily data there seems to be little point in both having a Lévy process and a second, short memory, component of volatility.

7 Illustration based on realised variances

7.1 The data and realised variances

In this Section we will use the theory of RV in the context of time-changed Lévy processes to estimate various volatility models. We start out with a discussion of high frequency data, together with its relevant stylised facts. Then we estimate OU based models and their superposition extensions. We then compare their fit to those built out of log-OU processes and long-memory models.

To illustrate some of the empirical features of RV we have used a similar return dataset employed by Andersen, Bollerslev, Diebold, and Labys (2001), although we have made slightly different adjustments to deal with some missing data. These are described in detail in Barndorff-Nielsen and Shephard (2002a). This series records the United States Dollar/ German Deutsche Mark series. It covers the ten year period from 1st December 1986 until 30th November 1996. The original dataset records every 5 minutes the most recent mid-quote to appear on the Reuters screen. It has been kindly supplied to us by Olsen and Associates in Zurich, who document their path breaking work in this area in Dacorogna, Gencay, Müller, Olsen, and Pictet (2001).

Figure 4(a) shows the implied daily returns over this period and Figure 4(b) the corresponding correlogram for squared daily returns. The correlations tend to be quite small, although with a preponderance of positive numbers. Figure 4(c) shows the RV for each day based on the full 5 minute dataset. Hence for this series $M = 288$, while $\hbar$ represents one day. The time series shows changing level of the variability of the series. Figure 4(d) shows the correlogram of the RVs. The correlogram starts at around 0.55 and quickly falls to around zero at lag 100, although at longer lags the correlations tend to be positive.

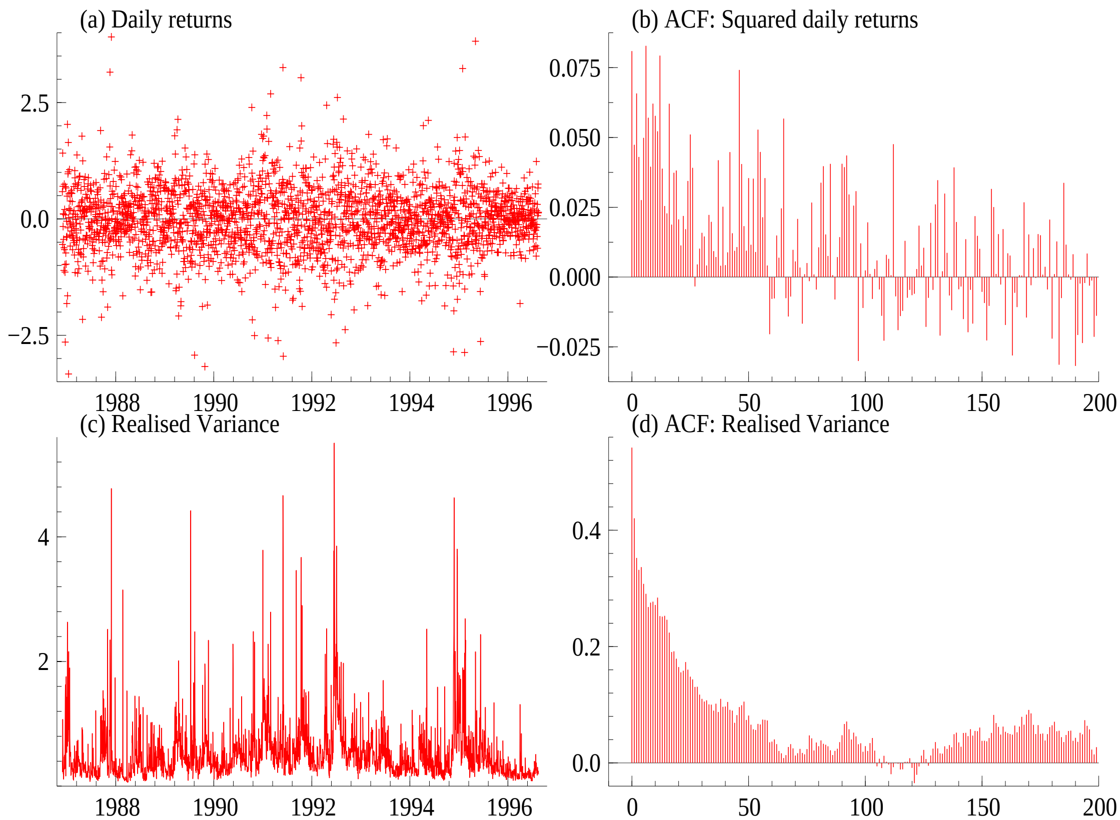

Figure 4: *DM against the Dollar, based on the Olsen dataset. (a) Implied daily returns. (b) Correlogram for the squared daily returns. (c) Daily realised variance based on the 5 minute returns. (d) Correlogram for the realised variance for the DM series. Code:* `realised_quasi_track.ox`.

