Indirect estimation of $\alpha$-stable stochastic volatility models

Marco J. Lombardi*  Giorgio Calzolari
European Central Bank  University of Florence

Abstract

The $\alpha$-stable family of distributions constitutes a generalization of the Gaussian distribution, allowing for asymmetry and thicker tails. Its many useful properties, including a central limit theorem, are especially appreciated in the financial field. However, estimation difficulties have up to now hindered its diffusion among practitioners. In this paper we propose an indirect estimation approach to stochastic volatility models with $\alpha$-stable innovations that exploits, as auxiliary model, a GARCH(1,1) with $t$-distributed innovations. We consider both cases of heavy-tailed noise in the returns or in the volatility. The approach is illustrated by means of a detailed simulation study and an application to currency crises.

1 Introduction

Heavy-tailedness of asset returns is one of the most prominent stylized facts in finance: studies questioning the Gaussian random-walk hypothesis and suggesting the use of $\alpha$-stable distributions for the modelling of financial returns started appearing in the sixties, following the seminal works by Mandelbrot (1963) and Fama (1965). The features and the analytic properties of $\alpha$-stable distributions are especially appreciated in the financial field: the fact that the family is closed under linear combination helps in portfolio analysis and risk management; the possibility to accommodate for skewness and heavy tails allows to appropriately measure risk, avoiding to underestimate the probability of extreme losses; finally, the presence of a central limit theorem constitutes a theoretical basis which should lead to prefer the $\alpha$-stable family over other heavy-tailed alternatives: since asset returns are commonly thought of as the result of the aggregation of the asset allocation decisions of the market participants, the resulting distributions should arise, in the limit, from a central limit theorem.

However, practical application of models based on $\alpha$-stable distributions has been hindered by estimation difficulties: the $\alpha$-stable density function cannot be expressed in a closed form except for very few cases. This difficulty, coupled with the fact that moments of order greater than or equal to $\alpha$ do not exist whenever $\alpha \neq 2$, has made impossible the use of standard estimation methods such as maximum likelihood and

*Corresponding author; email marco.lombard@ecb.int.
Researchers have thus proposed alternative estimation procedures, mainly based on quantiles (McCulloch 1986) or on the empirical characteristic function (Koutrouvelis 1980, Kogon & Williams 1998); those methods however can only estimate the parameters of the distributions, so that dealing with more complex models (both linear and nonlinear) based on \( \alpha \)-stable disturbances would require a two-step estimation approach. In the recent years the availability of fast computing machines has made possible to employ computationally-intensive estimation; in particular, likelihood-based inference has been carried out by approximating the density with the FFT of the characteristic function (Mittnik, Doganoglu & Chenyao 1999) or with numerical quadrature (Nolan 1997); indirect approaches have been proposed by Lombardi & Calzolari (2004) and Garcia, Renault & Veredas (2004). The Bayesian approach has also benefited from the introduction of modern computers: simulation-based MCMC methods have been proposed by Buckle (1995), Qiou & Ravishanker (1998), Lombardi (2004) and Casarin (2004b).

Several studies have highlighted that the heavy-tailedness of asset returns can be the consequence of conditional heteroscedasticity (Engle 1982). ARCH models have thus become very popular, given their ability to account for volatility clustering and, implicitly, heavy-tailedness at the same time. The introduction of this alternative way to deal with heavy-tailedness, coupled with the above-mentioned estimation difficulties, has somehow dampened the academic interest in \( \alpha \)-stable distributions. A notable exception was an interesting analysis of the relation between GARCH models and \( \alpha \)-stable distributions proposed by de Vries (1991). However, it must be remarked that, in practice, GARCH models are seldom able to accommodate for the excess of kurtosis: the standardized residuals are often found to be still leptokurtic. Thus, practitioners often use GARCH models with \( t \)-distributed innovations (Fiorentini, Sentana & Calzolari 2003), although it has to be remarked that GARCH models with \( \alpha \)-stable innovations have been proposed (McCulloch 1985 and Liu & Brorsen 1995).

A widely employed alternative to ARCH-type models is represented by stochastic volatility models (Taylor 1986); their close relationship with continuous-time diffusions makes them particularly appreciated in what they can bridge the most recent results of the theoretical finance literature. However, even in the simplest Gaussian SV case, the estimation is complicated by the latent structure of the model; indirect estimation approach were proposed by Gallant, Hsieh & Tauchen (1997), Fiorentini, Sentana & Calzolari (2004) and Monfardini (1998).

In this paper we show how stochastic volatility models with \( \alpha \)-stable innovations can be estimated using an indirect estimation approach. Traditionally, heavy tails in the setting of SV models have been accounted for using \( t \)-distributed innovations (Chib, Nardari & Shephard 2002); a recent exception is Casarin (2004a), where a simulation-based Bayesian approach to a stochastic volatility model with symmetric \( \alpha \)-stable noise is proposed. Our opinion is that the use of \( \alpha \)-stable distributions should be preferred: first because of the presence of the generalized central limit theorem, second because of the availability of formal option pricing schemes. For example, Hurst, Platen & Rachev (1999) consider log-symmetric prices for the assets, and in McCulloch (2003) and Carr & Wu (2003) the symmetry assumption is relaxed; Cartea & Howison (2005) also consider the more appealing case of time-varying volatility with \( \alpha \)-stable shocks.
As auxiliary model, we will employ a GARCH model with skew-$t$-distributed innovations; this in a sense mimics the approach followed by Fiorentini et al. (2004) for the indirect estimation of stochastic volatility models. We will first examine the compliance of the auxiliary model with the conditions required to ensure consistency; then, a detailed simulation study aimed at assessing the properties of the estimators will be conducted. An application to exchange rate crises will conclude the paper.

2 \( \alpha \)-Stable distributions

The \( \alpha \)-stable family of distributions is identified by means of the characteristic function

\[
\phi_1(t) = \begin{cases} 
\exp \left\{ i\delta t - \gamma |t|^\alpha \left[ 1 - i\tilde{\beta} \text{sgn}(t) \tan \frac{\pi \alpha}{2} \right] \right\} & \text{if } \alpha \neq 1 \\
\exp \left\{ i\delta t - \gamma |t| \left[ 1 + i\tilde{\beta} \frac{2}{\pi} \text{sgn}(t) \ln |t| \right] \right\} & \text{if } \alpha = 1
\end{cases}
\]

which depends on four parameters: \( \alpha \in (0, 2] \), measuring the tail thickness (thicker tails for smaller values of the parameter), \( \tilde{\beta} \in [-1, 1] \) determining the degree and sign of asymmetry, \( \gamma > 0 \) (scale) and \( \delta \in \mathbb{R} \) (location). The distribution will be denoted as \( S_1(\alpha, \tilde{\beta}, \gamma, \delta) \).

