Likelihood-based comparison of stable Paretian and competing models: Evidence from daily exchange rates

Efthymios G. Tsionas

Department of Economics
Athens University of Economics and Business
76 Patission Street, 104 34 Athens, Greece.
Tel. (++301) 8203 392, Internet: tsionas@aueb.gr

Abstract

Considering alternative models for exchange rates has always been a central issue in applied research. Despite this fact, formal likelihood-based comparisons of competing models are extremely rare. In this paper, we apply the Bayesian marginal likelihood concept to compare GARCH, stable, stable GARCH, stochastic volatility, and a new stable Paretian stochastic volatility model for seven major currencies. Inference is based on combining Monte Carlo methods with Laplace integration. The empirical results show that neither GARCH nor stable models are clear winners, and a GARCH model with stable innovations is the model best supported by the data.

Keywords: Stable distribution; Bayesian analysis; Unit roots; GARCH; Stochastic volatility; Exchange rates.

JEL codes : C11, C13

1. Introduction

Since Mandelbrot (1963) introduced the stable Paretian distributions to explain the behavior of stock market price changes, interest on estimation, inference, and comparison with competing models has grown rapidly, and still remains an active area of research. Several authors claim that other *iid* models may be more appropriate (see for example Blattberg and Gonedes, 1974, and Tucker, 1992) while others use the stability-under-addition test (see Hall, Brorsen and Irwin, 1989), the Paretian tail index (Akrigay and Booth, 1988, and Jansen and de Vries, 1991) or empirical moments (Lau, Gribbin, and Wingender, 1990) to conclude that the stable Paretian hypothesis is not appropriate for their data.

Lau and Lau (1993) and McCulloch (1997) rightly reply that such tests are not powerful enough to claim with reasonable certainty that the stable Paretian hypothesis is wrong. On the other hand, recent competitors of stable specifications include the GARCH family of models. The situation becomes more complicated because, as Ghose and Kroner (1995) argue, many of the properties of stable models are shared by GARCH models, implying that volatility clustering could cause many of the findings of fat-tailed stable distributions. Groenendijk, Lucas, and de Vries (1995) show that stable and stationary ARCH processes are partially nested with respect to tail shapes, thus reinforcing the conclusions of Ghose and Kroner (1995).

Formal comparison of stable with ARCH-like processes is a relatively new idea. Koutrouvelis and Meintanis (1999) use a discrimination test due to Csorgo (1987) based on the empirical characteristic function (ecf), and apply the procedures to daily and monthly stock returns, as well as certain financial ratios. Although the ecf, the Paretian index, the stability-under-addition test or the behavior of empirical moments leads to convenient tests, likelihood-based inference and testing would be preferable on the grounds of asymptotic efficiency. Although likelihood-based inference is typical for GARCH models, this is not the case for the stable model because the density is not available in closed form. For the symmetric stable Paretian model, the problems are presented in Tsionas (1999) where Bayesian Monte Carlo inference methods are presented. It is clear that on grounds of asymptotic efficiency, we cannot truly trust tests that are not based on (even approximate) likelihood functions, and the issue of model comparison will not be settled unless likelihood tests are developed and applied to real life data sets. For these reasons, formal likelihood-based comparison of GARCH and stable models has not been considered in previous research.

One may argue that it is always possible to use an encompassing model like the stable-GARCH of Liu and Brorsen (1995), and choose the stable model if certain zero restrictions on the GARCH parameters are not rejected by the data. The problem with such tests is that some GARCH parameters may be statistically significant (thus rejecting the stable Paretian hypothesis in its pure form) but this does not imply that the stable model is rejected in terms of fit: Indeed, the stable model is more parsimonious relative to a stable-GARCH, so one cannot reject *a priori* the possibility that it finally performs better in terms of probabilities to observe the data. Such considerations, although reasonable, cannot be accounted for by existing tests, and require examining the likelihood function.

