Aggregation and model construction for volatility models

Ole E. Barndorff-Nielsen
Department of Mathematical Sciences,
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark.
oebn@mi.aau.dk

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

April 1998

Abstract

In this paper we will rigourously study some of the properties of continuous time stochastic volatility models. We have five main results: (i) the stochastic volatility class can be linked to Cox process based models of tick-by-tick financial data; (ii) we characterise the moments, autocorrelation function and spectrum of squared returns; (iii) based only on discrete time returns, we give a simple consistent and asymptotically normally distributed estimator of continuous time volatility models without any simulation or discretisation error. Furthermore, we review a new class of Ornstein-Uhlenbeck processes of volatility, introduced in a companion paper, which allows (iv) the discrete time returns to be simulated without any form of discretisation error, (v) explicit modelling of correlation structures and allow analytic calculations of the properties of returns.

Contents

1 Introduction

Continuous time stochastic volatility (SV) models have had a substantial impact on theoretical financial economics and econometric theory and practice. A review of the literature is given in Ghysels, Harvey, and Renault (1996). A standard model for the evolution of an asset price in the literature is where $x^*(t)$ is the log-price and $w(t)$ is Brownian motion, then $x^*(t)$ follows the solution to a linear stochastic differential equation (SDE) of the form

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

where $t \geq 0$ and where $\sigma^2(t)$, the instantaneous volatility, is latent. Such models, by appropriate design of the stochastic process for $\sigma^2(t)$, allow aggregated returns $\{y_n\}$ measured over a period Δ, where

$$y_n = \int_{(n-1)\Delta}^{n\Delta} \mathrm{d}x^*(t) = x^*(n\Delta) - x^*\left\{(n-1)\Delta\right\}, \quad \Delta > 0,$$

to be heavy-tailed, exhibit volatility clustering and aggregate to Gaussianity as Δ gets large. These are the main features of asset returns surveyed by, for example, Campbell, Lo, and MacKinlay (1997, pp. 17-21). Common models for $\sigma^2(t)$ include an Ornstein-Uhlenbeck process with Brownian motion increments written for $\log\sigma^2(t)$ (e.g. Hull and White (1987)), an ARCH diffusion (Nelson (1990)) and a square root process (e.g. Heston (1993)).

In this paragraph we will assume $\{\sigma^2(t)\}$ is independent of the $\{w(t)\}$. Then whatever the model for σ^2 it follows that

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

where

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

This implies integrated volatility plays a crucial role in continuous time volatility models. Unfortunately, existing models of volatility do not allow an easy treatment of integrated volatility and so most researchers tend to resort to discretisation approximations even to simulate return sequences.

In this paper we will rigourously study some of the properties of continuous time stochastic volatility models. We have five main results: (i) stochastic volatility class can be linked to Cox process based models, put forwarded independently by Rogers and Zane (1998) and Rydberg and Shephard (1998), of tick-by-tick financial data; (ii) the moments, autocorrelation function and spectrum of $\{y_n^2\}$ are characterised; (iii) based only on discrete time returns, we give a simple consistent and asymptotically normally distributed estimator of continuous time volatility models without any simulation or discretisation error. Furthermore, we review a new class of Ornstein-Uhlenbeck processes of $\sigma^2(t)$ introduced in a companion paper Barndorff-Nielsen and Shephard (1998), which allows (iv) the $\{y_n\}$ to be simulated without any form of discretisation error as well as providing (v) simple and interpretable dynamic structures for the volatility allowing explicit modelling of correlation structures and analytic calculations of the properties of returns.

The structure of the paper is as follows. In Section 2 we will formally introduce our notation and consider some of the basic mathematical properties of $x^*(t)$. In this section we also connect stochastic volatility models with the recent work on tick-by-tick models where point processes have been used to model the time between trades (see, for example, Engle and Russell (1998), Ghysels, Jasiak, and Gourieroux (1998), Rydberg and Shephard (1998) and Rogers and Zane (1998)). In particular we show that we can derive the SV class of processes as a limit of a general class of Cox processes for tick-by-tick data. In Section 3 we will study the moments of y_n and derive simple expressions for the autocorrelation and spectrum of the squared returns. The results can be used to provide simple consistent and asymptotically normally distributed estimators of the continuous time models using discrete time return data. In Section 4 we will review our recent work on the construction of continuous time models for $\sigma^2(t)$. Particularly novel about our work is that the models will be derived directly on the positive half-line as Ornstein-Uhlenbeck type processes. In particular we argue for the use of background driving Lévy processes, rather than Brownian motion, as the forcing mechanisms for these models.

Finally, throughout the paper we shall use the following notation for cumulant and Laplace transforms of a random variate x with associated probability measure ν.

$$\mathbb{C}\left(\zeta x\right) = \mathrm{C}\left(\zeta \ddagger x\right) = \log \mathrm{E}\left(e^{i\zeta x}\right)$$

$$\mathrm{L}(ux) = \mathrm{L}\left(u \ddagger x\right) = \int_0^\infty e^{ux}\nu(\mathrm{d}x)$$

$$\bar{\mathrm{L}}(ux) = \mathrm{L}(-ux) = \int_0^\infty e^{-ux}\nu(\mathrm{d}x)$$

$$\mathrm{K}(ux) = \mathrm{K}\left(u \ddagger x\right) = \log \mathrm{L}(ux)$$

$$\bar{\mathrm{K}}(ux) = \bar{\mathrm{K}}\left(u \ddagger x\right) = \mathrm{K}\left(-u \ddagger x\right)$$

The terms Wiener process and Brownian motion are used synonymously.

2 Basic model type

2.1 Notation and characterisation

Most of the models we shall consider arise by refinement or approximation of the following basic model type. Let w be a Wiener process, let σ be a positive and (strictly) stationary stochastic process with caglad (continuous from left, with limits from the right) sample paths, and suppose that w and σ are adapted to one and the same filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}\}$. By ξ, ω^2 and r we denote, respectively, the mean and, when they exist, the variance and the autocorrelation function of the process $\sigma^2(t)$. Furthermore, let

$$\sigma^{2*}(t) = \int_0^t \sigma^2(u)\mathrm{d}u \tag{1}$$

The process σ^{2*} has continuous sample paths. We may now introduce a process $x^*(t)$ by the stochastic differential equation

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

with solution

$$x^*(t) = \int_0^t \sigma(u)\mathrm{d}w(u) + \beta\sigma^{2*}(t) + \mu t \tag{3}$$

where the integral is defined in the Itô sense. Note that σ and w may be dependent and that, at this stage, no assumptions are made about existence of moments of $x^*(t)$ beyond the second order. The parameter β expresses a possible asymmetry of the process $x^*(t)$, the distribution of $x^*(t)$ being symmetric around μt if $\beta = 0$. Likewise μ reflects the possible drift in the log-price. Typically we will set $\mu = 0$ in our mathematical development for ease of exposition as the drift raises no new issues.

If $\beta = \mu = 0$, i.e.

$$x^*(t) = \int_0^t \sigma(u)\mathrm{d}w(u) \tag{4}$$

then $x^*(t)$ is a continuous local martingale (cf., for instance, Protter (1992, theorem 30, p. 143)) and its quadratic variation is $\sigma^{*2}(t)$, i.e. we have

$$[x^*](t) = \text{p-}\lim_{r\to\infty} \sum\{x^*(t_{i+1}^r) - x^*(t_i^r)\}^2 = \sigma^{2*}(t) \tag{5}$$

for any sequence of partitions $t_0^r = 0 < t_1^r < ... < t_{m_r}^r = t$ with $\sup_i\{t_{i+1}^r - t_i^r\} \to 0$ for $r \to \infty$. Note also that, since $\sigma^{2*}(t)$ is continuous, the two terms in (5) constitute the Doob-Meyer decomposition of $x^*(t)$. The quadratic variation of volatility models has recently been highlighted by Andersen and Bollerslev (1998) in order to approximately estimate integrated volatility in foreign exchange markets. It is clear that for any finite r such a quadratic variation estimator will be an unbiased estimator of the integrated volatility, with its variance falling as $r \to \infty$. However, in practice this style of continuous time model will be a poor approximation to the sample paths of the price process when we look at very fine time intervals, which tend to have some form of discreteness. Hence there is a bias/variance trade-off. One way of studying this problem would be to use the tick-by-tick models we study in a subsection below.

Traditionally there are broadly two forms of models for σ^2 used in the literature which enforce the volatility to be stochastic and positive (positive stochastic processes also appear in the term structure of interest rates literature — see for example, Cox, Ingersoll, and Ross (1985)). The most common is where the logarithm of σ^2 follows a Gaussian Ornstein-Uhlenbeck process

$$\mathrm{d}\log\sigma^2(t) = -\lambda\left\{\log\sigma^2(t) - \mu\right\}\mathrm{d}t + \varsigma\mathrm{d}b(t), \quad \lambda > 0, \tag{6}$$

where $b(t)$ is a Brownian motion. This process has a marginal distribution for instantaneous returns which is a normal mixed with a log-normal, which possess all of its moments. This process for σ^2 has been used by, for example, Wiggins (1987), Chesney and Scott (1989) and Melino and Turnbull (1990) in the context of option pricing, while a discrete time version of this model was put forward by Taylor (1982) and has been studied from an econometric viewpoint by, for example, Harvey, Ruiz, and Shephard (1994), Jacquier, Polson, and Rossi (1994) and Kim, Shephard, and Chib (1998). This process has some advantages as it has a simple strong solution, while

$$r_{\log}(u) = \mathrm{cor}\left\{\log\sigma^2(t+u), \log\sigma^2(t)\right\} = \exp\left(-\lambda|u|\right), \quad \lambda > 0.$$

However, it is not obvious how to work with σ^{2*} in this framework without making discretisation errors.

The other commonly used process is the 'constant elasticity of variance' process

$$\mathrm{d}\sigma^2(t) = -\lambda\left\{\sigma^2(t) - \overline{\sigma}^2\right\}\mathrm{d}t + \gamma\left\{\sigma(t)^2\right\}^d\mathrm{d}b(t), \quad d \geq 1/2.$$

This general structure, which is always covariance (and strictly) stationary if $\lambda > 0$, has been recently highlighted by Meddahi and Renault (1996) who strongly argue that it provides a great deal of tractability in terms of studying temporal aggregation using different information sets. A principle advantage of this style of model is that

$$\begin{aligned}\sigma^2\left\{\Delta(n+1)\right\} &= \overline{\sigma}^2\left(1 - e^{-\lambda\Delta}\right) + e^{-\lambda\Delta}\sigma^2(\Delta n) \\ &\quad + e^{-\lambda\Delta}\gamma\int_{\Delta n}^{\Delta(n+1)} e^{\lambda(u-\Delta n)}\left\{\sigma(t)^2\right\}^d\mathrm{d}b(t)\end{aligned}$$

implying generically that whatever the value of d and γ,

$$r(u) = cor\left\{\sigma^2(t+u), \sigma^2(t)\right\} = \exp(-\lambda|u|).$$

However, in general this process does not possess an analytic strong solution which allows us to work out the marginal density of this process, nor can we simulate from it (or its integrated version) without incurring a discretisation error.

