L_1-*Statistical Procedures and Related Topics*
IMS Lecture Notes – Monograph Series (1997) Volume 31

The marginal distributions of returns and volatility

Simon R. Hurst and Eckhard Platen

The Australian National University, Canberra, Australia.

Abstract: The aim of the paper is to identify a distribution that best fits empirical asset returns. By assuming a normal-variance mixture distribution for the returns, the distribution of volatility is implied.

Key words: Asset price model, leptokurtosis, quadratic variation process, mixture distributions, Student t distribution.

AMS subject classification: Primary 90A12; secondary 60J30, 62P20.

1 Introduction

The modelling of the stochastic process followed by the price of an asset is an important part of financial analysis. An understanding of this process is the first step to the pricing of derivative securities and general risk management. It is therefore important to identify a model for asset price processes which is consistent with their major empirical properties, such as heavy tailed return distributions, volatility clustering, long memory and persistence after volatility shocks. Previous approaches have typically concentrated on specific models, e.g. ARCH, and not succeeded so far to jointly model all of the major empirical properties. To attack this problem systematically we first study the marginal distributions of returns and volatility for market price indexes. Only after that we feel a substantial effort can be made to identify further evolutionary properties of volatility and asset price processes.

In this paper we compare various distributions to model the leptokurtic marginal distribution of asset returns. The distributions considered are: the normal (or Gaussian); the stable; the normal-lognormal mixture of Clark (1973); the generalised hyperbolic which include the Student t, the normal-inverse Gaussian mixture of Barndorff-Nielsen (1995), the hyperbolic of Eberlein and Keller (1995) and Küchler et al. (1995) and the variance-gamma (normal-gamma mixture) of Madan and Seneta (1990). These distributions are all mixtures of the normal distribution and differ only by their mixing volatility distributions.

It is crucial for any kind of serious risk analysis and management to emphasise the importance of correctly modelling the tail probabilities of returns. This is the reason why we will focus our comparative analysis on the identification of typical tail properties of index returns. Tests are performed on price indexes to directly determine a best marginal distribution for returns from the above mentioned alternatives. This distribution indirectly determines a best marginal distribution for the volatility. The best distribution for the index returns, with respect to the likelihood ratio test, turns out to be the Student t distribution. This distribution implies an inverted gamma distribution for the squared volatility. The Anderson and Darling (1952) test is also used to identify specifically the tail properties and additionally supports this result.

2 The class of generalised lognormal asset price models

Consider the *class of generalised lognormal models* for the asset price process $\mathbf{S} = \{S(t),\ t \geqslant 0\}$ given by the Ito process with stochastic differential equation

$$dS(t) = \mu(t)\, S(t)\, dt + \sigma(t)\, S(t)\, dW(t), \tag{1}$$

for $0 \leqslant t_0 \leqslant t < \infty$. The stochastic process $\mathbf{W} = \{W(t),\ t \geqslant 0\}$ represents the noise process which is assumed to be a *standard Wiener process* on a filtered probability space $(\Omega, \mathcal{F}, \underline{\mathcal{F}} = \{\mathcal{F}_t\}_{t \geqslant 0}, P)$ fulfilling the usual conditions. We also have the *drift process* $\mu = \{\mu(t),\ t \geqslant 0\}$ and the non-negative *volatility process* $\sigma = \{\sigma(t),\ t \geqslant 0\}$. These two processes may be constant, deterministic time-dependent or stochastic. In general we assume that they are $\underline{\mathcal{F}}$-adapted, right-continuous with left hand limits and that a unique, strong solution for (1) exists. The explicit solution of (1) for the asset price $\mathbf{S}$ has the form

$$S(t) = S(t_0) \exp\left\{ \int_{t_0}^{t} (\mu(u) - \frac{1}{2}\, \sigma(u)^2)\, du + \int_{t_0}^{t} \sigma(u)\, dW(u) \right\}, \tag{2}$$

for $0 \leqslant t_0 \leqslant t < \infty$. Throughout this paper we will stay within the class of generalised lognormal asset price models.

Let us define the returns of the asset price process $\mathbf{S}$. We denote $r_\Delta(t)$ to be the time t (continuously compounded) return of the asset price $\mathbf{S}$ for the interval $[t, t + \Delta)$. It is defined as

$$r_\Delta(t) = \log S(t + \Delta) - \log S(t). \tag{3}$$

When the drift μ and the volatility σ are constant for all times t, the asset price model, defined by equation (1) or (2), is called the *classical lognormal model*. The Gaussian assumption of the theoretical return distribution of this classical model seriously restricts its shape, especially the tail thickness. Asset returns are usually observed to have leptokurtic empirical distributions. That is, they have heavier tails and have a more pronounced peak around the mode than a normal distribution.