7.2 Superposition model

Table 6 shows the fitted results for the OU and superposition of OU processes for squared daily returns. This corresponds to $M = 1$. This suggests the move from the Brownian time-deformation model to the Lévy version improves the fit of the model, but not dramatically in terms of quasi-likelihood fit. The Lévy motion has better variance tracking properties than the corresponding estimated Brownian models. In both the Brownian and Lévy models we have that the components of the superpositions are well separated with one component representing quite persistent shocks to the volatility, while the other component is more rapidly reverting. In both of these cases the components have roughly equal weight. A move to a fourth component does not improve the quasi-likelihood.

QML estimators						Empirical	Implied	Empirical
ξ	ω^2	κ_4	λ	w	$\log L_Q$	$\mathrm{E}\left([y_M^*]_i\right)$	$\mathrm{V}\left([y_M^*]_i\right)$	$\mathrm{V}\left([y_M^*]_i\right)$
.644	.104		.0170		-3582.3	.504	1.14	1.11
.504	.0731	.772	.0311		-3579.0	.504	1.11	1.11
.504	.0773	.753	(.0146,.0686)	(.479,.520)	-3578.7	.504	1.11	1.11

Table 6: Quasi-likelihood fit to the DM data using squared daily data. This is the same as using the realised variance using $M = 1$. Code: realised_quasi_track.ox

Table 7 gives the corresponding results using $M = 288$. Before we discuss the details of the estimated model we can see that although the average value of $[y_M^*]_i$ is very close to that of y_i^2 for the daily data, the variance of $[y_M^*]_i$ is much lower than the corresponding y_i^2. This reflects the fact that the RV is a much more accurate estimator of the integrated variance. It provides a much more informative basis for estimating the parameters of these models.

Table 7 shows that the fit of the model continually improves as we increase the complexity of the model. Now, the estimated parameters change quite a lot with the Brownian based model. However, at all stages the model is well tracked, with the implied and empirical expectations and variances of the RVs being comparable. Here we will discuss in some detail the DM example, as this is typical of the results for the other rates.

For the single OU DM case Figure 7(a) shows that the estimated model has little memory in it and provides a poor match to the empirical observations. The reason for this is that the three parameter model is not sufficiently flexible to simultaneously fit the average value of RV, its variance and its decay in the autocorrelation function. In effect, the quasi-likelihood chooses to highly weight the mean, variance and very short lags in the autocorrelation function. This then almost entirely neglects higher lags in the autocorrelation function.

The second order Brownian based model is much more reasonable, although it tends to under weight the longer lags in the acf. When we fitted the third order superposition model one of the component has a value for λ which is above 200. This is basically almost instantly mean reverting and so can be thought to proxy very short term deviations from the local Gaussian assumption. Many researchers would model this using a jump type process.

In the Lévy case the estimated model based on the OU process is quite poor, but is much better than the Brownian version. Figure 7(b) shows the memory of the acf last to around 50 lags, which is still of course too short. When we move to the second order model the implied acf is similar to that implied by the third order Brownian model. This seems a reasonable description of the empirical correlogram. The move to a third order model allows the implied