The purpose of this paper is to take up inference in stable Paretian models using likelihood-based models, and use the likelihood function to develop model comparison with the GARCH model, the stochastic volatility model, the stable GARCH model, and a new stable stochastic volatility model. For the stochastic volatility, and the new stable stochastic volatility model, Monte Carlo methods have to be used to compute the likelihood function because of the presence of $T-$dimensional integrals with respect to the latent volatilities, where T is the (potentially very large) number of observations. For the stable stochastic volatility model proposed in this paper, two difficulties arise: The fact that the stable density function is not available in closed form, and the fact that latent volatilities have to be integrated out of the likelihood function using numerical methods.

Inference procedures are based on Bayesian methods. This suggests a natural way to approach the problem of model comparison, namely the computation and presentation of marginal likelihood, and posterior odds ratios. Although Monte Carlo methods have to be used to evaluate the likelihood functions of stochastic volatility or stable stochastic volatility models, the estimation procedure is made practical by restricting the use of Monte Carlo methods to computation of multivariate integrals with respect to latent volatilities. Inference for the structural parameters is subsequently taken up using Laplace integration to compute marginal posterior moments. This combination of Laplace integration with Monte Carlo methods, results in practical computations that require only maximization of likelihood functions, it can be used routinely even for large data sets, as well as minimal programming efforts. It should be noted that Markov Chain Monte Carlo methods could be used to obtain marginal posterior moments. For the stochastic volatility model, these methods are explored in Jacquier, Polson and Rossi (1994) and Geweke (1996). For the symmetric stable Paretian model, see Tsionas (1999). Such methods are relatively difficult to program or use routinely in practice, and most often require fine-tuning for each data set. On the contrary, Monte Carlo-Laplace integration is straightforward, and requires only the ability to code the likelihood function, draw random numbers, and maximize numerically the simulated likelihood. Relative to Liu and Brorsen (1995), the Bayesian methods proposed in this paper provide accurate approximations to small-sample posterior distributions, and make it unnecessary to rely upon asymptotics.

Model comparison and Bayesian inference methods are applied to daily data of seven major exchange rate series. An important issue in applied research is whether exchange rate series contain a unit root. For Bayesian analysis of unit roots in exchange rates, see Schotman and van Dijk (1991), and Tsionas (1999). The remainder of the paper is organized as follows. The econometric models are presented in section 2. Bayesian inference procedures are described in section 3. The results are analyzed in section 4. The final section offers some concluding remarks.

2. The models

Let y_t denote the logarithmic change of an exchange rate series, *i.e.* $y_t = \log(S_t / S_{t-1})$ where S_t denotes the rate of a currency against a base, for example the U.S. dollar. Following standard practice, it is assumed that y_t follows a first order autoregressive scheme. The **GARCH** model (Bollerslev, 1986, and Bollerslev, Chou and Kroner, 1992) is given by

$$y_t = \mu + \rho y_{t-1} + u_t \tag{1}$$

$$u_t \sim N(0, h_t) \tag{2}$$

$$h_t = \psi_0 + \psi_1 h_{t-1} + \psi_2 u_{t-1}^2 \tag{3}$$

where the ψ_i's are constrained to be nonnegative. The major competing model is a symmetric stable process for the innovations u_t. Specifically, we have equation (1) where u_t has a **stable Paretian distribution** with characteristic exponent α, and scale parameter σ. Although stable densities do not have closed form expressions, stable random variables may be conveniently described by their characteristic function

$$\varphi(t) = \exp\{imt - |\sigma t|^{\alpha} + i\beta \operatorname{sgn}(t)\omega(\alpha,t)\} \tag{4}$$

where

$$\omega(\alpha,t) = \begin{cases} \tan(\pi\alpha/2), & \text{for } \alpha \neq 1 \\ \frac{2}{\pi}\log|t|, & \text{for } \alpha = 1 \end{cases} \tag{5}$$