We denote the *log asset price process* by $\mathbf{L} = \{L(t) = \log S(t),\ t \geqslant 0\}$. The *quadratic variation process* $\langle \mathbf{L} \rangle = \{\langle L \rangle(t),\ t \geqslant 0\}$ is given by (see e.g. Jacod and Shiryaev, 1987, §4e)

$$\langle L \rangle(t) = \int_0^t \sigma(u)^2\, du, \tag{4}$$

for a generalised lognormal model.

We define the *empirical quadratic variation process* $\langle \mathbf{L} \rangle_{\boldsymbol{\Delta}} = \{\langle L \rangle_\Delta(t),\ t \geqslant 0\}$, based on time steps of length Δ, of the log asset price process $\mathbf{L}$ to be

$$\langle L \rangle_\Delta(t) = \sum_{j=0}^{n-1} r_\Delta(j\,\Delta)^2, \tag{5}$$

where $n\,\Delta \leqslant t < (n+1)\,\Delta$, for some $n \in \mathbf{N}$, and $r_\Delta(\cdot)$ are the returns defined in (3). Note that the empirical quadratic variation process $\langle \mathbf{L} \rangle_{\boldsymbol{\Delta}}$ is an estimate for the true underlying quadratic variation process $\langle \mathbf{L} \rangle$. It converges (under rather general assumptions) P-a.s. to $\langle \mathbf{L} \rangle$ as $\Delta \to 0$.

The daily empirical quadratic variation processes of the log indexes are shown in Figure 1. This figure indicates that the quadratic variation processes $\langle \mathbf{L} \rangle$ are stochastic non-decreasing processes. Consequently it follows from (4) that the volatility σ is a stochastic process. One can say it is the stochastic nature of volatility which makes the distribution of returns leptokurtic. Below it will be generally shown that the theoretical return distribution in a generalised lognormal model is leptokurtic when the volatility σ is stochastic.

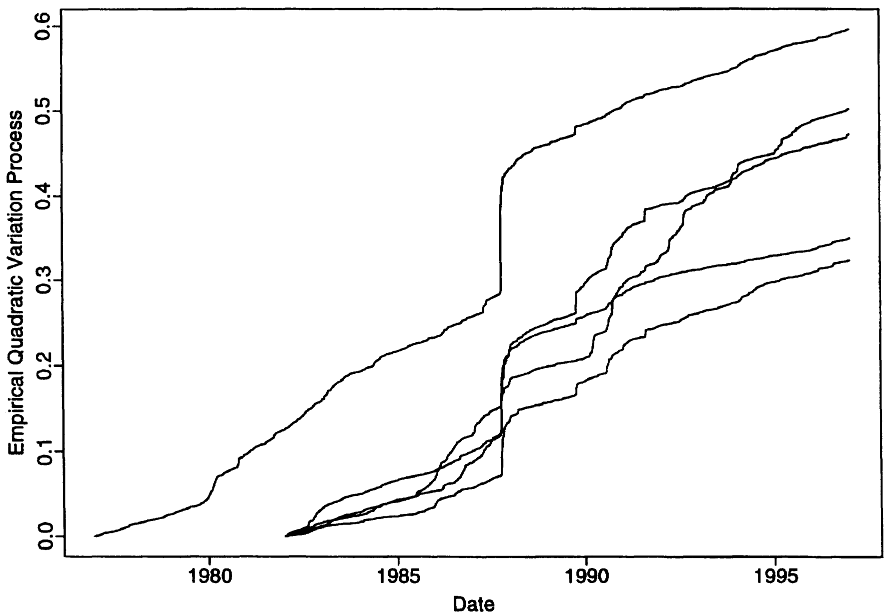

Figure 1: Daily empirical quadratic variation processes for the log market price indexes.

3 Mixing distributions

Consider the time discretisation

$$0 = t_0 \leqslant t_1 \leqslant t_2 \leqslant \dots, \tag{6}$$

where $t_i = i\Delta$ for all $i \in \mathbf{N}$ and $\Delta > 0$. Define i_t as the largest integer i such that t_i is less than or equal to t, i.e. $i_t = \max\{i \in \mathbf{N} : t_i \leqslant t\}$. We obtain the important class of discrete time generalised lognormal asset price models by keeping the drift μ and the volatility σ piece-wise constant over the discretisation intervals. A discrete time generalised lognormal model's log asset price process is then given by $\mathbf{L}_\Delta = \{L_\Delta(t) = L_\Delta(t_{i_t}),\ t \geqslant 0\}$, where

$$L_\Delta(t_{i+1}) = L_\Delta(t_i) + (\mu(t_i) - \frac{1}{2}\sigma(t_i)^2)\,\Delta + \sigma(t_i)\,\Delta W(t_i), \tag{7}$$

for $i \in \mathbf{N}$ and where $\Delta W(t_i)$, $i = 0, 1, 2, \dots$, are independent and identically distributed normal random variables with zero mean and variance Δ. The stochastic difference equation can be interpreted as that of an Euler approximation for a certain log asset price process $\mathbf{L}$ as defined earlier.