An important special case of the above setup is where $d = 1$, for this is the so called ARCH diffusion of Nelson (1990), which can be motivated as the continuous time limit of the discrete time GARCH process proposed by Bollerslev (1986) and Taylor (1986). Important references in this regard are Drost and Nijman (1993) and Drost and Werker (1996). In this special case σ^2's marginal density is inverse gamma with $(4\lambda/\gamma^2)+2$ degrees of freedom (see also Hurst and Platen (1997)). As $\lambda > 0$, it implies this process always has a marginal law which is inverse gamma with greater than two degrees of freedom. As a result ARCH diffusion models have heavier tailed (Student's t) marginal continuous time returns than the process generated by (6), although it should be noted that all ARCH diffusion models for returns possess their first two moments. A problem with ARCH diffusions is that it is not sufficiently flexible as they have only three parameters: one governs the average level of volatility ($\overline{\sigma}^2$), another the persistence of volatility (λ), the third the magnitude of the persistence of the volatility (γ). The marginal distribution of the σ^2 process is deduced by the interactions of these parameters and is not freely determined. Consequently ARCH diffusions typically do not fit the data very well, although they can produce typical autocorrelation functions for the square returns data, and have to be combined with jumps in order to become empirically reasonable.

Another important special case is where $d = 1/2$, the so called square-root diffusion (see Feller (1951) and Cox, Ingersoll, and Ross (1985) for its use in the context of interest rates). This model has been used in the option pricing literature by Gennotte and Marsh (1993) and Heston (1993). The advantage of the square root process is that the distribution of $\sigma^2\left\{\Delta\left(n+1\right)\right\}|\sigma^2\left(\Delta n\right)$ is a non-central chi-square and has an analytic moment generating function. Its unconditional marginal density is a gamma distribution — which has the disadvantage that it implies an unconditional returns density which does not fit the shape of empirical unconditional distributions of log returns. Further, it is again not clear that the existence of the strong solution for the square root process is enough for us to simulate from the integrated volatility without any discretisation error.

In this paper, in contrast, we are interested in processes where σ^2 satisfies a (nonanticipative) stochastic differential equation

$$\mathrm{d}\sigma^2(t) = \mu\left\{x^*(.), \sigma^2(.), t\right\}\mathrm{d}t + \chi\left\{x^*(.), \sigma^2(.), t\right\}\mathrm{d}z(t) \tag{7}$$

z being a Lévy process. The simplest form of these, which we study in some detail in Section 4, is the Ornstein-Uhlenbeck type

$$\mathrm{d}\sigma^2(t) = -\lambda\sigma^2(t)\mathrm{d}t + \mathrm{d}z(t) \tag{8}$$

where $\lambda > 0$. Here $z(t)$ will be a process with positive, independent and stationary increments. For the applications we have in mind it is important to realize that, in essence, (8) has a stationary solution with a fixed one-dimensional marginal law for σ^2 if and only if that law is selfdecomposable, cf. section 4 below.

Example 2.1 We shall particularly be considering the case where the one-dimensional marginal law of $\sigma^2(t)$ is inverse Gaussian or inverse gamma and where $\sigma^2(t)$ constitutes a process of Ornstein-Uhlenbeck type, or superpositions of such processes. In the case where σ^2 is covariance stationary, then $r(u) = \exp(-\lambda|u|)$. These parametric models imply instantaneous returns will be normal inverse Gaussian or Student's t.

□

Lemma 2.1 If the processes w and σ are independent and $\beta = 0$ then the cumulant transform of the finite dimensional distributions of the process

$$x^*(t) = \int_0^t \sigma(u)\mathrm{d}w(u)$$

are given by, with $\zeta_* = (\zeta_1, ..., \zeta_n)$, $t_* = (t_1, ..., t_n)$, $0 < t_1 < ... < t_n$,

$$\mathrm{C}\{\zeta_* \ddagger x^*(t_*)\} = \bar{\mathrm{K}}\{J/2 \ddagger \sigma^{2*}(t_*)\} \tag{9}$$

where $x^*(t_*) = \{x^*(t_1), ..., x^*(t_n)\}$, $\sigma^{2*}(t_*) = \{\sigma^{2*}(t_1), ..., \sigma^{2*}(t_n)\}$ and

$$J = \sum_{i=1}^{n} (\zeta_i + \zeta_{i+1} + ... + \zeta_n)^2 \left\{\sigma^{2*}(t_i) - \sigma^{2*}(t_{i-1})\right\} \tag{10}$$

In particular, for the one-dimensional marginal distributions we have

$$\mathrm{C}\{\zeta \ddagger x^*(t)\} = \bar{\mathrm{K}}\{\zeta^2/2 \ddagger \sigma^{2*}(t)\}. \tag{11}$$

□

Proof See Appendix.

Recall that the autocorrelation function of the process σ^2 is denoted by r. We now introduce the notation r^* for the cumulative autocorrelation function, i.e.

$$r^*(t) = \int_0^t r(u)\mathrm{d}u \tag{12}$$

and we let

$$R^*(t) = \int_0^t r^*(u)\mathrm{d}u \tag{13}$$

For use below we note that

$$\int_0^t \int_0^t r(u-v)\mathrm{d}u\mathrm{d}v = 2R^*(t) \tag{14}$$

and consequently, assuming that $\sigma^2(t)$ is square integrable,

$$\mathrm{Var}\{\sigma^{2*}(t)\} = 2\omega^2 R^*(t) \tag{15}$$

2.2 Subordination

Typically we think of the $\sigma^2(t)$ process as the instantaneous volatility of $x^*(t)$. However, it has another equally important interpretation, which is that its integral is a random clock for the asset price process. This well known result (see, for example, Conley, Hansen, Luttmer, and Scheinkman (1997)) is stated in Theorem 2.1. It follows directly from the Dubins-Schwarz representation theorem of continuous local martingales as time transformations of Wiener processes (see, for example, Rogers and Williams (1996, p. 64)).

Theorem 2.1 Consider the process

$$x^*(t) = \int_0^t \sigma(u)\mathrm{d}w(u) \tag{16}$$

and recall that

$$\sigma^{2*}(t) = \int_0^t \sigma^2(u)\mathrm{d}u \tag{17}$$

On the filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}\}$ there exists a Brownian motion b such that

$$x^*(t) = b\left\{\sigma^{2*}(t)\right\} \tag{18}$$

for all $t \geq 0$.

□

From the statistical point of view the identity of the processes $x^*(\cdot)$ and $b\{\sigma^{2*}(\cdot)\}$ implies that we can consistently think of x^* both as a stochastic volatility model and as a subordinated model.

There is an extensive literature on modelling economic phenomena in terms of differing clock speeds. Important work in macroeconomics on this topic includes Stock (1988), while in finance Clark (1973), Tauchen and Pitts (1983), Ghysels and Jasiak (1994) and Ané and Geman (1997) study the relationship between calender and financial time as specified by latent and observable variables such as volume. Ané and Geman (1997) is particularly interesting from our viewpoint. They report in Figure 1 a density estimator of the number of trades made on stocks which make up the S&P500 in one minute intervals. It has a shape which appears consistent with a mixture of two inverse Gaussian random variables with suitably chosen parameters. This will be consistent with the parametric volatility models which we develop later in our paper.

Recent work by Muller, Dacorogna, Olsen, Pictet, Schwarz, and Morgenegg (1990) and Guillaume, Dacorogna, Dave, Muller, Olsen, and Pictet (1997) has modelled financial time, which they call 'theta-time', as a function of quote changes and geographical information on closures in markets. One of their important findings is that theta-time can be best thought of as a stochastic process with components with differing rates of persistence

— reflecting the structure of the market which has traders which have different trading horizons. Finally, Engle and Russell (1998), Ghysels, Jasiak, and Gourieroux (1998), Rogers and Zane (1998) and Rydberg and Shephard (1998) model the time between price movements as a stochastic process.

2.3 Aggregational Gaussianity

It follows immediately from the subordination representation (18) that if $\sigma^2(t)$ is ergodic then, as $t \to \infty$,

$$t^{-1}\sigma^{2*}(t) \overset{a.s.}{\to} \xi,$$

implying $t^{-1/2}x^*(t)$ is asymptotically normal with mean 0 and variance ξ (i.e. the log returns tend to normality for long lags). Thus, in particular, no assumptions need to be made about existence of third or fourth moments of the prices or the returns. One implication of this results is that all ARCH diffusion models aggregate to Gaussianity (as they all possess their second moment). The convergence of $t^{-1/2}x^*(t)$ to normality will, however, be slow in case the process $\sigma^2(t)$ exhibits long range dependence. In that case one has that

$$r(t) \sim L(t)t^{-2\bar{H}}$$

as $t \to \infty$ and where $L(t)$ is a slowly varying function and $\bar{H} \in (0, \frac{1}{2})$. Typically then

$$\text{Var}\{t^{-1}\sigma^{2*}(t)\} = 2\omega^2 t^{-2} R^*(t) \sim L(t)t^{-2\bar{H}}$$

which tends to 0 slower than t^{-1}, and very slowly if $\bar{H}$ is close to 0. Hence asymptotic normality will be achieved but this will be a poor approximation unless t is very large.

2.4 Stochastic volatility and tick-by-tick data

Recent studies of tick-by-tick data by Engle and Russell (1998) have demonstrated that it is sometimes useful to think of financial data measured at very high frequencies as being generated by processes with discontinuous sample paths. Instead, focus could be placed on the time between trades (see also Ghysels, Jasiak, and Gourieroux (1998)). In this section we will discuss unifying this style of approach with the continuous time stochastic volatility models discussed above.

To enable us to present general results we will adopt the Rydberg and Shephard (1998) framework for tick-by-tick data. A special case of this model has recently been independently proposed by Rogers and Zane (1998), while an interesting alternative autoregressive style model has been independently studied by Russell and Engle (1998). We model the number of trades $N(t)$ up to time t as a Cox process (which is sometimes called a doubly stochastic point process) with random intensity $\lambda(t) = \lambda\sigma^2(t) > 0$. In general we write τ_i

as the time of the $i-th$ event and so $\tau_{N(t)}$ is the time of the last recorded event when we are standing at calender time t.

Then we model the current price as

$$x_{\lambda}^{*}(t)=\mu\tau_{N(t)}+\beta\sigma^{2*}\left\{\tau_{N(t)}\right\}+\frac{1}{\sqrt{\lambda}}\left\{y_{1}+...+y_{N(t)}\right\}, \tag{19}$$

where the $\{y_i\}$ are $NID(0,1)$ and $\sigma^{2*}(t)=\int_0^t\sigma^2(u)\mathrm{d}u$. We assume the Cox process and the $\{y_i\}$ are all completely independent. This model models prices as being discontinuous in time, jumping with the arrivals from the Cox process. Then we have the following result.

Theorem 2.1 For the price process (19), if $y_i\sim NID(0,1)$, $\sigma^{2*}(t)=\int_0^t\sigma^2(u)\mathrm{d}u$, $N(t)$ is a Cox process with random intensity $\lambda(t)=\lambda\sigma^2(t)>0$, then

$$\lim_{\lambda\uparrow\infty}x_{\lambda}^{*}(\cdot)\stackrel{L}{\rightarrow}x^{*}(\cdot).$$

Proof: Given in the Appendix.

The interpretation of this is that the tick-by-tick model of the price evolution will converge to a general standard stochastic volatility model as the amount of trading gets large and the average tick size becomes small. Of course for infrequently traded markets the models can be substantially different.

We should note that the requirement that the $\{y_i\}$ are $NID(0,1)$ can be relaxed to allow general sequences of $\{y_i\}$ which exhibit a central limit theorem for the sample average. This is particularly useful for it may be helpful in applied work to allow the price innovations to live on a lattice (see Rydberg and Shephard (1998)).