The parameters m and σ are location and scale parameters respectively, $\alpha \in (0,2]$ is the characteristic exponent, and $\beta \in [-1,1]$ is a skewness parameter. Parameter σ should not be confused with the standard deviation of a normal distribution. The shape parameter α measures tail thickness, which increases as α decreases from $\alpha = 2$, which corresponds to the normal distribution. For $\alpha = 1$, and $\beta = 0$ we obtain the Cauchy distribution. In this study, we will consider only the symmetric stable class, which corresponds to $\beta = 0$. The standard form obtains if we set $m = 0$, and $\sigma = 1$, in which case the characteristic function is simply $\varphi(t;\alpha) = \exp\{-|t|^{\alpha}\}$. The family of symmetric stable Paretian laws in standard form will be denoted by $S(\alpha)$. The density function can be obtained via Cramer's inversion theorem, and will be denoted by

$$f(u;\alpha) = \int_{-\infty}^{\infty} \exp(-itu)\varphi(t;\alpha)dt \tag{6}$$

A possible encompassing model is a GARCH model with stable innovations, namely in equation (1) u_t has a symmetric stable distribution with characteristic exponent α, and scale parameter h_t which is given by (3). This is called the **SGARCH** model, proposed by Liu and Brorsen (1995). An alternative to the GARCH model is the stochastic volatility process (Jacquier, Polson and Rossi, 1994) given by

$$y_t = \mu + \rho y_{t-1} + u_t \tag{7}$$

$$u_t \sim N(0, h_t) \tag{8}$$

$$h_t = \delta + \phi h_{t-1} + \varepsilon_t \tag{9}$$

$$\varepsilon_t \sim IN(0, \sigma_{\varepsilon}^2) \tag{10}$$

We call this the **SV** model. A generalization is provided if we allow the innovation u_t to be distributed according to a stable symmetric distribution with characteristic exponent α, and scale parameter h_t, which is described by a lognormal stochastic volatility model as in equations (6) and (7). This is the stable stochastic volatility model or **SSV** given by the following:

$$y_t = \mu + \rho y_{t-1} + u_t \tag{11}$$

$$u_t \mid h_t = h_t^{1/2} \xi_t \tag{12}$$

$$\xi_t \sim S(\alpha) \tag{13}$$

$$\log h_t = \delta + \phi \log h_{t-1} + \varepsilon_t \tag{14}$$

$$\varepsilon_t \sim IN(0, \sigma_\varepsilon^2) \tag{15}$$

Relative to Liu and Brorsen (1995) this model is more general, since volatility is allowed to be a random variable. The two models have the same number of parameters (instead of ψ_2 the SSV model has the new parameter σ_ε so the nesting can only be partial) but the SSV model allows for a higher degree of leptokurtosis, in the same manner that an SV model allows for thicker tails relative to a GARCH model.

3. Inference techniques

Estimation and inference techniques are based on Bayesian methods. The parameter vector is denoted by $\boldsymbol{\theta}$ and the data by $\mathbf{Y}$. The likelihood function of a given model is denoted by $L(\boldsymbol{\theta}; \mathbf{Y})$. The parameters are, of course, model specific but parameters like μ and ρ have a similar interpretation across models. Given a prior distribution $p(\boldsymbol{\theta})$, the posterior distribution is given by Bayes' theorem as

$$p(\boldsymbol{\theta} \mid \mathbf{Y}) \propto L(\boldsymbol{\theta}; \mathbf{Y}) p(\boldsymbol{\theta}) \tag{16}$$