Note that the discrete time log asset price process $\mathbf{L}_\Delta$ then converges to its continuous time limit $\mathbf{L}$ as the time step size Δ tends to zero (see e.g. Kloeden and Platen, 1992).

Consider now the random variable X. We denote its distribution function (d.f.) by $F_X(\cdot)$, its characteristic function (c.f.) by $\phi_X(\cdot)$, and its probability density function (p.d.f.) by $f_X(\cdot)$ if it exists. We also denote the nth moment of X by $m_{X,n} = \mathrm{E}(X^n)$, where $\mathrm{E}(\cdot)$ denotes the expectation operator. It is well known that the mean $\mu_X = \mathrm{E}(X)$, variance $\sigma_X^2 = \mathrm{E}((X-\mu_X)^2)$, skewness $\beta_X = \mathrm{E}((X-\mu_X)^3)/\sigma_X^3$ and kurtosis $\kappa_X = \mathrm{E}((X-\mu_X)^4)/\sigma_X^4$ are respectively measures of the location, variability, degree of asymmetry and tail thickness/peakness of the distribution of X. Note that the moments $m_{X,n}$, $n \in \mathbf{N}$, can be calculated by using the c.f. $\phi_X(\cdot)$ with the well known formula

$$m_{X,n} = (-i)^n \, \phi_X^{(n)}(0), \tag{8}$$

where $\phi_X^{(n)}(\cdot)$ denotes the nth derivative of the c.f. The mean, variance, skewness and kurtosis can be calculated from the moments.

The volatility σ is an unobservable quantity. As such, the quantification of the distribution of volatility can not be directly obtained. However we can obtain it indirectly via the distribution of returns. Equations (3) and (7) give the returns for the discrete time generalised lognormal model as

$$r_\Delta(t_i) = (\mu(t_i) - \frac{1}{2}\,\sigma(t_i)^2)\,\Delta + \sigma(t_i)\,\Delta W(t_i), \tag{9}$$

for $i \in \mathbf{N}$. The return $r_\Delta(t_i)$, conditioned on the random variable $\sigma(t_i)^2$, is a normal distributed random variable with mean $(\mu(t_i) - \frac{1}{2}\,\sigma(t_i)^2)\,\Delta$ and variance $\sigma(t_i)^2\,\Delta$. The drift coefficient $\mu(t_i)$ may depend on the random variable $\sigma(t_i)^2$ so we write the conditional mean as $\xi(t_i, \sigma(t_i)^2) = \mu(t_i) - \frac{1}{2}\,\sigma(t_i)^2$ to denote this possible dependence.

Let us assume some properties for the conditional mean $\xi(t_i, \sigma(t_i)^2)\,\Delta$ and the conditional variance $\sigma(t_i)^2\Delta$ of the return $r_\Delta(t_i)$ given in (9):

First we assume that the volatility σ is a stationary process (see e.g. Feller, 1966, §III.7). Then $\sigma(t)$ is identically distributed according to the invariant distribution of σ for any given time t. The stationarity is a reasonable assumption because if we look at the quadratic variation processes in Figure 1 we notice that the processes are similar over the entire time period. They could be described as having linear trends interspersed with strongly increasing periods (very volatile) which slowly revert back towards another linear trend with the same slope. This indicates that the volatility is a stationary process.

We also make the simplifying assumption that the mean has the structure

$$\xi(t_i, \sigma(t_i)^2)\,\Delta = (\eta + \varrho\,\sigma(t_i)^2)\,\Delta, \tag{10}$$

for some constants η and ϱ.

These two assumptions imply the following relations for the unconditional distribution of returns $r_\Delta(\cdot)$. It follows that their d.f. is

$$F_{r_\Delta}(x) = \int_0^\infty \Phi\left(\frac{x - (\eta + \varrho\,u)\,\Delta}{\sqrt{u\,\Delta}}\right) dF_{\sigma^2}(u) \quad (x \in \mathbf{R}), \tag{11}$$

where $\Phi(\cdot)$ is the standard normal d.f. The corresponding p.d.f. is

$$f_{r_\Delta}(x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{u\,\Delta}} \exp\left\{-\frac{(x - (\eta + \varrho\,u)\,\Delta)^2}{2\,u\,\Delta}\right\} dF_{\sigma^2}(u) \quad (x \in \mathbf{R}), \tag{12}$$

if it exists, and the corresponding c.f. is

$$\phi_{r_\Delta}(\theta) = \int_0^\infty \exp\left\{i\,(\eta + \varrho\,u)\,\Delta\,\theta - \frac{1}{2}\,u\,\Delta\,\theta^2\right\} dF_{\sigma^2}(u) \quad (\theta \in \mathbf{R}). \tag{13}$$

The unconditional distribution is called a *mixture distribution* and the distribution of σ^2 is called the *mixing distribution* (see e.g. Feller, 1966, §II.5).

