

## STATIONARY-INCREMENT STUDENT AND VARIANCE-GAMMA PROCESSES

RICHARD FINLAY \* \*\* AND

EUGENE SENETA, \* \*\*\* *University of Sydney*

### Abstract

A continuous-time model with stationary increments for asset price  $\{P_t\}$  is an extension of the symmetric subordinator model of Heyde (1999), and allows for skewness of returns. In the setting of independent variance-gamma-distributed returns the model resembles closely that of Madan, Carr, and Chang (1998). A simple choice of parameters renders  $\{e^{-rt} P_t\}$  a familiar martingale. We then specify the activity time process,  $\{T_t\}$ , for which  $\{T_t - t\}$  is asymptotically self-similar and  $\{\tau_t\}$ , with  $\tau_t = T_t - T_{t-1}$ , is gamma distributed. This results in a skew variance-gamma distribution for each log price increment (return)  $X_t$  and a model for  $\{X_t\}$  which incorporates long-range dependence in squared returns. Our approach mirrors that for the (symmetric) Student process model of Heyde and Leonenko (2005), to which the present work is intended as a complement and a sequel. One intention is to compare, partly on the basis of fitting to data, versions of the general model wherein the returns have either (symmetric)  $t$ -distributions or variance-gamma distributions.

*Keywords:* Skew-stationary process; skewness; martingale; subordinator; volatility; long-range dependence;  $t$ -distribution; variance-gamma distribution; Laplace distribution; self-similarity; heavy tail; minimum  $\chi^2$  estimation; efficient market

2000 Mathematics Subject Classification: Primary 60G10

Secondary 60E99; 62-07; 62M10

### 1. Introduction

The paradigm model for asset price movements, geometric Brownian motion, has afforded the financial world numerous insights into how markets function, as well as spawning a multibillion-dollar global derivatives industry. The model, however, is relatively simple, giving log returns as independent and identically distributed normal random variables, as opposed to typical asset price data, which display the following characteristics (see, for example, Heyde and Liu (2001) and the references therein):

- (i) a leptokurtic distribution (kurtosis greater than three) – higher peaks above the mean, and thicker tails, than a normal distribution;
- (ii) a heteroscedastic time series (time-dependent variance), unlike the geometric Brownian motion model;
- (iii) a long-range dependence structure in squared returns, violating the independence assumption;

---

Received 21 March 2005; revision received 28 March 2006.

\* Postal address: School of Mathematics and Statistics F07, University of Sydney, Sydney, NSW 2006, Australia.

\*\* Email address: richardf@maths.usyd.edu.au

\*\*\* Email address: eseneta@maths.usyd.edu.au

- (iv) occasionally skewed distributions, as opposed to the symmetry of a normal distribution;
- (v) little or no autocorrelation present in returns, at least past one or two lags.

To deal with some of these issues, Heyde (1999) introduced a subordinator model (a model which gives asset prices as geometric Brownian motion driven by some nondecreasing stochastic ‘activity time’ or ‘market time’ process) based on fractal activity time. Heyde and Leonenko (2005) expanded on this model to allow for  $t$ -distributed asset price increments, which incorporate self-similarity (a scaling property) and long-range dependence (LRD).

We consider the variance-gamma (VG) process introduced in Madan and Seneta (1990), show that it can incorporate similar features to the Heyde and Leonenko  $t$ -distributed process, and thus advance the view that it can be regarded as a direct competitor in many respects. In Section 2 we introduce a general subordinator asset price model. In Section 3 we deal with option pricing. In Section 4 we compare the model which allows for skewness with the symmetric model, and discuss volatility and long-range dependence. Section 5 is about the consequences of specification of the distribution of increments over unit time of the activity time process as either gamma or inverse gamma. In Section 6 we discuss the concept of self-similarity and properties that flow from it. In Section 7 we construct a process,  $\{T_t\}$ , for which  $\{T_t - t\}$  is asymptotically self-similar and  $\{\tau_t\}$  has LRD with a marginal gamma distribution. This leads to skew VG-distributed asset price returns which exhibit LRD. In Section 8 we compare the symmetric  $t$ -distribution and the symmetric VG distribution, and fit both to data.

## 2. The general asset price model

Let  $P_t$  be the price of a risky asset at some time  $t \geq 0$ , and assume that  $P_t$  follows subordinated geometric Brownian motion. Specifically, we have

$$P_t = P_0 e^{\mu t + \theta T_t + \sigma B(T_t)},$$

where  $\mu$ ,  $\theta$ , and  $\sigma > 0$  are constants and  $\{B(t)\}$  is standard Brownian motion independent of  $\{T_t\}$ , which is a positive, nondecreasing random process with stationary, but not necessarily independent, increments and  $T_0 = 0$ .

We denote the identically distributed increments over unit time by  $\tau_t = T_t - T_{t-1}$ , and these form a discrete-time stationary process. The corresponding log increments of  $P_t$ , which can be interpreted as continuously compounded returns, are given by

$$X_t = \log P_t - \log P_{t-1} = \mu + \theta \tau_t + \sigma (B(T_t) - B(T_{t-1})). \quad (2.1)$$

Without loss of generality, we take  $E(\tau_t) = 1$ , since any scaling can be absorbed into  $\theta$  and  $\sigma$  as required (assuming that  $E(\tau_t) < \infty$ ).

When  $t$  is a positive integer, as is appropriate when sampling at equally spaced points of time, the joint probability structure of the  $X_t$  is the same as for the model

$$X_t = \mu + \theta \tau_t + \sigma \tau_t^{1/2} W_t, \quad t = 1, 2, \dots, \quad (2.2)$$

where the  $W_t$ ,  $t = 1, 2, \dots$ , are NID(0, 1) random variables, independent also of the process  $\{\tau_t\}$ ,  $t = 1, 2, \dots$ . This may be seen by appealing to joint characteristic functions using (2.1), and first conditioning on  $T_i$ ,  $i = 1, 2, \dots$ .