While the characteristic function (1) has a quite manageable expression and can straightforwardly produce several interesting analytic results, it unfortunately has a major drawback for what concerns estimation and inferential purposes: it is not continuous with respect to the parameters, having a pole at \( \alpha = 1 \).

An alternative way to write the characteristic function that overcomes this problem, due to Zolotarev (1986), is the following:

\[
\phi_0(t) = \begin{cases} 
\exp \left\{ i\delta_0 t - \gamma |t|^\alpha \left[ 1 + i\tilde{\beta} \tan \frac{\pi \alpha}{2} \text{sgn}(t) \left( |\gamma| \frac{1}{|t|^{1-\alpha}} - 1 \right) \right] \right\} & \text{if } \alpha \neq 1 \\
\exp \left\{ i\delta_0 t - \gamma |t| \left[ 1 + i\tilde{\beta} \frac{2}{\pi} \text{sgn}(t) \ln(\gamma |t|) \right] \right\} & \text{if } \alpha = 1
\end{cases}
\]

In this case, the distribution will be denoted as \( S_0(\alpha, \tilde{\beta}, \gamma, \delta_0) \). The formulation of the characteristic function is, in this case, more cumbersome, and the analytic properties have less intuitive meaning; but it is much more useful for statistical purposes and, unless otherwise stated, we will refer to it in the following. The only parameter that needs to be “translated” according to the following relationship is \( \delta \):

\[
\delta_0 = \begin{cases} 
\delta_1 + \tilde{\beta} \gamma \tan \frac{\pi \alpha}{2} & \text{if } \alpha \neq 1 \\
\delta_1 + \tilde{\beta} \frac{2}{\pi} \gamma \ln \gamma & \text{if } \alpha = 1
\end{cases}
\]

On the basis of the above equations, a \( S_1(\alpha, \tilde{\beta}, 1, 0) \) distribution corresponds to a \( S_0(\alpha, \tilde{\beta}, 1, -\tilde{\beta} \gamma \tan \frac{\pi \alpha}{2}) \), provided that \( \alpha \neq 1 \).

Unfortunately, (1) and (2) cannot be analytically inverted to yield a closed-form density function except for very few cases: \( \alpha = 2 \), corresponding to the normal distribution, \( \alpha = 1 \) and \( \tilde{\beta} = 0 \), yielding the Cauchy distribution, and \( \alpha = \frac{1}{2}, \tilde{\beta} = \pm 1 \) for the Lévy distribution. We remark that, in the case of the normal distribution, \( \tilde{\beta} \) becomes unidentified.
Despite the computational burden associated with the evaluation of the probability density function, stably distributed pseudo-random numbers can be straightforwardly simulated using the algorithm proposed in Chambers, Mallows & Stuck (1976) and Chambers, Mallows & Stuck (1987). Let $W$ be a random variable with exponential distribution of mean 1 and let $U$ be an uniformly distributed random variable on $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Furthermore, let $\zeta = \arctan \left( \frac{\tilde{\beta} \tan \frac{\pi}{2}}{\alpha} \right)$. Then

$$Z = \begin{cases} \frac{\sin \alpha (\zeta + U)}{\sqrt{\cos \alpha \zeta \cos U}} \left[ \frac{\cos (\alpha \zeta + \alpha U - U)}{W} \right]^{\frac{1}{\alpha}} & \text{if } \alpha \neq 1 \\ 2 \pi \left[ \left( \frac{\pi}{2} + \tilde{\beta} U \right) \tan U - \tilde{\beta} \ln \frac{\pi}{2} W \cos U \right] & \text{if } \alpha = 1 \end{cases}$$

has $S_0(\alpha, \tilde{\beta}, 1, 0)$ distribution. Random numbers for the general case containing also the location and scale parameters $\delta$ and $\gamma$ may be straightforwardly obtained exploiting the fact that, if $X \sim S_1(\alpha, \tilde{\beta}, \gamma, \delta)$, then $Z = \frac{X - \delta}{\gamma} \sim S_1(\alpha, \tilde{\beta}, 1, 0)$. Similarly, random numbers from an $\alpha$-stable distribution expressed in parametrization (1) can be readily obtained using (3). In what follows, we will often omit the subscript and the parameters $\gamma$ and $\delta$; we will use the shorthand notation $S(\alpha, \tilde{\beta}) = S_1(\alpha, \tilde{\beta}, 1, 0)$.

### 2.1 $\alpha$-Stable stochastic volatility models

Stochastic volatility models have been studied extensively by Taylor (1986) as an alternative to ARCH models. Their main advantage is that they can be regarded as the discrete time analogue of the continuous time stochastic processes for instantaneous log volatility frequently used in the theoretical finance literature. A standard stochastic volatility model is composed of a latent volatility equation and of an observed return equation:

$$\begin{align*}
\ln h_t &= \delta + \varphi \ln h_{t-1} + \sigma_h v_t, \\
r_t &= w_t \sqrt{h_t};
\end{align*}$$

in the most simple case, the noise terms $v_t$ and $w_t$ are assumed to be Gaussian and uncorrelated. The latent structure of the model makes inference troublesome, as the likelihood cannot be expressed in closed form. The estimation is therefore carried out using QML and the Kalman filter (Harvey, Ruiz & Shephard 1994), Bayesian MCMC techniques (Jacquier, Polson & Rossi 1994, Kim, Shephard & Chib 1998) or indirect approaches (Gallant et al. 1997, Fiorentini et al. 2004, Monfardini 1998).

The Gaussian assumption is often unsatisfactory for applied purposes, as observed series tend to be heavy-tailed and display discontinuities. Therefore, several studies (Chib et al. 2002, Jacquier, Polson & Rossi 2004) have considered the possibility to
employ $t$-distributed innovations in the return equation. Nevertheless, the use of $t$ distributions is arbitrary and in a sense dampens most of the theoretical appeal of SV models: the resulting continuous-time process is not anymore a Brownian motion, which is a requirement of most of the theoretical finance literature.