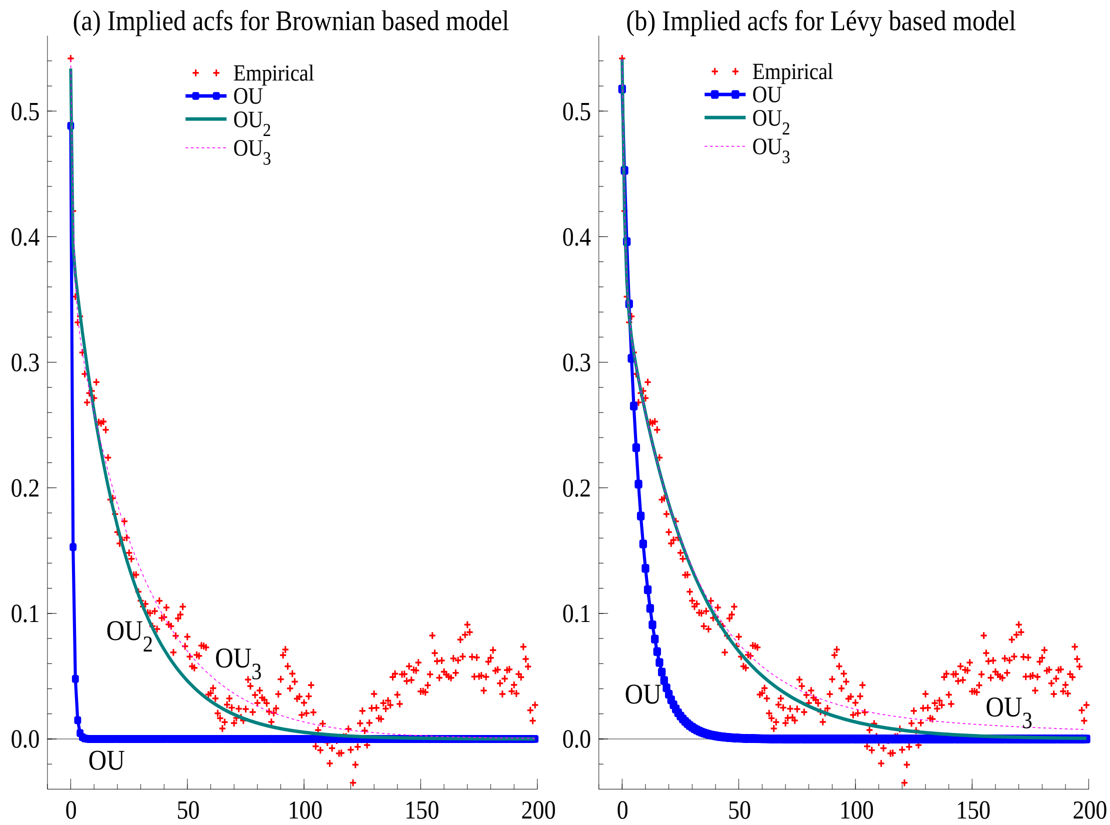

Figure 5: *Empirical and implied autocorrelation functions for the realised variances for the DM against the Dollar using $M = 288$. (a) Brownian motion based model. (b) Lévy based model. Code:* `realised_quasi_track.ox`*.*

acf to have more memory at long lags. This additional component has a value of λ which is 0.007 and a weight of 0.03. Although this may make some difference to very long forecasts, which can been seen from the fitted autocorrelation function drawn in Figure 7(b), it does not really impact the fit of the model in terms of the quasi-likelihood.

Although the Brownian and Lévy based models are rather similar for large order superposition models, for short order models the Lévy based model is preferable. It is more stable as we change the model, adding new OU components.

Very similar results hold when we fit the model using $M = 144$. These results are given in Table 7.

Again as we take M down so that $M = 72$ the parameters do not move very much. Importantly, for the simplest Brownian model the estimated value of λ is quite large, which means the fitted model has very little memory in it. This changes quite a lot when we allow for Lévy effects for then the estimated value of λ falls dramatically down to around 0.1. When the superposition model is added the estimated values of λ are around 0.03 and 1. This allows a component of

M	QML estimators						Empir	Implied
$(\widehat{V[y_M^*]_i})$	ξ	ω^2	κ_4	λ	w	$\log L_Q$	$\mathrm{E}[y_M^*]_i$	$\mathrm{V}[y_M^*]_i$
288	.530	.328		1.16		-1356.2	.528	.235
(.237)	.527	.140	.190	.133		-1222.4	.528	.237
	.526	.385		(2.85,.0432)	(.741,.258)	-1195.5	.528	.236
	.525	.193	.130	(.975,.0328)	(.543,.456)	-1190.0	.528	.237
	.525	4.68		(203,.973,.0328)	(.958,.0223,.0188)	-1190.0	.528	.237
	.523	.193	.131	(.983,.0361,.00728)	(.541,.420,.0380)	-1189.9	.528	.237
144	.511	.373		1.43		-1552.2	.509	.253
(.255)	.507	.121	.260	.0851		-1398.9	.509	.255
	.507	.482		(3.88,.0430)	(.796,.203)	-1382.7	.509	.254
	.506	.174	.193	(.970,.0330)	(.495,.504)	-1378.4	.509	.255
	.506	4.48		(.966,155,.0330)	(.0191,.961,.0195)	-1378.4	.509	.255
	.504	.174	.193	(.982,.0367,.00880)	(.493,.458,.0475)	-1378.3	.509	.255
72	.491	.393		1.42		-1671.5	.488	.275
(.277)	.487	.125	.296	.0996		-1547.2	.488	.276
	.487	.486		(3.48,.0437)	(.802,.197)	-1528.8	.488	.276
	.486	.188	.213	(.990,.0342)	(.545,.454)	-1524.6	.488	.276
	.486	4.67		(.988,199,.0342,)	(.0219,.959,.0183)	-1524.6	.488	.276
	.486	.188	.213	(.992,.0114,.0349)	(.544,.0101,.445)	-1524.6	.488	.276
2	.738	.158		.0104		-3123.1	.491	.860
(.774)	.489	.0735	.794	.0410		-3114.2	.491	.774
	.723	.178		(.00003,.0420)	(.589,.410)	-3116.5	.491	.878
	.488	.0750	.792	(.0560,.0143)	(.758,.241)	-3114.1	.491	.774
1	.644	.104		.0170		-3582.3	.504	1.14
(1.11)	.504	.0731	.772	.0311		-3579.0	.504	1.11
	.639	.109		(.0633,.00390)	(.506,.494)	-3580.8	.504	1.14
	.504	.0773	.753	(.0686,.0146)	(.520,.479)	-3578.7	.504	1.11

Table 7: Quasi-likelihood fit to the DM data using the realised variance, for a variety of values of M. Code: realised_quasi_track.ox

the variance process which has a great deal of memory.