3 Returns

3.1 Various moments

Now consider the model type (3) and assume that w and σ are independent. Let $\Delta>0$ and, for $n=1,2,...$, write

$$\begin{aligned}y_n &= \int_{(n-1)\Delta}^{n\Delta}\left\{\sigma(u)\mathrm{d}w(u)+\beta\sigma^2(u)+\mu\mathrm{d}u\right\}\\ &= \int_{(n-1)\Delta}^{n\Delta}\sigma(u)\mathrm{d}w(u)+\beta\left[\sigma^{2*}(n\Delta)-\sigma^{2*}\left\{(n-1)\Delta\right\}\right]+\mu\Delta \end{aligned} \tag{20}$$

In the financial context the y_n are the log asset returns over time periods of length Δ. We have

$$\mathrm{E}\{y_n\}=\beta\Delta\xi+\mu\Delta \tag{21}$$

$$\mathrm{Var}\{y_n\}=\Delta\xi+2\beta^2\omega^2R^*(\Delta) \tag{22}$$

and

$$\mathrm{E}\left[\{(y_n - \mathrm{E}(y_n)\}^3\right] = 6\beta\omega^2 R^*(\Delta) + \beta^3\omega^3 \int_0^\Delta \int_0^\Delta \int_0^\Delta r(u,v,w)\mathrm{d}u\mathrm{d}v\mathrm{d}w$$

where $r(u,v,w)$ denotes the third order normalized autocumulant function of the σ^2 process. Recall $R^*(.)$ is defined in (13). In the rest of this section we assume that $\beta = \mu = 0$. Then, for $s = 1, 2, ...,$

$$\mathrm{cor}\{y_n, y_{n+s}\} = 0$$

while for the series of squared returns y_n^2 we have

$$\mathrm{E}\{y_n^2\} = \Delta\xi \tag{23}$$

$$\mathrm{Var}\{y_n^2\} = 6\omega^2 R^*(\Delta) + 2\Delta^2\xi^2$$

and

$$\begin{aligned}
\mathrm{Cov}\{y_n^2, y_{n+s}^2\} &= \mathrm{E}\{y_n^2 y_{n+s}^2\} - \mathrm{E}\{y_n^2\}^2 \\
&= \mathrm{E}\left\{\int_{(n-1)\Delta}^{n\Delta} \sigma(u)^2\mathrm{d}u \int_{(n+s-1)\Delta}^{(n+s)\Delta} \sigma(v)^2\mathrm{d}v\right\} - \Delta^2\xi^2 \\
&= \int_{(n-1)\Delta}^{n\Delta} \int_{(n+s-1)\Delta}^{(n+s)\Delta} \{\mathrm{E}\{\sigma(u)^2\sigma(v)^2\} - \xi^2\}\mathrm{d}u\mathrm{d}v \\
&= \int_{(n-1)\Delta}^{n\Delta} \int_{(n+s-1)\Delta}^{(n+s)\Delta} \mathrm{Cov}\{\sigma(u)^2\sigma(v)^2\}\mathrm{d}u\mathrm{d}v \\
&= \omega^2 \int_{(n-1)\Delta}^{n\Delta} \int_{(n+s-1)\Delta}^{(n+s)\Delta} r(v-u)\mathrm{d}u\mathrm{d}v \\
&= \omega^2 \int_0^{\Delta} \int_{s\Delta}^{(s+1)\Delta} r(v-u)\mathrm{d}u\mathrm{d}v \\
&= \omega^2 \int_0^{\Delta} \{r^*((s+1)\Delta - u) - r^*(s\Delta - u)\}\mathrm{d}u
\end{aligned} \tag{24}$$

i.e.

$$\mathrm{Cov}\{y_n^2, y_{n+s}^2\} = \omega^2 \diamondsuit R^*(\Delta s)$$

where

$$\diamondsuit R^*(s) = R^*(s+\Delta) - 2R^*(s) + R^*(s-\Delta) \tag{25}$$

Consequently the correlation is

$$\mathrm{cor}\{y_n^2, y_{n+s}^2\} = q^{-1}\Delta^{-2}\diamondsuit R^*(\Delta s) \tag{26}$$

where

$$q = 6\Delta^{-2}R^*(\Delta) + 2(\xi/\omega)^2 \tag{27}$$

Recall that the spectral density $\bar{f}$ of a time series with autocorrelation function $\rho(s)$, say, is defined by $\bar{f} = (2\pi)^{-1} f$ where

$$f(\psi) = \sum_{s=-\infty}^{\infty} \rho(s) \cos(s\psi)$$

for $\psi \in [0, 2\pi]$. For notational convenience we shall refer to both $\bar{f}$ and f as the spectral density. Thus, using the notation (25), the spectral density of the series y_n^2 may be written as

$$f(\psi) = 1 + 2q^{-1}\Delta^{-2} \sum_{s=1}^{\infty} \Diamond R^*(\Delta s) \cos(s\psi) \tag{28}$$

Example 3.1 Suppose that

$$r(u) = e^{-\lambda|u|} \tag{29}$$

for some $\lambda > 0$. (This is true, in particular, if the process $\sigma^2(t)$ is of Ornstein-Uhlenbeck type, see further in section 4.) Then, for $u > 0$,

$$r^*(u) = \int_0^u r(v)\mathrm{dv} = \lambda^{-1}(1 - e^{-\lambda u}), \tag{30}$$

so

$$R^*(u) = \int_0^u r^*(v)\mathrm{dv} = \lambda^{-2}\left(\lambda u + e^{-\lambda u} - 1\right). \tag{31}$$

This implies

$$\Diamond R^*(\Delta s) = \lambda^{-2}(1 - e^{-\lambda\Delta})^2 e^{-\lambda\Delta(s-1)} \tag{32}$$

which falls exponentially with s. Hence (26) takes the form

$$\mathrm{cor}\{y_n^2, y_{n+s}^2\} = c e^{-\lambda\Delta(s-1)} \tag{33}$$

where

$$c = q^{-1}(\lambda\Delta)^{-2}(1 - e^{-\lambda\Delta})^2 \tag{34}$$

and

$$q = 6(\lambda\Delta)^{-2}(e^{-\lambda\Delta} - 1 + \lambda\Delta) + 2(\xi/\omega)^2 \tag{35}$$

Note that $0 < c < 1$ and that (33) implies that y_n^2 follows a constrained ARMA(1,1) process. This implies y_n is weak GARCH(1,1) in the sense of Drost and Nijman (1993). Andersen and Bollerslev (1997b, p. 137) have fitted GARCH(1,1) models to (seasonally adjusted) equity and exchange rate returns measured using a variety of values of Δ and found that the above aggregation results broadly describe the fit of the various GARCH models.

For $\Delta \to \infty$ we have $q \to 2(\xi/\omega)^2$ and

$$\mathrm{cor}\{y_n^2, y_{n+s}^2\} \sim \{2(\xi/\omega)^2\}^{-1}(\lambda\Delta)^{-2} e^{-\lambda\Delta(s-1)} \tag{36}$$

Thus the effect of $\Delta \to \infty$ is to reduce the constant in front of the exponential and to increase the slope of the damping down in the autocorrelation function. We also note that the spectral density corresponding to (33) is given by

$$\begin{aligned} f(\psi) &= 1 + 2c\frac{\cos\psi - e^{-\lambda\Delta}}{1 - 2e^{-\lambda\Delta}\cos\psi + e^{-2\lambda\Delta}} &(37)\\ &= 1 + ce^{\lambda\Delta}\frac{2e^{-\lambda\Delta}\cos\psi - 2e^{-2\lambda\Delta}}{1 - 2e^{-\lambda\Delta}\cos\psi + e^{-2\lambda\Delta}} \\ &= 1 - ce^{\lambda\Delta} + ce^{\lambda\Delta}\frac{1 - e^{-2\lambda\Delta}}{1 - 2e^{-\lambda\Delta}\cos\psi + e^{-2\lambda\Delta}} &(38)\end{aligned}$$

Letting $\rho = \exp(-\lambda\Delta)$ we may rewrite this as

$$1 - c\rho^{-1} + c\rho^{-1}a(\psi;\rho)$$

where

$$a(\psi;\rho) = \frac{1-\rho^2}{1 - 2\rho\cos\psi + \rho^2}$$

equals the autocorrelation function for an autoregressive process of order 1 with regression coefficient ρ.

□

Example 3.2 More generally, suppose that $r(u)$ is a weighted sum of exponentials

$$r(u) = w_1 \exp\left(-\lambda_1 |u|\right) + ... + w_m \exp\left(-\lambda_m |u|\right) \tag{39}$$

where the w_i are positive and sum to 1. This type of model has the interpretation that the (squared) volatility follows a weighted sum of Ornstein-Uhlenbeck processes with different persistence rates. Hence some of the components of the volatility may represent short term variation in the process while others represent long term movements. Alternative empirical models of this, written directly in discrete time, are discussed by Engle and Lee (1992) and Dacorogna, Muller, Olsen, and Pictet (1997), the latter type termed HARCH models. Under (39) it is, by linearity, straightforward to obtain the corresponding expressions for the right hand sides of (33) and (37). In particular, for $m = 2$ we obtain, using 0 and 1 as indices instead of 1 and 2,

$$\begin{aligned} \mathrm{cor}\{y_n^2, y_{n+s}^2\} &= q^{-1}[w_0(\lambda_0\Delta)^{-2}\{1 - \exp(-\lambda_0\Delta)\}^2 \exp(-\lambda_0\Delta(s-1)) \\ &\quad + w_1(\lambda_1\Delta)^{-2}\{1 - \exp(-\lambda_1\Delta)\}^2 \exp(-\lambda_1\Delta(s-1))] \end{aligned} \tag{40}$$

where

$$\begin{aligned} q &= 6\{w_0(\lambda_0\Delta)^{-2}(\exp(-\lambda_0\Delta) - 1 + \lambda_0\Delta) \\ &\quad + w_1(\lambda_1\Delta)^{-2}(\exp(-\lambda_1\Delta) - 1 + \lambda_1\Delta)\} \\ &\quad + 2(\xi/\omega)^2 \end{aligned} \tag{41}$$

□

Example 3.3 By choosing the weights and damping factors in (39) appropriately and letting $m \to \infty$ it is possible to construct tractable models with long range or quasi long range dependence. For an informal discussion of this see Cox (1991). In particular, there exists (Barndorff-Nielsen (1998b)) a model such that $x^*(t)$ has stationary increments with normal inverse Gaussian laws and for which, for $\lambda > 0$,

$$r(u) = (1 + \lambda |u|)^{-2(1-H)} \tag{42}$$

with $H \in (\frac{1}{2}, 1)$ being the long memory parameter ($\overline{H} = (1 - H)$ from section 2). Note that $r(u)$ may be, for $u > 0$, reexpressed as

$$r(u) = \Gamma \{2(1 - H)\}^{-1} \int_0^\infty x^{2(1-H)-1} e^{-x} e^{-\lambda u x} \mathrm{d}x \tag{43}$$

showing that (42) is a limiting version of (39). For $r(u)$ of the form (42) we have

$$r^*(u) = \lambda^{-1}(2H - 1)^{-1}(1 + \lambda u)^{2H-1} \tag{44}$$

and, writing $v = \lambda^{-2}\{2H(2H - 1)\}^{-1}$

$$R^*(u) = v(1 + \lambda u)^{2H}. \tag{45}$$

This implies

$$\diamondsuit R^*(\Delta s) = v \left\{ (1 + \lambda\Delta (s + 1))^{2H} - 2(1 + \lambda\Delta s)^{2H} + (1 + \lambda\Delta (s - 1))^{2H} \right\},$$

which seems difficult to usefully simplify without approximation.