The central objects of interest are marginal posterior moments of the form

$$E\left[\mathbf{f}(\boldsymbol{\theta}) \mid \mathbf{Y}\right] = \frac{\int_{\boldsymbol{\Theta}} g(\boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta}}{\int_{\boldsymbol{\Theta}} p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta}} \tag{17}$$

where $g(\boldsymbol{\theta})$ is a vector function of the parameters, and $\boldsymbol{\Theta}$ is the parameter space. The integral in the denominator is the marginal likelihood of the data. For models like GARCH, stable or stable-GARCH, these integrals cannot be evaluated analytically. For the SV and SVGARCH or SSV models we have the additional complexity that the likelihood function cannot be expressed in closed form due to the presence of $T-$dimensional integrals with respect to the latent variances, where T denotes the dimensionality of $\mathbf{Y}$. Therefore, we have to resort to numerical methods of inference. In this paper, the Laplace integration procedure is used (Tierney and Kadane, 1986 and Tierney, Kass and Kadane, 1989). If we define

$$-T \cdot h(\boldsymbol{\theta}) = \log p(\boldsymbol{\theta}) + \log L(\boldsymbol{\theta}; \mathbf{Y}) \tag{18}$$

and

$$-T \cdot h^{*}(\boldsymbol{\theta}) = \log p(\boldsymbol{\theta}) + \log L(\boldsymbol{\theta}; \mathbf{Y}) \tag{13}$$

for a positive function of interest $g(\boldsymbol{\theta})$, then the Laplace approximation to the posterior expectation

$$E\left[g(\boldsymbol{\theta}) \mid \mathbf{Y}\right] = \frac{\int g(\boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta}}{\int p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta}} \tag{19}$$

is given by

$$\widetilde{E}\left[g(\boldsymbol{\theta}) \mid \mathbf{Y}\right] = \left(\frac{|\widetilde{\boldsymbol{\Sigma}}^{*}|}{|\widetilde{\boldsymbol{\Sigma}}|}\right)^{-1/2} \exp\left\{-T \cdot \left[h^{*}(\widetilde{\boldsymbol{\theta}}^{*}) - h(\widetilde{\boldsymbol{\theta}})\right]\right\} \tag{20}$$

where $\widetilde{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}^{*}$ are the maximizers of $-T \cdot h(\boldsymbol{\theta})$ and $-T \cdot h^{*}(\boldsymbol{\theta})$ respectively, $\widetilde{\boldsymbol{\Sigma}} = T \cdot \frac{\partial^2 h(\widetilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$, and $\widetilde{\boldsymbol{\Sigma}}^{*} = T \cdot \frac{\partial^2 h(\widetilde{\boldsymbol{\theta}}^{*})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$. Also, for the approximation of marginal posterior densities we have the following result. If $\boldsymbol{\theta} = [\gamma, \boldsymbol{\lambda}]$ then

$$\widetilde{p}(\gamma \mid \mathbf{Y}) \propto \left| -\frac{\partial^2 h_{\rho}(\widetilde{\boldsymbol{\lambda}}_{\gamma})}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}'} \right|^{-1/2} \cdot \exp\left\{-T \cdot h_{\gamma}(\widetilde{\boldsymbol{\lambda}}_{\gamma})\right\} \tag{21}$$

where $-Th_\gamma(\boldsymbol{\lambda}) = \log p(\gamma, \boldsymbol{\lambda}) + \log L(\gamma, \boldsymbol{\lambda}; \mathbf{Y})$, and $\tilde{\boldsymbol{\lambda}}_\gamma$ is the maximizer of this function with respect to $\boldsymbol{\lambda}$ for fixed γ. The approximation of log-marginal likelihood is based on the following simple idea. Since

$$p(\boldsymbol{\theta} \mid \mathbf{Y}) = K(\mathbf{Y})^{-1} L(\boldsymbol{\theta}; \mathbf{Y}) p(\boldsymbol{\theta}) \tag{22}$$

where $K(\mathbf{Y})$ is the normalizing constant of the posterior distribution, and $\int_{\boldsymbol{\Theta}} p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta} = 1$, we have $K(\mathbf{Y}) = \dfrac{L(\boldsymbol{\theta}; \mathbf{Y}) p(\boldsymbol{\theta})}{p(\boldsymbol{\theta} \mid \mathbf{Y})}$ identically in $\boldsymbol{\theta}$. Therefore,

$$K(\mathbf{Y}) = \frac{L(\boldsymbol{\theta}^*; \mathbf{Y}) p(\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}^* \mid \mathbf{Y})} \tag{23}$$

