

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

OLE E. BARNDORFF-NIELSEN

*Department of Mathematical Sciences, University of Aarhus,
Ny Munkegade, DK-8000 Aarhus C, Denmark*
oebn@imf.au.dk

NEIL SHEPHARD

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

First circulated April 2003. This draft April 2004

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 realised variance is an inconsistent estimator of the time-change, however 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, as 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. We will see that in terms of contributions to the mean square error, the square of the bias of the realised variances is dominated by its variance. The moments of realised variances 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 (2004). To understand this class of processes

we start with the standard setup in asset pricing models. We let log-prices Y_t , for time $t \geq 0$, follow a semimartingale process

$$Y = A + M.$$

Here A is a finite variation process, which informally means that the sum of the absolute values of the increments of this process measured over very small time intervals is finite. Further, M is a local martingale — which is a convenient generalisation of a martingale in financial economics. Without loss of generality, we assume that $A_0 = 0$ and $M_0 = 0$, which implies that $Y_0 = 0$. 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 follow Carr, Geman, Madan, and Yor (2003), Geman, Madan, and Yor (2003) and Carr and Wu (2004) 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 Cont and Tankov (2004). Here τ^* is a time-change (that is a process with non-decreasing paths) such that, for all t , $\tau_t^* < \infty$. Finally, we will assume that $z \perp\!\!\!\perp \tau^*$ which means that the two processes are independent. This rules out leverage type effects, which is clearly regrettable. This structure means that $A_t = \mu t + \tau_t^* \mathbf{E}(Z_1)$ while $M_t = Z_{\tau_t^*} - \tau_t^* \mathbf{E}(Z_1)$.

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. Further, and 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

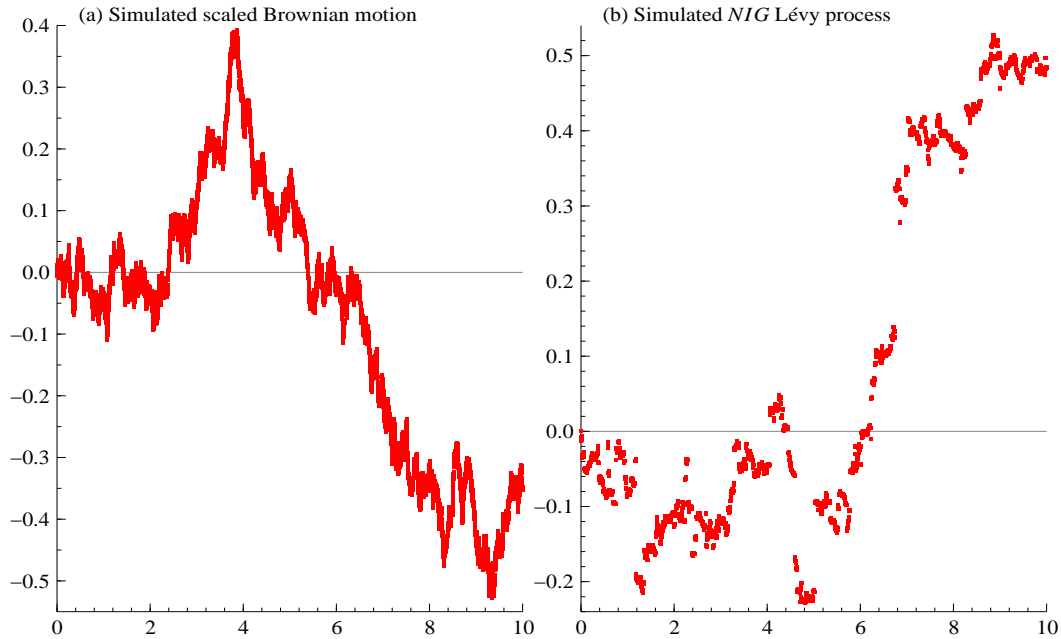


Figure 1: (a) Sample path of $\sqrt{0.02}$ times standard Brownian motion. (b) Sample path of a $NIG(0.2,0,0,10)$ Lévy process. Thus the increments of both processes have the same variance. Code: `levy_graphs.ox`.

stable processes in the econometric literature. Such stable 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 typically 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

$$\tau_t^* = \int_0^t \tau_u du,$$

where τ is a non-negative process. This means that τ^* has a continuous, but not necessarily differentiable, sample paths. Under this assumption M has a continuous sample path with probability one if and only if Z is Brownian motion. In the Brownian motion case this process is equivalent to a stochastic volatility process.

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 functions 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

$$y_i = Y_{i\hbar} - Y_{(i-1)\hbar}, \quad i = 1, 2, \dots,$$

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 mostly focus on the case where we additionally have M intra- \hbar high frequency observations during each \hbar time period. The time-gap between these observations will be $\delta = \hbar/M$. The j -th intra- \hbar return for the i -th period will be calculated as

$$y_{j,i} = Y_{(i-1)\hbar+j\delta} - Y_{(i-1)\hbar+(j-1)\delta}, \quad j = 1, \dots, M. \quad (1)$$

High frequency returns allow us to compute

$$[Y_\delta]_i = \sum_{j=1}^M y_{j,i}^2, \quad (2)$$

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 $\text{Var}(y_i|\tau_i) = \tau_i \text{Var}(Z_1)$, where $\tau_i = \tau_{i\hbar}^* - \tau_{(i-1)\hbar}^*$. That is it estimates the variance of y 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 (2002), while some of this work was extended by Meddahi (2002b). See also Barndorff-Nielsen and Shephard (2004a) for a discussion of the multivariate case and Andersen, Bollerslev, and Diebold (2004) for surveys of this area. Bai, Russell, and Tiao (2000) provide some simulation evidence for the effects of jumps on realised variances.

The notation $[Y_\delta]_i$ is designed to reflect the fact that this quantity is based on the Y process using δ spaced 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.

1.3 Some literature on jumps in financial economics

Being able to understand the empirical relevance of jumps is important as it has implications for risk management and asset allocation. A stream of recent papers in financial econometrics has addressed this issue using low frequency return data (e.g. the parametric models of Eraker, Johannes, and Polson (2003), Andersen, Benzoni, and Lund (2002), Chernov, Gallant, Ghysels, and Tauchen (2003) and the Markovian, non-parametric analysis of Aït-Sahalia (2002), Johannes (2003) and Bandi and Nguyen (2003)) and options data (e.g. Bates (1996), Carr and Wu (2004) and the review by Garcia, Ghysels, and Renault (2004)). Barndorff-Nielsen and Shephard (2004b) and Barndorff-Nielsen and Shephard (2003) introduced realised bipower variation which can be scaled to consistently estimate the quadratic variation of the continuous component of prices in the presence of jumps. Their results suggest that the realised bipower variation of a time-changed finite activity Lévy process would converge to zero as the time gaps between observations goes to zero. Other work along this line is Mancini (2003a) and Mancini (2003b). She introduces a jump threshold into the estimator of the quadratic variance. The absolute value of this threshold goes to zero as the number of observations within each day goes to infinity. Following Barndorff-Nielsen and Shephard (2004b), Woerner (2003) has studied the robustness of realised power variation $\delta^{1-r/2} \sum_{j=1}^{\lfloor t/\delta \rfloor} |y_{j,i}|^r$ to an infinite numbers of jumps in finite time periods showing that the robustness property of realised power variation goes through in that case. A related paper is Aït-Sahalia (2004), which shows that maximum likelihood estimation can disentangle a homoskedastic diffusive component from a purely discontinuous infinite activity Lévy component of prices. Outside the likelihood framework, the paper also studies the optimal combinations of moment functions for the generalised method of moment estimation of homoskedastic jump-diffusions.

1.4 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. Readers whose interest focuses solely on realised variance can start by reading Section 3 immediately after Section 2.1, without going through the other subsections of Section 2. In Section 4 we use the properties of realised variances 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 estimating various short memory and long memory Lévy and Brownian motion based SV models. Section 6 concludes. We provide

an Appendix which contains the proofs of the main results in the paper.