One way of approximating this expression is to think of it with fixed Δ and letting s increase to infinity. Write then, for $s > 1$,

$$\begin{aligned} \diamondsuit R^*(\Delta s) &= v(1 + \lambda\Delta s)^{2H} \left[\left\{ \frac{1 + \lambda\Delta (s + 1)}{1 + \lambda\Delta s} \right\}^{2H} - 2 + \left\{ \frac{1 + \lambda\Delta (s - 1)}{1 + \lambda\Delta s} \right\}^{2H} \right] \\ &= v(1 + \lambda\Delta s)^{2H} \left[\left\{ 1 + \frac{\lambda\Delta}{1 + \lambda\Delta s} \right\}^{2H} - 1 + \left\{ 1 - \frac{\lambda\Delta}{1 + \lambda\Delta s} \right\}^{2H} - 1 \right]. \end{aligned}$$

But if we use two Taylor expansions of $\left\{ (1 + x)^{2H} - 1 \right\}$ about $x = 0$, the first terms disappear while all the odd terms in the expansion cancel. This leaves for large s

$$\begin{aligned} \diamondsuit R^*(\Delta s) &\sim v(1 + \lambda\Delta s)^{2H} \times 2 \frac{2H(2H - 1)}{2} \left(\frac{\lambda\Delta}{1 + \lambda\Delta s} \right)^2 \\ &= v'(1 + \lambda\Delta s)^{2H-2}. \end{aligned}$$

This implies, for large s,

$$\mathrm{cor}\{y_n^2, y_{n+s}^2\} \sim v''(1 + \lambda\Delta s)^{2H-2}$$

which is directly inherited from the original long memory model for volatility (42). The role of the stochastic volatility and aggregation is to introduce a constant in front of the slowly decaying factor (as with all stochastic volatility models) and to change the decay rate from λ to $\Delta\lambda$. Importantly, this theory suggests that aggregation of returns does not change the long memory parameter at all. That is estimating the long memory parameter for yearly returns should give the same parameter as estimating the model using daily or 10 minute returns. Andersen and Bollerslev (1997a) have already found, using fractionally integrated GARCH models, that this result holds empirically for speculative returns data.

This argument generalises in a number of directions. Clearly we could weight and add short and long memory components by writing, in a simple form,

$$r(u) = w_1(1 + \lambda_1|u|)^{-2(1-H)} + w_2 \exp\left(-\lambda_2|u|\right),$$

where the w_i are positive and sum to one. This again allows analytic calculations of the correlations of the squares and the spectrum.

The framework also allows for multifractal behaviour where

$$r(u) = \sum_{i=1}^{m} w_i(1 + \lambda_i|u|)^{-2(1-H_i)}, \quad H_i \in (\frac{1}{2}, 1), \quad \lambda_i > 0,$$

where the w_i are positive and sum to one. Multifractal behaviour is familiar in the turbulence literature but seems unstudied in finance.

□

Example 3.4 Suppose the log of the volatility follows a Gaussian Ornstein-Uhlenbeck process

$$\mathrm{d}\log\sigma^2(t) = -\lambda\left\{\log\sigma^2(t) - \mu\right\}\mathrm{d}t + \varsigma \mathrm{d}b(t), \quad \lambda > 0,$$

where $b(t)$ is a Wiener process. Writing the autocorrelation function of the log-volatility as $r_{\log}(u) = \exp(-\lambda|u|)$ and $\xi_{\log} = \mathrm{Var}\left\{\log\sigma^2(t)\right\}$, then using the properties of log-normal that

$$\begin{aligned}
\mathrm{cov}\left\{\sigma^2(t+u), \sigma^2(t)\right\} &= \mathrm{E}\exp\left\{\log\sigma^2(t+u) + \log\sigma^2(t)\right\} - \mathrm{E}\left\{\exp\log\sigma^2(t)\right\}^2 \\
&= \exp\left[2E\log\sigma^2(t+u) + \xi_{\log} + \mathrm{cov}\left\{\log\sigma^2(t+u), \log\sigma^2(t)\right\}\right] \\
&\quad - \exp\left\{2E\log\sigma^2(t+u) + \xi_{\log})\right\} \\
&= \mathrm{E}\left\{\sigma^2(t)\right\}^2\left[\exp\left\{\xi_{\log} r_{\log}(u)\right\} - 1\right],
\end{aligned}$$

it follows that

$$r(u) = \mathrm{cor}\left\{\sigma^2(t+u), \sigma^2(t)\right\} = \frac{\exp\left\{\xi_{\log} r_{\log}(u)\right\} - 1}{\exp\left\{\xi_{\log}\right\} - 1}.$$

The implication is that

$$\begin{aligned}
\{\exp\{\xi_{\log}\}-1\} r_{\sigma^2}^*(u) &= \int_0^u [\exp\{\xi_{\log} r_{\log}(v)\}-1]\,\mathrm{d}v \\
&= \sum_{j=1}^{\infty} \frac{(\xi_{\log})^j}{j!} \int_0^u \exp(-\lambda j|v|)\mathrm{d}v \\
&= \sum_{j=1}^{\infty} \frac{(\xi_{\log})^j}{j!} j^{-1} r_{\log}^*(ju).
\end{aligned}$$

As a result

$$R_{\sigma^2}^*(u) = \frac{1}{\{\exp\{\xi_{\log}\}-1\}} \sum_{j=1}^{\infty} \frac{(\xi_{\log})^j}{j!} j^{-2} R_{\log}^*(ju).$$

Of course for this model

$$R_{\log\sigma^2}^*(u) = \lambda^{-2}\left(\lambda u + e^{-\lambda u} - 1\right),$$

and

$$\diamondsuit R_{\log\sigma^2}^*(\Delta s) = \lambda^{-2}\left(1-e^{-\lambda\Delta}\right)^2 e^{-\lambda\Delta(s-1)}.$$

As a result

$$\begin{aligned}
\diamondsuit R_{\sigma^2}^*(\Delta s) &= \frac{1}{\{\exp\{\xi_{\log}\}-1\}} \sum_{j=1}^{\infty} \frac{(\xi_{\log})^j}{j!} j^{-2} \diamondsuit R_{\log}^*(j\Delta s) \\
&= \frac{\lambda^{-2}(1-e^{-\lambda\Delta})^2}{\{\exp\{\xi_{\log}\}-1\}} \sum_{j=1}^{\infty} \frac{(\xi_{\log})^j}{j!} j^{-2} e^{-\lambda\Delta(js-1)} \\
&= \frac{\lambda^{-2}(1-e^{-\lambda\Delta})^2}{\{\exp\{\xi_{\log}\}-1\}} e^{\lambda\Delta} \sum_{j=1}^{\infty} \frac{(\xi_{\log} e^{-\lambda\Delta s})^j}{j!} j^{-2}
\end{aligned}$$

This is useful as

$$\mathrm{cor}\{y_n^2, y_{n+s}^2\} \propto \diamondsuit R_{\sigma^2}^*(\Delta s) \propto \sum_{j=1}^{\infty} \frac{\left(\xi_{\log} e^{-\lambda\Delta s}\right)^j}{j! j^2}.$$

The implication is that when the log-volatility follows a Gaussian Ornstein-Uhlenbeck process the squared aggregated returns do not follow a weak GARCH process, although it could be well approximated by a weighted sum of such processes.

The analysis we have presented above extends to where we add together a number of uncorrelated exponentiated Gaussian Ornstein-Uhlenbeck processes. Finally the argument extends to where the log-volatility is fractional Brownian motion.

□

3.2 Quasi-likelihood estimation of dynamics

Suppose that $\sigma^2(t)$ is covariance stationary, then the $\{y_n^2\}$ are covariance stationary and have a spectrum for the squared returns which is

$$f(\psi) = 1 + \frac{2}{q\Delta^2} \sum_{s=1}^{\infty} \diamondsuit R^*(s\Delta)\cos(s\psi), \quad \text{where} \quad q = 6\Delta^{-2}R^*(\Delta) + 2(\xi/\omega)^2. \tag{46}$$

This is parameterized by $\xi = \mathrm{E}\{\sigma^2(t)\}, \omega = \mathrm{Var}\{\sigma^2(t)\}$ and the autocorrelation function $r(u)$ which in turn determines $R^*(\Delta)$ through (13) and $\diamondsuit R^*(s\Delta)$ through (25). It does not depend on the other more refined distributional properties of the $\sigma^2(t)$ process. Simple analytic expressions for $\diamondsuit R^*(s\Delta)$ were given in Examples 3.1,2,3 for typical autocorrelation functions for the volatility process,

A scaled version of the spectrum can be estimated by the periodogram. Define the sample Fourier coefficients of the squared data

$$\begin{aligned} A(\psi_p) &= \sqrt{\tfrac{2}{T}} \textstyle\sum_{n=0}^{T-1} y_n^2 \cos(n\psi_p), \qquad \psi_p = \tfrac{2\pi p}{T}, \\ B(\psi_p) &= \sqrt{\tfrac{2}{T}} \textstyle\sum_{n=0}^{T-1} y_n^2 \sin(n\psi_p), \qquad p = 0, ..., \left[\tfrac{T}{2}\right]. \end{aligned} \tag{47}$$

Notice $B(0) = 0$, while for T even $B(\psi_{T/2}) = 0$. Then we know that if $\{y_n^2\}$ is covariance stationary and has short memory, then for large samples (see, for example, Brillinger and Rosenblatt (1967)) $A_0 \xrightarrow{p} 2\xi\Delta$ while

$$A(\psi) \xrightarrow{d} N\left\{0, \frac{f(\psi)}{q\Delta}\right\} \quad \text{and} \quad B(\psi) \xrightarrow{d} N\left\{0, \frac{f(\psi)}{q\Delta}\right\}, \quad \psi \in (0, 2\pi). \tag{48}$$

The implication is that the periodogram

$$I(\psi) = A(\psi)^2 + B(\psi)^2 \xrightarrow{d} \chi_1^2 \frac{f(\psi)}{q\Delta}. \tag{49}$$

Further, it is known that asymptotically $I\left\{\psi_{(2)}\right\}$ and $I\left\{\psi_{(1)}\right\}$ are asymptotically uncorrelated if $\psi_{(1)} \neq \psi_{(2)}$.

If we write down a parametric model for $r(u)$ then it is possible to estimate the model directly using the Whittle quasi-likelihood applied to the $\{y_n^2\}$ whatever the value of Δ. As the $\{y_n^2\}$ are not Gaussian such a procedure is going to be inefficient, although the resulting estimator is well known to be consistent and asymptotically normal under short-range dependence models for $r(u)$ (see, for example, Rice (1979)). An elegant exposition of quasi-likelihoods and an application to point processes is given in Chandler (1997).

Typically the quasi-likelihood is written down having estimated $\Delta\xi$ by simply the average of the $\{y_n^2\}$. Hence we will work with the periodogram on the $\left\{y_n^2 - \overline{y^2}\right\}$ and will ignore the zero frequency in the periodogram. The resulting quasi-likelihood is

$$l_q\left\{r, \omega; y_1^2, ..., y_T^2\right\} = const + \frac{[T/2]}{2}\log q - \frac{1}{2}\sum_{p=1}^{[T/2]} \log f(\psi_p) - \frac{q\Delta}{2}\sum_{p=1}^{[T/2]} \frac{I(\psi_p)}{f(\psi_p)}. \tag{50}$$

Of course this method is likely to behave poorly if the volatility process is close to losing its fourth moment. In particular, in such cases the spectrum is likely to be close to being flat, even if there is a great deal of dependence in volatility, and so the quasi-likelihood function will be poorly behaved. In such cases full likelihood methods, based on distributional models of $\sigma^2(t)$, are particularly attractive. This is discussed in Barndorff-Nielsen and Shephard (1998).