In order to account for possible discontinuities which may result from surprise events, it has also been proposed to introduce jump components in the observation equation or in the variance equation; the role of jumps in the structure of the model is discussed in detail by Eraker, Johannes & Polson (2003). The decision to include jumps in the observation or in the state equation (or in both) is not neutral and has different effects on the pattern of the volatility. Most of the research up to now has concentrated on heavy tails and jumps in the observation equation. This implies that shocks are transient and, contrary to what would happen in a GARCH framework, have no subsequent impact on volatility. Therefore, in a Gaussian framework, in order to achieve a rapid increase in volatility similar to that observed in real datasets, one would need an unlikely long sequence of positive innovations in $v_t$. Instead, jumps in volatility have been widely documented (Bates 2000): in order to account for that, one can either introduce jumps or consider distributions with heavier tails.

On the other hand, it has to be remarked that in some situations the observed jumps impact only partially on the volatility (take, for example, the black Friday market crash), therefore it may be more reasonable to assume that discontinuities do come from the observable equation.

In this setting, $\alpha$-stable distributions are peculiar in what they generate processes with discontinuities (McCulloch 1978) that can, in a way, be interpreted as jumps. It is also interesting to remark that the inclusion of a jump component yields an unconditional mixture of normals representation; this parallels with the fact that symmetric $\alpha$-stable distributions can actually be represented as scale mixtures of normals.

For what concerns jumps in returns, it is reasonable to assume that they should be symmetric, therefore we will employ symmetric $\alpha$-stable distributions. Jumps in volatility are instead naturally skewed to the right, therefore we will fix $\beta = 1$. Since $\alpha$-stable random variables with $\beta = 1$ have nonzero positive density on $\mathbb{R}$ whenever $\alpha > 1$, this assumption does not imply that innovations in the volatility are always positive. Instead, we will experience positive jumps that tend to be less frequent as $\alpha$ approaches 2, since the skewness parameter $\beta$ loses relevance as we move towards a Gaussian distribution.

In the following, we will examine both the above-mentioned specifications; in formal terms, the models of interest will be:

$$\ln h_t = \delta + \varphi \ln h_{t-1} + \sigma_h v_t, \quad v_t \sim \mathcal{N}(0, 1), \quad (6)$$

$$r_t = w_t \sqrt{h_t}, \quad w_t \sim S(\alpha, 0);$$

and

$$\ln h_t = \delta + \varphi \ln h_{t-1} + \sigma_h v_t, \quad v_t \sim S(\alpha, 1), \quad (7)$$

$$r_t = w_t \sqrt{h_t}, \quad w_t \sim \mathcal{N}(0, 1).$$

To convince the reader that $\alpha$-stable distributions can appropriately model jumps and generate plausible patterns of volatility, we report (Figure 1) two simulated paths...
of volatility and returns generated by allowing $\alpha$-stable innovations in, respectively, the observable and the state equation. For the noise in the observable equation, $\alpha$ was fixed to 1.9 (with $\tilde{\beta} = 0$), whereas in the state equation we employed 1.7 (with $\tilde{\beta} = 1$). The parameters of the volatility equation were fixed to $\delta = -0.15$, $\varphi = 0.98$, $\sigma_h = 0.06$. We observe that the inclusion of the heavy-tailed noise in the observation equation generates spikes in the pattern of returns, but is unable to yield a realistic pattern of the volatility. On the other hand, the employment of the $\alpha$-stable noise in the volatility equation yields realistic patterns of volatility, with shocks that rapidly increase the volatility which then fade away smoothly. This also implies a more visible volatility clustering pattern in the returns, but of course rules out transient shocks.

To sum up, the use of $\alpha$-stable distributions should be in our opinion preferable over other heavy-tailed distributions and/or jumps for three different reasons:

- The presence of the central limit theorem should justify the arising of such a kind of distributions in the volatility setting: if (log-)volatility is thought of as the stream of news arriving to the market, then it is natural to assume it to be composed of a large number of individual contributions of each event;

- In continuous time, the $\alpha$-stable generalization of the Brownian motion (the
α-Stable Lévy motion) has been studied extensively (Samorodnitsky & Taqqu 1994, Janicki & Weron 1994). Finance theorists have widely exploited the α-stable assumption and several models are available, ranging from asset allocation to option pricing (see, for an excellent survey, McCulloch (1996));

- Unlike Brownian motions, Lévy motions are not almost surely continuous, but instead they are almost surely dense with discontinuities (McCulloch 1978); this means they can actually account for empirically-observed jumps without having to include a separate jump component.

3 The indirect estimation approach

The indirect estimation (Gouriéroux, Monfort & Renault 1993) is an inferential approach which is suitable for situations where the estimation of the statistical model of interest is too difficult to be performed directly while it is straightforward to produce simulated values from the same model. It was first motivated by econometric models with latent variables, but it can be applied in virtually every situation in which the direct maximization of the likelihood function turns out to be difficult.

The principle underlying “indirect inference” (Gouriéroux et al. 1993) is very simple: suppose we have a sample of \( T \) observations \( y \) and a model whose likelihood function \( L^*(y; \theta) \) is difficult to handle and maximize; the model could also depend on a matrix of explanatory variables \( X \). The maximum likelihood estimate of \( \theta \in \Theta \), given by

\[
\hat{\theta} = \arg \max_{\theta \in \Theta} \sum_{t=1}^{T} \ln L^*(\theta; y_t),
\]

is thus unavailable. Let us now take an alternative model, depending on a parameter vector \( \zeta \in Z \), which will be indicated as auxiliary model, easier to handle, and suppose we decide to use it in the place of the original one. Since the model is misspecified, the quasi-ML estimator

\[
\hat{\zeta} = \arg \max_{\zeta \in Z} \sum_{t=1}^{T} \ln \tilde{L}(\zeta; y_t),
\]

is not necessarily consistent: the idea is to exploit simulations performed under the original model to correct for inconsistency.