2 Some cumulants of returns

2.1 Background material on cumulants of Lévy processes

We will compute various moments of

$$Y_t = \mu t + Z_{\tau_t^*},$$

via the corresponding cumulants. To start, recall that cumulants are derived via the cumulant function, which is $\log \mathbb{E} e^{i\theta X}$ for some arbitrary random variable X . Then the j -th cumulant is defined as (assuming it exists)

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

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

$$\mu'_j = \mathbb{E} (X^j), \quad j = 1, 2, \dots,$$

then

$$\begin{aligned} \mu'_1 &= \kappa_1 \\ \mu'_2 &= \kappa_2 + \kappa_1^2 \\ \mu'_3 &= \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3 \\ \mu'_4 &= \kappa_4 + 4\kappa_1\kappa_3 + 3\kappa_2^2 + 6\kappa_1^2\kappa_2 + \kappa_1^4. \end{aligned} \tag{3}$$

Often we will be interested in the cumulants and moments of processes as a function of t . The cumulants and moments of Z_t will be written as $\kappa_{j,t}$ and $\mu'_{j,t}$, respectively.

A characterising feature of Lévy processes is that, so long as they exist,

$$\kappa_{j,t} = t\kappa_j, \quad j = 1, 2, \dots$$

This follows from the fact that

$$\mathbb{E} \{ e^{i\zeta Z_t} \} = \{ \mathbb{E} e^{i\zeta Z_1} \}^t,$$

which implies

$$\log \mathbb{E} \{ e^{i\zeta Z_t} \} = t \log \mathbb{E} e^{i\zeta Z_1}.$$

If we ignore drift for a moment, the only Lévy process without jumps is Brownian motion. Hence non-Gaussian increments and jumps are synonymous.

Suppose Z is Lévy process with $\kappa_4 < \infty$. Then

$$\begin{aligned} \mu'_{1,t} &= \kappa_1 t, \\ \mu'_{2,t} &= \kappa_2 t + \kappa_1^2 t^2, \\ \mu'_{3,t} &= \kappa_3 t + 3\kappa_1\kappa_2 t^2 + \kappa_1^3 t^3, \\ \mu'_{4,t} &= \kappa_4 t + (3\kappa_2^2 + 4\kappa_1\kappa_3) t^2 + 6\kappa_1^2\kappa_2 t^3 + \kappa_1^4 t^4. \end{aligned}$$

Note $\mu'_{3,t}$ and $\mu'_{4,t}$ have terms involving just t and so dominates for small t . In the case where

$$Z_t = \beta t + \sigma W_t,$$

with W being standard Brownian motion, then

$$\kappa_{1,t} = \beta t, \quad \kappa_{2,t} = \sigma^2 t \quad \text{and} \quad \kappa_{j,t} = 0 \text{ for } j > 2.$$

This implies that

$$\begin{aligned} \mu'_{1,t} &= \beta t, \\ \mu'_{2,t} &= \sigma^2 t + \beta^2 t^2, \\ \mu'_{3,t} &= 3\beta\sigma^2 t^2 + \beta^3 t^3, \\ \mu'_{4,t} &= 3\sigma^4 t^2 + 6\beta^2\sigma^2 t^3 + \beta^4 t^4. \end{aligned}$$

Now $\mu'_{3,t}$ and $\mu'_{4,t}$ have leading terms of order t^2 , so differ from some Lévy processes. The result on $\mu'_{4,t}$ will turn out as an essential observation for the behaviour of realised volatility for time-changed Lévy processes.

2.2 Fourth order cumulant and Brownian motion

Before we look at the properties of the time-change processes, we take a small technical detour. This can be skipped on first reading without losing the thread of the paper.

Suppose Z is a symmetric Lévy process with $\kappa_4 < \infty$. We now show that if $\kappa_4 = 0$ then Z must be scaled Brownian motion plus drift. This is a rather surprising result (as this does not hold for general random variables, that is variables which do not have to be infinitely divisible) and was communicated to us by Ken-iti Sato, who we thank for allowing us to state the result here.

Recall the Lévy-Khintchine representation for the cumulant function of Z_1 is, for a symmetric Lévy process

$$\begin{aligned} C\{\zeta \dagger Z_1\} &= \log E\left(e^{i\zeta Z_1}\right) \\ &= \mu i\zeta - \frac{\sigma^2}{2}\zeta^2 - \frac{1}{2} \int_R \left(1 - e^{i\zeta x} + i\zeta x \mathbf{1}_b(x)\right) \eta(dx) \\ &= \mu i\zeta - \frac{\sigma^2}{2}\zeta^2 + \int_{x \geq 0} \{\cos(\zeta x) - 1\} \eta(dx), \end{aligned}$$

where $b = [-1, 1]$ and η is associated symmetric Lévy measure (η must be symmetric as Z_1 is symmetric). Then

$$\frac{\partial^4 C\{\zeta \dagger Z_1\}}{\partial \zeta^4} = \int_{x \geq 0} x^4 \cos(\zeta x) \eta(dx).$$

Equation ζ to zero we have that, so long as it exists,

$$\kappa_4 = \frac{\partial^4 C\{\zeta \dagger Z_1\}}{\partial \zeta^4} \Big|_{\zeta=0} = \int_{x \geq 0} x^4 \eta(dx).$$

Hence if this is zero, then η must be zero and so there is no jump component in Z . This implies, by the Lévy-Khintchine representation, that Z must be essentially scaled Brownian motion plus drift. As a result one can test a Lévy process for jumps by first testing for symmetry and then, conditioning on symmetry, testing whether $\kappa_4 = 0$. In principle this is rather straightforward to do as it can be carried out using returns recorded at frequency and without detailed knowledge of a hypothesised form of the Lévy measure.

2.3 Conditional cumulants and moments of log-prices

Recall log-prices obey $Y_t = \mu t + Z_{\tau_t^*}$, where $Z \perp\!\!\!\perp \tau^*$. The cumulants of $Y_t | \tau_t^*$ are, when they exist,

$$\tau_t^* \kappa_j. \quad (4)$$

The relationship between cumulants and uncentred moments (3) can be used 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*

$$\begin{aligned} \mathbb{E}(Y_t | \tau_t^*) &= \kappa_1 \tau_t^*, & \mathbb{E}(Y_t^2 | \tau_t^*) &= \kappa_2 \tau_t^* + \kappa_1^2 \tau_t^{*2}, \\ \mathbb{E}(Y_t^3 | \tau_t^*) &= \kappa_3 \tau_t^* + 3\kappa_1 \kappa_2 \tau_t^{*2} + \kappa_1^3 \tau_t^{*3}, \\ \mathbb{E}(Y_t^4 | \tau_t^*) &= \kappa_4 \tau_t^* + (4\kappa_1 \kappa_3 + 3\kappa_2^2) \tau_t^{*2} + 6\kappa_1^2 \kappa_2 \tau_t^{*3} + \kappa_1^4 \tau_t^{*4}. \end{aligned} \quad (5)$$

Further,

$$\text{Var}(Y_t | \tau_t^*) = \kappa_2 \tau_t^*, \quad (6)$$

$$\text{Cov}(Y_t^2, Y_t | \tau_t^*) = \kappa_3 \tau_t^* + 2\kappa_1 \kappa_2 \tau_t^{*2}, \quad (7)$$

$$\text{Var}(Y_t^2 | \tau_t^*) = \kappa_4 \tau_t^* + 2(2\kappa_1 \kappa_3 + \kappa_2^2) \tau_t^{*2} + 4\kappa_1^2 \kappa_2 \tau_t^{*3}. \quad (8)$$

$$\mathbb{E}\left\{(Y_t^2 - \kappa_2 \tau_t^*)^2 | \tau_t^*\right\} = \kappa_4 \tau_t^* + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \tau_t^{*2} + 4\kappa_1^2 \kappa_2 \tau_t^{*3} + \kappa_1^4 \tau_t^{*4}.$$

□

Proof. Given in the Appendix.

2.4 Unconditional moments of log-prices

To derive unconditional moments of Y the second order properties of τ^* need to be found.