3.3 Multivariate versions

A simple q-dimensional version $x^*(t) = \left\{x_1^*(t), ..., x_q^*(t)\right\}$ of the process $x^*(t)$ considered in lemma 2.1 and theorem 2.1 is obtained by letting

$$x_i^*(t) = \beta_i \int_0^t \sigma_0(u)\mathrm{d}w_0(u) + \int_0^t \sigma_i(u)\mathrm{d}w_i(u)$$

Here $\beta_1, ..., \beta_q$ are unknown parameters and $\sigma_0, \sigma_1, ..., \sigma_q$ and $w_0, w_1, ..., w_q$ are $2(q+1)$ processes such that $\sigma_0, \sigma_1, ..., \sigma_q$ are square integrable and stationary while $w_0, w_1, ..., w_q$ are mutually independent Wiener processes. (Note that the present β-s have a meaning quite different from the β in (2).) For simplicity we also assume that $\sigma_0, \sigma_1, ..., \sigma_q$ are independent of $w_0, w_1, ..., w_q$, but no assumption is made at this stage about independence or dependence among $\sigma_0, \sigma_1, ..., \sigma_q$. The process $x^*(t)$ is a continuous q-dimensional local martingale. It constitutes a factor style model with a common, but differently scaled, stochastic volatility model and individual stochastic volatility models for each series. It generalizes straightforwardly to allow for two or more factors. This style of model is in keeping with the latent factor models of Diebold and Nerlove (1989), King, Sentana, and Wadhwani (1994) and Shephard and Pitt (1998).

Writing $\sigma = (\sigma_0, \sigma_1, ..., \sigma_q)$ and

$$\sigma_0^{2*}(t) = \int_0^t \sigma_0^2(u)\mathrm{d}u \quad \text{and} \quad \sigma_i^{2*}(t) = \int_0^t \sigma_i^2(u)\mathrm{d}u$$

$(i = 1, ..., q)$ we find

$$\mathrm{E}\{x_i^*(s)x_i^*(t) \mid \sigma\} = \beta_i^2\sigma_0^{2*}(s) + \sigma_i^{2*}(s)$$

while for $i \neq j$

$$\mathrm{E}\{x_i^*(s)x_j^*(t) \mid \sigma\} = \beta_i\beta_j\sigma_0^{2*}(s)$$

Furthermore, for $0 < t_1 < ... < t_n$ we have

$$\begin{aligned} \zeta_1 x^*(t_1) + ... + \zeta_n x^*(t_n) &= (\zeta_1 + ... + \zeta_n)x^*(t_1) \\ &\quad +(\zeta_2 + ... + \zeta_n)(x^*(t_2) - x^*(t_1)) \\ &\quad +... \\ &\quad +(\zeta_{n-1} + \zeta_n)(x^*(t_n) - x^*(t_{n-1})) \\ &\quad +\zeta_n x^*(t_n) \end{aligned} \tag{51}$$

Hence, letting $\beta = (\beta_1, ..., \beta_q)$ we find

$$\zeta_1 x^*(t_1) + ... + \zeta_m x^*(t_m) \mid \sigma \sim N_q(0, J)$$

where

$$J = \sum_{i=1}^{n} (\zeta_i + \zeta_{i+1} + ... + \zeta_n)^2 \{\Sigma^*(t_i) - \Sigma^*(t_{i-1})\}$$

and

$$\Sigma^*(t) = \beta^\top \beta \sigma_0^{2*}(t) + \mathrm{diag}\{\sigma_i^{2*}(t)\}$$

It follows that

$$\mathrm{C}\{\zeta_* \ddagger x^*(t_*)\} = \bar{K}\{J/2 \ddagger \sigma\}$$

generalizing (11). Now, for $i = 1, ..., q$, let

$$\widetilde{\sigma}_i^2(u) = \beta_i^2 \sigma_0^2(u) + \sigma_i^2(u)$$

and

$$\widetilde{\sigma}_i^{2*}(u) = \int_0^t \widetilde{\sigma}_i^2(u) \mathrm{d}u$$

and define a stopping time τ_{it} by

$$\widetilde{\sigma}_i^{2*}(\tau_{it}) = t$$

In extension of theorem 2.1 we have that

$$b_i(t) = x_i^*(\tau_{it})$$

is a Brownian motion and that

$$x_i^*(t) = b_i \left\{ \widetilde{\sigma}_i^{2*}(t) \right\}$$

is a representation of $x_i^*(t)$ by subordination. However, the law of the q-dimensional process

$$\left[b_1 \left\{ \widetilde{\sigma}_1^{2*}(t) \right\}, ..., b_q \left\{ \widetilde{\sigma}_q^{2*}(t) \right\} \right]$$

is not identical to that of $x^*(t)$. The quadratic variation and covariation processes of $x^*(t)$ take the form

$$[x_i^*](t) = \beta_i^2 \sigma_0^{2*}(t) + \sigma_i^{2*}(t)$$

$$[x_i^*, x_j^*](t) = \beta_i \beta_j \sigma_0^{2*}(t) \quad (i \neq j)$$

and the basic moments are

$$\mathrm{E}\{x_i^*(t)\} = 0$$

$$\mathrm{Var}\{x_i^*(t)\} = t(\beta_i^2 \xi_0 + \xi_i)$$

$$\mathrm{Cov}\{x_i^*(s), x_j^*(t)\} = \min\{s, t\} \beta_i \beta_j \xi_0 \quad (i \neq j)$$

where $\xi_0 = \mathrm{E}\{\sigma_0^2(t)\}$, $\xi_i = \mathrm{E}\{\sigma_i^2(t)\}$ $(i = 1, ..., q)$. Furthermore, for the returns

$$\begin{aligned} y_{in} &= x_i^*(n\Delta) - x_i^*((n-1)\Delta) \\ &= \int_{(n-1)\Delta}^{n\Delta} \{\beta_i \sigma_0(u)\mathrm{d}w_0(u) + \sigma_i(u)\mathrm{d}w_i(u)\} \end{aligned} \tag{52}$$

we obtain

$$\mathrm{E}\{y_{in}\} = 0$$

$$\mathrm{E}\{y_{in}^2\} = \Delta(\beta_i^2\xi_0 + \xi_i)$$

and, assuming independence of the $\sigma_1, ..., \sigma_q$, the basic moments are

$$\mathrm{cor}\{y_{in}, y_{jn}\} = \frac{\beta_i\beta_j}{\{(\beta_i^2 + \xi i/\xi_0)(\beta_j^2 + \xi_j/\xi_0)\}^{1/2}}$$

$$\mathrm{Var}\{y_{in}^2\} = 6\{\beta_i^4\omega_0^2 R_0^*(\Delta) + \omega_i^2 R_i^*(\Delta)\} + 2\Delta^2(\beta_i^2\xi_0 + \xi_i)^2$$

$$\mathrm{Cov}\{y_{in}^2, y_{in+s}^2\} = \beta_i^4\omega_0^2 \diamondsuit R_0^*(\Delta s) + \omega_i^2 \diamondsuit R_i^*(\Delta s)$$

$$\mathrm{Cov}\{y_{in}^2, y_{jn}^2\} = 2\beta_i^2\beta_j^2\{3\omega_0^2 R_0^*(\Delta) + \Delta^2\xi_0^2\}$$

$$\mathrm{Cov}\{y_{in}^2, y_{jn+s}^2\} = \beta_i^2\beta_j^2\omega_0^2 \diamondsuit R_0^*(\Delta s)$$

here $i \neq j$ and $s > 0$.

4 Ornstein-Uhlenbeck type volatilities

4.1 Motivation

In this section we will develop some new models for the volatility process σ^2. These dynamic models will be linear, allowing a great deal of analytic tractability. Our basic approach is to specify a marginal distribution D for σ^2 and then ask if it is possible to construct an Ornstein-Uhlenbeck type process

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

with a marginal distribution D restricted to the positive halfline and where z is a Lévy process? If it is, then we work out the behaviour of the always positive increments for the process. One of the main advantages of this type of model is that we can use the Barndorff-Nielsen (1998a, p. 50-1) result that

$$\sigma^{2*}(t) = \int_0^t \sigma^2(u)\mathrm{d}u = \lambda^{-1}\{z(\lambda t) - \sigma^2(t) + \sigma^2(0)\}, \tag{53}$$

which implies

$$\begin{aligned} \sigma_{n+1}^2 &= \sigma^{2*}\{(n+1)\Delta\} - \sigma^{2*}(n\Delta) \\ &= \lambda^{-1}\left[z\{\lambda\Delta(n+1)\} - z(\lambda\Delta n) + \sigma^2(\Delta n) - \sigma^2\{\Delta(n+1)\}\right] \\ &\stackrel{L}{=} \lambda^{-1}\left\{\left(1 - e^{-\lambda\Delta}\right)\sigma^2(\Delta n) + \int_0^{\lambda\Delta}\left(1 - e^{-\lambda\Delta}e^s\right)\mathrm{d}z(s + \Delta n)\right\}. \end{aligned}$$

to study integrated volatility. In particular if assume $\mathrm{E}\{\mathrm{d}z(s)\} = \varpi \mathrm{d}s$ exists then the minimum mean square forecast of integrated volatility will be

$$\begin{aligned}
\mathrm{E}\left\{\sigma_{n+1}^2|\sigma^2(\Delta n)\right\} &= \lambda^{-1}\left\{\left(1-e^{-\lambda\Delta}\right)\sigma^2(\Delta n) + \varpi\int_0^{\lambda\Delta}\left(1-e^{-\lambda\Delta}e^s\right)\mathrm{d}z(s+\Delta n)\right\} \\
&= \lambda^{-1}\left\{\left(1-e^{-\lambda\Delta}\right)\sigma^2(\Delta n) + \varpi\left(\lambda\Delta - 1 + e^{-\lambda\Delta}\right)\right\} \\
&= \varpi\Delta + \frac{\left(1-e^{-\lambda\Delta}\right)}{\lambda}\left\{\sigma^2(\Delta n) - \varpi\right\}.
\end{aligned}$$

As $\lambda \downarrow 0$ then this forecast converges to $\Delta\sigma^2(\Delta n)$ the most recent instantaneous volatility, while as $\lambda \to \infty$ the forecast becomes $\Delta\varpi$ reflecting the long term average of the process. Likewise, if we assume $\mathrm{E}\{\mathrm{d}z(s) - \varpi\mathrm{d}s\}^r = \mu_r \mathrm{d}s$ exists we have that $\mathrm{E}\left\langle\left[\sigma_{n+1}^2 - \mathrm{E}\left\{\sigma_{n+1}^2|\sigma^2(\Delta n)\right\}\right]^r|\sigma^2(\Delta n)\right\rangle$ equals

$$\begin{cases}
\lambda^{-r}\mu_r\left\{\int_0^{\lambda\Delta}\left(1-e^{-\lambda\Delta}e^s\right)^r\mathrm{d}s\right\}, & r = 2, 3, \\
\lambda^{-4}\left[\mu_4\int_0^{\lambda\Delta}(1-e^{-\lambda\Delta}e^s)^4\mathrm{d}s + 3\mu_2^2\left\{\int_0^{\lambda\Delta}(1-e^{-\lambda\Delta}e^s)^2\mathrm{d}s\right\}^2\right], & r = 4
\end{cases}.$$

Of course another useful result is that if σ^2 is covariance stationary then

$$r(u) = \mathrm{cor}\left\{\sigma^2(t+u), \sigma^2(t)\right\} = \exp(-\lambda|u|).$$

Our paper will detail results for D chosen to be generalized inverse Gaussian. Special cases of this family of densities are: inverse gamma, gamma, inverse Gaussian and positive hyperbolic.

As we noted in the second section, there are diffusion alternatives to the models we are designing which will have the same autocorrelation function of squared aggregated returns and similar unconditional densities. However, those models will not be OU type processes, but instead have complicated non-linear dynamic structures. This hinders our understanding of integrated volatility as such models do not in general obey such simple linear results as (53).

4.2 Background driving Lévy process

We now review some recent work by Barndorff-Nielsen and Shephard (1998) which studies the construction of OU type processes with fixed marginal densities. That paper in turn develops some ideas from Barndorff-Nielsen (1998a). We recall that a (homogeneous) Lévy process z is a stochastic process with independent stationary increments. Without essential restriction it is assumed that $z(0) = 0$ and that z has cadlag (right continuous with left limits) sample paths.