Feller (1966), §XVII.3(i), gives the general representation of the c.f. $\phi_{\sigma^2}(\cdot)$ for the non-negative random variable σ^2. Then the moments $m_{\sigma^2,n}$, $n \in \mathbf{N}$, can be calculated by using (8) and this representation of the c.f. $\phi_{\sigma^2}(\cdot)$. The mean μ_{σ^2}, variance $\sigma^2_{\sigma^2}$, skewness β_{σ^2} and kurtosis κ_{σ^2} can consequently be calculated. It is easily shown from these values that $\mu_{\sigma^2}, \sigma^2_{\sigma^2}, \beta_{\sigma^2} > 0$ and $\kappa_{\sigma^2} > 3$, if σ^2 is not deterministic. That is, the random variable σ^2 is positively skewed and leptokurtic.

Let us now compute the mean μ_{r_Δ}, variance $\sigma^2_{r_\Delta}$, skewness β_{r_Δ} and kurtosis κ_{r_Δ} for the return $r_\Delta(\cdot)$ as important measures required to understand the distributional properties of asset prices. Formulae for these measures are obtained via equations (8) and (13) giving

$$\begin{aligned}
\mu_{r_\Delta} &= \eta\,\Delta + \varrho\,\mu_{\sigma^2}\,\Delta, \\
\sigma^2_{r_\Delta} &= \mu_{\sigma^2}\,\Delta + \varrho^2\,\sigma^2_{\sigma^2}\,\Delta^2, \\
\beta_{r_\Delta} &= \frac{3\,\varrho\,\sigma^2_{\sigma^2}\,\Delta^2 + \varrho^3\,\beta_{\sigma^2}\,\sigma^3_{\sigma^2}\,\Delta^3}{\sigma^3_{r_\Delta}}, \\
\kappa_{r_\Delta} &= \frac{3\,(\mu^2_{\sigma^2} + \sigma^2_{\sigma^2})\,\Delta^2 + 6\,\varrho^2\,(\mu_{\sigma^2}\,\sigma^2_{\sigma^2} + \beta_{\sigma^2}\,\sigma^3_{\sigma^2})\,\Delta^3 + \varrho^4\,\kappa_{\sigma^2}\,\sigma^4_{\sigma^2}\,\Delta^4}{\sigma^4_{r_\Delta}}
\end{aligned} \tag{14}$$

It now easily follows that if σ^2 is stochastic, i.e. $\sigma^2_{\sigma^2} > 0$, then $\kappa_{r_\Delta} > 3$. We also get that $\text{sign}(\beta_{r_\Delta}) = \text{sign}(\varrho)$, by using (14) and the above fact that σ^2

is positively skewed and leptokurtic, i.e. $\beta_{\sigma^2} > 0$, $\kappa_{\sigma^2} > 3$. The returns $r_\Delta(\cdot)$ are therefore leptokurtic and skewed in the direction of the sign of ϱ. The above relations show that the stochastic nature of volatility implies a leptokurtic distribution for the returns.

We state that in our analysis of price indexes in Section 5 the empirical return distributions are fairly symmetrical. Consequently to simplify the analysis, we assume that the distribution of asset returns is *symmetric*. This means we explicitly assume that $\varrho = 0$ from this point onward.

4 Marginal distributions of asset returns

As already mentioned we focus our analysis on the marginal distribution of asset returns. The marginal distributions for the returns in the models which we examine in this section are mixtures of the normal distribution. They differ by their mixing distribution σ^2. As discussed above, this implies a leptokurtic distribution for the returns if σ^2 is random. The models can be characterised by *either* its mixing distribution for σ^2 *or* equivalently by its mixture distribution for r_Δ, since each are related by equations (11), (12) and (13).

Below we briefly characterise the different models by only giving the probability density function $f_{r_\Delta}(\cdot)$ or the characteristic function $\phi_{r_\Delta}(\cdot)$ for the marginal distribution of the returns $r_\Delta(\cdot)$. A more detailed treatment can be found in Hurst, Platen and Rachev (1996).

It must be emphasised that we do not intend to similarly model the asset price process as having independent increments as the following models do. The i.i.d. modelling assumption is inconsistent with the well known properties of volatility clustering, long-memory and persistence for the asset prices. We will interpret the result more correctly as an identification of the marginal distribution for asset returns.