for some point $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$. In large samples, the denominator can be approximated by a normal density, namely

$$p(\boldsymbol{\theta}^* \mid \mathbf{Y}) \approx (2\pi)^{-m/2} \mid \overline{\boldsymbol{\Sigma}} \mid^{-1/2} \exp\left\{-\tfrac{1}{2}(\boldsymbol{\theta}^* - \overline{\boldsymbol{\theta}})' \overline{\boldsymbol{\Sigma}}^{-1} (\boldsymbol{\theta}^* - \overline{\boldsymbol{\theta}})\right\} \tag{24}$$

where m is the dimensionality of the parameter space, $\overline{\boldsymbol{\theta}} = \int_{\boldsymbol{\Theta}} \boldsymbol{\theta} p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta}$ is the posterior mean, and $\overline{\boldsymbol{\Sigma}} = \int_{\boldsymbol{\Theta}} (\boldsymbol{\theta} - \overline{\boldsymbol{\theta}})(\boldsymbol{\theta} - \overline{\boldsymbol{\theta}})' p(\boldsymbol{\theta} \mid \mathbf{Y}) d\boldsymbol{\theta}$ is the posterior covariance matrix. If we choose $\boldsymbol{\theta}^* = \overline{\boldsymbol{\theta}}$, we have the following simple approximation to the log-marginal likelihood:

$$\log K(\mathbf{Y}) = \log L(\boldsymbol{\theta}^*; \mathbf{Y}) + \log p(\boldsymbol{\theta}^*) - \log p(\boldsymbol{\theta}^* \mid \mathbf{Y}) \approx \log L(\overline{\boldsymbol{\theta}}; \mathbf{Y}) + \log p(\overline{\boldsymbol{\theta}}) + \tfrac{m}{2} \log(2\pi) + \tfrac{1}{2} \log \mid \overline{\boldsymbol{\Sigma}} \mid \tag{25}$$

This approximation requires only the first two posterior moments, and the ability to evaluate the log-likelihood and the log-prior kernels at the posterior mean. For the GARCH, SGARCH and the stable model these are reasonable requirements. For the GARCH model, the log-likelihood is available in closed form. For the stable model, this is not the case but the fast and accurate approximation due to McCulloch (1994) may be used to evaluate the density of symmetric stable laws.

The model with the largest value of log-marginal likelihood can be selected on the grounds of best support from the data. Posterior odds can be computed easily based on the log-marginal likelihood. Consider two models, M_1 and M_2 characterized by parameters θ_1 and θ_2, with prior distributions $p_1(\theta_1)$ and $p_2(\theta_2)$, likelihood functions $L_1(\theta_1; Y)$ and $L_2(\theta_2; Y)$, and kernel posterior distributions

$p_1(\theta_1 \mid Y)$ and $p_2(\theta_2 \mid Y)$, where $p_i(\theta_i \mid Y) \propto L_i(\theta_i \mid Y)p_i(\theta_i)$, $i = 1,2$. The log-marginal likelihood for the two models can be computed using the approximation in (25), resulting in $K_i(\mathbf{Y})$. The Bayes factor in favor of model M_1 and against model M_2 is given by

$$B_{12} = \frac{K_1(\mathbf{Y})}{K_2(\mathbf{Y})} \tag{26}$$

If we have prior information about the relative probabilities of the two models, these can be incorporated into the analysis resulting in the posterior odds ratio $POR_{12} = B_{12}\dfrac{\pi_1}{\pi_2}$, where $\pi_i = P(M_i)$ is the prior probability of model i ($i = 1,2$).