Finally, Table 7 gives the corresponding result for $M = 2$. This is in line with the analysis based on squared daily data. This shows that we need the Lévy effects in order to produce variance tracking.

7.3 Log-OU models

The corresponding results for the estimated log-OU based time-deformation models are given in Table 8. This Table shows that the model is quite unstable as M changes in the Brownian case. The Lévy version of the model does not vary much as M alters. It also has a much higher quasi-likelihood, although for small values of M there is not much difference between the Brownian and Lévy based models.

The Table indicates that for large M the Brownian LNOU model fits much better in terms of the quasi-likelihood criteria than the OU based models, although the Lévy version of the model is only slightly superior. The fact that introducing Lévy effects reduces the difference between the two sets of models is interesting.

M	QML estimators $\xi_{\log}$	$\omega^2_{\log}$	κ_4	λ	$\log L_Q$
288	.553	.882		.000456	-1271.1
	-.844	.409	.187	.115	-1219.4
144	.405	.784		.000565	-1433.3
	-.870	.385	.257	.0748	-1397.2
72	.278	.631		.00117	-1582.6
	-.932	.428	.292	.0884	-1545.3
24	.0253	.510		.00223	-1940.9
	-.981	.412	.444	.0599	-1914.3
6	-1.16	.911		1.14	-2370.0
	-.901	.326	.490	.0383	-2290.2
2	-1.14	.902		1.23	-3148.7
	-.848	.269	.791	.0346	-3114.2
1	-.552	.224		.0145	-3582.1
	-.810	.254	.768	.0263	-3578.9

Table 8: Quasi-likelihood estimation of LNOU SV process using realised variances based on different values of M, the number of intra-day observations. Code: realised_quasi_track.ox

7.4 Long memory models

Table 9 gives the RV estimates of the gamma-long memory process using the Brownian and Lévy based models. The Table includes the results when we add in short memory OU components to the variance process.

The results are much less stable with respect to M than we saw in Table 7 on the Lévy based model whose spot variance was constructed by a superposition of OU processes. The Lévy based models still tend to have higher long-memory parameter H than the Brownian versions, while the addition of the OU components tend to mean that H drifts higher. Throughout the long-memory component has a high weight, the only exception being the Lévy based model using daily data where the long memory component is close to being irrelevant.

	QML estimators							
	ξ	ω^2	κ_4	α	λ	H	w	$\log L_Q$
288	.512	2.75		.000122		.171		-1197.2
	.515	2.75		.000456	.0278	.235	(.980,.0199)	-1192.3
	.524	.667		.0438	(1.15,.0296)	.402	(.883,.00460,.111)	-1191.5
	.495	.525	.0801	.00564		.113		-1195.3
	.523	.605	.0110	.0483	.0291	.386	(.891,.108)	-1191.5
144	.505	4.13		.000112		.200		-1389.7
	.502	4.14		.000589	.0294	.298	(.983,.0165)	-1379.6
	.476	.287	.183	.0706		.0919		-1383.3
	.505	.258	.149	.262	.0290	.470	(.708,.291)	-1379.3
72	.489	3.92		.000119		.194		-1532.9
	.479	3.93		.000476	.0297	.272	(.984,.0153)	-1526.4
	.461	.513	.157	.00515		.109		-1529.8
	.485	.621	.0652	.0475	(.0306,.0293)	.400	(.884,.115,small)	-1525.9
2	.679	.225		.174		.0563		-3117.7
	.722	.176		.180	(.0493,.0410)	.0146	(.661,.123,.214)	-3116.3
	.480	.0834	.797	14.4		.470		-3114.7
1	.640	.109		9.44		.227		-3580.5
	.490	.145	.548	.581		.128		-3580.7
	.505	.0739	.768	2.09	.0308	.357	(.0244,.975)	-3579.0

Table 9: Estimates of a pure gamma long memory variance model using squared daily returns from a variety of exchange rates. Also given are the estimates with added OU components. Code: quasi_track.ox