Remark 1 (*Barndorff-Nielsen and Shephard (2001)*) Suppose τ is covariance stationary with ξ , ω^2 and r being, respectively, its mean, variance and autocorrelation function. Then

$$\mathbb{E}(\tau_t^*) = \xi t, \quad \text{and} \quad \text{Var}(\tau_t^*) = 2\omega^2 r_t^{**} \quad (9)$$

where

$$r_t^* = \int_0^t r_u du \quad \text{and} \quad r_t^{**} = \int_0^t r_u^* du. \quad (10)$$

□

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

Proposition 2 If $\kappa_4 < \infty$, τ is covariance stationary and $\mathbb{E}(\tau_t^4) < \infty$, then

$$\mathbb{E}(Y_t) = \kappa_1 t \xi,$$

$$\mathbb{E}(Y_t^2) = \kappa_2 t \xi + \kappa_1^2 \left\{ 2\omega^2 r_t^{**} + (t\xi)^2 \right\},$$

$$\mathbb{E}(Y_t^3) = \kappa_3 t \xi + 3\kappa_1 \kappa_2 \left\{ 2\omega^2 r_t^{**} + (t\xi)^2 \right\} + \kappa_1^3 \mathbb{E}(\tau_t^{*3}),$$

$$\mathbb{E}(Y_t^4) = \kappa_4 t \xi + (4\kappa_1 \kappa_3 + 3\kappa_2^2) \left\{ 2\omega^2 r_t^{**} + (t\xi)^2 \right\} + 6\kappa_1^2 \kappa_2 \mathbb{E}(\tau_t^{*3}) + \kappa_1^4 \mathbb{E}(\tau_t^{*4}).$$

Consequently

$$\text{Var}(Y_t) = \kappa_2 t \xi + 2\kappa_1^2 \omega^2 r_t^{**},$$

$$\text{Cov} \left\{ (Y_t^2 - \kappa_2 \tau_t^*), \kappa_2 \tau_t^* \right\} = \kappa_1^2 \kappa_2 \mathbb{E}(\tau_t^{*3}),$$

$$\mathbb{E} \left\{ (Y_t^2 - \kappa_2 \tau_t^*)^2 \right\} = \kappa_4 t \xi + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \left\{ 2\omega^2 r_t^{**} + (t\xi)^2 \right\} + 4\kappa_1^2 \kappa_2 \mathbb{E}(\tau_t^{*3}) + \kappa_1^4 \mathbb{E}(\tau_t^{*4}).$$

□

Proof. This follows by the application of iterative expectations and (9).

Example 1 (*Barndorff-Nielsen and Shephard (2001)*) In the special case of $\kappa_1 = \kappa_3 = \kappa_4 = 0$

$$\mathbb{E}(Y_t) = 0, \quad \mathbb{E}(Y_t^2) = \kappa_2 t \xi, \quad \mathbb{E}(Y_t^3) = 0,$$

$$\text{Cov} \left\{ (Y_t^2 - \kappa_2 \tau_t^*), \kappa_2 \tau_t^* \right\} = 0,$$

$$\mathbb{E} \left\{ (Y_t^2 - \kappa_2 \tau_t^*)^2 \right\} = 2\kappa_2^2 \left\{ 2\omega^2 r_t^{**} + (t\xi)^2 \right\}.$$

□

When $\kappa_1 \neq 0$, the 3rd and 4th moments of log-prices require us to calculate the corresponding moments of τ^* . Proposition 3 is helpful in thinking about this when t is small.

Proposition 3 *If $\alpha_j = E\{|\tau_t - \xi|^j\}$ is finite for some natural number j then*

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

In fact,

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

□

Proof. Given in the Appendix. □

If we combine Propositions 2 and 3, then we have immediately the following result. Some of these results are derived in the special case of time-change Brownian motion by Meddahi (2002a) and Barndorff-Nielsen and Shephard (2002).

Proposition 4 *If $\kappa_4 < \infty$ and τ is stationary and $E\tau_t^4 < \infty$, then for $t \downarrow 0$,*

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

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

while

$$\text{Cov} \left[\{Y_t^2 - \kappa_2 \tau_t^*\}, \kappa_2 \tau_t^* \right] = O(t^3),$$

$$E \left\{ (Y_t^2 - \kappa_2 \tau_t^*)^2 \right\} = \kappa_4 t \xi + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \left\{ 2\omega^2 r_t^{**} + (t\xi)^2 \right\} + O(t^3).$$

□

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 $h > 0$,

$$y_i = Y_{hi} - Y_{h(i-1)}, \quad i = 1, 2, \dots, n.$$

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

$$\tau_i = \tau_{hi}^* - \tau_{h(i-1)}^*, \quad i = 1, 2, \dots, n.$$

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

$$y_1 = Y_{\hbar}.$$

Hence Proposition 2 directly computes

$$\mathbb{E}(y_i), \quad \mathbb{E}(y_i^2), \quad \mathbb{E}(y_i^3), \quad \mathbb{E}(y_i^4), \quad \text{Var}(y_i^2), \quad (11)$$

and

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

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 $s \neq 0$,

$$\text{Cov}(\tau_i, \tau_{i+s}) = \omega^2 \diamond r_{\hbar s}^{**} \quad (12)$$

where

$$\diamond r_{\hbar s}^{**} = r_{(s+1)\hbar}^{**} - 2r_{\hbar s}^{**} + r_{(s-1)\hbar}^{**}. \quad (13)$$

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

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

The above results imply the following.

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

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

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

and

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

Further,

$$\text{Cov}(y_i, y_{i+s}) = \kappa_1^2 \omega^2 \diamond r_{\hbar s}^{**},$$

while for $\hbar \downarrow 0$, to the order indicated,

$$\text{Cov}(y_i, y_{i+s}^2) = \kappa_1 \kappa_2 \omega^2 \diamond r_{\hbar s}^{**} + O(\hbar^3),$$

$$\text{Cov}(y_i^2, y_{i+s}^2) = \kappa_2^2 \omega^2 \diamond r_{\hbar s}^{**} + O(\hbar^3),$$

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

□

Proof. Given in the Appendix.

3 Cumulants of realised variance

3.1 Setting the scene: the theory of quadratic variation

Our major motivation for the development of the above results is in their use in understanding the properties of realised variance. Before we discuss this we will recall some general results on semimartingales.

It is well known that if Y is a semimartingale then the probability limit of $[Y_\delta]_i$ as $\delta \downarrow 0$ can be found from the quadratic variation process. 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

$$[Y]_t = \text{p-lim}_{M \rightarrow \infty} \sum_{j=1}^M (Y_{t_j} - Y_{t_{j-1}})^2, \quad (14)$$

for any sequence of partitions $t_0 = 0 < t_1 < \dots < t_M = t$ with $\sup_j \{t_{j+1} - t_j\} \rightarrow 0$ for $M \rightarrow \infty$. Here 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 τ^* and $[Y]$ in the case where $Y = W_{\tau^*}$ 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 $M \rightarrow \infty$

$$[Y_\delta]_i \xrightarrow{p} [Y]_{hi} - [Y]_{h(i-1)} = [Y]_i.$$

Example 2 Suppose A and τ^* have continuous sample paths and $M_t = W_{\tau_t^*}$, a SV process, then it is very well known that $[Y]_t = \tau_t^*$.

This example shows that in the Brownian motion time-change case $[Y_\delta]_i \xrightarrow{p} \kappa_2 \tau_i$. In the more general case of a time-changed Lévy process, $M = Z_{\tau^*}$, then

$$[M]_t \neq \kappa_2 \tau_t^*, \quad \text{so} \quad [Y_\delta]_i \xrightarrow{p} [Y]_i \neq \kappa_2 \tau_i.$$

3.2 Basic results on realised variance

3.2.1 Mean and variance of $[Y_\delta]_i - \kappa_2 \tau_i$

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 $[Y_\delta]_i$. Here we study the properties of the RV error $[Y_\delta]_i - \kappa_2 \tau_i$. To do this it is helpful to note that Barndorff-Nielsen and Shephard (2002) proved that as $\delta \downarrow 0$ so

$$2M^2 r_\delta^{**} \rightarrow \hbar^2.$$

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

$$\begin{aligned} \mathbb{E}([Y_\delta]_i - \kappa_2 \tau_i) &= \kappa_1^2 \{2\omega^2 M r_\delta^{**} + M^{-1} \hbar^2 \xi^2\} \\ \text{Var}([Y_\delta]_i - \kappa_2 \tau_i) &= \kappa_4 \hbar \xi + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \{2\omega^2 M r_\delta^{**} + M^{-1} \hbar^2 \xi^2\} + O(M^{-2}), \\ \text{Cov}\{([Y_\delta]_i - \kappa_2 \tau_i), \kappa_2 \tau_i\} &= O(M^{-2}) \end{aligned}$$

and

$$\text{Cov}\{([Y_\delta]_i - \kappa_2 \tau_i), ([Y_\delta]_{i+s} - \kappa_2 \tau_{i+s})\} = O(M^{-3}).$$

□

Proof. Given in the Appendix.

In the $\kappa_1 = \kappa_4 = 0$ and $\kappa_2 = 1$ case this reproduces the Barndorff-Nielsen and Shephard (2002) result on the behaviour of realised variance. Further, in the limit as $\delta \downarrow 0$ so

$$\text{Var}([Y]_i - \kappa_2 \tau_i) = \kappa_4 \hbar \xi.$$

The above means our Proposition 6 tells us the following.