The stationary process σ^2 is of Ornstein-Uhlenbeck type if it is representable as

$$\sigma^2(t) = e^{-\lambda t}\sigma^2(0) + \int_0^t e^{-\lambda(t-s)}\mathrm{d}z(\lambda s) \tag{54}$$

for some Lévy process $z = \{z(t) : t \geq 0\}$ and where λ is a positive number. It then satisfies a stochastic differential equation

$$\mathrm{d}\sigma^2(t) = -\lambda\sigma^2(t)\mathrm{d}t + \mathrm{d}z(\lambda t) \tag{55}$$

The process $z(t)$ is termed the *background driving Lévy process* (BDLP) corresponding to the process $\sigma^2(t)$.

4.3 Existence

In essence, given a one-dimensional distribution D (not necessarily restricted to the positive halfline) there exists a stationary process of Ornstein-Uhlenbeck type whose one-dimensional marginal law is D if and only if D is *selfdecomposable*, i.e. if and only if the characteristic function ϕ of D satisfies $\phi(\zeta) = \phi(c\zeta)\phi_c(\zeta)$ for all $\zeta \in \mathbf{R}$ and all $c \in (0,1)$ and for some family of characteristic functions $\{\phi_c : c \in (0,1)\}$. This restriction does, however, still leave a great flexibility in the choice of D. The precise statement of existence is as follows, cf. Wolfe (1982) and Jurek and Vervaat (1983) (see also Barndorff-Nielsen, Jensen, and Sørensen (1998)).

Theorem 4.1 Let ϕ be the characteristic function of a random variable x. If x is selfdecomposable, i.e. if

$$\phi(\zeta) = \phi(c\zeta)\phi_c(\zeta) \tag{56}$$

for all $\zeta \in \mathbf{R}$ and all $c \in (0,1)$, then there exists a stationary stochastic process $x(t)$ and a Lévy process $z(t)$ such that $x(t) \stackrel{L}{=} x$ and

$$x(t) = e^{-\lambda t}x(0) + \int_0^t e^{-\lambda(t-s)}\mathrm{d}z(\lambda s) \tag{57}$$

for all $\lambda > 0$.

Conversely, if $x(t)$ is a stationary stochastic process and $z(t)$ is a Lévy process such that $x(t) \stackrel{L}{=} x$ and $x(t)$ and $z(t)$ satisfy the equation (57) for all $\lambda > 0$ then x is selfdecomposable.

□

4.4 Lévy densities

Suppose we choose a probability distribution D on the positive halfline which is self-composable. Then there exists a strictly stationary Ornstein-Uhlenbeck process

$$\sigma^2(t) = e^{-\lambda t}\sigma^2(0) + \int_0^t e^{-\lambda(t-s)}\mathrm{d}z(\lambda s). \tag{58}$$

Then the Lévy density w of $z(1)$ is related to the Lévy density u of $\sigma^2(t)$ by the formula

$$w(x) = -u(x) - xu'(x) \tag{59}$$

(this presupposes that u is differentiable) and, letting

$$\overline{Q}(x) = \int_x^\infty w(y)\mathrm{d}y \tag{60}$$

we have, moreover,

$$\overline{Q}(x) = xu(x) \tag{61}$$

cf. Barndorff-Nielsen (1998a).

4.5 Integrals of the BDLP

The fundamental result that

$$\sigma^{2*}(t) = \lambda^{-1}\{z(\lambda t) - \sigma^2(t) + \sigma^2(0)\},$$

implies

$$\sigma^2_{n+1} = \lambda^{-1}\left[z\left\{\lambda\Delta\left(n+1\right)\right\} - z\left(\lambda\Delta n\right) + \sigma^2(\Delta n) - \sigma^2\left\{\Delta\left(n+1\right)\right\}\right].$$

Thus we can simulate sequences of integrated volatilities by simulating the bivariate process of the BDLP and the instantaneous volatility. In particular by defining $\rho = \exp\left(-\lambda\Delta\right)$ and noting

$$\begin{array}{rcl} \sigma^2\left\{\Delta\left(n+1\right)\right\} & = & \rho\sigma^2(\Delta n) + w_{1,n+1}, \\ z\left\{\lambda\Delta\left(n+1\right)\right\} & = & z\left(\lambda\Delta n\right) + w_{2,n+1}. \end{array}$$

where the innovations

$$w_n \overset{L}{=} \left\{ \begin{array}{c} \rho\int_0^{\lambda\Delta} e^s \mathrm{d}z(s) \\ \int_0^{\lambda\Delta} \mathrm{d}z(s) \end{array} \right\}, \tag{62}$$

we have that a convenient representation of integrated volatility. In particular

$$\sigma^2_{n+1} = \lambda^{-1}\left\langle\left[w_{2,n+1} + \sigma^2(\Delta n) - \left\{\rho\sigma^2(\Delta n) + w_{1,n+1}\right\}\right]\right\rangle$$

We can use the following result to simulate the innovations. Suppose f is a positive and integrable function on $[0,\lambda]$ then

$$\int_0^\lambda f(s)\mathrm{d}z(s) \overset{L}{=} \sum_{i=1}^\infty \bar{Q}^{-1}(a_i^*/\lambda) f(\lambda r_i)$$

where $\{a_i^*\}$ and $\{r_i\}$ are two independent sequences of random variables with the r_i independent copies of a uniform random variable r on $[0, 1]$ and $a_1^* < ... < a_i^* < ...$ as the arrival times of a Poisson process with intensity 1. Further,

$$\bar{Q}^{-1}(x) = \inf\{y > 0 : \bar{Q}(y) \leq x\}.$$

This follows from work of Marcus (1987) and Rosinski (1991). A thorough exposition with self-contained proofs is given in Barndorff-Nielsen and Shephard (1998). A seeming difficulty with the infinite sum representation of the integral is that it involves $\bar{Q}^{-1}$ which has to be inverted numerically. However, this is straightforward as it is always convex. In practice this infinite sum has proven to be quickly convergent allowing many hundreds of draws to be made per second on a fast PC.

A convenient feature of these infinite sum representations is that if we fix $\{a_i^*\}$ and $\{r_i\}$ then the simulations we draw from these integrals will be differentiable with respect to parameters which index $\bar{Q}$ and λ. This is extremely convenient for some modern forms of simulation based econometric estimators rely on this type of smoothness assumption. In particular the indirect inference estimator of Smith (1993), which was generalized by Gourieroux, Monfort, and Renault (1993) and recast into EMM by Gallant and Tauchen (1996), can be used on these models due to this feature of the simulator.

In the next subsection we will give some results for the case where D is the generalized inverse Gaussian distribution. For more extensive results see Barndorff-Nielsen and Shephard (1998) who discuss general methods for deriving these types of results.

4.6 Generalized inverse Gaussian distributions

4.6.1 General case

The generalized inverse Gaussian (GIG) marginal law means $\sigma^2(t) \sim GIG(\overline{\lambda}, \delta, \gamma)$ has a density of

$$\frac{(\gamma/\delta)^{\overline{\lambda}}}{2K_{\overline{\lambda}}(\delta\gamma)} x^{\overline{\lambda}-1} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\}, \quad x > 0,$$

where $K_{\overline{\lambda}}$ is a modified Bessel function of the third order. Note that when δ or γ are 0, the norming constant in the formula for the density of the generalized inverse Gaussian distribution has to be interpreted in the limiting sense, using the well-known results that for $x \downarrow 0$ we have

$$K_\lambda(x) \sim \begin{matrix} -\log x & \text{if} \quad \lambda = 0 \\ \Gamma(|\lambda|)2^{|\lambda|-1}x^{-|\lambda|} & \text{if} \quad \lambda \neq 0 \end{matrix}$$

Special cases of the GIG density are: (i) the inverse Gaussian law, where $\overline{\lambda} = -\frac{1}{2}$, (ii) the positive hyperbolic law where $\overline{\lambda} = 1$, (iii) and the inverse chi-squared (inverse gamma) law with ν degrees of freedom which occurs when $\gamma = 0$, $\overline{\lambda} = -v/2$ and $\delta = \sqrt{\nu}$, (iv) gamma, where $\delta = 0$ and $\overline{\lambda} > 0$. Of course if $\sigma^2 \sim GIG(\overline{\lambda}, \delta, \gamma)$ and is independent of $\varepsilon \sim N(0,1)$, then $\sigma\varepsilon$ is the generalized hyperbolic with density

$$\frac{(\gamma/\delta)^{\overline{\lambda}}}{\sqrt{2\pi}\gamma^{\overline{\lambda}-1/2}K_{\overline{\lambda}}(\delta\gamma)} \left(\sqrt{\delta^2 + x^2}\right)^{\overline{\lambda}-1/2} K_{\overline{\lambda}-1/2}\left(\gamma\sqrt{\delta^2 + x^2}\right). \tag{63}$$

Hence a continuous time volatility model built using a volatility model of OU type with GIG marginals will have generalized hyperbolic marginals. Special cases of this include the normal inverse Gaussian distribution, the hyperbolic and the Student t.

It is known that the $GIG(\overline{\lambda}, \delta, \gamma)$ law is self-decomposable (Halgreen (1979)). The following theorem, which is in Barndorff-Nielsen and Shephard (1998), computes the Lévy measure.

Theorem 4.1 The Lévy measure of the generalized inverse Gaussian distribution is absolutely continuous with density

$$u(x) = \left[\delta^2 \int_0^\infty e^{-x\xi} g_{\overline{\lambda}}(2\delta^2\xi)\mathrm{d}\xi + \max\{0, \overline{\lambda}\}\lambda x^{-1}\right] \exp\left(-\gamma^2 x/2\right) \tag{64}$$

where

$$g_\lambda(x) = \left[(\pi^2/2)x\left\{J^2_{|\overline{\lambda}|}(\sqrt{x}) + N^2_{|\overline{\lambda}|}(\sqrt{x})\right\}\right]^{-1}$$

and $J_{\overline{\lambda}}$ and $N_{\overline{\lambda}}$ are Bessel functions.

□

For further information on these Bessel functions see, for instance, Gradstheyn and Ryzhik (1965, pp. 958-71). We have reproduced the proof of Barndorff-Nielsen and Shephard (1998) in the Appendix for the convenience of the reader. We will now discuss various special cases of this result.

4.6.2 Inverse Gaussian distribution

The inverse Gaussian marginal law means $\sigma^2(t) \sim IG(\delta, \gamma)$ whose density is

$$\frac{\delta}{\sqrt{2\pi}} e^{\delta\gamma} x^{-3/2} \exp\left\{-\frac{1}{2}\left(\delta^2 x^{-1} + \gamma^2 x\right)\right\}, \quad x > 0,$$

where the parameters δ and γ satisfy $\delta > 0$ and $\gamma \geq 0$. This is the distribution of the first passage time to level δ of Brownian motion with drift γ and unit diffusion coefficient.

When the law of $\sigma^2(t)$ is $IG(\delta, \gamma)$ we find the upper tail integral (recalling $\overline{Q}(x) = xu(x)$) is

$$\overline{Q}(x) = \frac{\delta}{\sqrt{2\pi}} x^{-1/2} \exp\left(-\frac{1}{2}\gamma^2 x\right).$$

4.6.3 Positive hyperbolic distribution

The density of the positive hyperbolic distribution is

$$\frac{(\gamma/\delta)}{2K_1(\delta\gamma)} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\}, \quad x > 0,$$

where the parameters δ and γ satisfy $\delta > 0$ and $\gamma \geq 0$. The corresponding distribution for $\sigma\varepsilon$, where ε is standard normal and independent of σ, i.e. the hyperbolic, and the associated Lévy process have been suggested for high frequency financial data by Eberlein and Keller (1995), who model the log price as having independent and stationary increments.