8 Filtered estimators of the time-change

8.1 Theory and application

Durbin's method also automatically delivers the sequence of best linear unbiased estimator of $\kappa_2\tau_i$ given $y_1^2, \ldots, y_{i-1}^2$, written as

$$\kappa_2\widehat{\tau}_{i|i-1} = \mathrm{P}_L\left(\kappa_2\tau_i|y_1^2, \ldots, y_{i-1}^2\right)$$

and the associated mean square error f_i for $i = 1, 2, \ldots, n$ in $O(n^2)$. Likewise, for the more specialised case of the Kalman filter, all these quantities are also produced in $O(n)$. Figure 6(a) shows, for the German DM series (selected as it had quite a difference between the BM and Lévy based models), a plot of $|y_i|$ and $\sqrt{\kappa_2\widehat{\tau}_{i|i-1}}$ for the BM based time-deformation process with Markovian variance process, as well as the associated Lévy based version. These are plotted for only a short section of the time series at the start of the sample period. The one-step ahead predictions are seen to respond to large absolute returns. The BM and Lévy based models give roughly similar results, with the BM based model slightly higher throughout the sample. Figure 6(b) shows $\sqrt{\kappa_2\widehat{\tau}_{i|i-1}}$ for the Lévy based model, depicting the result for the entire sample. The one-step ahead volatility forecasts vary considerably throughout the sample ranging in value from around 0.55 to 1.4.

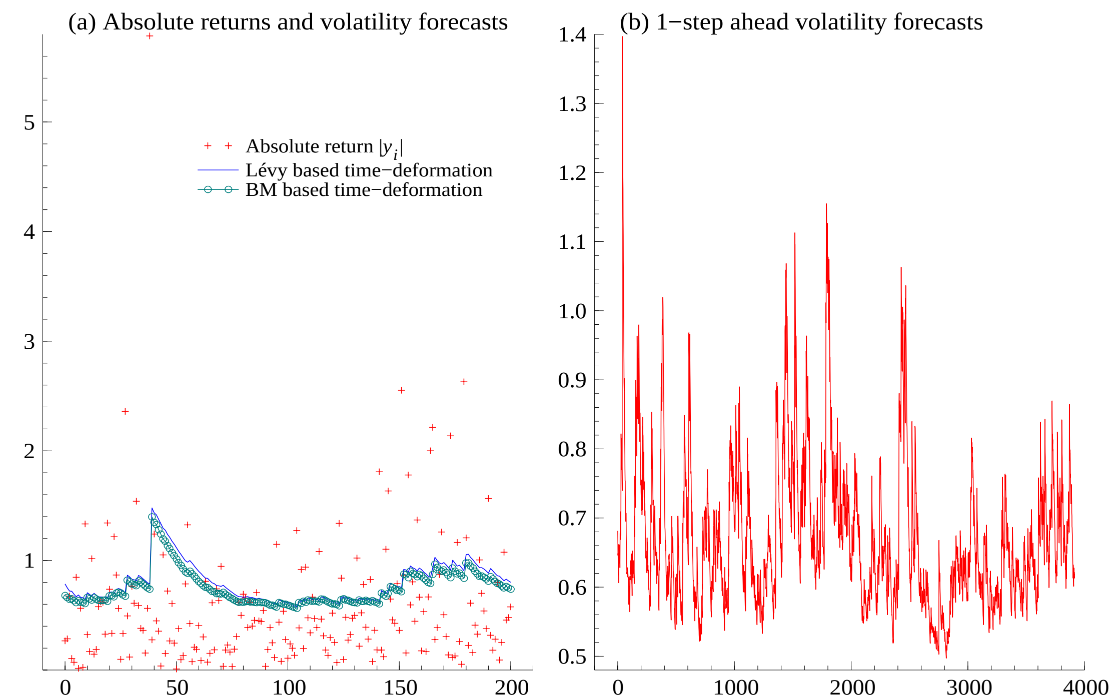

Figure 6: *One-step ahead forecast of the volatility using the Kalman filter. Based on the BM time-deformed model and the Lévy based time-deformed model. Code:* `quasi.ox`.

When we employ superposition based models for the spot variance then we can write

$$\tau(t) = \sum_{j=1}^{J} \tau^j(t),$$

where the components $\tau^j(t)$ are uncorrelated. Then

$$\kappa_2 \widehat{\tau}_{i|i-1} = \sum_{j=1}^{J} \mathrm{P}_L\left(\kappa_2 \tau_i^j | y_1^2, \ldots, y_{i-1}^2\right),$$

where $\mathrm{P}_L\left(\kappa_2 \tau_i^j | y_1^2, \ldots, y_{i-1}^2\right)$ can be computed using the Kalman filter. This decomposition of the forecast is potentially interesting for the time series plots of $\mathrm{P}_L\left(\kappa_2 \tau_i^j | y_1^2, \ldots, y_{i-1}^2\right)$ would be expected to be very different over superscript j. Figure 7 shows these terms for the German DM series. Importantly we see the first component has much less memory than the second, while its size is typically much larger. Importantly the picture shows that the estimates of the components can go negative, even though $\widehat{\tau}_{i|i-1}$ is always strictly positive.