4.1 The Mandelbrot and Fama logstable model

Mandelbrot (1963, 1967) and Fama (1963, 1965) proposed returns to be distributed with an α-stable distribution. This occurs when the stationary distribution of σ^2 is a maximally skewed $\alpha/2$-stable distribution with $\alpha \in (0,2)$ (see e.g. Mandelbrot and Taylor, 1967). The c.f. of the returns $r_\Delta(\cdot)$ is

$$\phi_{r_\Delta}(\theta) = \exp\{i\,\eta\,\Delta\,\theta - c^\alpha \Delta\,|\theta|^\alpha\} \qquad (\theta \in \mathbf{R}). \tag{15}$$

The parameter α is called the *index of stability* and is a shape parameter for the distribution, the smaller α, the larger the tail thickness. This model also implies infinite variance and infinite kurtosis for the return distribution.

4.2 The Clark model

Clark (1973) proposed the change in the *asset price* process to be distributed with a normal-lognormal mixture distribution. Prices have to be positive and represent growth processes therefore we are interested in the change in the *log asset price* process, i.e. returns. We modify Clark's model here and propose the returns to be distributed with a normal-lognormal mixture distribution. In this modified version of the Clark model the stationary distribution of σ^2 is a lognormal distribution. The p.d.f. of the returns $r_\Delta(\cdot)$ is

$$f_{r_\Delta}(x) = \frac{1}{2\pi\,\varphi\sqrt{\Delta}} \cdot$$

$$\int_0^\infty u^{-\frac{3}{2}} \exp\left\{-\frac{1}{2}\left(\frac{(x-\eta\,\Delta)^2}{u\,\Delta} + \frac{\left(\log u - \log c^2 + \frac{1}{2}\varphi^2\right)^2}{\varphi^2}\right)\right\} du \quad (x \in \mathbf{R}), \tag{16}$$

The parameter φ is a shape parameter for the distribution. The return distribution for this model has kurtosis $\kappa_{r_\Delta} = 3\exp(\varphi^2)$.

4.3 The log symmetric generalised hyperbolic model

Various authors (e.g. Praetz, 1972; Blattberg and Gonedes, 1974; Madan and Seneta, 1990; Barndorff-Nielsen 1995; Eberlein and Keller 1995; Küchler et al., 1995) have proposed returns to be distributed within the class of the generalised hyperbolic distributions. We consider the more restrictive *symmetric* generalised hyperbolic distributions for the return distribution. These distributions result when the stationary distribution of σ^2 is a generalised inverse Gaussian distribution. The p.d.f. of the returns $r_\Delta(\cdot)$ is

$$f_{r_\Delta}(x) = \frac{1}{\delta\sqrt{\Delta}\,K_\lambda(\alpha\,\delta)}\sqrt{\frac{\alpha\,\delta}{2\,\pi}} \cdot$$

$$\left(1 + \frac{(x-\eta\,\Delta)^2}{\delta^2\Delta}\right)^{\frac{1}{2}\left(\lambda-\frac{1}{2}\right)} K_{\lambda-\frac{1}{2}}\left(\alpha\,\delta\sqrt{1 + \frac{(x-\eta\,\Delta)^2}{\delta^2\Delta}}\right) \quad (x \in \mathbf{R}). \tag{17}$$

where $K_\lambda(\cdot)$ is the modified Bessel function of the third kind with index λ, $\lambda \in \mathbf{R}$ and $\alpha, \delta \geqslant 0$. In addition $\alpha \neq 0$ if $\lambda \geqslant 0$ and $\delta \neq 0$ if $\lambda \leqslant 0$. The parameters λ and $\bar{\alpha} = \alpha\,\delta$ are invariant shape parameters. The return distribution for this model has kurtosis $\kappa_{r_\Delta} = 3K_\lambda(\bar{\alpha})\,K_{\lambda+2}(\bar{\alpha})/K_{\lambda+1}(\bar{\alpha})^2$.

In the following we briefly consider special parameterisations of this model which have previously been considered by other authors.

4.4 The log Student t model

Praetz (1972) and Blattberg and Gonedes (1974) proposed that the returns should be distributed with a Student t distribution with degrees of freedom $\nu > 0$. This occurs when the shape parameters $\lambda = -\frac{1}{2}\nu < 0$ and $\bar{\alpha} = 0$, i.e. $\alpha = 0$, and the parameter $\delta = c\sqrt{\nu}$. The p.d.f. of the returns $r_\Delta(\cdot)$ is

$$f_{r_\Delta}(x) = \frac{\Gamma(\frac{1}{2}\nu + \frac{1}{2})}{c\sqrt{\pi\,\nu\,\Delta}\,\Gamma(\frac{1}{2}\nu)}\left(1 + \frac{(x-\eta\,\Delta)^2}{c^2\,\nu\,\Delta}\right)^{-\frac{1}{2}\nu-\frac{1}{2}} \qquad (x \in \mathbf{R}). \tag{18}$$

The degrees of freedom ν is the shape parameter for the distribution. The return distribution for this model has kurtosis $\kappa_{r_\Delta} = 3(\nu-2)/(\nu-4)$, for $\nu > 4$, and is infinite otherwise.