We now express some properties of the covariance structure of the process  $\{X_t\}$  in terms of properties of the process  $\{\tau_t\}$ , assuming finiteness of moments as necessary. For integer  $k \geq 1$ , we have

$$\text{cov}(X_t, X_{t+k}) = \theta^2 \text{cov}(\tau_t, \tau_{t+k}). \quad (2.3)$$

If  $\theta = 0$  then all such autocorrelations of  $\{X_t\}$  are 0, irrespective of the autocorrelation structure of the process  $\{\tau_t\}$ . Next consider  $\text{cov}(X_t^2, X_{t+k}^2)$  for integer  $k \geq 1$ :

$$\begin{aligned} \text{cov}(X_t^2, X_{t+k}^2) &= \text{cov}((\mu + \theta\tau_t + \sigma\tau_t^{1/2}W_1)^2, (\mu + \theta\tau_{t+k} + \sigma\tau_{t+k}^{1/2}W_2)^2) \\ &= (\sigma^4 + 4\theta^2\mu^2 + 4\theta\mu\sigma^2) \text{cov}(\tau_t, \tau_{t+k}) + \theta^4 \text{cov}(\tau_t^2, \tau_{t+k}^2) \\ &\quad + (\theta^2\sigma^2 + 2\theta^3\mu)(\text{cov}(\tau_t^2, \tau_{t+k}) + \text{cov}(\tau_t, \tau_{t+k}^2)). \end{aligned} \tag{2.4}$$

The above expression reduces to

$$\text{cov}(X_t^2, X_{t+k}^2) = \sigma^4 \text{cov}(\tau_t, \tau_{t+k}) \tag{2.5}$$

for  $\theta = 0$ , irrespective of the size of  $\mu$ . From (2.4) and (2.5) it is clear that structural dependence properties expressed by the autocovariances for  $\{\tau_t\}$  imply those for  $\{X_t^2\}$ .

For a  $\{\tau_t\}$  with dependence structure,  $\{X_t\}$  displays conditional heteroscedasticity (time-dependent variance). Let  $\mathcal{F}_t = \sigma(\{B(T_i), T_i\}, i = 1, 2, \dots, t)$ , which can be thought of as information available up to  $t$ . Then

$$\text{var}(X_t \mid \mathcal{F}_{t-1}) = \theta^2 \text{var}(\tau_t \mid \mathcal{F}_{t-1}) + \sigma^2 \text{E}(\tau_t \mid \mathcal{F}_{t-1}).$$

Under the restricted model with  $\theta = 0$ , the above expression reduces to  $\text{var}(X_t \mid \mathcal{F}_{t-1}) = \sigma^2 \text{E}(\tau_t \mid \mathcal{F}_{t-1})$  (see Heyde and Liu (2001)).

The general form of the continuous-time model for  $P_t$  considered in this section arose (see Seneta (2004)) as an extension of the model in Heyde (1999) and as a generalization of the model which allows for skewness in the same way, but in the setting of independent, VG-distributed increments, in Madan *et al.* (1998). The initial treatment (Madan and Seneta (1990)) was for independent increments with a symmetric VG distribution. The process with independent, symmetric  $t$ -distributed increments originated in Praetz (1972). More historical details may be found in Seneta (2004).

### 3. Option pricing

We may price a European call option in an arbitrage-free way under our model for  $\{P_t\}$ ,  $t \geq 0$ , if the discounted price process,  $\{e^{-rt}P_t\}$ ,  $t \geq 0$ , is a martingale under a suitable choice of parameters of  $\{P_t\}$ , where  $r > 0$  is a fixed but arbitrary small number (the interest rate). This is achieved with the choice  $\mu = r$  and  $\theta = -\sigma^2/2$ : for  $\mathcal{F}_s = \sigma(\{B(u), u \leq T_s\}, \{T_u, u \leq s\})$  and  $\mathcal{F}_* = \sigma(\{B(u), u \leq T_s\}, \{T_u, u \leq s\}, T_t)$  we have

$$\begin{aligned} \text{E}(P_t \mid \mathcal{F}_s) &= P_0 \text{E}(\exp\{rt - \frac{1}{2}\sigma^2(T_t - T_s + T_s) + \sigma(B(T_t) - B(T_s) + B(T_s))\} \mid \mathcal{F}_s) \\ &= P_0 e^{rs - \sigma^2 T_s / 2 + \sigma B(T_s)} e^{r(t-s)} \text{E}(\exp\{-\frac{1}{2}\sigma^2(T_t - T_s) + \sigma(B(T_t) - B(T_s))\} \mid \mathcal{F}_s) \\ &= P_s e^{r(t-s)} \text{E}(\text{E}(\exp\{-\frac{1}{2}\sigma^2(T_t - T_s) + \sigma(B(T_t) - B(T_s))\} \mid \mathcal{F}_*) \mid \mathcal{F}_s) \end{aligned} \tag{3.1}$$

$$\begin{aligned} &= P_s e^{r(t-s)} \text{E}(\exp\{-\frac{1}{2}\sigma^2(T_t - T_s) + \frac{1}{2}\sigma^2(T_t - T_s)\} \mid \mathcal{F}_s) \tag{3.2} \\ &= P_s e^{r(t-s)} \text{E}(1 \mid \mathcal{F}_s) \\ &= P_s e^{r(t-s)}. \end{aligned}$$

Here (3.1) follows from the ‘tower’ (repeated expectation) property of conditional expectation, since  $\mathcal{F}_s \subset \mathcal{F}_*$ , and (3.2) follows from the properties of moment generating functions of normal random variables.

This martingale requires no knowledge of the characteristic function of  $X_t$ , unlike more common constructions such as the mean-correcting martingale (see Schoutens (2003, pp. 79–80).

Heyde and Leonenko (2005, Section 5.2) reported the same pricing equation for the risk-neutral setting as that resulting from the preceding argument, even though they begin with the restricted (symmetric) version of our model for  $P_t$  for the real-world situation (that is, with  $\theta = 0$ ).

#### 4. Comparing the symmetric and skewed models. Volatility

Is it appropriate to use the general model (2.1) for real-world analysis of asset movements? The preceding discussion indicates that in the risk-neutral universe of asset pricing it is clearly appropriate. However, even in the real-world there seems, at least sometimes, to be a skewness in the distribution of returns suggested by the histogram obtained from such data. We need to comment on whether the generalization of the symmetric model which we are proposing is an appropriate way to allow for this skewness, compared with the properties of the symmetric model.