When the law of $\sigma^2(t)$ is positive hyperbolic we find the upper tail integral is

$$\overline{Q}(x) = \left\{\delta^2 x \int_0^\infty e^{-x\xi} g_1(2\delta^2\xi)\mathrm{d}\xi + 1\right\} \exp\left(-\gamma^2 x/2\right).$$

4.6.4 Inverse gamma distribution

We now look at Ornstein-Uhlenbeck processes with an inverse gamma marginal law. The implication is that instantaneous returns will be Student's t, which is the same as the marginal distribution of an ARCH diffusion (see section 2) when the degrees of freedom is strictly greater than two. The reciprocal gamma distribution (i.e. the law of the reciprocal of a gamma variate) has density

$$\{2^\nu \Gamma(\nu)\}^{-1} x^{-\nu-1} \exp\left(-\frac{1}{2}x^{-1}\right), \quad \nu > 0.$$

We denote this distribution by $\Gamma^{-1}(\nu, \frac{1}{2})$.

When $\sigma^2(t) \sim \Gamma^{-1}(\nu, \frac{1}{2})$ then the corresponding upper tail integral is

$$\overline{Q}(x) = \frac{1}{2}\int_0^\infty \exp\left(-\frac{1}{2}x\xi\right) g_\nu(\xi)\mathrm{d}\xi \tag{65}$$

where

$$g_\nu(x) = 2\left[\pi^2 x \left\{J_\nu^2(\sqrt{x}) + N_\nu^2(\sqrt{x})\right\}\right]^{-1}$$

and J_ν and N_ν are Bessel functions.

4.6.5 Gamma distribution

We now look at Ornstein-Uhlenbeck processes with a gamma marginal law. This has the density

$$\frac{1}{\Gamma\left(\overline{\lambda}\right)} x^{\overline{\lambda}-1} \exp\left(-x\right), \quad x > 0.$$

This has the corresponding upper tail integral of

$$\overline{Q}(x) = \overline{\lambda} e^{-x},$$

which has the convenient property that it can be analytically inverted. In particular

$$\bar{Q}^{-1}(a_i^*/\lambda) = \max\left\{0, -\log\left(\frac{a_i^*}{\lambda\bar{\lambda}}\right)\right\} = \max\left\{0, -\log a_i^* + \log\left(\lambda\bar{\lambda}\right)\right\}.$$

Discussion The positive hyperbolic and inverse Gaussian densities are reasonably unfamiliar to most econometricians and so it maybe helpful to compare their implications with that implied by the inverse gamma. We can see that the inverse Gaussian and (a special case of) the inverse gamma both arrive from the GIG distribution when $\overline{\lambda} = -1/2$. This is an important special case as it allows instantaneous returns to be Cauchy and so possess no moments. More generally the densities for the implied instantaneous returns are always inside the class of generalized hyperbolic densities (63).

For large returns $x \to \pm\infty$ we have that the density (63) of the generalized hyperbolic distribution behaves as

$$c\left|x\right|^{\bar{\lambda}-1} e^{-\gamma|x|}$$

provided $\gamma > 0$. This follows from the result that as $x \to \infty$,

$$K_{\overline{\lambda}}(x) \sim \sqrt{\frac{\pi}{2}} x^{-1/2} e^{-x},$$

and holds whatever the value of $\overline{\lambda}$. We can have $\gamma = 0$ only if $\bar{\lambda} < 0$ and then the density reduces to the Student law

$$c(\delta^2 + x^2)^{\bar{\lambda}-1/2}.$$

4.7 Alternative modelling approach

Instead of specifying a model for $\sigma^2(t)$ and working out the density for the BDLP, it is possible to go the other way and construct the model through the BDLP. Of course there are constraints on valid BDLPs which must be satisfied. In this subsection we give a simple valid construction which allows easy simulation and analytic results for the implied density of $\sigma^2(t)$.

Suppose the BDLP z has a Lévy density w with tail integral

$$\overline{Q}(x) = cx^{-\varepsilon}(1+x)^{-\beta}\exp\left(-\frac{1}{2}\gamma^2 x\right)$$

where c is a positive constant, $0 \leq \varepsilon < 1$, $0 \leq \beta$, $0 \leq \gamma$ and $\max\{(\beta - 1), \gamma\} > 0$. Then

$$\begin{aligned} w(x) &= c\{\varepsilon x^{-1} + \beta(1+x)^{-1} + \frac{1}{2}\gamma^2\}x^{-\varepsilon}(1+x)^{-\beta}\exp\left(-\frac{1}{2}\gamma^2 x\right) \\ &= c\{\varepsilon x^{-1} + \beta(1+x)^{-1} + \frac{1}{2}\gamma^2\}\overline{Q}(x) \qquad (66) \end{aligned}$$

and, since the Lévy density u of $\sigma^2(t)$ satisfies $xu(x) = \overline{Q}(x)$, we have

$$u(x) = cx^{-1-\varepsilon}(1+x)^{-\beta} \exp\left(-\frac{1}{2}\gamma^2 x\right)$$

The integral $\int_0^\infty \min\{1, x^2\}u(x)\mathrm{d}x$ is finite and hence u is indeed a Lévy density. Furthermore, $xu(x)$ is decreasing so that $\sigma^2(t)$ is selfdecomposable.

Note that for $\varepsilon = \frac{1}{2}$ and $\beta = 0$ we recover the IG law for $\sigma^2(t)$.

If $\gamma = 0$, implying $\beta > 1$, then for the moments of $\sigma^2(t)$ we have

$$\mathrm{E}\left[\left\{\sigma^2(t)\right\}^\nu\right] < \infty \quad \text{if and only if} \quad \nu < \beta + \varepsilon$$

Furthermore, the j-th order cumulant of σ^2 ($j < \beta + \varepsilon$) is

$$\kappa_j = cB(j - \varepsilon, \beta + \varepsilon - j)$$

where $B(x, y)$ denotes the beta function, i.e.

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

It follows that

$$\kappa_j = \frac{j - \varepsilon - 1}{\beta + \varepsilon - j}\kappa_{j-1}$$

which should be helpful for estimating β and ε.

Finally, we are currently working on generalising this approach of specifying models through BDLPs to the multivariate case. This raises a number of interesting technical and modelling issues. If successful this work will allow us to write down general multivariate Ornstein-Uhlenbeck processes with multivariate increments, which are independent and stationary but dependent across series.

5 Conclusion

We have presented five basic results which provide a sounder theoretical underpinning of the use of continuous time stochastic volatility models. In particular we have presented some simple results on moments of SV models and a new class of volatility models. This second result is particularly attractive as it allows the construction of a wide class of positive continuous time processes which may have broad application in economics. We have given general methods for simulating such processes and have discussed broad methods for their construction.

6 Appendix: a lemma and some proofs

Proof of Lemma 2.1 Conditionally on the process σ we have

$$\zeta_1 x^*(t_1) + ... + \zeta_n x^*(t_n) \mid \sigma \sim N\left\{0, \int_0^\infty I(u;\zeta_*,t_*)\sigma^2(u)\mathrm{d}u\right\}$$

where

$$\begin{aligned} I(u;\zeta_*,t_*) &= \left\{\sum_{i=1}^n \zeta_i 1_{[0,t_i]}(u)\right\}^2 \\ &= \sum_{i=1}^n \zeta_i(\zeta_i + 2\zeta_{i+1} + ... + 2\zeta_n)1_{[0,t_i]}(u) \end{aligned}$$

It follows that

$$\begin{aligned} \mathrm{C}\{\zeta_* \ddagger x^*(t_*)\} &= \log \mathrm{E}\left[\exp\{-\frac{1}{2}\int_0^\infty I(u;\zeta_*,t_*)\sigma^2(u)\mathrm{d}u\}\right] \\ &= \log \mathrm{E}\left\{\exp\{-J/2\}\right\} \\ &= \bar{\mathrm{K}}\{J/2 \ddagger \sigma^{2*}(t_*)\} \end{aligned}$$

□

Lemma 6.1 Let $N(t)$ be a Cox process with random intensity $\lambda(t) = \lambda\sigma^2(t) > 0$. We write τ_i as the time of the $i - th$ event and so $\tau_{N(t)}$ is the time of the last recorded event when we are standing at calender time t. Then for $\lambda \to \infty$ we have that $\tau_{N(t)} \xrightarrow{p} t$.
Proof: It suffices to show that for every $\varepsilon > 0$ we have that

$$\Pr(\text{no event in } [t-\varepsilon, t]) \to 0 \text{ as } \lambda \to \infty.$$

Now, via conditioning on the intensity process we find, for every $\delta > 0$,

$$\begin{aligned} \Pr(\text{no event in } [t-\varepsilon,t]) &= E\left\{\Pr(\text{no event in } [t-\varepsilon,t] \,|\lambda(.))\right\} \\ &= E\left[\exp\left\{-\int_{t-\varepsilon}^t \lambda(s)ds\right\}\right] \\ &= E\left[\exp\left\{-\lambda\int_{t-\varepsilon}^t \sigma^2(s)ds\right\}\right] \\ &= E\left[\exp\left\{-\lambda\left\{\sigma^{2*}(t) - \sigma^{2*}(t-\varepsilon)\right\}\right\}\right] \\ &= E\left[1_{\sigma^{2*}(t)-\sigma^{2*}(t-\varepsilon)>\delta}\exp\left\{-\lambda\left\{\sigma^{2*}(t) - \sigma^{2*}(t-\varepsilon)\right\}\right\}\right] \\ &\quad +E\left[1_{\sigma^{2*}(t)-\sigma^{2*}(t-\varepsilon)\le\delta}\exp\left\{-\lambda\left\{\sigma^{2*}(t) - \sigma^{2*}(t-\varepsilon)\right\}\right\}\right] \\ &\le \Pr\left\{\sigma^{2*}(t) - \sigma^{2*}(t-\varepsilon) \le \delta\right\} + e^{-\delta\lambda} \end{aligned}$$

Consequently

$$\lim_{\lambda\uparrow\infty}\sup \Pr(\text{no event in } [t-\varepsilon,t]) \le \Pr\left\{\sigma^{2*}(t) - \sigma^{2*}(t-\varepsilon) \le \delta\right\}$$

and since this holds for all $\delta > 0$ the conclusion of the Lemma follows.
$\square$

Proof of Theorem 2.1 It is helpful to rewrite the process as

$$x^*_\lambda(t) = -\mu\left\{t - \tau_{N(t)}\right\} + \beta\left[\sigma^{2*}(t) - \sigma^{2*}\left\{\tau_{N(t)}\right\}\right] + \beta\sigma^{2*}(t) + \mu t + \frac{1}{\sqrt{\lambda}}\left\{y_1 + ... + y_{N(t)}\right\}.$$

We obtain from Lemma 6.1 and the continuity of $\sigma^{2*}(.)$ that the limiting behaviour in the distribution of $x^*_\lambda(t)$, as $\lambda \to \infty$, is the same as that of

$$\overline{x}^*_\lambda(t) = \mu t + \beta\sigma^{2*}(t) + \frac{1}{\sqrt{\lambda}}\left\{y_1 + ... + y_{N(t)}\right\}.$$