The first step consists of computing the quasi maximum likelihood estimate of \( \zeta \), which will be denoted as \( \hat{\zeta} \). Next, one simulates a set of \( S \) vectors of size \( T \) from the original model on the basis of an arbitrary parameter vector \( \hat{\theta}^{(0)} \). Let us denote each one of those vectors as \( y_s^{\star}(\hat{\theta}^{(0)}) \). The simulated values are then estimated using the auxiliary model, yielding

\[
\tilde{\zeta}(\hat{\theta}^{(0)}) = \arg \max_{\zeta \in Z} \sum_{s=1}^{S} \sum_{t=1}^{T} \ln \tilde{L} \left[ \zeta; y_s^{\star}(\hat{\theta}^{(0)}) \right].
\]
The idea is to numerically update the initial guess \( \hat{\theta}^{(0)} \) in order to minimize the distance

\[
\left[ \hat{\zeta} - \tilde{\zeta}(\theta) \right] \Omega \left[ \hat{\zeta} - \tilde{\zeta}(\theta) \right],
\]

where \( \Omega \) is a symmetric nonnegative matrix defining the metric.

An alternative but similar approach, leading to the so-called EMM (Gallant & Tauchen 1996), considers directly the score function of the auxiliary model:

\[
\sum_{t=1}^{T} \frac{\partial \ln \tilde{L}(\zeta; y_t)}{\partial \zeta},
\]

which is clearly zero for the quasi-maximum likelihood estimator of \( \beta \). The idea is to make as close as possible to zero the score computed on the simulated observations, namely

\[
\arg \min_{\theta} \left\{ \sum_{s=1}^{S} \sum_{t=1}^{T} \frac{\partial \ln \tilde{L}(\zeta; y_{st}(\theta))}{\partial \zeta} \right\} \Xi \left\{ \sum_{s=1}^{S} \sum_{t=1}^{T} \frac{\partial \ln \tilde{L}(\zeta; y_{st}(\theta))}{\partial \zeta} \right\},
\]

where \( \Xi \) is a symmetric nonnegative definite matrix. This approach is especially useful when an analytic expression for the gradient of the auxiliary model is available, since it allows us to avoid the numerical optimization routine for the computation of the \( \hat{\zeta}(\theta) \)s.

Indirect estimators are consistent and asymptotically normal under certain regularity conditions. The most difficult one to establish is that the binding function, that is the function that maps the parameter (sub-)space of the auxiliary model onto the parameter space of the true model, is one-to-one. In general, the binding function cannot be expressed analytically and the above condition needs to be verified numerically. It is clear that the choice of the auxiliary model is crucial for the successful implementation of the algorithm. For further details concerning the desirable properties of the auxiliary model, we refer to Heggland & Frigessi (2004) for the “just-identified” case. For a particular and “over-identified” choice of the auxiliary model producing efficient estimators, we refer to Gallant & Tauchen (1996).

Once one manages to specify an adequate auxiliary model, indirect estimators for the parameters of \( \alpha \)-stable distributions can be readily implemented and exploited by relying on the pseudo-random number generator of Chambers et al. (1976).

### 3.1 Indirect estimation of stochastic volatility models

We will consider two different stochastic volatility models: the former allowing for heavy tails in the observation equation and the latter with heavy-tailed innovations in the state equation.

The characteristic exponent \( \alpha \) will be constrained on \((1, 2)\), as values smaller than 1 are very unlikely in a financial setting. The persistence parameter \( \varphi \) will also be constrained on \((-1, 1)\) to ensure stationarity of the volatility.

Fiorentini et al. (2004) employ a GARCH with Student’s \( t \) innovations as an over-identified auxiliary model for the indirect estimation of a Gaussian stochastic volatility model. This choice could be fruitful also in our case, as the degrees of freedom of the \( t \)
distribution are naturally linked to the characteristic exponent \( \alpha \) in determining the tail thickness.

We will express the \( t \) distribution in terms of \( \eta = \nu^{-1} \), where \( \nu \) are the degrees of freedom, and we constrain \( \eta \) on \((0.01, 1)\). On the lower bound, the auxiliary distribution is very close to the Gaussian, while on the upper bound it corresponds to an \( \alpha \)-stable distribution with \( \alpha = 1 \). It is important to remark that here, contrary to Fiorentini et al. (2004), we do not employ the standardized version of the \( t \) distribution and we allow for values of \( \eta \) that give rise to distributions with infinite variance \((\eta \geq 0.5)\). This choice was made with the goal of avoiding the degrees of freedom to clash on their lower bound because of the heavier-tailedness of the true model. The only caveat is that \( h_t \) cannot be interpreted anymore as a conditional variance: when \( \nu < 2 \) the variance is infinite and when \( \nu \geq 2 \) the conditional variance will be determined as \( \frac{h_t}{\nu-2} \). Therefore, we will see \( h_t \) as a scale parameter. The fact that the auxiliary model may produce infinite variances is not a limitation in our opinion: in the case of \( \alpha \)-stable noise in the observation equation (6), the variance of the returns is infinite in the true model as well and the volatility equation yields only a time-varying scale parameter; when the \( \alpha \)-stable noise is employed in the variance equation, on the other hand, the returns do have finite variance – given by \( \sqrt{h_t} \) – but the variance of the volatility, which is related to the fourth moment of returns, is infinite because of the \( \alpha \)-stable innovations, and the parameter \( \sigma_h \) would not be interpreted as the variance of the (log-)volatility, but rather as a simple scale coefficient.

To sum up, the auxiliary model will be:

\[
\begin{align*}
  h_t &= \omega + \psi r_{t-1}^2 + \beta h_{t-1} \\
  r_t &= u_t \sqrt{h_t}, \quad u_t \sim t_{1/\eta};
\end{align*}
\]

In order to ensure non-negativity of the conditional variance in the auxiliary model, we will impose \( \omega > 0, \psi > 0 \) and \( \beta > 0 \) (Nelson & Cao 1992); to safeguard against weird asymptotic behavior of the QML estimator (Lumsdaine 1996) we will also enforce the constraint \( \psi + \beta < 1 \). In the case \( \psi = 0 \), it turns out (Andrews 1999) that \( \beta \) gets asymptotically unidentified; we will therefore impose \( \psi > \varepsilon \), with \( \varepsilon \) arbitrarily small.