It is readily seen that  $E(B(T_t) \mid \{B(T_i), T_i\}, i = 1, \dots, t - 1) = B(T_{t-1})$ , which may be interpreted (cf. Campbell *et al.* (1997, pp. 23–24)) as the efficient markets hypothesis holding for the symmetric version of model (2.1), since  $\{B(T_i), i = 1, 2, \dots\}$  is a martingale and, hence,  $\{X_t\}$  is a martingale difference sequence (Heyde and Liu (2001) expressed this in a related, continuous-time way). The striking property that  $\text{cov}(X_t, X_{t+k}) = 0$  may consequently be regarded as a manifestation of the efficient markets hypothesis.

For the general model we see from (2.3) and (2.4) that, for  $k \geq 1$ ,

$$\text{corr}(X_t, X_{t+k}) = \frac{\theta^2 \text{cov}(\tau_t, \tau_{t+k})}{\theta^2 \text{var}(\tau_t) + \sigma^2} \leq \theta^2 \frac{\text{corr}(\tau_t, \tau_{t+k})}{\sigma^2} \text{var}(\tau_t) \leq \theta^2 \frac{\text{var}(\tau_t)}{\sigma^2}$$

when  $\text{corr}(\tau_t, \tau_{t+k}) > 0$ . From estimation procedures applied to the asymmetric VG distributional model for returns (to be discussed in a later section), for three sets of observations  $\{X_t\}$ , each of approximate length  $N = 2500$ , evaluation of the right-hand side of this equation resulted in a bound of about 0.04. The smallness of this value provides some support for our model from these data sets.

It could be argued that the efficient markets hypothesis in any case would imply the statistical independence of returns, which would imply that  $\text{cov}(X_t^r, X_{t+k}^r) = 0$ ,  $r = 1, 2, \dots$ . However, there is evidence that for  $r = 2$  this is not the case, and (2.5) is supported for the symmetric model. It may also be argued that real-world markets are not entirely efficient, in support of our general model.

We note that there is an increasing body of evidence, based on data analysis continuing from that of Heyde and Liu (2001), which suggests that, at least when  $\theta = 0$ ,  $E(X_t^4)$  may be infinite. This corresponds to  $\text{var}(\tau_t)$  being infinite, in which case the above argument is not applicable.

In the probabilistic guise of (2.2), when  $\theta = 0$  and  $\{\tau_t\}$ ,  $t = 1, 2, \dots$ , are independent and identically distributed, the model is the familiar random volatility model (see, for example, Taylor (1994)). Thus, in our model, in the case  $\theta = 0$  it is natural to interpret  $\sigma \tau_t^{1/2}$  as the volatility at time  $t$  and, hence, the stationary process  $\{\sigma \tau_t^{1/2}\}$  as describing a stationary random process of stochastically dependent volatilities. This can be seen from an equivalent standpoint by noting that the conditional distribution of the right-hand side of (2.2), given  $V \equiv \tau_t$ , is  $N(\mu, \sigma^2 V)$ . In the general model (2.1), then, where the conditional distribution of the right-hand side is  $N(\mu + \theta V, \sigma^2 V)$ , we may think of the size and direction of the fluctuation in

random mean as being driven by volatility. We recognize the (marginal) distribution described by the above expression as being that of the normal mixed distributions (terminology due to Barndorff-Nielsen *et al.* (1982)).

From the symmetric case, in which  $\theta = 0$ , it is clear that the structure of the distribution of  $V$ , which is a focus of this paper, is central also to the question of the heaviness of tails (leptokurtosis) of the distribution of  $X_t$ . The standard coefficient of kurtosis in this symmetric case is (see Heyde and Liu (2001))  $\kappa = \text{kurtosis}(X_t) = 3(1 + \text{var}(\tau_t)) > 3$ . In the case of general  $\theta$ , if we ignore terms of order  $\theta^2, \theta^3, \dots$ , we find that for small  $\theta$  the coefficients of skewness and kurtosis are approximately

$$\beta = \frac{3\theta \text{var}(\tau_t)}{\sigma}, \quad \kappa = 3(1 + \text{var}(\tau_t)).$$

However, even in the symmetric case, in comparing the effect on kurtosis of competing specific distributions for  $\tau_t$ , we shall need to make in the sequel a more detailed study of tail behaviour.

In the case  $\theta = 0$ , as we have seen, the autocorrelations  $\text{corr}(X_t, X_{t+k})$  are all 0. One of the motivations for investigating this (symmetric) model of Heyde (1999) and Heyde and Liu (2001) was apparently to reflect the fact that asset returns, while displaying a sample autocorrelation function plot characteristic of white noise, no longer do so in sample autocorrelation function plots of squared returns and absolute values of returns. This, however, should be the case if the process  $\{X_t\}$  is indeed a sequence of independent, identically distributed random variables. The corresponding physical interpretation, accordingly, is that LRD of, say, the process  $\{X_t^2\}$  is being empirically manifested, and a plausible model for  $\{X_t\}$  should have the capability of reflecting this long-term dependence as well as the weak autocorrelation structure of the values  $X_t$  themselves.

Indeed, inasmuch as the values  $X_t^2$  are themselves measures of volatility, such a model would impart a desirable dependence structure on the evolution of volatilities.

An investigation of how to express LRD mathematically is contained in a paper, published prior to those of Heyde already mentioned, by Heyde and Yang (1997). An overview of the subject of LRD can be found in Beran (1994, pp. 41–43). We shall say that LRD holds for a stationary process on the integers if it holds in the Allen sense (see, for example, Heyde and Yang (1997)), and denote this by  $\text{LRD}(AV)$ ; this criterion is expressible in terms of the autocovariances  $\{\gamma_k\}$ ,  $k \geq 1$ , and, in the event that these covariances are ultimately nonnegative, amounts to the divergence of their sum,  $\sum_{k=1}^{\infty} \gamma_k$ . Thus, certainly in the symmetric case ( $\theta = 0$ ),  $\text{LRD}(AV)$  of squared returns follows from  $\text{LRD}(AV)$  of  $\{\tau_t\}$  on account of (2.5), which is in keeping with the interpretation of  $\{\sigma \tau_t^{1/2}\}$  as a sequence of (dependent) volatilities.

To specify the distribution of  $X_t$  we need to choose a distribution for  $\tau_t$ . We defer the question of the existence of dependence models in continuous time with specific distributions for  $\tau_t$  and, hence,  $X_t$ , and turn to the distributional questions. In what follows, ‘ $\stackrel{D}{=}$ ’ denotes equality in distribution.