- The bias of $[Y_\delta]_i$ as an estimator of $\kappa_2 \tau_i$. We see this is $O(M^{-1})$.
- The variability of $[Y_\delta]_i - \kappa_2 \tau_i$, the RV error. When $\kappa_4 > 0$ this involves a term which does not disappear as $M \rightarrow \infty$. This is important as this captures the inconsistency of $[Y_\delta]_i$ 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 square of the bias which is of $O(M^{-2})$. Hence in terms of the mean square error the bias has little impact in practice.
- The $\text{Cov}\{([Y_\delta]_i - \kappa_2 \tau_i), \kappa_2 \tau_i\}$ tells us that the time-change and RV error are, to a high order, uncorrelated.
- The $\text{Cov}\{([Y_\delta]_i - \kappa_2 \tau_i), ([Y_\delta]_{i+s} - \kappa_2 \tau_{i+s})\}$ 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 $[Y_\delta]_i$

To carry out inference on these types of models it is helpful to know the second order properties of $[Y_\delta]_i$.

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

$$\mathbb{E}([Y_\delta]_i) = \kappa_2 \hbar \xi + \kappa_1^2 (2\omega^2 M r_\delta^{**} + M^{-1} \hbar^2 \xi^2),$$

$$\text{Var}([Y_\delta]_i) = 2\omega^2 \kappa_2^2 r_\hbar^{**} + \kappa_4 \hbar \xi + (4\kappa_1 \kappa_3 + 2\kappa_2^2) (2\omega^2 M r_\delta^{**} + M^{-1} \hbar^2 \xi^2) + O(M^{-2}),$$

while, for $s \neq 0$,

$$\begin{aligned} \text{Cov}([Y_\delta]_i, [Y_\delta]_{i+s}) &= \kappa_2^2 \text{Cov}(\tau_i, \tau_{i+s}) + O(M^{-2}) \\ &= \kappa_2^2 \omega^2 \diamond r_{\hbar s}^{**} + O(M^{-2}). \end{aligned}$$

□

Proof. Given in the Appendix.

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

$$\mathbb{E}([Y_\delta]_i) \simeq \kappa_2 \hbar \xi, \quad \text{Cov} \{([Y_\delta]_i - \kappa_2 \tau_i), \kappa_2 \tau_i\} \simeq 0, \quad \text{Cov} \{([Y_\delta]_i - \kappa_2 \tau_i), \kappa_2 \tau_{i+s}\} \simeq 0,$$

$$\text{Var}([Y_\delta]_i) \simeq 2\omega^2 \kappa_2^2 r_\hbar^{**} + \kappa_4 \hbar \xi + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \{2\omega^2 M r_\delta^{**} + M^{-1} \hbar^2 \xi^2\},$$

$$\text{Cov}([Y_\delta]_i, [Y_\delta]_{i+s}) \simeq \kappa_2^2 \omega^2 \diamond r_{\hbar s}^{**},$$

3.3 Some implications for autocorrelation

Notice that jumps have little effect on $\mathbb{E}([Y_\delta]_i)$ and $\text{Cov}([Y_\delta]_i, [Y_\delta]_{i+s})$. Taken together these results mean that

$$\text{Cor}([Y_\delta]_i, [Y_\delta]_{i+s}) = \frac{\kappa_2^2 \omega^2 \diamond r_{\hbar s}^{**}}{2\omega^2 \kappa_2^2 r_\hbar^{**} + \kappa_4 \hbar \xi + 4\kappa_2^2 \omega^2 M r_\delta^{**} + 2M^{-1} \kappa_2^2 (\hbar \xi)^2}.$$

Thus the autocorrelation function of $[Y_\delta]_i$ is monotonically decreasing in κ_4 . Further, as $M \rightarrow \infty$ and assuming $\kappa_2 = 1$ to simplify the exposition

$$\begin{aligned} \text{Cor}([Y_\delta]_i, [Y_\delta]_{i+s}) &\rightarrow \frac{\omega^2 \diamond r_{\hbar s}^{**}}{2\omega^2 r_\hbar^{**} + \kappa_4 \hbar \xi} \\ &= \text{Cor}([Y]_i, [Y]_{i+s}) \\ &\leq \frac{\diamond r_{\hbar s}^{**}}{2r_\hbar^{**}} \\ &= \text{Cor}(\tau_i, \tau_{i+s}). \end{aligned}$$

Of course the equality, in the inequality, is obtained only in the Brownian case. Otherwise the autocorrelation amongst the $[Y_\delta]_i$ 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 (2004) and Andersen, Bollerslev, and Meddahi (2002) 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 $\kappa_4 \rightarrow \infty$, while fixing $\kappa_2 = 1$ to simplify the result. In this case, holding M fixed,

$$\text{Cor}([Y_\delta]_i, [Y_\delta]_{i+s}) \xrightarrow{\kappa_4 \rightarrow \infty} 0,$$

as the variance of the squares of the returns becomes infinity.

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

$$\begin{aligned} \text{Cor}([Y_\delta]_i, \kappa_2 \tau_i) &= \frac{\text{Var}(\kappa_2 \tau_i)}{\sqrt{\text{Var}(\kappa_2 \tau_i) \{ \text{Var}(\kappa_2 \tau_i) + \text{Var}(u_{M:i}) \}}} \\ &= \frac{2\omega^2 \kappa_2^2 r_h^{**}}{\sqrt{2\omega^2 \kappa_2^2 r_h^{**} \left\{ 2\omega^2 \kappa_2^2 r_h^{**} + \kappa_4 \bar{h} \xi + 4\kappa_2^2 \omega^2 M r_\delta^{**} + 2M^{-1} \kappa_2^2 (\bar{h} \xi)^2 \right\}}} \\ &\rightarrow \frac{r_h^{**}}{\sqrt{r_h^{**} \left\{ r_h^{**} + \frac{\kappa_4}{2\omega^2 \kappa_2^2} \bar{h} \xi \right\}}}, \end{aligned}$$

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

In the Brownian case Meddahi (2002b) has argued that we should replace $[Y_\delta]_i$ by the regression estimator

$$\widehat{\kappa_2 \tau_i} = (1 - \beta_M) \text{E}([Y_\delta]_i) + \beta_M [Y_\delta]_i, \quad \text{where} \quad \beta_M = \frac{\text{Var}(\kappa_2 \tau_i)}{\text{Var}([Y_\delta]_i)} \in [0, 1].$$

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

$$\begin{aligned} \beta_M &= \frac{\text{Var}(\kappa_2 \tau_i)}{\text{Var}([Y_\delta]_i)} \\ &= \frac{2\omega^2 \kappa_2^2 r_h^{**}}{2\omega^2 \kappa_2^2 r_h^{**} + \kappa_4 \bar{h} \xi + 4\kappa_2^2 \omega^2 M r_\delta^{**} + 2M^{-1} \kappa_2^2 (\bar{h} \xi)^2} \\ &\rightarrow \frac{r_h^{**}}{r_h^{**} + \frac{\kappa_4}{2\omega^2 \kappa_2^2} \bar{h} \xi} \\ &= \beta \leq 1. \end{aligned}$$

Now

$$\widehat{\kappa_2 \tau}_i \xrightarrow[M \rightarrow \infty]{p} (1 - \beta) \kappa_2 \hbar \xi + \beta [Y]_i.$$

Again, inevitably, this is an inconsistent estimator of $\kappa_2 \tau_i$ for $Z \neq W$. Andersen, Bollerslev, and Meddahi (2004) and Andersen, Bollerslev, and Meddahi (2002) have studied in detail other properties of these estimators in the Brownian motion time-deformation case.

4 Quasi-likelihood estimation based on $[Y_\delta]_i$

4.1 Basic principles

The theory we have been developing shows that if $\kappa_4 < \infty$ and τ is second order stationary then we can compute the unconditional mean and covariance of $\underline{[Y_\delta]} = ([Y_\delta]_1, \dots, [Y_\delta]_n)'$ as a function of the parameters of the model. We can argue exactly when $\kappa_1 = \kappa_3 = 0$ or asymptotically approximately for large M when this does not hold. We write the parameter vector as θ . Here n denotes the sample size. It makes sense to constrain $\kappa_2 = 1$ when the model is estimated using just the second order properties of $[Y_\delta]_i$ for otherwise the model will be unidentified as κ_2 and ξ play the same role in determining $\text{Var}([Y_\delta]_i)$.