The responsiveness of the estimators of the variance process to news can be measured in a

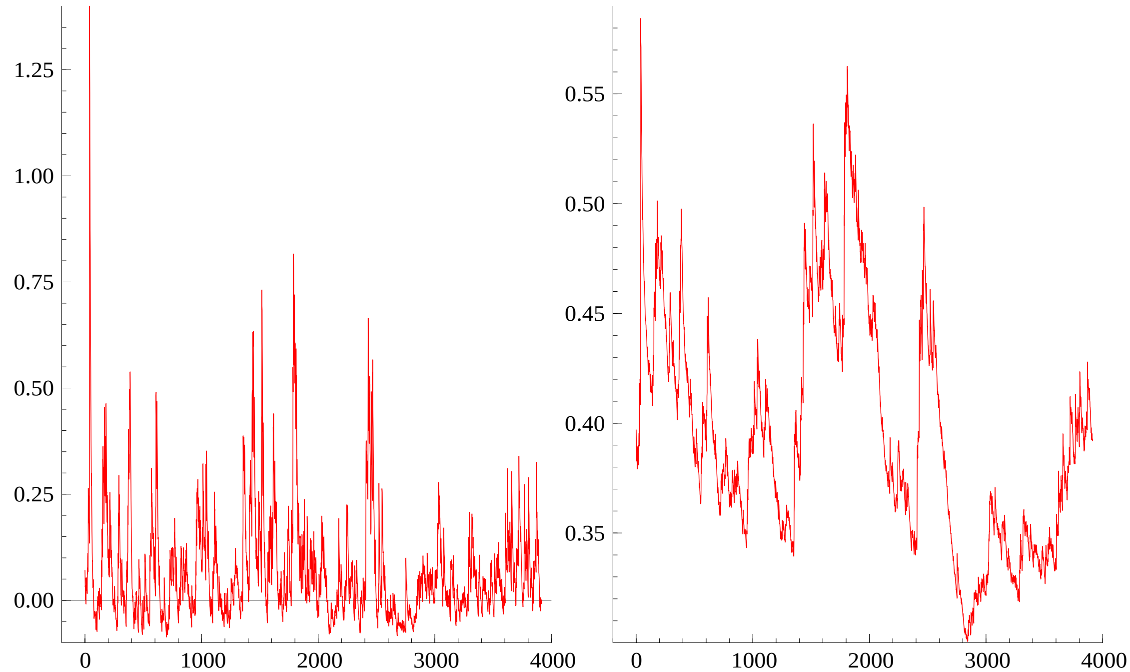

Figure 7: *One-step ahead forecast of the components of variance using the Kalman filter. Based on the BM time-deformed Lévy based time-deformed model. Code:* `quasi.ox`.

number of ways. One possibility is to plot

$$\mathrm{P}_L\left(\kappa_2\tau_{i+1}|y_1^2,\ldots,y_i^2\right)-\mathrm{P}_L\left(\kappa_2\tau_i|y_1^2,\ldots,y_{i-1}^2\right),$$

against y_i. We favour a plot of

$$news_i=\mathrm{P}_L\left(\kappa_2\tau_i|y_1^2,\ldots,y_i^2\right)-\mathrm{P}_L\left(\kappa_2\tau_i|y_1^2,\ldots,y_{i-1}^2\right)$$

against y_i. The difference between the two measures of news seem, in the experiments we have carried out, to be very minor indeed. The picture of $news_i$ against y_i shows how the estimator of the variance at time i is changed as we record y_i, implying a sensible measure of the impact of news on volatility estimation. Engle and Ng (1993) and Engle (2002) have developed rather different ideas in this regard.

Figure 8 shows a graph for $news_i$ against y_i. It demonstrates that the estimated variance changes in a roughly quadratic manner with y_i. The curve is not entirely regular, with variation around the overall trend. The reason for this is that the Kalman filter implies that

$$news_i=K_i\left\{y_i^2-\mathrm{P}_L\left(\kappa_2\tau_i|y_1^2,\ldots,y_{i-1}^2\right)\right\},$$

where K_i is the so-called Kalman gain (which quickly converges to a steady state for moderately