### 5. Specifying the distribution

We say that  $\tau_t \stackrel{D}{=} \Gamma(\alpha, \alpha)$  has a marginal gamma distribution (with parameters chosen so that  $E(\tau_t) = 1$ ), when its probability density function (PDF) is of the form

$$f_{\Gamma}(x) = \frac{\alpha^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\alpha x}, \quad x > 0.$$

In this case  $X_t$  will have a marginal skew VG distribution (see Madan *et al.* (1998)) with PDF

$$f_{VG}(x) = \sqrt{\frac{2}{\pi}} \frac{\alpha^\alpha e^{(x-\mu)\theta/\sigma^2}}{\sigma \Gamma(\alpha)} \left( \frac{|x - \mu|}{\sqrt{\theta^2 + 2\alpha\sigma^2}} \right)^{\alpha-1/2} K_{\alpha-1/2} \left( \frac{|x - \mu| \sqrt{\theta^2 + 2\alpha\sigma^2}}{\sigma^2} \right).$$

The case in which  $\theta = 0$  was introduced in a financial context by Madan and Seneta (1990). This distribution was called VG because  $X_t | V \stackrel{D}{=} N(\mu, \sigma^2 V)$ , where  $V$  has a gamma distribution as above. The VG distribution is sometimes also known as the Bessel  $K$ -function distribution (see Johnson *et al.* (1994, pp. 50–51)). (The function  $K_\eta(\omega)$  is defined below.)

We say that  $\tau_t \stackrel{D}{=} R\Gamma(\beta, \beta - 1)$ , where  $\beta > 1$  has marginal inverse gamma distribution (again with parameters chosen so that  $E(\tau_t) = 1$ ), when its PDF is of the form

$$f_{R\Gamma}(x) = \frac{(\beta - 1)^\beta}{\Gamma(\beta)} x^{-\beta-1} e^{-(\beta-1)/x}, \quad x > 0.$$

In this case  $X_t$  will have a marginal skew  $t$ -distribution with PDF (see Sørensen and Bibby (2003))

$$f_t(x) = \begin{cases} \frac{\Gamma(\beta + \frac{1}{2})}{\sqrt{2\sigma^2(\beta - 1)}\sqrt{\pi}\Gamma(\beta)} \left( 1 + \left( \frac{x - \mu}{\sqrt{2\sigma^2(\beta - 1)}} \right)^2 \right)^{-(\beta+1/2)}, & \theta = 0, \\ \sqrt{\frac{2}{\pi}} \frac{(\beta - 1)^\beta e^{(x-\mu)\theta/\sigma^2}}{\sigma \Gamma(\beta)} \left( \frac{\theta^2}{2(\beta - 1)\sigma^2 + (x - \mu)^2} \right)^{(\beta+1/2)/2} \\ \times K_{\beta+1/2} \left( \frac{|\theta| \sqrt{2(\beta - 1)\sigma^2 + (x - \mu)^2}}{\sigma^2} \right), & \theta \neq 0. \end{cases}$$

The number of degrees of freedom,  $\nu$ , is defined by  $\nu = 2\beta$ , this being a direct generalization of the concept for the classical  $t$ -distribution. In the symmetric case (in which  $\theta = 0$ ), the distribution is the scaled  $t$ -distribution and is well known. It is a slight generalization of the classical Student  $t$ -distribution, and is at the heart of the paper of Heyde and Leonenko (2005). Seneta (2004, Section 6) discussed its parallelism and duality with the (symmetric) VG distribution. Here  $X_t | V \stackrel{D}{=} N(\mu + \theta V, \sigma^2 V)$ , where  $V$  has an inverse gamma distribution as above (and, thus,  $1/V$  has a gamma distribution, but not with unit expectation).

Finally, we say that  $\tau_t \stackrel{D}{=} GIG(\alpha, \beta, \gamma)$  has marginal generalized inverse Gaussian distribution when its PDF is of the form

$$f_{GIG}(x) = \frac{(\gamma/\beta)^{\alpha/2}}{2K_\alpha(\sqrt{\beta\gamma})} x^{\alpha-1} e^{-(\beta/x + \gamma x)/2}, \quad x > 0.$$

Due to the complexity of the generalized inverse Gaussian distribution, we cannot impose the unit expectation constraint for  $\tau_t$  explicitly to reduce the number of parameters.

In the above,  $\alpha \in \mathbb{R}$ ,  $\beta > 0$ , and  $\gamma > 0$ . In this case  $X_t$  will have a marginal skew generalized hyperbolic (GH) distribution (see Barndorff-Nielsen (1977)) with PDF

$$f_{GH}(x) = \left( \frac{\gamma}{\beta} \right)^{\alpha/2} \left( \frac{\beta\sigma^2 + (x - \mu)^2}{\gamma\sigma^2 + \theta^2} \right)^{\alpha/2-1/4} \times K_{\alpha-1/2} \left( \sqrt{\left( \gamma + \frac{\theta^2}{\sigma^2} \right) \left( \beta + \frac{(x - \mu)^2}{\sigma^2} \right)} \right) e^{(x-\mu)\theta/\sigma^2} \left( \sqrt{2\pi\sigma^2} K_\alpha(\sqrt{\beta\gamma}) \right)^{-1}.$$

Wherever it occurs above,  $K_\eta(\omega)$ , given by  $K_\eta(\omega) = \frac{1}{2} \int_0^\infty z^{\eta-1} e^{-(\omega/2)(z+1/z)} dz$  for  $\omega > 0$  and  $\eta \in \mathbb{R}$ , is a modified Bessel function of the third kind (see Erdélyi *et al.* (1954)) with index  $\eta$ . As  $\beta \rightarrow 0$  for  $\alpha > 0$  and  $\gamma \rightarrow 0$  for  $\alpha < 0$ , we obtain as limiting cases the gamma and inverse gamma distributions, respectively, for the distributions of  $\tau_t$ , and correspondingly the (skewed) VG distribution and (skewed)  $t$ -distribution from the skew GH distribution.

In Section 8 we compare, in particular, the properties of the two symmetric distributions. Both have the normal distribution as a limiting case, but are perhaps illuminated in general by noting that the (symmetric)  $t$ -distribution is a generalization of the Cauchy distribution (the case  $\beta = \frac{1}{2}$  in the above), while the symmetric VG distribution is a generalization of the Laplace distribution (also sometimes called the double exponential distribution), which results from letting  $\theta = 0$  and  $\alpha = 1$  in  $f_{VG}(x)$ , since  $K_{1/2}(\omega) = e^{-\omega}(\pi/(2\omega))^{1/2}$  for  $\omega > 0$ .