Sections of the binding functions for models (6) and (7) and are plotted in Figure 2 and 3 and signal that such auxiliary model could be appropriate. In particular, we observe that the persistence parameter \( \varphi \) is, as expected, linked to both \( \beta \) and \( \psi \). Also the scale of the innovations in the volatility equation depend on the ARCH and GARCH coefficients, albeit in a different fashion: an increase in \( \sigma_h \) causes an increase in \( \beta \) and a decrease in \( \psi \). This effect is caused by the fact that bigger innovations in the volatility in the SV model cannot be captured by the ARCH coefficient, which is related to innovations in the returns, and need to be discounted by the past volatility dynamics which is controlled via the GARCH coefficient.

We also remark that the binding function seems to behave quite similarly for both models; some differences arise with respect to the sections \( \alpha - \nu \) and \( \delta - \omega \). The relation between the tail-thickness parameters is more smooth in model (7) and the degrees of freedom of the auxiliary model do not appear to explode to \( \infty \). This could be caused by the fact that, when allowing for heavy tails in the volatility equation, also very small
shocks may, by means of the multiplicative effect, yield a pattern in the returns that can be captured only via tails heavier than the Gaussian.

3.2 Simulation results

In order to evaluate the performance of the proposed estimator, we have conducted a simulation exercise; all the results are based on a set of 1000 replications of the indirect estimation, with \( S = 10 \). We will use three different sample sizes: 1000, 3000 and 5000 observations.

We will use two sets parameters in the stochastic volatility equation with various choices of the characteristic index \( \alpha \), namely one that roughly matches typical values obtained on weekly returns (\( \delta = -0.7, \varphi = 0.9, \sigma_h = 0.35 \)) and one roughly matching daily returns (\( \delta = -0.15, \varphi = 0.98, \sigma_h = 0.06 \)); the simulation design is therefore similar to that of Jacquier et al. (1994) and Fiorentini et al. (2004).

As remarked by Jacquier et al. (1994) in the SV setting, a parameter set yielding a low signal-to-noise-ratio (SNR) is bound to give rise to estimation difficulties. In the \( \alpha \)-stable case, however, some caveats are necessary. In the Gaussian SV model, the (power) SNR is:

\[
\text{SNR}_{2} = \frac{\text{Var}(h_t)}{\text{Var}(r_t)} = \frac{\text{Var}(h_t)}{E(h_t)^2}.
\]

Employing the above expression under model (6) yields a SNR of zero, whereas for model (7) the SNR becomes infinite. An alternative which is often use in signal pro-
cessing is to consider

$$\text{SNR}_A = \frac{E|h_t|}{E|r_t|}.$$  

When the $\alpha$-stable noise is in the observation equation (6), this choice would lead to a nonzero quantity as long as $\alpha > 1$. In the other case, however, the numerator would be the expected value of a log-stable random variable, which is unfortunately infinite also for $\alpha > 1$.

In Tables 1 and 2 we report, respectively, the outcome of the simulation experiment on the model (6) with the first and the second parameter set and for different values of $\alpha$. We remind that in this case we only consider symmetric distributions, therefore $\tilde{\beta}$ is fixed to 0. First of all, we remark that in the case with high SNR (Table 1), the estimation went smoothly and the results are very satisfactory. While the parameter $\alpha$ is estimated with strong precision, the parameters in the volatility equation have larger standard errors, especially in the presence of heavier tails. This is of course consistent with the impact of heavy tails on the SNR we have discussed above. For what concerns the execution time, we notice a weird effect: for smaller values of $\alpha$ the algorithm takes less time to estimate larger sample sizes. This is related to the fact that, when heavy tails are involved, having more observations implies a clearer picture of the tail behavior: therefore, working with smaller sample sizes may imply convergence difficulties. On the other hand, we remark that, not surprisingly, slow convergence arises also when $\alpha$ is very close to 2. In that case, since no heavy-tailed behavior has to be observed, the sample size constitutes an hinderance and considerably increases the computational time.

Figure 3: Profiles of the binding function for model (7) for various parameter values. The solid line corresponds to “daily parameters” and the dotted line to “weekly parameters”.