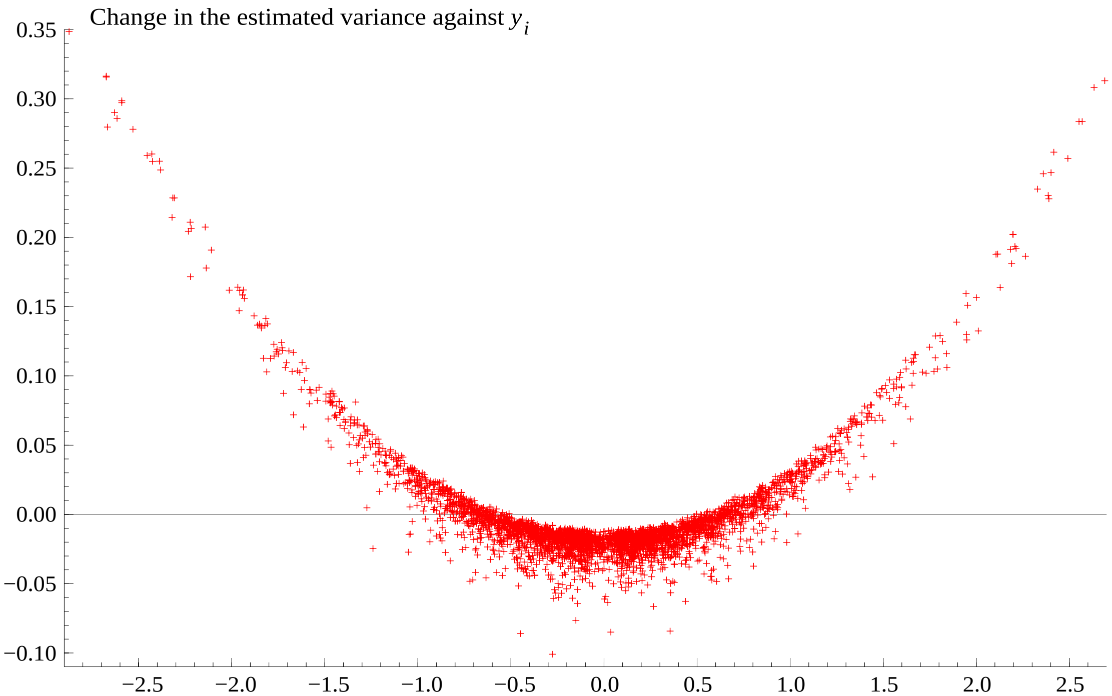

Figure 8: *Graph of* $news_i$ *against* y_i *for the BM time-deformation model for the German DM series. Code:* `quasi.ox`.

large i). Hence the movement recorded in $news_i$ is not affine in y_i^2, rather it is proportional to the linear innovation $y_i^2 - \mathrm{P}_L\left(\kappa_2 \tau_i | y_1^2, \ldots, y_{i-1}^2\right)$.

This discussion suggests we could be interested in plotting $news_i$ against $y_i^2 - \mathrm{P}_L\left(\kappa_2 \tau_i | y_1^2, \ldots, y_{i-1}^2\right)$, the innovation. Of course, this would just be linear, reflecting the steady state Kalman gain.

8.2 Limitations of linear filters

These best linear estimators have many advantages: they are simple and fast to compute, they are optimal in the linear sense and have associated measures of uncertainty, etc. However, they only provide a partial solution to the filtering problem. In principle a more complete filtering solution would yield the density of $\tau_i | y_1, \ldots, y_{i-1}$ or $\tau_i | y_1, \ldots, y_i$ or, indeed, $\tau_i | y_1, \ldots, y_n$.

However, the linear approach does not even provide any of the moments of these densities. Indeed, in general, there is nothing stopping the estimators from becoming negative. The lack of a full posterior density makes it impossible to construct valid confidence intervals of τ_i or use the estimator to imply meaningfully properties of derived quantities such as the vital actual volatility$\sqrt{\kappa_2 \tau_i}$. In order to do this we can use particle filters, but this is beyond the scope of this paper. We refer the reader to the literature on this topic, see for example, Gordon, Salmond, and Smith (1993), Pitt and Shephard (1999a) and Doucet, de Freitas, and Gordon

(2001). Work on using particle filters for time-changed processes include Kim, Shephard, and Chib (1998), Pitt and Shephard (1999a), Pitt and Shephard (1999b), Pitt and Shephard (2001), Barndorff-Nielsen and Shephard (2001) and Johannes, Polson, and Stroud (2002).

9 Conclusion

In this paper we have generalised the usual SV model to a time-changed Lévy process. The effect of allowing for the possibility of jumps is that the probability limit of RV, the increments to quadratic variation, is no longer the increments to the time-change. In fact, outside the Brownian motion, even if we observe the exact path of the price process we cannot recover the time-change.

Even though the RV is an inconsistent estimator of the time-change, it is an almost unbiased one. We characterise the variability of the difference between RV and the time-change. This allows us to use the time series of realised variances to estimate the parameters of models of time-change. Further, the time series can be used to produce forecasts of future time-changes of the Lévy processes.

10 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`), which is 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 thank Nour Meddahi for encouragement on this topic and Frank Gerhard and Matthias Winkel for their comments on an early draft.

References

Alizadeh, S., M. Brandt, and F. Diebold (2002). Range-based estimation of stochastic volatility models. *Journal of Finance 57*, 1047–1091.

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

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2003). Parametric and nonparametric measurement of volatility. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland. Forthcoming.

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.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2002a). Analytic evaluation of volatility forecasts. Unpublished paper: Department of Economics, Duke University.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2002b). Correcting the errors: A note on volatility forecast evaluation based on high-frequency data and realized volatilities. Unpublished paper: Department of Economics, Duke University.

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

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

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

Barndorff-Nielsen, O. E. and N. Shephard (2002b). Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. Unpublished paper: Nuffield College, Oxford, Economics working paper 2002-W13.