In the case when the  $\tau_t$  can be taken to be independent as well as exponentially distributed, the  $\{T_t\}$ ,  $t = 1, 2, \dots$ , process may be understood to be the sequence of time points at which events of a Poisson process occur. The consequences of such an attractive interpretation of the  $\{T_t\}$  process accords with the results of the plethora of ‘Laplace process’ studies currently appearing in the literature.

### 6. Self-similarity and long-range dependence

Suppose that  $\{T_t\}$ , with  $T_0 = 0$ , is a positive process in continuous time with stationary increments, and that  $\{T_t - t\}$  is self-similar with Hurst index,  $H$ , satisfying  $\frac{1}{2} < H < 1$ ; that is,  $E(T_t) = t$  and for any  $c > 0$ ,  $T_{ct} - ct \stackrel{D}{=} c^H(T_t - t)$ , meaning that, for  $c = 1/t$ ,  $T_t \stackrel{D}{=} t + t^H(T_1 - 1)$ . Then with  $\sigma_1^2 = \text{var}(T_1)$ , for  $t > s$  we have

$$\sigma_1^2 t^{2H} = \text{var}(T_t) = \sigma_1^2 (t - s)^{2H} + \sigma_1^2 s^{2H} + 2 \text{cov}(T_s, T_t - T_s),$$

since the increments  $\{\tau_t\}$  are stationary and  $T_0 = 0$ . Then, for  $t > s$ ,

$$\text{cov}(T_s, T_t) = \frac{1}{2} \sigma_1^2 (t^{2H} + s^{2H} - (t - s)^{2H})$$

and, thus,

$$\text{cov}(\tau_t, \tau_{t+k}) = \frac{1}{2} \sigma_1^2 ((k + 1)^{2H} + (k - 1)^{2H} - 2k^{2H}).$$

Hence, we find that  $\rho_k$ , the autocorrelation function of  $\{\tau_t\}$ , is given by

$$\rho_k = \frac{1}{2} ((k + 1)^{2H} + (k - 1)^{2H} - 2k^{2H}) \tag{6.1}$$

$$\begin{aligned} &= H(2H - 1)k^{2H-2} + \sum_{i=2}^\infty \binom{2H}{2i} k^{2H-2i} \\ &\geq H(2H - 1)k^{2H-2}. \end{aligned} \tag{6.2}$$

From (6.1) it follows that  $\rho_k$  increases with  $H$ , with  $H = \frac{1}{2}$  resulting in  $\rho_k = 0$  while  $H = 1$  gives  $\rho_k = 1$ . Hence, the value of  $H$  determines the strength of dependence between increments. Equation (6.2) further implies that, for  $H > \frac{1}{2}$ ,

$$\sum_{k=1}^\infty \rho_k \geq H(2H - 1) \sum_{k=1}^\infty k^{2H-2} = \infty,$$

indicating long-range dependence. Hence, if  $\{T_t - t\}$  is self-similar with  $\frac{1}{2} < H < 1$ , then  $\{\tau_t\}$ ,  $t = 1, 2, \dots$ , is LRD(AV).

Now, the continuous-time process,  $\{T_t\}$ , of our general model is additionally supposed to be nondecreasing, and Heyde and Leonenko (2005) noted (between their Equations (5.4) and (5.5)) that such a process cannot be exactly self-similar. However there is growing evidence (see Heyde (1999) and Heyde and Liu (2001)) that at least asymptotic self-similarity of  $T_t - t$ , to a good degree of approximation, is supported by financial data. Via (2.5), this renders the stochastic process  $\{X_t^2\}$  LRD(AV).

Perhaps the most important consequence of the existence of such a (dependent-stationary-increment) process  $\{T_t\}$  is that  $T_t$  can be taken to have the distribution of  $t + t^H(T_1 - 1) = t + t^H(\tau_1 - 1)$ . As was shown in Heyde and Leonenko (2005, Section 5.2), this distribution is needed in evaluating the call price of an option expiring at time  $t$  by taking the expectation of the Black–Scholes pricing formula evaluated at  $T_t$ .

In the next section we focus on constructing such a process  $\{T_t\}$ , the distribution of whose increments is VG.

### 7. Specifying the correlation structure

In this section we show that a continuous-time *correlated* process of the type described in Section 6, having skew VG-distributed returns, actually exists.

We draw on Heyde and Leonenko (2005, Sections 3.3 and 5.1) for many of the ideas used. Their Section 3.3 initiates the successful construction of stationary Student and inverse gamma processes with given (monotonic) correlation function in continuous time. We note that although Heyde and Leonenko (2005) constructed a LRD symmetric  $t$ -model via a self-similar process,  $\{T_t - t\}$ , in which  $\tau_t$  was inverse gamma distributed, the extension to a skew  $t$ -process is trivial. Note, however, that their construction requires the integer number of degrees of freedom of the symmetric  $t$ -distribution to satisfy  $\nu > 4$ . This implies that the fourth absolute moment must be finite. Yet, as we shall see, successful fitting sometimes produces fewer than four degrees of freedom. Recall that we need  $\tau_t \stackrel{D}{=} R\Gamma(\beta, \beta - 1)$  with  $\beta > 1$  and, in fact,  $\beta = \nu/2$ .

The mathematics turns out to be somewhat simpler than in the Heyde and Leonenko (symmetric)  $t$ -distributed  $\{X_t\}$  case, due to the fact that the consequences of having a gamma distribution for  $\tau_t$  are easier to handle than the consequences of having an inverse gamma distribution. It is still necessary to have a positive integer-valued parameter  $\nu$ , such that  $\tau_t \stackrel{D}{=} \Gamma(\alpha, \alpha)$  with  $\alpha = \nu/2$ .

Before proceeding, however, we note that other continuous-time approaches to the construction of stationary LRD (symmetric) Student processes are possible. One such approach was described in Heyde and Leonenko (2005, Section 3.2).