11
Table 1: Monte Carlo mean and standard error (in parentheses) for the first set of parameters ($\delta = -0.7$, $\varphi = 0.9$, $\sigma_h = 0.35$) and various values of $\alpha$, model (6). The rows “Time” report the average time to convergence (in seconds) of one iteration.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$T = 1000$</th>
<th>$T = 3000$</th>
<th>$T = 5000$</th>
<th>$T = 1000$</th>
<th>$T = 3000$</th>
<th>$T = 5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>1.8279</td>
<td>1.8145</td>
<td>1.8096</td>
<td>1.9104</td>
<td>1.9052</td>
<td>1.9040</td>
</tr>
<tr>
<td></td>
<td>(0.0848)</td>
<td>(0.0551)</td>
<td>(0.0447)</td>
<td>(0.0593)</td>
<td>(0.0384)</td>
<td>(0.0304)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>-1.0493</td>
<td>-0.8785</td>
<td>-0.8351</td>
<td>-0.8682</td>
<td>-0.7528</td>
<td>-0.7449</td>
</tr>
<tr>
<td></td>
<td>(1.0936)</td>
<td>(0.7286)</td>
<td>(0.6373)</td>
<td>(0.6778)</td>
<td>(0.3757)</td>
<td>(0.3076)</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.8502</td>
<td>0.8744</td>
<td>0.8807</td>
<td>0.8762</td>
<td>0.8924</td>
<td>0.8936</td>
</tr>
<tr>
<td></td>
<td>(0.1559)</td>
<td>(0.1042)</td>
<td>(0.0909)</td>
<td>(0.0964)</td>
<td>(0.0535)</td>
<td>(0.0439)</td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>0.4038</td>
<td>0.3816</td>
<td>0.3720</td>
<td>0.3737</td>
<td>0.3590</td>
<td>0.3577</td>
</tr>
<tr>
<td></td>
<td>(0.2599)</td>
<td>(0.1829)</td>
<td>(0.1569)</td>
<td>(0.1710)</td>
<td>(0.1071)</td>
<td>(0.0891)</td>
</tr>
<tr>
<td>Time</td>
<td>6.5613</td>
<td>6.2942</td>
<td>5.7529</td>
<td>3.1315</td>
<td>3.0858</td>
<td>2.7685</td>
</tr>
<tr>
<td></td>
<td>(0.0440)</td>
<td>(0.0304)</td>
<td>(0.0249)</td>
<td>(0.0279)</td>
<td>(0.0177)</td>
<td>(0.0147)</td>
</tr>
<tr>
<td></td>
<td>-0.8019</td>
<td>-0.7349</td>
<td>-0.7299</td>
<td>-0.7201</td>
<td>-0.6967</td>
<td>-0.7017</td>
</tr>
<tr>
<td></td>
<td>(0.4879)</td>
<td>(0.2687)</td>
<td>(0.2148)</td>
<td>(0.2898)</td>
<td>(0.1772)</td>
<td>(0.1380)</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.8856</td>
<td>0.8950</td>
<td>0.8958</td>
<td>0.8974</td>
<td>0.9005</td>
<td>0.8998</td>
</tr>
<tr>
<td></td>
<td>(0.0696)</td>
<td>(0.0382)</td>
<td>(0.0306)</td>
<td>(0.0411)</td>
<td>(0.0251)</td>
<td>(0.0196)</td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>0.3609</td>
<td>0.3564</td>
<td>0.3555</td>
<td>0.3377</td>
<td>0.3437</td>
<td>0.3465</td>
</tr>
<tr>
<td></td>
<td>(0.1234)</td>
<td>(0.0806)</td>
<td>(0.0661)</td>
<td>(0.0792)</td>
<td>(0.0535)</td>
<td>(0.0426)</td>
</tr>
<tr>
<td>Time</td>
<td>4.1063</td>
<td>4.5971</td>
<td>4.7299</td>
<td>4.9718</td>
<td>10.8701</td>
<td>15.9124</td>
</tr>
</tbody>
</table>
Table 2: Monte Carlo mean and standard error (in parentheses) for the second set of parameters ($\delta = -0.15$, $\varphi = 0.98$, $\sigma_h = 0.06$) and various values of $\alpha$, model (6). The rows “Time” report the average time to convergence (in seconds) of one iteration.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 1.8$</th>
<th></th>
<th></th>
<th>$\alpha = 1.9$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T = 1000$</td>
<td>$T = 3000$</td>
<td>$T = 5000$</td>
<td>$T = 1000$</td>
<td>$T = 3000$</td>
<td>$T = 5000$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.8238</td>
<td>1.8171</td>
<td>1.8152</td>
<td>1.9105</td>
<td>1.9087</td>
<td>1.9079</td>
</tr>
<tr>
<td></td>
<td>(0.0697)</td>
<td>(0.0474)</td>
<td>(0.0433)</td>
<td>(0.0527)</td>
<td>(0.0340)</td>
<td>(0.0280)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>-0.5966</td>
<td>-0.5100</td>
<td>-0.4879</td>
<td>-0.6110</td>
<td>-0.4651</td>
<td>-0.4154</td>
</tr>
<tr>
<td></td>
<td>(1.5701)</td>
<td>(1.4438)</td>
<td>(1.4170)</td>
<td>(1.3140)</td>
<td>(1.0989)</td>
<td>(1.0787)</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.9206</td>
<td>0.9320</td>
<td>0.9351</td>
<td>0.9187</td>
<td>0.9381</td>
<td>0.9448</td>
</tr>
<tr>
<td></td>
<td>(0.2086)</td>
<td>(0.1924)</td>
<td>(0.1884)</td>
<td>(0.1748)</td>
<td>(0.1459)</td>
<td>(0.1431)</td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>0.1006</td>
<td>0.0880</td>
<td>0.0831</td>
<td>0.1001</td>
<td>0.0871</td>
<td>0.0807</td>
</tr>
<tr>
<td></td>
<td>(0.1625)</td>
<td>(0.1415)</td>
<td>(0.1411)</td>
<td>(0.1390)</td>
<td>(0.1215)</td>
<td>(0.1125)</td>
</tr>
<tr>
<td>Time</td>
<td>30.401</td>
<td>51.324</td>
<td>55.424</td>
<td>26.958</td>
<td>41.921</td>
<td>49.806</td>
</tr>
</tbody>
</table>

When the lower SNR is concerned (Table 2), the performance of the indirect estimator worsens dramatically. While the performance of the estimator for $\alpha$ remains satisfactory, we have found a considerable distortion and inefficiency for the estimators of the parameters in the volatility equation. As we have previously remarked, smaller values of $\alpha$ do impact negatively on the performance of the estimator as they imply an even smaller SNR. A control experiment aimed at assessing the asymptotic properties of the estimator confirmed indeed its consistency, but it took over 100000 observations to achieve reasonable accuracy. In this case, it is interesting to remark that the weird effect of sample size on computational time highlighted above is not present: smaller values of $\alpha$ – implying a lower SNR – and larger sample sizes always imply an increase in the computational burden.

When one moves to consider model (7), the situation is completely changed. We report only the “difficult” case (low SNR, parameter values $\delta = -0.15$, $\varphi = 0.98$, $\sigma_h = 0.06$, $\beta = 1$ and various values of $\alpha$). In this case, as we have pointed out, an heavier-tailed noise yields a clearer signal and therefore facilitates the estimation. In actual facts, we observe that in this case, the estimators of the volatility parameters