4.5 The Barndorff-Nielsen log normal\\inverse Gaussian model

Barndorff-Nielsen (1995) proposed returns to be distributed with a normal-inverse Gaussian mixture distribution. This occurs when the shape parameter $\lambda = -\frac{1}{2}$. The p.d.f. of the returns $r_\Delta(\cdot)$ is

$$f_{r_\Delta}(x) = \frac{\sqrt{\bar{\alpha}}\,\exp\{\bar{\alpha}\}}{c\sqrt{\Delta}\,\pi}\cdot$$

$$\left(1 + \frac{(x-\eta\,\Delta)^2}{\bar{\alpha}\,c^2\Delta}\right)^{-\frac{1}{2}} K_1\left(\bar{\alpha}\sqrt{1 + \frac{(x-\eta\,\Delta)^2}{\bar{\alpha}\,c^2\Delta}}\right) \qquad (x \in \mathbf{R}). \tag{19}$$

The parameter $\bar{\alpha}$ is the shape parameter for the distribution. The return distribution for this model has kurtosis $\kappa_{r_\Delta} = 3(1 + 1/\bar{\alpha})$.

4.6 The log hyperbolic model

Eberlein and Keller (1995) and Küchler et al. (1995) proposed returns to be distributed with a hyperbolic distribution. This occurs when the shape parameter $\lambda = 1$. The p.d.f. of the returns $r_\Delta(\cdot)$ is

$$f_{r_\Delta}(x) = \frac{1}{2\,\delta\sqrt{\Delta}\,K_1(\bar{\alpha})}\exp\left\{-\bar{\alpha}\sqrt{1 + \frac{(x-\eta)^2}{\delta^2\Delta}}\right\} \qquad (x \in \mathbf{R}). \tag{20}$$

The parameter $\bar{\alpha}$ is the shape parameter for the distribution. The return distribution for this model has kurtosis $\kappa_{r_\Delta} = 3K_1(\bar{\alpha})\,K_3(\bar{\alpha})/K_2(\bar{\alpha})^2$.

4.7 The Madan and Seneta log variance gamma model

Madan and Seneta (1990) proposed returns to be distributed with a normal-gamma mixture distribution. This occurs when the shape parameters $\lambda > 0$ and $\bar{\alpha} = 0$, i.e. $\delta = 0$. The p.d.f. of the returns $r_\Delta(\cdot)$ is

$$f_{r_\Delta}(x) = \frac{\sqrt{\lambda}}{c\sqrt{\Delta\pi}\,\Gamma(\lambda)\,2^{\lambda-1}} \cdot$$

$$\left(\sqrt{2\,\lambda}\,\frac{|x-\eta\,\Delta|}{c\sqrt{\Delta}}\right)^{\lambda-\frac{1}{2}} K_{\lambda-\frac{1}{2}}\left(\sqrt{2\,\lambda}\,\frac{|x-\eta\,\Delta|}{c\sqrt{\Delta}}\right) \qquad (x \in \mathbf{R}). \qquad (21)$$

The parameter λ is the shape parameter for the distribution. The return distribution for this model has kurtosis $\kappa_{r_\Delta} = 3\,(1 + 1/\lambda)$.

5 Analysis of major world market indexes

The empirical analysis will be performed on market indexes from the United States of America, Japan, Germany, Switzerland and Australia. The Australian index is calculated by Datastream International and the other indexes are calculated by Morgan Stanley Capital International. The data for these indexes are daily data for the 15 years from the beginning of 1982 to the end of 1996, except for 20 years of Australian index data which start from the beginning of 1977. We note that all of these indexes include the stock market crash of October 1987.

In Table 1 we display the results of our analysis. Under the heading of *Empirical Model* we show the total number of daily returns n and the sample measure of kurtosis $\widehat{\kappa}_{r_\Delta}$. Also included in this table is the sample measure of kurtosis $\widehat{\kappa}^\star_{r_\Delta}$ corresponding to the data with the largest absolute return removed.

The two sample measures of kurtosis, κ_{r_Δ} and $\widehat{\kappa}^\star_{r_\Delta}$, indicate that the index returns are highly leptokurtic and hence are very heavy tailed. In fact they are so large that higher moments (including the fourth moment) may be unbounded for the index returns, i.e. be infinite. In this case the sample measure of kurtosis would be unstable. This is what we observe by removing one extreme observation from the sample, the sample kurtosis changes significantly for each index. We also observe this property by a plot of the sample measure of kurtosis against the sample size which is not shown here. It is therefore important in our analysis to concentrate on the *entire distribution* and not just on a *single statistic* (in particular the possibly unbounded sample kurtosis) to identify a good or best model, or more precisely a best marginal distribution. We will therefore base our final judgement on the likelihood ratio test.