We can define a Gaussian realised quasi-likelihood function

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

which allows us to define

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

using the fact that the score has zero expectation under the second order properties of the model.

In particular, for short-memory models if we write

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

then we would expect from the general theory of method of moments estimation

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

The asymptotic theory of this type of quasi-likelihood estimator is worked out by Dunsmuir (1979). It will be asymptotically equivalent to an estimator defined via the Whittle likelihood. Typically, in theory, we have to estimate \mathcal{J} using spectral matrix methods (e.g. Newey and West (1987)). Empirical evidence suggests the daily contributions to the score vector are close to being uncorrelated but that the standard estimator of \mathcal{J} is typically quite a lot larger than the estimator of \mathcal{I} even when M is large.

Clearly $\tilde{\theta}_{RQ}$ is suboptimal as $\underline{[Y_\delta]}$ is not Gaussian, although it is likely to be less poorly behaved than a quasi-likelihood based on (y_1^2, \dots, y_n^2) when M is large. Indeed we would expect it to become more efficient as M increases as the added noise $[Y_\delta]_i - \kappa_2 \tau_i$ becomes smaller as M increases.

Usually n is quite large in financial economics when $\hbar = 1$ represents one day and so computing \log_{RQ} is onerous due to the need to compute the inverse of the $n \times n$ matrix $\text{Cov}(\underline{[Y_\delta]})$, which is typically an $O(n^3)$ operation. However, stationarity of τ means that $\underline{[Y_\delta]}$ is itself stationary and so $\text{Cov}(\underline{[Y_\delta]})$ must be Toeplitz. The Durbin algorithm¹ (e.g. Golub and Van Loan (1989, p. 187) and Doornik and Ooms (2003, p. 29)) can be used to compute $\log L_{RQ}$ in $O(n^2)$. This works with the Choleski decomposition of the inverse of $\text{Cov}(\underline{[Y_\delta]})$. We write $\text{Cov}(\underline{[Y_\delta]})$ as LDL' , where L is lower triangular and D is diagonal, so

$$\log L_{RQ}(\theta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |D| - \frac{1}{2} e' e, \quad \text{where } e = L^{-1} D^{-1/2} \left\{ \underline{[Y_\delta]} - \text{E}(\underline{[Y_\delta]}) \right\}.$$

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 $[Y_\delta]_i$,

$$\widehat{[Y_\delta]_{i|i-1}} = P_L([Y_\delta]_i | [Y_\delta]_1, \dots, [Y_\delta]_{i-1})$$

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

$$e_i = [Y_\delta]_i - \widehat{[Y_\delta]_{i|i-1}}.$$

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

$$\tau_i = \xi + x \varpi_i, \quad \varpi_{i+1} = \mathcal{T} \varpi_i + v_i, \tag{15}$$

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 of finite dimension and v_i is a zero mean, weak white noise process. Then we can compute $\log L_{RQ}$ 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

$$[Y_\delta]_i = \kappa_2 \tau_i + u_i,$$

¹Most matrix languages have functions which carry out the Durbin algorithm, e.g. in Ox it is `pacf`.

where the properties of u_i were given in Proposition 6, can be combined with (15) to put $[Y_\delta]_i$ into a linear state space representation. In particular if we write the best linear predictor of ϖ_i using $[Y_\delta]_1, \dots, [Y_\delta]_{i-1}$ as $a_{i|i-1}$ and the associate mean square error matrix as $P_{i|i-1}$, then

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

and so²

$$\log_{RQ}(\theta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^n \log f_i - \frac{1}{2} \sum_{i=1}^n f_i^{-1} ([Y_\delta]_i - \kappa_2 \widehat{\tau}_{i|i-1})^2. \quad (16)$$

Informal checks suggest that when $n = 3,000$ and the dimension of the state space is two then the Kalman filter computes $\log L_{RQ}$ 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 (15) 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. Jacquier, Polson, and Rossi (1994), Kim, Shephard, and Chib (1998)), (ii) simulation based indirect inference (e.g. Gouriou, Monfort, and Renault (1993), Gallant and Tauchen (1996)), (iii) generalised method of moments (e.g. Andersen and Sørensen (1996)), (iv) simulation based maximum likelihood using importance sampling (e.g. Durham and Gallant (2002)).

Previous related work on estimating parameters of SV models using realised variances includes Bollerslev and Zhou (2002) and Barndorff-Nielsen and Shephard (2002). 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 (2002) used the above quasi-likelihood in the case where the model can be handled by the Kalman filter.

Finally, Durbin's method also automatically delivers the sequence of best linear unbiased estimators of $\kappa_2 \tau_i$ given $[Y_\delta]_1, \dots, [Y_\delta]_{i-1}$, written as

$$\kappa_2 \widehat{\tau}_{i|i-1} = P_L (\kappa_2 \tau_i | [Y_\delta]_1, \dots, [Y_\delta]_{i-1})$$

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

and the associated mean square error f_i for $i = 1, 2, \dots, n$ in $O(n^2)$. Likewise, for the more specialised case of the Kalman filter, all these quantities are also produced in $O(n)$. 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 $\tau_i | [Y_\delta]_1, \dots, [Y_\delta]_{i-1}$ or $\tau_i | [Y_\delta]_1, \dots, [Y_\delta]_i$ or, indeed, $\tau_i | [Y_\delta]_1, \dots, [Y_\delta]_n$. 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).

4.2 Numerical aspects

4.2.1 Single factor model

We start with the simplest model for τ where

$$r_s = \text{Cor}(\tau_t, \tau_{t+s}) = \exp(-\lambda s).$$

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 $[Y_\delta]_i$ 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 (12) implies that

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

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

$$(\tau_i - \bar{h}\xi) = \phi(\tau_{i-1} - \bar{h}\xi) + e_i + \theta e_{i-1}.$$

The autoregressive root is $\phi = e^{-\lambda h}$. Barndorff-Nielsen and Shephard (2002) noted this but then worked out θ and $\text{Var}(e_i)$ 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 $\text{Var}(\tau_i)$ and $\text{Cor}(\tau_i, \tau_{i-1})$, which is the case in this class of models.

Lemma 1 (after a related result by Meddahi (2003) on the process for y_i^2). Straightforwardly $\text{Var}(e_i) (1 + \theta^2 + 2\theta\phi) = \text{Var}(\tau_i) (1 - \phi^2)$. Now write $c_i = \tau_i - \phi\tau_{i-1}$, then $c_i = e_i + \theta e_{i-1}$ so that

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

and

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

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

$$\begin{aligned} \rho_1 &= \frac{\text{Cov}(c_i, c_{i-1})}{\text{Var}(c_i)} \in \left[0, \frac{1}{2}\right], \\ \theta &= \frac{1 - \sqrt{1 - 4\rho_1^2}}{2\rho_1} \in [0, 1]. \end{aligned}$$

Remark 2 Note that

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

which only depends on λ and h . 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 $h = 1$.

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

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

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.

4.2.2 J factor model

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

$$r_s = \sum_{j=1}^J w_j \exp(-\lambda_j s) \quad \text{where} \quad \sum_{j=1}^J w_j = 1 \quad \text{and} \quad w_j \geq 0. \quad (17)$$

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 (2002) and Chernov, Gallant, Ghysels, and Tauchen (2003).

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 $\log L_{RQ}$ 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

$$\tau_i = \tilde{h}\xi = (1 \ 0 \ 1 \ 0 \dots 1 \ 0) \varpi_i,$$

and

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

Hence the Kalman filter can be used to rapidly compute $\log L_{RQ}$.

4.2.3 Log-normal OU process

An important class of models is where we put $\log \tau$ as a $N(\xi_{\log}, \omega_{\log}^2)$ -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

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

while

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

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

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

where

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

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 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.

4.2.4 A long memory model

Suppose we parameterise r

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

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 rigorous 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 Joyeux (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), Meddahi (2001), Comte, Coutin, and Renault (2003) and Gloter and Hoffmann (2004). 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 $\varepsilon^*(t; \lambda) = \lambda^{-2}(e^{-\lambda t} - 1 + \lambda t)$, we have that

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

Example 3 Let π correspond to a $\Gamma(2H, \alpha)$ distribution. 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

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

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

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

which is straightforward to calculate and so we can use Durbin's method to compute the quasi-likelihood.

5 Illustration based on realised variances

5.1 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 using daily realised variances 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 return dataset similar to that 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 (2002). This series records the United States Dollar/ German Deutsch 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 2(a) shows the implied daily returns over this period and Figure 2(b) the corresponding correlogram for squared daily returns. The correlations tend to be quite small, although with a preponderance of positive numbers. Figure 2(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 2(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.