Let  $\{\eta_i(t), t \in \mathbb{R}\}$ ,  $i = 1, \dots, \nu$ , be independent and identically distributed stationary Gaussian processes with zero mean, unit variance, and monotonic correlation function  $\rho_\eta(\tau)$ ,  $\tau \in \mathbb{R}$ . Define the stationary process  $\{\chi_\nu^2(t), t \geq 0\}$  by  $\chi_\nu^2(t) = \frac{1}{2}(\eta_1^2(t) + \dots + \eta_\nu^2(t))$ . Thus, for each  $t$ ,  $\chi_\nu^2(t) \stackrel{D}{=} \Gamma(\nu/2, 1)$  is a scaled  $\chi_\nu^2$  random variable and, so,

$$\begin{aligned} E(\chi_\nu^2(t)) &= \frac{\nu}{2}, & \text{var}(\chi_\nu^2(t)) &= \frac{\nu}{2}, \\ \text{cov}(\chi_\nu^2(t), \chi_\nu^2(t + \tau)) &= \frac{\nu}{4} \text{cov}(\eta_1^2(t), \eta_1^2(t + \tau)) = \frac{\nu}{2} \rho_\eta^2(\tau), \end{aligned} \tag{7.1}$$

implying that

$$\text{corr}(\chi_\nu^2(t), \chi_\nu^2(t + \tau)) = \rho_\eta^2(\tau). \tag{7.2}$$

Here (7.1) follows from the independence of  $\eta_i(t)$  and  $\eta_j(t)$ ,  $i \neq j$ , and the fact that, for  $(X, Y)$  bivariate normal with zero mean, unit variance and correlation coefficient  $\rho$ , and  $Z_1, Z_2 \stackrel{D}{=} \text{NID}(0, 1)$ , we have  $(X, Y) \stackrel{D}{=} (Z_1, \rho Z_1 + \sqrt{1 - \rho^2} Z_2)$ , implying that

$$E(X^2 Y^2) = E(\rho^2 Z_1^4 + 2\rho\sqrt{1 - \rho^2} Z_1^3 Z_2 + (1 - \rho^2) Z_1^2 Z_2^2) = 2\rho^2 + 1. \tag{7.3}$$

If we now take  $T_t = (2/\nu)(\chi_\nu^2(1) + \dots + \chi_\nu^2(t))$ , then we have

$$\tau_t = T_t - T_{t-1} = \frac{2}{\nu} \chi_\nu^2(t) \stackrel{D}{=} \Gamma(\alpha, \alpha)$$

for  $\alpha = \nu/2$ . Hence, we can choose  $\alpha$  to be any integer multiple of  $\frac{1}{2}$ . The autocorrelation function of  $\{\tau_t\}$  is  $\rho_\eta^2(\tau)$ , by (7.2). It follows that if we take  $\rho_\eta(\tau) = (1/(1 + |\tau|^2))^\gamma$ ,  $0 < \gamma < \frac{1}{2}$  (as in Heyde and Leonenko (2005)), then the  $\{\tau_t\}$  process is LRD(AV). Furthermore,

$$\text{var}\left(\sum_{s=1}^n (\eta^2(s) - 1)\right) = 2 \sum_{s=1}^n \sum_{s^*=1}^n \frac{1}{(1 + |s - s^*|^2)^\gamma} \tag{7.4}$$

$$\begin{aligned} &\sim 2 \int_0^n \int_0^n \frac{ds ds^*}{(1 + |s - s^*|^2)^\gamma} \\ &= 2n^{2-2\gamma} \int_0^1 \int_0^1 \frac{du dv}{(1/n^2 + |u - v|^2)^\gamma} \\ &\sim n^{2-2\gamma} c(\gamma) \quad \text{as } n \rightarrow \infty. \end{aligned} \tag{7.5}$$

Here (7.4) follows from (7.3), (7.5) follows by letting  $u = s/n$  and  $v = s^*/n$ , and

$$c(\gamma) = 2 \int_0^1 \int_0^1 \frac{1}{|x - y|^{2\gamma}} dy dx = 4 \int_0^1 \int_0^x \frac{1}{(x - y)^{2\gamma}} dy dx = \frac{2}{(1 - 2\gamma)(1 - \gamma)}$$

for  $0 < \gamma < \frac{1}{2}$ . Hence,  $\text{var}(\sum_{s=1}^n (\eta^2(s) - 1)) \sim c(\gamma)n^{2-2\gamma}$  as  $n \rightarrow \infty$ . Now, from Taqqu (1975, Theorem 6.1, p. 298, and Proposition 6.1, p. 299) we take the following result. By  $[x]$  we denote the integer part of  $x$ .

**Proposition 7.1.** *For a stationary Gaussian sequence,  $X_i$ , with  $E(X_i) = 0$  and  $\text{var}(X_i) = 1$ , such that  $\rho_X(\tau) \sim \tau^{-\gamma} L(\tau)$  as  $\tau \rightarrow \infty$ , with  $0 < \gamma < \frac{1}{2}$  and  $L$  slowly varying,*

$$Z_n(t) = \frac{1}{n^{1-\gamma}} \sum_{i=1}^{[nt]} (X_i^2 - 1)$$

converges weakly, as  $n \rightarrow \infty$ , to a process,  $R(t)$ , which has the following properties, for  $H = 1 - \gamma$ :

1.  $R(t)$  has strictly stationary increments;
2.  $R(t)$  is  $H$ -self-similar;
3.  $E(R(t)) = 0$  and  $E(|R(t)|^\alpha) < \infty$  for  $\alpha \leq 1/H$ ;
4.  $R(t)$  is separable and almost surely continuous.

Furthermore, the characteristic function of  $R(t)$  admits the representation

$$\phi(u) = \exp\left\{\frac{1}{2} \sum_{k=2}^{\infty} \frac{(2iu)^k}{(k+1)!} S_{\gamma}\right\}, \tag{7.6}$$

$$S_{\gamma} = \int_0^1 \cdots \int_0^1 \frac{dx_1 \cdots dx_k}{(|x_1 - x_2| |x_2 - x_3| \cdots |x_{k-1} - x_k| |x_k - x_1|)^{\gamma}},$$

which is valid for small values of  $|u|$ .