It can be shown that all of the models in Section 4 include the classical lognormal model as a specific or limiting case. Consequently we can test if each model is significantly better than the classical lognormal model by using the likelihood ratio test (see e.g. Rao, 1973, §6e.2). Define the likelihood ratio

$$\Lambda = \frac{\mathfrak{L}_{lognormal}}{\mathfrak{L}_{other}}, \tag{22}$$

where $\mathfrak{L}_{lognormal}$ is the likelihood value of the classical lognormal model and $\mathfrak{L}_{other}$ is the likelihood value of the other model we are testing. The asymptotic distribution of $-2\log\Lambda$ is chisquare with degrees of freedom equal to the difference in the number of parameters between the two models. Large values of $-2\log\Lambda$ indicate that the model under consideration is significantly better (explains more) than the classical lognormal model. We choose the model with the significantly [1] largest value of $-2\log\Lambda$ to be the *best* model. Intuitively, this is the model which has the largest probability for the returns and therefore is adding the most information to the classical lognormal model with the minimum number of parameters.

Another way of comparing the models, especially the tail properties, is to statistically determine how close the empirical distribution function and the model's distribution function are for the returns. We use the Anderson and Darling (1952) test here. This test increases the power of the more commonly used Kolmogorov test in the tails of the distribution by using the properly *weighted* test statistic given by

$$AD = \max_{x\in\mathbf{R}} \frac{|F_e(x) - F_m(x)|}{\sqrt{F_m(x)\,(1 - F_m(x))}}, \tag{23}$$

where $F_e(\cdot)$ is the empirical distribution function and $F_m(\cdot)$ is the model's distribution function. A good (bad) fit is indicated by a small (large) difference between the two distribution functions and hence a small (large) value of the test statistic AD.

The parameters for each model of Section 4 are estimated by the maximum likelihood method. For each model we display in Table 1 the estimated shape parameter(s), the corresponding kurtosis κ_{r_Δ}, the likelihood ratio test value $-2\log\Lambda$ and the Anderson-Darling test statistic AD.

For all of the models the likelihood ratio test values $-2\log\Lambda$ are extremely large indicating that all of the models from Section 4 are significantly better than the classical lognormal model. In Table 1 we have highlighted the most significant likelihood ratio value $-2\log\Lambda$ for each in-

[1]The log symmetric generalised hyperbolic model has an extra parameter than all of the other models and so it has to be accounted for correctly.

dex. It is clear that the log Student t model is the best model for all five indexes.

Table 1: Results for the quantification of the marginal distributions of returns and volatility.

Model	Statistic	Country: Australia	Germany	Japan	Switzerland	USA
Empirical	n	5052	3761	3723	3761	3803
	$\widehat{\kappa}_{r_\Delta}$	125.7994	16.7455	21.4201	26.037	93.5112
	$\widehat{\kappa}^{\star}_{r_\Delta}$	9.4216	11.1454	11.1512	17.5189	11.3116
Normal	κ_{r_Δ}	3	3	3	3	3
	AD	2.67e+81	1.47e+14	4.34e+19	3.88e+18	1.45e+59
Stable	α	1.8055	1.7237	1.6099	1.6931	1.6878
	κ_{r_Δ}	∞	∞	∞	∞	∞
	$-2\log\Lambda$	1433.665	828.7447	1008.4832	1132.1436	1180.4462
	AD	**0.045991**	0.061149	0.057997	**0.060315**	**0.056682**
Clark	φ	0.8452	0.9182	1.0643	0.9833	1.0004
	κ_{r_Δ}	6.1291	6.9712	9.3127	7.889	8.1616
	$-2\log\Lambda$	1416.0134	806.4511	1034.6007	1101.4388	1229.0101
	AD	43.745	0.29045	0.22173	0.35477	3.5355
Symmetric Generalised Hyperbolic	λ	-2.2721	-1.9253	-1.4109	-1.753	-1.7799
	$\bar{\alpha}$	2.6803e-06	9.3886e-07	0.20115	1.7939e-06	5.8455e-07
	κ_{r_Δ}	14.0143	263.349	19.1107	6581.6199	6135.9287
	$-2\log\Lambda$	1449.143	836.0898	1048.1408	1138.6009	1231.9909
	AD	0.37820	0.057845	0.054036	0.081052	0.12035
Student t	ν	4.5441	3.8506	3.0687	3.5061	3.5598
	κ_{r_Δ}	14.0268	∞	∞	∞	∞
	$-2\log\Lambda$	**1449.143**	**836.0898**	**1047.256**	**1138.6009**	**1231.9909**
	AD	0.37838	**0.057845**	**0.030761**	0.081066	0.12073
Normal Inverse Gaussian	$\bar{\alpha}$	0.97355	0.80127	0.52894	0.6519	0.6359
	κ_{r_Δ}	6.0815	6.7441	8.6717	7.602	7.7178
	$-2\log\Lambda$	1388.7137	800.6066	1033.4658	1092.0354	1209.8897
	AD	7808.0	0.71372	0.62053	1.0496	105.02
Hyperbolic	$\bar{\alpha}$	0.72732	0.57797	0.25477	0.42248	0.28722
	κ_{r_Δ}	5.1335	5.3104	5.7425	5.5127	5.6981
	$-2\log\Lambda$	1343.293	755.1277	970.3204	1023.5112	1173.6493
	AD	6.54e+06	8.5788	26.189	31.885	88645
Variance Gamma	λ	1.481	1.3375	1.0212	1.1912	1.1225
	κ_{r_Δ}	5.0256	5.243	5.9376	5.5185	5.6726
	$-2\log\Lambda$	1322.8124	739.5928	965.7311	1007.2743	1165.7439
	AD	3.15e+07	11.613	19.632	36.099	1.00e+05