For the GARCH model, the likelihood function is given by

$$L(\boldsymbol{\theta};\mathbf{Y}) = (2\pi)^{-T/2}\prod_{t=1}^{T} h_t^{-1/2}\exp\left\{-\frac{(y_t - \mu - \rho y_{t-1})^2}{2h_t}\right\} \tag{27}$$

where $h_t = \psi_0 + \psi_1 h_{t-1} + \psi_2 u_{t-1}^2$ as in (3). Obtaining the maximum likelihood estimator, and performing the optimizations necessary for Laplace integration, is straightforward. For the stable model, the likelihood function is

$$L(\boldsymbol{\theta};\mathbf{Y}) = \sigma^{-T}\prod_{t=1}^{T} f\left(\frac{y_t - \mu - \rho y_{t-1}}{\sigma};\alpha\right) \tag{28}$$

where $f(u;\alpha)$ is the density function of symmetric stable laws in standard form. Although the density is not available in closed form, the accurate and fast approximation of McCulloch (1994) may be used to compute $f(u;\alpha)$. This approximation is based on interpolating between normal and Cauchy laws, and using spline functions for the approximating errors. Given this approximation, estimation by maximum likelihood, and Laplace integration is straightforward. The only disturbing fact is that, during the course of iterations, α may become negative or greater than its maximum possible value 2. This can be remedied by using a reparameterization that will be introduced later.

The likelihood function for the stable-GARCH model is

$$L(\boldsymbol{\theta};\mathbf{Y}) = (2\pi)^{-T/2}\prod_{t=1}^{T} h_t^{-1} f\left(\frac{y_t-\mu-\rho y_{t-1}}{h_t};\alpha\right) \tag{29}$$

where $h_t = \psi_0 + \psi_1 h_{t-1} + \psi_2 u_{t-1}^2$ as in (3). Given McCulloch's (1994) approximation for the density of standard symmetric stable laws, estimation by maximum likelihood, and Laplace integration is, again, straightforward.

For the SV and SSV models, estimation ceases to be so tractable. For the SV model, the likelihood function is

$$L(\boldsymbol{\theta};\mathbf{Y}) = (2\pi)^{-T}\sigma_\varepsilon^{-T}\int_0^\infty \ldots \int_0^\infty \prod_{t=1}^{T} h_t^{-1}\exp\left\{-\frac{(y_t-\mu-\rho y_{t-1})^2}{2h_h}\right\} h_t^{-1}\exp\left\{-\frac{(\log h_t-\delta-\phi\log h_{t-1})^2}{2\sigma_\varepsilon^2}\right\} dh_1 \ldots dh_T \tag{30}$$

To obtain the maximum likelihood estimator, a simulation approach is used resulting in the simulated maximum likelihood (SML) estimator. Let $\{h_t^{(s)}; t=1,..,T\}$ be a realization from the SV process ($s=1,..,S$), namely $\log h_t^{(s)} = \delta + \phi \log h_{t-1}^{(s)} + \varepsilon_t^{(s)}$, $t=1,..,T$, with a possible choice for the initial condition being $h_0^{(s)} = \delta/(1-\phi)$. Then we have the following simulation-based approximation to the likelihood function:

$$\widetilde{L}_S(\boldsymbol{\theta};\mathbf{Y}) = S^{-1}\sum_{s=1}^{S}(2\pi)^{-T/2}\prod_{t=1}^{T} h_t^{(s)^{-1}}\exp\left\{-\frac{(y_t-\mu-\rho y_{t-1})^2}{2h_t^{(s)}}\right\} \tag{31}$$

Provided the same set of random numbers $\{\varepsilon_t^{(s)}; t=1,..,T\}$ is used, the simulated likelihood function is a smooth function of the parameter vector, and standard methods may be used for maximization. See Danielsson (1994), and Danielsson and Richard (1993).