Further, for the characteristic function of $\overline{x}^*_\lambda(t)$ we find that

$$\begin{aligned} E\left[\exp\left\{i\xi\overline{x}^*_\lambda(t)\right\}\right] &= \exp(i\xi t\mu)\, E\left\langle \exp\left\{i\xi\beta\sigma^{2*}(t)\right\} E\exp\left[i\xi\frac{1}{\sqrt{\lambda}}\left\{y_1 + ... + y_{N(t)}\right\}\right] |\lambda(.)\right\rangle \\ &= \exp(i\xi t\mu)\, E\left[\exp\left\{i\xi\beta\sigma^{2*}(t)\right\} E\exp\left\{i\xi\sqrt{\frac{N(t)}{\lambda}}\overline{y}_{N(t)}\right\} |\lambda(.)\right], \end{aligned}$$

where $\overline{y}_{N(t)} = \sqrt{\frac{1}{n}}(y_1 + ... + y_n)$. Trivially, conditionally on $\lambda(.)$ we have that $N(t)/\lambda \overset{a.s.}{\to} \sigma^{2*}(t)$ as $\lambda \to \infty$ and $\overline{y}_{N(t)} \sim N(0,1)$ exactly. Thus

$$\begin{aligned} \lim_{\lambda\uparrow\infty} E\left[\exp\left\{i\xi x^*_\lambda(t)\right\}\right] &= \lim_{\lambda\uparrow\infty} E\left[\exp\left\{i\xi\overline{x}^*_\lambda(t)\right\}\right] \\ &= \lim_{\lambda\uparrow\infty} \exp(i\xi t\mu)\, E\left[\exp\left\{i\xi\left(\beta\sigma^{2*}(t) + \sigma^*(t)u\right)\right\}\right], \end{aligned}$$

where $u \sim N(0,1)$ and is independent of $\sigma^{2*}(t)$. That is the limiting distribution of $x^*_\lambda(t)$ is the same as the law of $x^*(t)$. This argument is easily extended to convergence of all finite dimensional distributions of $x^*_\lambda(t)$, i.e. $x^*_\lambda(\cdot) \overset{L}{\to} x^*(\cdot)$.
$\square$

Proof of Theorem 4.1 Let $z \sim GIG(\overline{\lambda}, \delta, \gamma)$. From Halgreen (1979) we have that if $\overline{\lambda} \leq 0$ then

$$\bar{K}\{\theta \ddagger z\} = -\delta^2 \int_{\gamma^2/2}^{\infty} g_{\overline{\lambda}}\{2\delta^2(y - \gamma^2/2)\}\log(1 - \theta/y)\mathrm{d}y$$

Differentiating both sides of this equation with respect to θ and transforming the integral by setting $\xi = y - \gamma^2/2$ we obtain

$$\begin{aligned} \frac{\partial\bar{K}\{\theta \ddagger z\}}{\partial\theta} &= -\delta^2 \int_0^\infty g_{\overline{\lambda}}\{2\delta^2\xi\}(\gamma^2/2 - \theta + \xi)^{-1}\mathrm{d}\xi \\ &= -\delta^2 \int_0^\infty g_{\overline{\lambda}}\{2\delta^2\xi\} \int_0^\infty \exp\left\{-(\gamma^2/2 - \theta + \xi)x\right\}\mathrm{d}x\mathrm{d}\xi \\ &= -\int_0^\infty e^{\theta x}u(x)\mathrm{d}x \end{aligned}$$

where

$$u(x) = \delta^2 \int_0^\infty e^{-x\xi} g_{\overline{\lambda}}\{2\delta^2\xi\} \mathrm{d}\xi \exp\left(-\gamma^2 x/2\right)$$

For $\overline{\lambda} > 0$ the expression for $q(\xi)$ follows from the convolution formula

$$GIG(\overline{\lambda}, \delta, \gamma) = GIG(-\overline{\lambda}, \delta, \gamma) * \Gamma(\overline{\lambda}, \gamma^2/2)$$

where $\Gamma(\overline{\lambda}, \phi)$ is the gamma distribution with probability density

$$\frac{\phi^{\overline{\lambda}}}{\Gamma(\overline{\lambda})} x^{\overline{\lambda}-1} e^{-\phi x}$$

and corresponding Lévy density

$$\overline{\lambda} x^{-1} e^{-\phi x}$$

□

7 Acknowledgements

This paper represents around a half of a longer piece of work discussed at numerous conferences and departmental seminars under the title 'Continuous time volatility: model construction and inference.' The other half of that original piece of work is now presented in Barndorff-Nielsen and Shephard (1998). We are grateful to a number of people for their comments on earlier revisions. In particular we would like to thank Gary Chamberlain and Jan Pedersen for various discussions, while the comments of Torben Andersen on a summary of 'Continuous time volatility: model construction and inference' were particularly helpful. We also thank Chris Rogers for allowing us to quote his unpublished paper on high frequency data. NS thanks the ESRC for their financial support through the grant 'Estimation via simulation in econometrics,' while we both thank the Centre for Analytical Finance at Aarhus University for financial support.

References

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

Andersen, T. G. and T. Bollerslev (1997b). Intraday periodicity and volatility persistence in financial markets. *J. Empirical Finance 4*, 115–58.

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Unpublished paper: Kellogg School of Management, Northwestern University.

Ané, T. and H. Geman (1997). Transaction clock and asset price dynamics: Clark's model revisited. Unpublished paper.

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

Barndorff-Nielsen, O. E. (1998b). Superposition of Ornstein-Uhlenbeck type processes. In preparation.

Barndorff-Nielsen, O. E., J. L. Jensen, and M. Sørensen (1998). Some stationary processes in discrete and continuous time. *Advances in Applied Probability 30*. Forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (1998). Ornstein-Uhlenbeck volatility models. In preparation: Nuffield College, Oxford.

Bollerslev, T. (1986). Generalised autoregressive conditional heteroskedasticity. *J. Econometrics 51*, 307–327.

Brillinger, D. R. and M. Rosenblatt (1967). Asymptotic theory of estimates of kth order spectra. In *Spectral Analysis of Time Series*, pp. 153–88. New York: Wiley.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton, New Jersey: Princeton University Press.

Chandler, R. E. (1997). A spectral method for estimating parameters in ainfall models. *Bernoulli 3*, 301–322.

Chesney, M. and L. O. Scott (1989). Pricing European options: a comparison of the modified Black-Scholes model and a random variance model. *J. Financial and Qualitative Analysis 24*, 267–84.

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

Conley, T. G., L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman (1997). Short-term interest rates as subordinated diffusions. *Review of Financial Studies 10*, 525–77.

Cox, D. R. (1991). Long-range dependence, non-linearity and time irreversibility. *J. Time Series Analysis 12*, 329–35.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–407.

Dacorogna, M. M., U. A. Muller, R. B. Olsen, and O. V. Pictet (1997). Modelling short term volatility with GARCH and HARCH. In *Nonlinear Modelling of High Frequency Financial Time Series*. Chichester: Wiley.

Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *J. Applied Econometrics 4*, 1–21.

Drost, F. C. and T. E. Nijman (1993). Temporal aggregation of GARCH processes. *Econometrica 61*, 909–927.

Drost, F. C. and B. Werker (1996). Closing the GARCH gap: continuous time GARCH modelling. *J. Econometrics 74*, 31–57.

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*, 281–99.

Engle, R. F. and G. G. J. Lee (1992). A permanent and transitory component model of stock return volatility. Unpublished paper: Department of Economics, University of California at San Diego.

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

Feller, W. (1951). Two singular diffusion problems. *Annals of Mathematics 54*, 173–82.

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

Gennotte, G. and T. A. Marsh (1993). Variations in economic uncertainty and risk premiums on capital assets. *European Economic Review 37*, 1021–1041.

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, Universite de Montreal.

Ghysels, E., J. Jasiak, and C. Gourieroux (1998). Stochastic volatility duration models. Unpublished paper: Pennsylvania State University. Presented at Second international conference on high frequency data in finance, Zurich, Switzerland, April.

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

Gradstheyn, I. and I. W. Ryzhik (1965). *Table of Integrals. Series and Products*. New York: Academic Press.

Guillaume, D. M., M. M. Dacorogna, R. R. Dave, U. A. Muller, R. B. Olsen, and O. V. Pictet (1997). From the bird's eye view to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics 2*, 95–130.

Halgreen, C. (1979). Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions. *Z. Wahrscheinlichkeitstheorie verw. Gebiete 47*, 13–17.

Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies 61*, 247–64.

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. *J. Finance 42*, 281–300.

Hurst, S. R. and E. Platen (1997). The marginal distribution of returns and volatility. In Y. Dodge (Ed.), *L_1-Statistical Procedures and Related Topics, Volume 31 of IMS Lecture Notes – Monograph Series*. Hayward, California: Institute of Mathematical Statistics.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models (with discussion). *J. Business and Economic Statist. 12*, 371–417.

Jurek, Z. and W. Vervaat (1983). An integral representation for selfdecomposable Banach space valued random variables. *Z. Wahrscheinlichkeitstheorie verw. Gebiete 62*, 247–262.

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

King, M., E. Sentana, and S. Wadhwani (1994). Volatility and links between national stock markets. *Econometrica 62*, 901–933.

Marcus, M. B. (1987). ξ-radial processes and random Fourier series. *Mem. Amer. Math. Soc. 368*.

Meddahi, N. and E. Renault (1996). Aggregation and marginalization of GARCH and stochastic volatility models. Unpublished paper: CREST-INSEE.

Melino, A. and S. M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *J. Econometrics 45*, 239–265.

Muller, U. A., M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, and C. Morgenegg (1990). Statistical study of foreign exchange rates. *J. Banking and Finance 14*, 1189–1208.

Nelson, D. B. (1990). ARCH models as diffusion approximations. *J. Econometrics 45*, 7–38.

Protter, P. (1992). *Stochastic Integration and Differential Equations*. New York: Springer.

Rice, J. (1979). On the estimation of the parameters of a power spectrum. *J. Multivariate Analysis 9*, 378–92.

Rogers, L. C. G. and D. Williams (1996). *Diffusions, Markov Processes and Martingales. Volume 2, Ito Calculus.* Chichester: Wiley.

Rogers, L. C. G. and O. Zane (1998). Designing and estimating models of high frequency data. Unpublished paper: Department of Mathematics, University of Bath. Presented at Workshop on Mathematical Finance, University of Bremen, Germany, February.

Rosinski, J. (1991). On a class of infinitely divisible processes represented as mixtures of Gaussian processes. In S. Cambanis, G. Samorodnitsky, and M. S. Taqqu (Eds.), *Stable Processes and Related Topics*, pp. 27–41. Basel: Birkhauser.

Russell, J. R. and R. F. Engle (1998). Econometric analysis of discrete-valued, irregularly-spaced financial transactions data using a new autoregressive conditional multinomial models. Unpublished paper: Graduate School of Business, University of Chicago. Presented at Second international conference on high frequency data in finance, Zurich, Switzerland, April.

Rydberg, T. H. and N. Shephard (1998). A modelling framework for the prices and times of trades made on the NYSE. In preparation: Nuffield College, Oxford. Presented at Workshop on Mathematical Finance, University of Bremen, Germany, February.

Shephard, N. and M. K. Pitt (1998). Analysis of time varying covariances: a factor stochastic volatility approach. Unpublished paper: Nuffield College, Oxford. Invited presentation, Sixth Valencia International Meeting on Bayesian Statistics, Valencia, Spain, May.

Smith, A. A. (1993). Estimating nonlinear time series models using simulated vector autoregressions. *J. Applied Econometrics 8*, S63–S84.

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

Tauchen, G. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*, 485–505.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes — a study of the daily sugar prices 1961-75. In O. D. Anderson (Ed.), *Time Series Analysis: Theory and Practice, 1.* Amsterdam: North-Holland.

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

Wiggins, J. B. (1987). Option values under stochastic volatilities. *J. Financial Economics 19*, 351–372.

Wolfe, S. J. (1982). On a continuous analogue of the stochastic difference equation $x_n = \rho x_{n-1} + b_n$. *Stoch. Proc. Appl. 12*, 301–312.