Mixed results are obtained when we use the Anderson and Darling test. Smaller AD values indicate a better fit. There is a mixture between the log Student t model and the log stable model as to which model is the best. We note that the October 1987 crash return is having a great influence on the test statistic and therefore seems to bias towards heavy tailed distributions. We consider this test only as an additional check for identifying the appropriate tail properties. From our point of view the likelihood ratio test provides the most objective basis for a comparison between the marginal return distributions.

Based on the above results we conclude that the three parameter Student t distribution is the best marginal distribution for index returns. It is closely followed by the four parameter symmetric generalised hyperbolic distribution, which for all but the Japanese index turns out to be exactly the Student t distribution. The stable, Clark and normal\\inverse Gaussian distributions can be described as distributions which do not explain the tail properties accurately enough. The stable distribution overestimates the tail thickness whereas the Clark and normal\\inverse Gaussian distributions both underestimate it. The hyperbolic and variance gamma distributions are poor and dramatically underestimate the tail properties.

Some readers may argue that for each index the large negative return caused by the stock market crash of October 1987 is an outlier and is therefore influencing our results. That is, in favour of a model with very heavy tails opposed to one with less heavy tails. We would like to point out that from the paper on extreme value theory for asset price returns by Longin (1996) the stock market crash of October 1987 is dismissed as an outlier, i.e. it is consistent with the rest of the data. For principle reasons we do not like to exclude the extreme events such as stock market crashes from our samples because it is this feature (namely tail heaviness) we are explicitly emphasising to model in a consistent way. However to provide a view on the robustness of our results we removed the large negative return caused by the stock market crash of October 1987 and repeated our study. As to be expected by removing extreme events, the distributions with thinner tails improve whereas the stable distribution, which has the heaviest tails, gets worse (the results can be obtained from the authors). The Student t distribution is still the best distribution from the alternatives considered.

6 Conclusion

The Student t distribution has been shown above to be the best marginal distribution for index returns, with respect to the likelihood ratio test. It implies an inverted gamma distribution for the marginal distribution of the squared volatility σ^2. This distributional property can now be exploited to identify possible dynamics of the volatility process σ and hence the evolution of the asset price process.

References

[1] Anderson, T. W. and D. A. Darling (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math.*

Statist. **23,** 193–212.

[2] Barndorff-Nielsen, O. E. (1995). Normal Inverse Gaussian processes and the modelling of stock returns. Research Report 300, University of Aarhus.

[3] Blattberg, R. C. and N. Gonedes (1974). A comparison of the stable and Student distributions as statistical models for stock prices. *Journal of Business* **47**, 244–280.

[4] Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* ***41,*** 135–159.

[5] Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli* **1,** 281–299.

[6] Fama, E. F. (1963). Mandelbrot and the stable paretian hypothesis. *Journal of Business* **36,** 420–429.

[7] Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business* **38,** 34–105.

[8] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*, Volume 2. New York: Wiley.

[9] Hurst, S. R., E. Platen, and S. T. Rachev (1996). Subordinated market index models: A comparison. Research Report FMRR 002–96, The Australian National University. *Financial Engineering and the Japanese Markets.* To appear.

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

[11] Kloeden, P. E. and E. Platen (1992). *Numerical Solution of Stochastic Differential Equations.* Berlin: Springer-Verlag.

[12] Küchler, U., K. Neumann, M. Sørensen, and A. Streller (1995). Stock returns and hyperbolic distributions. Technical report, Humboldt University of Berlin.

[13] Longin, F. M. (1996). The asymptotic distribution of extreme stock market returns. *Journal of Business* **69,** 383–408.

[14] Madan, D. B. and E. Seneta (1990). The variance gamma (V.G.) model for share market returns. *Journal of Business* ***63,*** 511–524.

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

[16] Mandelbrot, B. (1967). The variation of some other speculative prices. *Journal of Business* **40,** 393–413.

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

[18] Praetz, P. D. (1972). The distribution of share price changes. *Journal of Business* ***45,*** 49–55.

[19] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications* (2nd ed.). New York: Wiley.