---

1Results are not reported here for the sake of brevity, but are available upon request.
Table 3: Monte Carlo mean and standard error (in parentheses) for the second set of parameters ($\delta = -0.15$, $\varphi = 0.98$, $\sigma_h = 0.06$) and various values of $\alpha$, model (7). The rows “Time” report the average time to convergence (in seconds) of one iteration.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 1.5$</th>
<th></th>
<th>$\alpha = 1.7$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T = 1000$</td>
<td>$T = 3000$</td>
<td>$T = 5000$</td>
<td></td>
<td>$T = 1000$</td>
<td>$T = 3000$</td>
<td>$T = 5000$</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.5454</td>
<td>1.5119</td>
<td>1.5041</td>
<td></td>
<td>1.7083</td>
<td>1.6826</td>
<td>1.6992</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3568)</td>
<td>(0.3216)</td>
<td>(0.2984)</td>
<td></td>
<td>(0.3546)</td>
<td>(0.3111)</td>
<td>(0.2743)</td>
<td></td>
</tr>
<tr>
<td>$\delta$</td>
<td>-0.2747</td>
<td>-0.2252</td>
<td>-0.1961</td>
<td></td>
<td>-0.2235</td>
<td>-0.1614</td>
<td>-0.1604</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5756)</td>
<td>(0.3133)</td>
<td>(0.2757)</td>
<td></td>
<td>(0.3240)</td>
<td>(0.1163)</td>
<td>(0.1046)</td>
<td></td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.9616</td>
<td>0.9626</td>
<td>0.9663</td>
<td></td>
<td>0.9691</td>
<td>0.9779</td>
<td>0.9783</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0753)</td>
<td>(0.0695)</td>
<td>(0.0662)</td>
<td></td>
<td>(0.0437)</td>
<td>(0.0145)</td>
<td>(0.0132)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>0.0950</td>
<td>0.0806</td>
<td>0.0712</td>
<td></td>
<td>0.0754</td>
<td>0.0667</td>
<td>0.0667</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1181)</td>
<td>(0.0935)</td>
<td>(0.0854)</td>
<td></td>
<td>(0.0570)</td>
<td>(0.0466)</td>
<td>(0.0437)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>13.969</td>
<td>46.351</td>
<td>61.503</td>
<td></td>
<td>11.819</td>
<td>36.843</td>
<td>53.529</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 1.9$</th>
<th></th>
<th>$\alpha = 1.95$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T = 1000$</td>
<td>$T = 3000$</td>
<td>$T = 5000$</td>
<td></td>
<td>$T = 1000$</td>
<td>$T = 3000$</td>
<td>$T = 5000$</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.7783</td>
<td>1.8053</td>
<td>1.8468</td>
<td></td>
<td>1.8020</td>
<td>1.8373</td>
<td>1.8762</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3204)</td>
<td>(0.2632)</td>
<td>(0.2134)</td>
<td></td>
<td>(0.3060)</td>
<td>(0.2418)</td>
<td>(0.1877)</td>
<td></td>
</tr>
<tr>
<td>$\delta$</td>
<td>-0.2761</td>
<td>-0.1715</td>
<td>-0.1614</td>
<td></td>
<td>-0.2961</td>
<td>-0.1782</td>
<td>-0.1631</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3913)</td>
<td>(0.0925)</td>
<td>(0.0670)</td>
<td></td>
<td>(0.3532)</td>
<td>(0.0999)</td>
<td>(0.0674)</td>
<td></td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.9628</td>
<td>0.9770</td>
<td>0.9784</td>
<td></td>
<td>0.9603</td>
<td>0.9761</td>
<td>0.9782</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0537)</td>
<td>(0.0121)</td>
<td>(0.0087)</td>
<td></td>
<td>(0.0461)</td>
<td>(0.0132)</td>
<td>(0.0089)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>0.0639</td>
<td>0.0574</td>
<td>0.0586</td>
<td></td>
<td>0.0617</td>
<td>0.0556</td>
<td>0.0560</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0448)</td>
<td>(0.0305)</td>
<td>(0.0248)</td>
<td></td>
<td>(0.0426)</td>
<td>(0.0273)</td>
<td>(0.0220)</td>
<td></td>
</tr>
</tbody>
</table>

are much more precise, whereas the additional uncertainty carried by the heavy-tailed noise is discounted in larger standard errors for the estimator of $\alpha$.

3.3 An empirical application

In this subsection, we will apply the models we have introduced to the analysis of two currency crises. This is a rather new type of application for $\alpha$-stable distributions, but we see it as very illustrative. A recent paper by Hartmann, Straetmans & de Vries (2004) actually points out that, when the exchange rate fundamentals are heavy-tailed, currency crises tend to spread across different countries. Although their analysis does not involve any kind of nonlinearity in returns, it shall nevertheless be considered as a good argument supporting the use of heavy-tailed distributions, as most of the currency crises that targeted a specific country eventually spread to its neighbors.

From a practical perspective, the patterns of volatility and returns generated during currency crises are very interesting: in some cases, monetary authorities who let the exchange rate float inside a certain band are forced to switch to a free floating regime, and it may take quite a while before the monetary authority can attempt to enforce a
new floating band. In this case, abandoning a managed floating regime has a strong impact of volatility, and such a shock will be absorbed as soon as the new floating band is adopted. In other occasions, the exchange rate can be allowed to float in a largely free fashion, and a currency crisis will result in a currency depreciation. This hasn’t necessarily an impact on volatility and the “depreciation return” may look like a single spike. A crisis of the first type may therefore generate a return and volatility pattern quite similar to that of model (7), whereas a spike in the returns without subsequent impact on volatility can be produced by model (6).

In January 1999, the Central Bank of Brazil was forced to abandon a managed depreciation regime by a speculative attack triggered by a 90-days debt moratorium announced by a provincial governor. The Central Bank also tried unsuccessfully to enforce a new and wider managed floating band, but was eventually forced to let its currency float freely; it took several months before the exchange rate stabilized (cf. Figure 4). The pattern of returns and volatility is therefore similar to that generated by model (7).

As an example of a crisis of the second type, we will consider the exchange rate of the British Pound against the US Dollar from January 1, 1990 to December 31, 1994. During September 1992, a speculative attack targeted the British Pound and some other European currencies, eventually forcing the monetary authorities to abandon the European crawling peg and to depreciate their currencies. If we consider the exchange rate the British Pound versus the US Dollar (Figure 5), however, no increase in volatility appears, as the Pound was already free floating. The depreciation is therefore a one-time event, and can be accounted for by model (6).

We have then estimated models (7) and (6) for both countries, and the results confirm our intuitions. As Brazil (Table 4) is concerned, the estimation of model (7) yields \( \hat{\alpha} = 1.71 \), whereas for model (6) we have \( \hat{\alpha} = 1.912 \): this confirms the presence of heavy-tailed behavior in the volatility innovations rather than in the returns. As the persistence of volatility is concerned, we observe that not allowing for heavy tails in the volatility equation yields a \( \hat{\varphi} = 0.999 \) which is very close to the stationarity bound; if model (7) is considered, however, the estimated value reduces to \( \hat{\varphi} = 0.973 \).
jumps in the volatility are not allowed, we also remark that the estimate of the parameter $\sigma_h$ is significantly higher, since the Gaussian evolution of the volatility requires a higher variance to accommodate for the dynamics. The estimation of the auxiliary model confirms indeed the results: we have obtained a strong degree of persistence ($\beta + \psi = 0.9603$) and quite heavy tails ($\eta = 0.2212$ corresponding to 4.5 degrees of freedom).