The same procedure may be applied for the SSV model, for which the likelihood function is

$$\widetilde{L}_S(\boldsymbol{\theta};\mathbf{Y}) = \int_0^\infty .. \int_0^\infty h_t^{-1} f\left(\frac{y_t-\mu-\rho y_{t-1}}{h_t};\alpha\right) h_t^{-1}(2\pi\sigma_\varepsilon^2)^{-1/2}\exp\left\{-\frac{(\log h_t-\delta-\phi\log h_{t-1})^2}{2\sigma_\varepsilon^2}\right\} dh_1 . dh_T \tag{32}$$

and the simulated likelihood function is

$$\widetilde{L}_S(\boldsymbol{\theta};\mathbf{Y}) = S^{-1}\sum_{s=1}^{S}\prod_{t=1}^{T} h_t^{(s)^{-1}} f\left(\frac{y_t-\mu-\rho y_{t-1}}{h_t^{(s)}};\alpha\right) \tag{33}$$

This computation is more difficult relative to the SV model, because the symmetric stable density $f(u;\alpha)$ has to be approximated. As before, this is accomplished using McCulloch's (1994) procedure.

For SML estimation, we use $S = 1{,}000$ simulations, and a conjugate gradients algorithm for numerical optimization of the simulated likelihood function. Flat priors are used on all parameters (and Jeffreys priors for scale parameters like σ and σ_ε) with the exception of GARCH parameters, which are restricted in the stationarity region, and the SV parameter ϕ, for which $|\phi| < 1$.

For the characteristic exponent, we use the reparameterization $\alpha = 2\exp(-p^2)$, with the prior $p \sim N(0,1/\underline{h}_\alpha^2)$ where $\underline{h}_\alpha$ is prior precision for the characteristic exponent parameter. This is necessitated by the fact that we need to restrict α in the interval $(0,2]$, otherwise ML or SML may, during the course of iterations, give estimates outside the possible range. Given the reasonable reparameterization $\alpha = 2\exp(-p^2)$, normal priors on the parameter p result in priors for α which are consistent with the following prior beliefs about this parameter for exchange rates provided prior precision is not excessively small: Although α is expected to be strictly less than 2, it is not far from this value, and normality receives fairly large prior weight.

Small values of prior precision result in priors heavily concentrated towards small values of α. As prior precision increases, the prior is heavily concentrated near $\alpha = 2$. I have tried several different values of prior precision (10^{-4}, 0.1, 1, 2, and 5) but posterior moments and marginal likelihood values were virtually unaffected by these choices, so results are reported only for $\underline{h}_\alpha = 10^{-4}$. This result, of course, means that the prior does not dominate the data. Some representative prior distributions are provided in Figure 1.

4. Data and empirical results

We use daily data on seven currencies, namely the Canadian dollar (CD), the French franc (FF), the German mark (DM), the Italian lira (IL), the Swiss franc (SF), the British pound (BP), and the Japanese yen (JY). The period of estimation is 1990 through 1999. The empirical results are summarized in Tables 1-5, which give posterior means and posterior standard deviations of the parameters. The most important parameter is ρ since unit root inferences are based on this parameter. More specifically, posterior moments of ρ depend critically on the stochastic specification. Rough unit root inference can be based on the posterior "t-ratio" $E(\rho \mid \mathbf{Y}) / \sqrt{Var(\rho \mid \mathbf{Y})}$ where $E(\rho \mid \mathbf{Y})$ is the posterior mean, and $\sqrt{Var(\rho \mid \mathbf{Y})}$ is the posterior standard deviation. In Tables 1-5, boldface figures denote that ρ is not compatible with asymptotic highest posterior density intervals that include zero: In these cases, the unit root hypothesis can be rejected.

Under a GARCH model, we can reject unit roots only for the IL (Table 1). Under a stochastic volatility specification, only the CD has a unit root (Table 2). For the stable model, unit roots are rejected for IL, BP and the JY (Table 3). For stable GARCH, we can additionally reject the unit root for the DM (Table 4). For the stable stochastic volatility model, we additionally reject the unit root for the FF (Table 5). Estimates of the GARCH model (Table 1), and the stable GARCH (Table 4) show substantial persistence, as $\psi_1 + \psi_2$ is fairly close to unity. For the stochastic volatility specification (Table 2), posterior means of ϕ are close to unity only for FF and DM, so there is a qualitative difference between GARCH and stochastic volatility models. Further differences emerge if we consider the SVGARCH model in Table 5, where estimates of ϕ are substantially less than unity.