A direct result of Proposition 7.1 is that

$$\frac{1}{n^{1-\gamma}} \sum_{s=1}^{\lfloor nt \rfloor} (\eta^2(s) - 1) \xrightarrow{w} R(t) \quad \text{as } n \rightarrow \infty,$$

where  $R(t)$  is the so-called ( $H$ -self-similar, strictly stationary-increment) Rosenblatt process, with  $H = 1 - \gamma \in (\frac{1}{2}, 1)$ . Here ‘ $\xrightarrow{w}$ ’ denotes the weak convergence of one stochastic process to another, which implies the convergence of finite-dimensional distributions. Consequently, as  $n \rightarrow \infty$  we have

$$\begin{aligned} \frac{1}{n^{1-\gamma}} (T_{\lfloor nt \rfloor} - \lfloor nt \rfloor) &= \frac{2}{n^{1-\gamma\nu}} \sum_{s=1}^{\lfloor nt \rfloor} \left( \chi_{\nu}^2(s) - \frac{\nu}{2} \right) \\ &= \frac{1}{\nu} \sum_{i=1}^{\nu} \left( \frac{1}{n^{1-\gamma}} \sum_{s=1}^{\lfloor nt \rfloor} (\eta_{\nu}^2(s) - 1) \right) \\ &\xrightarrow{w} \frac{1}{\nu} \sum_{i=1}^{\nu} R_i(t), \end{aligned}$$

where, for  $i = 1, \dots, \nu$ ,  $R_i(t)$  are independent Rosenblatt processes.

Hence, we have constructed a process  $\{\tau_i\}$  such that  $\{T_i - t\}$  (appropriately normed) is asymptotically  $H$ -self-similar, being weakly convergent to the sum of  $\nu$  independent Rosenblatt processes.

As a final point, we note that  $R(t)$  is symmetric, which follows from (7.6) since the characteristic function of  $R(t)$  satisfies

$$\phi(-u) = \exp\left\{\frac{1}{2} \sum_{k=2}^{\infty} \frac{(-2iu)^k}{(k+1)!} S_{\gamma}\right\} = \overline{\phi(u)}.$$

Hence,  $R(t) \stackrel{d}{=} -R(t)$  and, so, the limiting distributions of the self-similar  $\chi^2$ -process constructed here (that is,  $(1/\nu) \sum_{i=1}^{\nu} R_i(t)$ ) and the self-similar inverse  $\chi^2$ -process constructed in Heyde and Leonenko (2005) (that is,  $-(1/\nu) \sum_{i=1}^{\nu} R_i(t)$ ) do in fact coincide as  $n \rightarrow \infty$ .

Therefore, it seems that circumstances exist in which there is little to distinguish between the skew VG process and the skew  $t$ -process, as the increment processes that drive the two models, at least for the asymptotically self-similar examples that we constructed, converge to a common process. This is supported by fits of real data, as is demonstrated to a limited extent in the next section.

### 8. Comparing symmetric increment distributions

An important question for applications, and for us to address briefly here, is as follows: What are the relative strengths of the competing models for the increment distribution, the symmetric scaled  $t$ -distribution, and the symmetric VG distribution, the second and third of whose PDFs are respectively given by  $f_t(x)$  and  $f_{VG}(x)$  when  $\theta = 0$ ? Such a question for skew distributions in general, and in greater statistical detail for symmetric distributions, will be addressed elsewhere in a statistical setting.

Possibly the most important question is that of tail weight. To be precise, for  $\theta = \mu = 0$ , we have

$$P_t(|X_t| > x) \sim \text{const}(\beta, \sigma) x^{-2\beta},$$

$$P_{VG}(|X_t| > x) \sim \text{const}(\alpha, \sigma) x^{\alpha-1} e^{-x\sqrt{2\alpha/\sigma^2}}, \quad x \rightarrow \infty,$$

where  $\text{const}(\cdot, \cdot)$  denotes a constant depending on the stated arguments.

Thus, statistical techniques which indicate the nonexistence of higher moments (as in Heyde and Liu (2001)) may incline the potential user towards the  $t$ -distribution in preference to the VG distribution for which all moments exist.

However, Heyde and Kou (2004) demonstrated that in practice as many as 100 000 readings may be necessary to distinguish between exponential (Laplace distribution-type) and power law ( $t$ -distribution-type) tails. Now, as we have seen, the Laplace distribution is a special case of the symmetric VG distribution. Allowing for a power law modification of exponential decay as in the symmetric VG case thus makes distinction even more difficult.

For  $\alpha = 1$  (in which case  $X_t$  has a Laplace distribution) the above relation is exact and not just asymptotic. For  $\alpha > 1$  the PDF of the VG distribution is smooth and increasingly bell-shaped for large  $\alpha$ . For  $\frac{1}{2} < \alpha \leq 1$  the PDF is cusped at the origin, while for  $0 < \alpha \leq \frac{1}{2}$  it is unbounded at the origin. Decreasing the value of  $\alpha$  from a value greater than 1 has the effect of increasing probability near the origin and in the tails, at the expense of probability in the middle range (see Madan and Seneta (1990)). This may be seen clearly by numerical investigation of quantiles, similar to the investigation of Heyde and Kou (2004) after standardization (by the standard deviation). Thus, a symmetric VG distribution would seem to provide a desirable model in several respects, not least because one may reasonably expect mean-corrected returns of a not excessively volatile asset to be concentrated round 0, with occasional large fluctuations.

As a concrete application, we fit the normal distribution, the symmetric VG model, the scaled  $t$ -distribution, and the symmetric GH distribution to three data sets obtained from the wire service Bloomberg. They are the daily closing values of the Australian dollar–US dollar exchange rate between 14 October 1992 and 13 December 2003 ( $N = 2864$ ), of the Standard & Poor’s 500 Index (US market index) between 1 March 1994 and 10 December 2003 ( $N = 2465$ ), and the daily closing share price of News Corporation between 2 March 1994 and 2 August 2003 ( $N = 2473$ ).

We fit the models using a minimum  $\chi^2$  procedure. Our minimum  $\chi^2$  fitting involves numerically minimizing a statistic of the form

$$\chi_{\text{data}}^2(\mu, \sigma, \alpha, \beta, \gamma) = \sum_{i=1}^{100} \frac{(O_i^* - E_i^*)^2}{E_i^*} \tag{8.1}$$

with respect to the parameter values, where in our case  $i = 1, \dots, 100$ ,  $O_i^*$  is the total number of observations in the data set divided by 100, and  $E_i^*$  is the expected number of observations falling within the  $i$ th 1% sample quantile band. This procedure is evidently a weighted least-squares data fit, and has the advantage over a modified method of moments (see Seneta (2004)

TABLE 1: Minimum  $\chi^2$  values for  $\theta = 0$ .