The situation is reversed for the British Pound (Table 5). In this case, there is no visual evidence of jumps in the volatility and it is therefore not surprising that $\hat{\alpha} = 1.999$ for model (7). Instead, when one considers model (6), the spike corresponding to the speculative attack is accounted for by $\hat{\alpha} = 1.621$. It is also interesting to remark that the auxiliary model yields heavier tails than for Brazil ($\eta = 0.4181$ corresponding to 2.4 degrees of freedom). This may appear counterintuitive, but it has to be remarked that, for GARCH models, it is impossible to distinguish jumps in the returns from jumps in the volatility.

4 Conclusions

In this paper, we have considered two different stochastic volatility models that allow for $\alpha$-stable innovations in, respectively, the returns equation and the volatility equation. An indirect estimation approach has been proposed and its properties have been examined in a simulation study. The models under consideration imply two very different patterns of volatility, and we have considered their application to two typologies of currency crises. Of course, the example we report should not be considered as evidence in favor of the very simple stochastic mechanism implied by the SV model with respect to more complex and somehow structural approaches. We just point out that, allowing for $\alpha$-stable innovations in the returns or in the volatility equation may yield reasonable patterns. A similar issue is in a sense addressed by Davidson (2004), who shows that patterns very similar to those observed during currency crises can be obtained by allowing a sequence of shocks in a long-memory HYGARCH model.

In our analysis, we have considered jumps either in the state or in the observable
Table 4: Estimate and standard error (in parentheses) of the parameters of models (7) and (6) and the GARCH(1,1) auxiliary models for the returns of the exchange rate of the Brazilian Real versus the US Dollar; January 1, 1998 to December 31, 2004.

<table>
<thead>
<tr>
<th>Model (7), heavy-tailed volatility</th>
<th>α</th>
<th>δ</th>
<th>ϕ</th>
<th>σ_h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.7105</td>
<td>-0.0941</td>
<td>0.9731</td>
<td>0.1090</td>
</tr>
<tr>
<td></td>
<td>(0.3248)</td>
<td>(0.1104)</td>
<td>(0.0360)</td>
<td>(0.0759)</td>
</tr>
<tr>
<td>Model (6), heavy-tailed returns</td>
<td>α</td>
<td>δ</td>
<td>ϕ</td>
<td>σ_h</td>
</tr>
<tr>
<td></td>
<td>1.9123</td>
<td>0.0018</td>
<td>0.9985</td>
<td>0.1908</td>
</tr>
<tr>
<td></td>
<td>(0.0356)</td>
<td>(0.0062)</td>
<td>(0.0514)</td>
<td>(0.0224)</td>
</tr>
<tr>
<td>Auxiliary Model</td>
<td>η</td>
<td>ω</td>
<td>ψ</td>
<td>β</td>
</tr>
<tr>
<td></td>
<td>0.2212</td>
<td>0.0001</td>
<td>0.0999</td>
<td>0.8604</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.1243)</td>
<td>(0.1172)</td>
<td>(0.0698)</td>
</tr>
</tbody>
</table>

Table 5: Estimate and standard error (in parentheses) of the parameters of models (7) and (6) for the returns of the exchange rate of the British Pound versus the US Dollar; January 1, 1990 to December 31, 1994.

<table>
<thead>
<tr>
<th>Model (7), heavy-tailed volatility</th>
<th>α</th>
<th>δ</th>
<th>ϕ</th>
<th>σ_h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.9997</td>
<td>-1.1821</td>
<td>0.5582</td>
<td>0.7151</td>
</tr>
<tr>
<td></td>
<td>(0.0975)</td>
<td>(0.2217)</td>
<td>(0.0785)</td>
<td>(0.1031)</td>
</tr>
<tr>
<td>Model (6), heavy-tailed returns</td>
<td>α</td>
<td>δ</td>
<td>ϕ</td>
<td>σ_h</td>
</tr>
<tr>
<td></td>
<td>1.6558</td>
<td>0.0062</td>
<td>0.9958</td>
<td>0.0926</td>
</tr>
<tr>
<td></td>
<td>(0.1278)</td>
<td>(0.4008)</td>
<td>(0.2569)</td>
<td>(0.4695)</td>
</tr>
<tr>
<td>Auxiliary Model</td>
<td>η</td>
<td>ω</td>
<td>ψ</td>
<td>β</td>
</tr>
<tr>
<td></td>
<td>0.4184</td>
<td>0.0008</td>
<td>0.0021</td>
<td>0.9362</td>
</tr>
<tr>
<td></td>
<td>(0.0053)</td>
<td>(0.0262)</td>
<td>(0.1064)</td>
<td>(0.0224)</td>
</tr>
</tbody>
</table>
A natural extension will be, as suggested also by Eraker et al. (2003), to allow for jumps in both; this could be accomplished by using bivariate $\alpha$-stable distributions. In actual facts, since the normal distribution is a particular case of $\alpha$-stable distribution, such a model would nest both a standard Gaussian stochastic volatility framework and the models we have considered in this paper. Representing the vector $[v_t, w_t]$ by means of a bivariate $\alpha$-stable distribution would also allow to include correlation among the two innovations, as proposed for example by Jacquier et al. (2004) in the setting of $t$-distributed innovations. Simulating random numbers from a bivariate $\alpha$-stable distribution is straightforward (Modarres & Nolan 1994), therefore an indirect estimation approach could be fruitful also in this case. One possible shortcoming of this approach could be that the tail-thickness parameter $\alpha$ has to be the same for both terms of the noise: this could yield unrealistic results as jumps in the volatility are usually more frequent than in returns. Another possible approach could be therefore to employ independent noise terms and induce correlation by means of copula functions. In both cases, however, the specification of an appropriate auxiliary model might be less immediate. This will be the subject of future research.

References


Lumsdaine, R. L. (1996), ‘Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models’, *Econometrica* **64**, 575–596.