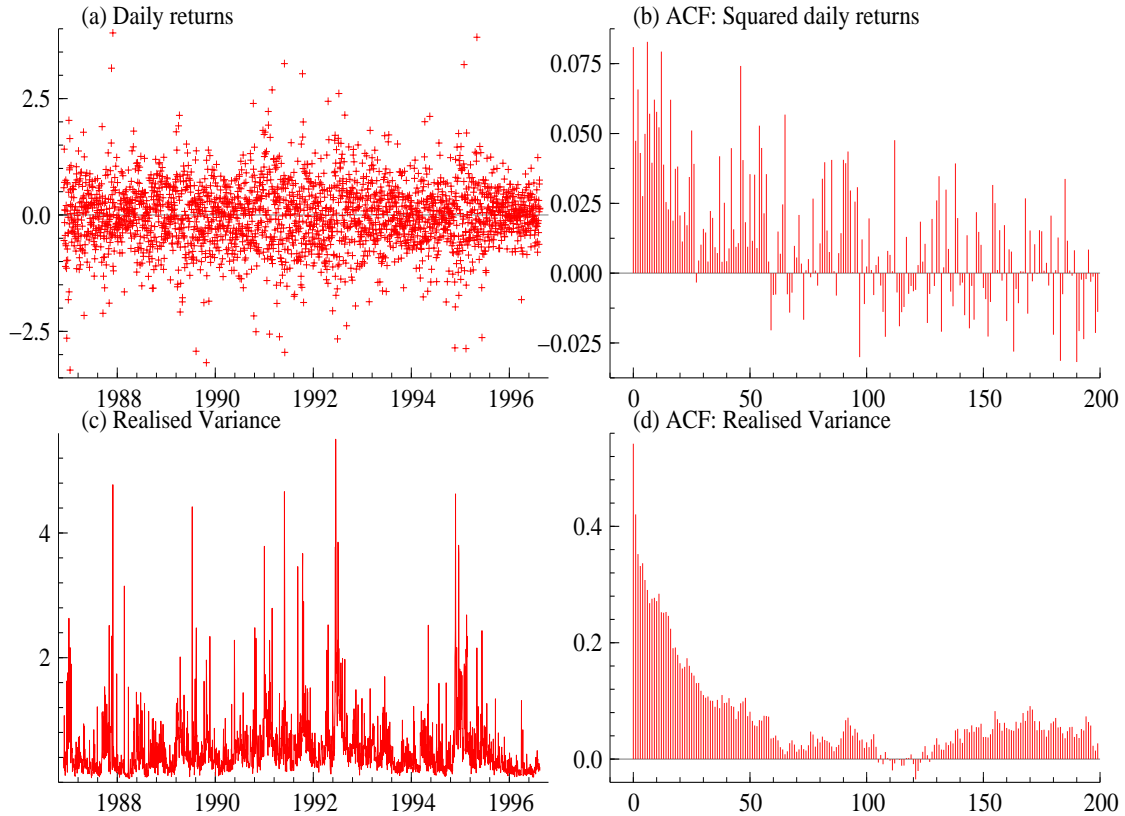


Figure 2: *DM against the Dollar, based on the Olsen dataset. (a) Implied daily returns. (b) Correlogram for the squared daily returns. (c) Daily realised variance based on the 5 minute returns. (d) Correlogram for the realised variance for the DM series. Code: realised_quasi_track.ox.*

5.2 Superposition model

Table 1 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 more general Lévy version improves the fit of the moments 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 third component does not improve the quasi-likelihood. Table 2 provides estimated standard errors for some of the estimates in this Table. It is designed to show how the estimated errors change as a function of M . We can see that κ_4 and ξ become more precisely determined as M increases, but the other parameters are not much effected.

Table 3 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 $[Y_\delta]_i$ is very close to that of y_i^2 for the daily data, the variance of $[Y_\delta]_i$ is much lower than the corresponding y_i^2 . This may reflect the fact that for time-change processes 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 3 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 fitted and empirical expectations and variances of the RVs being comparable.

For the single OU DM case Figure 3(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 $J = 2$ Brownian based model is much more reasonable, although it tends to under weight the longer lags in the acf. When we fitted the $J = 3$ 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 3(b) shows the memory of the acf last to around 50 lags, which is still of course too short. When we move to the $J = 2$ model the fitted acf is similar to that fitted by the $J = 3$ Brownian model. This seems a reasonable description of the empirical correlogram. The move to a $J = 3$ model allows the fitted 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 be seen from the

ξ	QML estimators				$\log L_{RQ}$	Empirical	Fitted	Empirical	BP_{25}
	ω^2	κ_4	λ	w		$E([Y_{\delta}]_i)$	$V([Y_{\delta}]_i)$	$V([Y_{\delta}]_i)$	
raw data									143
.644	.104		.0170		-3582.3	.504	1.14	1.11	19.5
.504	.0731	.772	.0311		-3579.0	.504	1.11	1.11	18.1
.504	.0773	.753	(.0146,.0686)	(.479,.520)	-3578.7	.504	1.11	1.11	17.1

Table 1: Quasi-likelihood fit to the DM data using the realised variances based on $M = 1$, which is the same as square returns. Raw data gives diagnostics of raw RVs. Code: realised_quasi_track.ox

Estimates of asymptotic standard errors				
M	ξ	ω^2	κ_4	λ
1	0.0483	0.0302	0.210	0.0150
2	0.0413	0.0248	0.159	0.0141
4	0.0418	0.0376	0.106	0.0205
72	0.0328	0.0397	0.0410	0.0399
144	0.0347	0.0350	0.0309	0.0303
288	0.0297	0.0375	0.0292	0.0507

Table 2: Standard errors for the quasi-likelihood estimator based on the realised variances using a variety of values of M . Code: quasi_small.ox

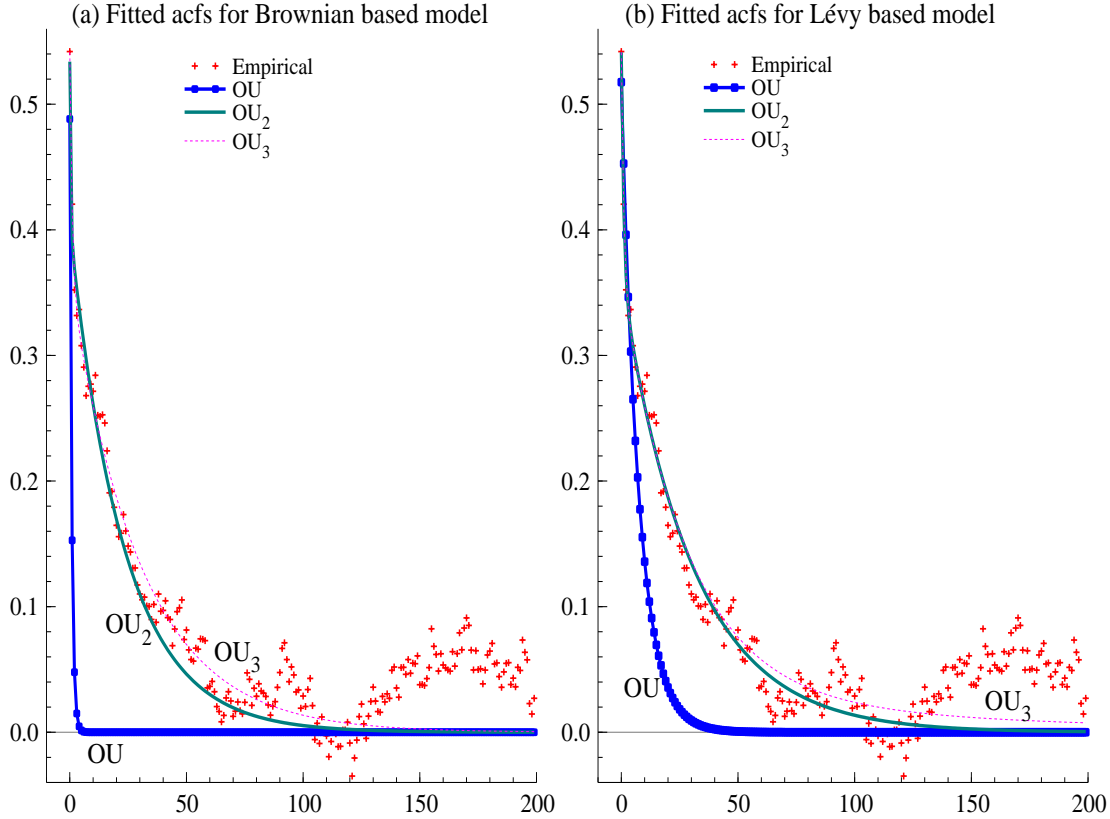


Figure 3: *Empirical and fitted autocorrelation functions for the realised variances for the DM against the Dollar using $M = 288$. (a) Brownian motion based model. (b) Lévy based model. Code: `realised_quasi_track.ox`.*

fitted autocorrelation function drawn in Figure 3(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 J superposition models, for small J 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 3.