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

Barndorff-Nielsen, O. E. and N. Shephard (2003). *Financial Volatility: Stochastic Volatility and Lévy based models.* Cambridge: Cambridge University Press. Forthcoming.

Bertoin, J. (1996). *Lévy Processes.* Cambridge: Cambridge University Press.

Bochner, S. (1949). Diffusion equation and stochastic processes. *Proc. National Academy of Science, USA 85*, 369–370.

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

Carr, P., H. Geman, D. B. Madan, and M. Yor (2003). Stochastic volatility for Lévy processes. *Mathematical Finance*. To appear.

Carr, P. and L. Wu (2003). Time-changed Lévy processes and option pricing. *Journal of Financial Economics*. Forthcoming.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2002). Alternative models of stock price dynamics. *Journal of Econometrics*. Forthcoming.

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.

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

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

Doornik, J. A. and M. Ooms (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Computational Statistics and Data Analysis 42*, 333–348.

Doucet, A., N. de Freitas, and N. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag.

Duffie, D. (1996). *Dynamic Asset Pricing Theory* (2 ed.). New Jersey: Princeton University Press.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Durham, G. and A. R. Gallant (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes (with discussion). *Journal of Business and Economic Statistics 20*, 297–338.

Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics 17*, 425–446.

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.

Engle, R. F. and V. Ng (1993). Measuring and testing the impact of news on volatility. *Journal of Finance 48*, 1749–1778.

Gallant, A. R., C. Hsu, and G. Tauchen (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance.

Gallant, A. R. and G. Tauchen (1996). Which moments to match. *Econometric Theory 12*, 657–81.

Geman, H., D. B. Madan, and M. Yor (2002). Stochastic volatility, jumps and hidden time changes. *Finance and Stochastics 6*, 63–90.

Geman, H., D. B. Madan, and M. Yor (2003). Time changes for Levy processes. *Mathematical Finance*. Forthcoming.

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

Ghysels, E. and J. Jasiak (1994). Stochastic volatility and time deformation: an application of trading volume and leverage effects. Unpublished paper: Department of Economics, Université de Montreal.

Golub, G. H. and C. F. Van Loan (1989). *Matrix Computations*. Baltimore: John Hopkins University Press.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993). A novel approach to nonlinear and non-Gaussian Bayesian state estimation. *IEE-Proceedings F 140*, 107–113.

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.

Granger, C. W. J. and R. Joyeau (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis 1*, 15–39.

Harvey, A. C. (1993). *Time Series Models* (2 ed.). Hemel Hempstead: Harvester Wheatsheaf.

Harvey, A. C. (1998). Long memory in stochastic volatility. In J. Knight and S. Satchell (Eds.), *Forecasting Volatility in Financial Markets*, pp. 307–320. Oxford: Butterworth-Heineman.

Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies 61*, 247–264.

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

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

Jacod, J. and A. N. Shiryaev (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag: Berlin.

Johannes, M., N. Polson, and J. Stroud (2002). Nonlinear filtering of stochastic differential equations with jumps. Unpublished paper: Graduate School of Business, Columbia University.

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.

Madan, D. B. and E. Seneta (1990). The VG model for share market returns. *Journal of Business 63*, 511–524.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business 36*, 394–419.

Mandelbrot, B. (1971). A fast fractional Gaussian noise generator. *Water Resource Research 7*, 543–553.

Mandelbrot, B. and H. Taylor (1967). On the distribution of stock price differences. *Operations Research 15*, 1057–1062.

Meddahi, N. (2001). An eigenfunction approach for volatility modeling. Unpublished paper: University of Montreal.

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

Meddahi, N. (2002b). ARMA representation of integrated and realized variances. Unpublished paper: University of Montreal.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3*, 125–144.

Newey, W. K. and K. D. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Pitt, M. K. and N. Shephard (1999a). Filtering via simulation: auxiliary particle filter. *Journal of the American Statistical Association 94*, 590–599.

Pitt, M. K. and N. Shephard (1999b). Time varying covariances: a factor stochastic volatility approach (with discussion). In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Bayesian Statistics 6*, pp. 547–570. Oxford: Oxford University Press.

Pitt, M. K. and N. Shephard (2001). Auxiliary variable based particle filters. In N. de Freitas, A. Doucet, and N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, pp. 273–293. New York: Springer-Verlag.

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

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

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.

Stock, J. H. (1988). Estimating continuous-time processes subject to time deformation: an application to postwar U.S. GNP. *Journal of the American Statistical Association 83*, 77–85.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes — a study of daily sugar prices 1961-79. In O. D. Anderson (Ed.), *Time Series Analysis: Theory and Practice, 1*, pp. 203–226. Amsterdam: North-Holland.

Winkel, M. (2002). The recovery problem for time-changed Lévy processes. MaPhySto Research Report 2001-37, Department of Mathematical Sciences, Aarhus University.