Model	A\$–US\$	S&P 500	News Corp
Normal	541.8	348.9	409.9
$t$	364.6	134.4	195.8
VG	343.8	130.9	194.6
GH	343.8	127.9	194.6

and Tjetjep and Seneta (2006)) applied to the same situation, in that the estimated parameters result in a very good general fit to the data, as measured quantitatively by the minimized  $\chi^2$  goodness-of-fit value (8.1).

Hsiao (1985) showed that in conventional large-sample applications of the minimum  $\chi^2$  procedure – assuming, in particular, independence – minimum  $\chi^2$  estimates have asymptotic covariances equal, to  $\mathcal{O}(n^{-1})$ , to those of maximum likelihood estimation. The ergodic hypothesis ensures that our minimum  $\chi^2$  procedure, which amounts to minimizing a distance (in the sense of (8.1)) between two histograms, can sensibly be used even when data are not independent.

To estimate the model parameters, we used MATLAB<sup>®</sup> 6 to numerically minimize (8.1) as a function of those parameters. We omit the estimated parameters here, only noting that the estimated GH models for the Australian dollar–US dollar exchange rate and News Corporation data are in effect actually VG models (i.e. the GH model is close to its limiting VG form). We do, however, present the termination values of the MATLAB algorithm in Table 1. These correspond to  $\chi^2$  goodness-of-fit statistics for each model.

The corresponding parameter estimates for  $(\beta, \alpha)$  for the  $t$ -distribution and VG distribution were respectively (2.43, 2.24), (1.54, 1.51), and (2.04, 1.90). Note that for the Standard & Poor's 500 Index data, the number of degrees of freedom for the  $t$ -distribution is  $\nu = 2 \times 1.54 < 4$ .

The main point to note is that, while the GH distribution, having one extra parameter and being a generalization of both the VG distribution and the  $t$ -distribution, is overall the slightly superior model, at least as measured by the empirical size of the minimum  $\chi^2$  statistic, the VG distribution in fact equals the GH distribution for the A\$–US\$ exchange rate and News Corporation data. This follows from the fact (not detailed here) that the fitted GH models for the A\$–US\$ exchange rate and News Corporation data are in fact nearly identical to the fitted VG models.

In conclusion, we make two further points. First, note that the figures for the normal (geometric Brownian motion) model are often over twice as large as the next worst model, indicating that we do indeed gain by considering subordinator models. Finally, given that the GH distribution improves only marginally on the VG distribution and the  $t$ -distribution, if at all, we postulate that in most cases it is sufficient to consider only the latter two models. The authors concede that, although the VG distribution seems marginally superior here, modelling of rare events may be better captured by the  $t$ -distribution.

### Acknowledgements

The authors are grateful to C. C. Heyde and N. N. Leonenko for access to and discussions of work which has now appeared as Heyde and Leonenko (2005). ES thanks Leonard Scott for stimulating discussions at the University of Virginia in October 2005, and the Department of Mathematics there for its hospitality.

## References

- BARNDORFF-NIELSEN, O. E. (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Z. Wahrscheinlichkeitsth.* **38**, 309–311.
- BARNDORFF-NIELSEN, O. E., KENT, J. AND SØRENSEN, M. (1982). Normal variance-mean mixtures and  $z$  distributions. *Internat. Statist. Rev.* **50**, 145–159.
- BERAN, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall, New York.
- CAMPBELL, J. Y., LO, A. W. AND MACKINLAY, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- ERDÉLYI, A., OBERHETTINGER, M. F. AND TRICOMI, F. G. (1954). *Tables of Integral Transforms. Based, in part, on Notes Left by Harry Bateman and Compiled by the Staff of the Bateman Manuscript Project*, Vol. 2. McGraw-Hill, New York.
- HEYDE, C. C. (1999). A risky asset model with strong dependence through fractal activity time. *J. Appl. Prob.* **36**, 1234–1239.
- HEYDE, C. C. AND KOU, S. G. (2004). On the controversy over tailweight of distributions. *Operat. Res. Lett.* **32**, 399–408.
- HEYDE, C. C. AND LEONENKO, N. N. (2005). Student processes. *Adv. Appl. Prob.* **37**, 342–365.
- HEYDE, C. C. AND LIU, S. (2001). Empirical realities for a minimal description risky asset model. The need for fractal features. *J. Korean Math. Soc.* **38**, 1047–1059.
- HEYDE, C. C. AND YANG, Y. (1997). On defining long-range dependence. *J. Appl. Prob.* **34**, 939–944.
- HSIAO, C. (1985). Minimum chi-square. In *Encyclopedia of Statistical Sciences*, Vol. 5, eds S. Kotz and N. L. Johnson, John Wiley, New York, pp. 518–522.
- JOHNSON, N. L., KOTZ, S. AND BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd edn. John Wiley, New York.
- MADAN, D. B. AND SENETA, E. (1990). The variance gamma (V.G.) model for share market returns. *J. Business* **63**, 511–524.
- MADAN, D. B., CARR, P. P. AND CHANG, E. C. (1998). The variance gamma process and option pricing. *Europ. Finance Rev.* **2**, 79–105.
- PRAETZ, P. D. (1972). The distribution of share price changes. *J. Business* **45**, 49–55.
- SCHOUTENS, W. (2003). *Lévy Processes in Finance*. John Wiley, London.
- SENETA, E. (2004). Fitting the variance-gamma model to financial data. In *Stochastic Methods and Their Applications* (J. Appl. Prob. Spec. Vol. **41A**), eds J. Gani and E. Seneta, Applied Probability Trust, Sheffield, pp. 177–187.
- SØRENSEN, M. AND BIBBY, M. (2003). Hyperbolic processes in finance. In *Handbook of Heavy Tailed Distributions in Finance*, ed. S. T. Rachev, Elsevier, Amsterdam, pp. 211–248.
- TAQQU, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitsth.* **31**, 287–302.
- TAYLOR, S. J. (1994). Modeling stochastic volatility: a review and comparative study. *Math. Finance* **4**, 183–204.
- TJETJEP, A. AND SENETA, E. (2006). Skewed normal variance-mean models for asset pricing and the method of moments. *Internat. Statist. Rev.* **74**, 109–126.