Again as we take M down to 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 the variance process which has a great deal of memory.

Finally, Table 3 gives the corresponding result for $M = 2$. This is in line with the analysis

M $(\widehat{V}[Y_\delta]_i)$	QML estimators						$\log L_{RQ}$	Empir	Fitted	BP_{25}
	ξ	ω^2	κ_4	λ	w	$E[Y_\delta]_i$		$V[Y_\delta]_i$		
288 (.237)	raw data									4689
	.530	.328		1.16			-1356.2	.528	.235	744
	.527	.140	.190	.133			-1222.4	.528	.237	82.2
	.526	.385		(2.85,.0432)	(.741,.258)		-1195.5	.528	.236	34.7
	.525	.193	.130	(.975,.0328)	(.543,.456)		-1190.0	.528	.237	23.1
	.525	4.68		(203,.973,.0328)	(.958,.0223,.0188)		-1190.0	.528	.237	23.1
	.523	.193	.131	(.983,.0361,.00728)	(.541,.420,.0380)		-1189.9	.528	.237	23.2
144 (.255)	raw data									3913
	.511	.373		1.43			-1552.2	.509	.253	843
	.507	.121	.260	.0851			-1398.9	.509	.255	57.7
	.507	.482		(3.88,.0430)	(.796,.203)		-1382.7	.509	.254	25.3
	.506	.174	.193	(.970,.0330)	(.495,.504)		-1378.4	.509	.255	16.9
	.506	4.48		(.966,155,.0330)	(.0191,.961,.0195)		-1378.4	.509	.255	17.0
	.504	.174	.193	(.982,.0367,.00880)	(.493,.458,.0475)		-1378.3	.509	.255	17.0
72 (.277)	raw data									3184
	.491	.393		1.42			-1671.5	.488	.275	678
	.487	.125	.296	.0996			-1547.2	.488	.276	70.0
	.487	.486		(3.48,.0437)	(.802,.197)		-1528.8	.488	.276	26.1
	.486	.188	.213	(.990,.0342)	(.545,.454)		-1524.6	.488	.276	24.2
	.486	4.67		(.988,199,.0342)	(.0219,.959,.0183)		-1524.6	.488	.276	24.2
	.486	.188	.213	(.992,.0114,.0349)	(.544,.0101,.445)		-1524.6	.488	.276	24.2
2 (.774)	raw data									249
	.738	.158		.0104			-3123.1	.491	.860	24.6
	.489	.0735	.794	.0410			-3114.2	.491	.774	23.2
	.723	.178		(.00003,.0420)	(.589,.410)		-3116.5	.491	.878	23.1
	.488	.0750	.792	(.0560,.0143)	(.758,.241)		-3114.1	.491	.774	22.8
1 (1.11)	raw data									143
	.644	.104		.0170			-3582.3	.504	1.14	19.5
	.504	.0731	.772	.0311			-3579.0	.504	1.11	18.5
	.639	.109		(.0633,.00390)	(.506,.494)		-3580.8	.504	1.14	18.1
	.504	.0773	.753	(.0686,.0146)	(.520,.479)		-3578.7	.504	1.11	17.1

Table 3: Quasi-likelihood fit to the DM data using the realised variance, for a variety of values of M . BP denotes Box-Pierce statistic, computed using 25 lags. The rows denoted raw data show the Box-Pierce statistic of the realised variances when no model is fitted. Code: realised_quasi_track.ox

based on squared daily data. This shows that we need the Lévy effects in order to produce variance tracking.

5.3 Log-OU models

The corresponding results for the estimated log-OU based time-deformation models are given in Table 4. 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.

5.4 Long memory models

Table 5 gives the RV estimates of the gamma-long memory process of Section 4.2.3 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 3 on the Lévy based model whose spot variance was constructed by a finite 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. When the OU components are added, then the estimated values of H seem in line with those estimated for FIGARCH models, e.g. Maheu (2003). 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.

6 Conclusion

In this paper we have studied the econometrics of 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

M	QML estimators				$\log L_{RQ}$	BP_{25}
	ξ_{\log}	ω_{\log}^2	κ_4	λ		
288	raw data					4689
	.553	.882		.000456	-1271.1	176
	-.844	.409	.187	.115	-1219.4	77.1
144	raw data					3913
	.405	.784		.000565	-1433.3	106
	-.870	.385	.257	.0748	-1397.2	54.6
72	raw data					3184
	.278	.631		.00117	-1582.6	121
	-.932	.428	.292	.0884	-1545.3	66.5
24	raw data					1810
	.0253	.510		.00223	-1940.9	64.9
	-.981	.412	.444	.0599	-1914.3	42.7
6	raw data					1002
	-1.16	.911		1.14	-2370.0	312
	-.901	.326	.490	.0383	-2290.2	16.8
2	raw data					249
	-1.14	.902		1.23	-3148.7	130
	-.848	.269	.791	.0346	-3114.2	23.3
1	raw data					142
	-.552	.224		.0145	-3582.1	19.3
	-.810	.254	.768	.0263	-3578.9	18

Table 4: Quasi-likelihood estimation of LNOU SV process using realised variances based on different values of M , the number of intra-day observations. BP denotes Box-Pierce statistic, computed using 25 lags. The rows denoted raw data show the Box-Pierce statistic of the realised variances when no model is fitted. Code: realised_quasi_track.ox

		QML estimators						$\log L_{RQ}$	BP_{25}
		ξ	ω^2	κ_4	α	λ	H		
288	raw data								4689
	.512	2.75		.000122		.171		-1197.2	35.9
	.515	2.75		.000456	.0278	.235	(.980,.0199)	-1192.3	27.0
	.524	.667		.0438	(1.15,.0296)	.402	(.883,.00460,.111)	-1191.5	25.9
	.495	.525	.0801	.00564		.113		-1195.3	30.0
	.523	.605	.0110	.0483	.0291	.386	(.891,.108)	-1191.5	25.9
144	raw data								3913
	.505	4.13		.000112		.200		-1389.7	41.2
	.502	4.14		.000589	.0294	.298	(.983,.0165)	-1379.6	19.1
	.476	.287	.183	.0706		.0919		-1383.3	23.0
	.505	.258	.149	.262	.0290	.470	(.708,.291)	-1379.3	18.9
72	raw data								3184
	.489	3.92		.000119		.194		-1532.9	38.8
	.479	3.93		.000476	.0297	.272	(.984,.0153)	-1526.4	27.6
	.461	.513	.157	.00515		.109		-1529.8	30.5
	.485	.621	.0652	.0475	(.0306,.0293)	.400	(.884,.115,small)	-1525.9	26.7
2	raw data								249
	.679	.225		.174		.0563		-3117.7	24.8
	.722	.176		.180	(.0493,.0410)	.0146	(.661,.123,.214)	-3116.3	23.2
	.480	.0834	.797	14.4		.470		-3114.7	23.7
1	raw data								142
	.640	.109		9.44		.227		-3580.5	18.5
	.490	.145	.548	.581		.128		-3580.7	18.1
	.505	.0739	.768	2.09	.0308	.357	(.0244,.975)	-3579.0	18.1

Table 5: 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. BP denotes Box-Pierce statistic, computed using 25 lags. The rows denoted raw data show the Box-Pierce statistic of the realised variances when no model is fitted. Code: `quasi_track.ox`

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. Empirical work suggests the time-change Lévy processes are empirical interesting generalisations of time-change Brownian motion models.

7 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, the comments from Frank Gerhard and Matthias Winkel on an early draft and the reports from three anonymous referees which improved the exposition. Finally, we thank Ken-iti Sato for answering our question on cumulants of Lévy processes.

8 Appendix

8.1 Proof of Proposition 1

The first four results follow from (4) together with (3). (6)-(8) follow from the use of these results with the standard formulae that, generically for some X and Y ,

$$\text{Var}(Y|X) = E(Y^2|X) - E(Y|X)^2, \quad \text{Cov}(Y^2, Y|X) = E(Y^3|X) - E(Y^2|X)E(Y|X),$$

and

$$\text{Cov}(Y^2, Y|X) = E(Y^3|X) - E(Y^2|X)E(Y|X).$$

Finally,

$$\begin{aligned} E\left\{(Y_t^2 - \kappa_2 \tau_t^*)^2 \mid \tau_t^*\right\} &= E\left\{(Y_t^4 - 2Y_t^2 \kappa_2 \tau_t^* + \kappa_2^2 \tau_t^{*2}) \mid \tau_t^*\right\} \\ &= E(Y_t^4 \mid \tau_t^*) - 2\kappa_2 \tau_t^* E(Y_t^2 \mid \tau_t^*) + \kappa_2^2 \tau_t^{*2} \\ &= E(Y_t^4 \mid \tau_t^*) + 3\kappa_2^2 \tau_t^{*2} - 2\kappa_1^2 \kappa_2 \tau_t^{*3} \\ &= \kappa_4 \tau_t^* + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \tau_t^{*2} + 4\kappa_1^2 \kappa_2 \tau_t^{*3} + \kappa_1^4 \tau_t^{*4}, \end{aligned}$$

as stated

□

8.2 Proof of Proposition 3

$$\begin{aligned} |\mathbb{E}\{(\tau_t^* - \xi t)^j\}| &\leq \mathbb{E} \left| \int_0^t \dots \int_0^t (\tau_{u_1} - \xi) \dots (\tau_{u_j} - \xi) du_1 \dots du_j \right| \\ &\leq \int_0^t \dots \int_0^t \mathbb{E} \{ |\tau_{u_1} - \xi) \dots (\tau_{u_j} - \xi)| \} du_1 \dots du_j. \end{aligned}$$

Further, we can use the result that for arbitrary random variables x_1, \dots, x_j we have

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

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

$$\begin{aligned} |\mathbb{E}\{(\tau_t^* - \xi t)^j\}| &\leq \int_0^t \dots \int_0^t \alpha_j^{1/j} \dots \alpha_j^{1/j} du_1 \dots du_j \\ &= \alpha_j \int_0^t \dots \int_0^t du_1 \dots du_j \\ &= \alpha_j t^j. \end{aligned}$$

□

8.3 Proof of Proposition 5

First

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

Second

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

Third

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

Finally

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

□

8.4 Proof of Proposition 6

It is useful to write

$$\tau_{j,i} = \tau_{(i-1)\hbar+j\delta}^* - \tau_{(i-1)\hbar+(j-1)\delta}^*, \quad j = 1, \dots, M.$$

Then we can decompose

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

Now Propositions 5, 4 and 1 tell us that

$$\begin{aligned}
\mathbf{E}(y_{j,i}^2 - \tau_{j,i}) &= \kappa_1^2 \left\{ 2\omega^2 r_\delta^{**} + (\delta\xi)^2 \right\} \\
\text{Var}(y_{j,i}^2 - \tau_{j,i}) &= \kappa_4 \delta\xi + (4\kappa_1 \kappa_3 + 2\kappa_2^2) \left\{ 2\omega^2 r_\delta^{**} + (\delta\xi)^2 \right\} \\
&\quad + 4\kappa_1^2 \kappa_2 \mathbf{E} \left\{ \tau_\delta^{*3} \right\} + \kappa_1^4 \mathbf{E} \left\{ \tau_\delta^{*4} \right\}, \\
\text{Cov}(y_{j,i}^2 - \tau_{j,i}, y_{j+s,i}^2 - \tau_{j+s,i}) &= O(M^{-4}).
\end{aligned}$$

This implies the desired result.

□

8.5 Proof of Proposition 7

We have that

$$[Y_\delta]_i = \kappa_2 \tau_i + ([Y_\delta]_i - \kappa_2 \tau_i).$$

Then (9) and Proposition 6 imply that

$$\mathbf{E}([Y_\delta]_i) = \kappa_2 \hbar \xi + \kappa_1^2 \left\{ 2\omega^2 M r_\delta^{**} + M^{-1} \hbar^2 \xi^2 \right\}.$$

Likewise

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

Finally,

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

□

References

- Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. *Journal of Finance* 57, 2075–2112.
- Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics*. Forthcoming.
- Alizadeh, S., M. Brandt, and F. Diebold (2002). Range-based estimation of stochastic volatility models. *Journal of Finance* 57, 1047–1091.
- Andersen, T. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239–1284.
- 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 (2004). 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. Correction published in 2003, volume 98, page 501.
- Andersen, T. G., T. Bollerslev, and N. Meddahi (2002). 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., T. Bollerslev, and N. Meddahi (2004). Analytic evaluation of volatility forecasts. *International Economic Review* 45. Forthcoming.
- Andersen, T. G. and B. Sørensen (1996). GMM estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Business and Economic Statistics* 14, 328–352.
- Back, K. (1991). Asset pricing for general processes. *Journal of Mathematical Economics* 20, 371–395.
- 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.
- Bandi, F. M. and T. H. Nguyen (2003). On the functional estimation of jump-diffusion models. *Journal of Econometrics* 116, 293–328.
- 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 (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* 64, 253–280.

- Barndorff-Nielsen, O. E. and N. Shephard (2003). Econometrics of testing for jumps in financial economics using bipower variation. Unpublished discussion paper: Nuffield College, Oxford.
- Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica* 72. Forthcoming.
- Barndorff-Nielsen, O. E. and N. Shephard (2004b). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1–48.
- Bates, D. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies* 9, 69–107.
- Bertoin, J. (1996). *Lévy Processes*. Cambridge: Cambridge University Press.
- Bochner, S. (1949). Diffusion equation and stochastic processes. *Proceedings of the National Academy of Science of the United States of America* 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* 13, 345–382.
- Carr, P. and L. Wu (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 113–141.
- Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative models of stock price dynamics. *Journal of Econometrics* 116, 225–257.
- Clark, P. K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica* 41, 135–156.
- Comte, F., L. Coutin, and E. Renault (2003). Affine fractional stochastic volatility models. Unpublished paper: University of Montreal.
- Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance* 8, 291–323.
- Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. London: Chapman and Hall.
- 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. J. 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.
- Dunsmuir, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its applications to models for a signal observed with noise. *Annals of Statistics* 7, 490–506.

- 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.
- Eraker, B., M. Johannes, and N. G. Polson (2003). The impact of jumps in returns and volatility. *Journal of Finance* 53, 1269–1300.
- Gallant, A. R., C. Hsu, and G. Tauchen (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *The Review of Economics and Statistics* 81, 617–631.
- Gallant, A. R. and G. Tauchen (1996). Which moments to match. *Econometric Theory* 12, 657–81.
- Garcia, R., E. Ghysels, and E. Renault (2004). The econometrics of option pricing. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. North Holland. Forthcoming.
- 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 Lévy 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.
- Gloter, A. and M. Hoffmann (2004). Stochastic volatility and fractional Brownian motion. *Stochastic Processes and Their Applications*. Forthcoming.
- 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. Joyeux (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-Heinemann.
- 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.
- Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics* 12, 371–417.
- Johannes, M. (2003). The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance*. Forthcoming.
- Johannes, M., N. G. 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. *The 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.
- Maheu, J. (2003). Can garch models capture the long-range dependence in financial market volatility? Unpublished paper: Department of Economics, University of Toronto.
- Mancini, C. (2003a). Estimation of the characteristics of jump of a general Poisson-diffusion process. Dipartimento di Matematica per le Decisioni, Università di Firenze.
- Mancini, C. (2003b). Statistics of a Poisson-Gaussian process. Dipartimento di Matematica per le Decisioni, Università di Firenze.
- Mandelbrot, B. (1971). A fast fractional Gaussian noise generator. *Water Resource Research* 7, 543–553.
- Meddahi, N. (2001). An eigenfunction approach for volatility modeling. Unpublished paper: University of Montreal.
- Meddahi, N. (2002a). Moments of continuous time stochastic volatility models. Unpublished paper: University of Montreal.
- Meddahi, N. (2002b). A theoretical comparison between integrated and realized volatilities. *Journal of Applied Econometrics* 17, 479–508.
- Meddahi, N. (2003). ARMA representation of integrated and realized variances. *The Econometrics Journal* 6, 334–355.
- 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. J. 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.
- Woerner, J. (2003). Estimation of integrated volatility in stochastic volatility models. Unpublished paper: OCIAM, Mathematical Institute, University of Oxford.