Log-marginal likelihood values for model comparison are provided in Table 6. Boldface figures denote the model with the largest value of the log-marginal likelihood. It is clear that the stable GARCH model performs best for all exchange rate series, except for SF where the stochastic volatility model performs best. These results imply that the competition between GARCH and stable models is not a very productive approach to the stochastic properties of exchange rates. Although both GARCH and stable models imply unconditional leptokurtosis, and certain theoretical results imply that they can be almost equivalent in the tails, empirically it makes a great deal of difference whether GARCH or a stable model is used. Our empirical results imply that GARCH and stable models could be used in a complementary way, not as competitors.

It is interesting to point out that differences in log-marginal likelihood values are enormous across models, so the preferred model receives the entire posterior weight. In other words, the posterior probability $P(M_i \mid \mathbf{Y})$ is practically unity for the preferred model M_i, which happens to be stable-GARCH most of the time. In this case, model comparison leads to fairly clear-cut model selection. If, for example, for compare, the stable-GARCH and stable models, the Bayes factor in favor of the first, and against the second model, is at least 10^{10} for each currency. This result has some practical importance because it implies that unit root inferences can be based entirely on the posterior distribution of the autoregressive coefficient ρ without the need to use model mixing in the computation of a "final" marginal posterior distribution unconditional on model uncertainty. It also implies that the data are fairly decisive in excluding all models from further consideration with the exception of the most preferred model, stable-GARCH.

Concluding remarks

The purpose of the paper was to consider formal likelihood-based comparisons of the stable Paretian model with several competitors, namely the GARCH, the GARCH with stable innovations, the stochastic volatility model, and a new specification based on the stochastic volatility model with stable innovations. For the last two models, likelihood-based inference is complicated because the likelihood function involves multidimensional integrals with respect to the latent volatilities. For the stable stochastic volatility model, this is additionally complicated by the fact that the stable density is not available in closed form. However, Monte Carlo methods can be used to approximate the multidimensional integrals, and then apply Laplace integration techniques to compute marginal posterior moments of the structural parameters, as well as marginal likelihood, and posterior odds ratios.

Formal likelihood-based comparison of the stable and competing models can be based on marginal likelihood. This involves an advance over currently available discrimination tests, which are based on the empirical characteristic function, the behavior of empirical moments etc. Such tests, although convenient and relatively easy to use, are disturbing because they are not based on the likelihood function and, therefore, cannot be efficient.

Likelihood comparison of all five models was considered for seven major currencies. The empirical results show the clear superiority of GARCH models with stable innovations. Although conventional wisdom holds that stable and GARCH models are competitors, this does not seem to be the case in practice, at least in exchange rates. This result has a plausible theoretical interpretation: Groenendijk, Lucas, and de Vries (1995) show that stable and stationary ARCH processes are *partially* nested with respect to tail shapes. This *partial* nesting does not seem to imply a great deal of overlap between the two models, so empirically both are useful, and even more useful is a nesting model with both GARCH effects and stable disturbances. This implies a degree of leptokurtosis that exceeds the degree afforded by either the GARCH or stable model. In other words, GARCH and stable models *alone* cannot be adequate descriptions of exchange rate series, and the partial nesting excludes some empirically important features. The superiority of the stable-GARCH model is compromised only for the Swiss franc, where the stochastic volatility is the model preferred by the data.

The stable-stochastic volatility model proposed in this paper does not appear to be a winner but this is not unreasonable. Although daily exchange rates are characterized by fat tails, the degree of leptokurtosis is not so large as to require a stable-stochastic volatility model, as opposed to a stable-GARCH model. This, of course, does not imply that the stable-stochastic volatility model could not be useful in daily stock returns. Evidence in Jacquier, Polson and Rossi (1994) points to the direction that a stochastic volatility model is more appropriate than the GARCH model, so it is not impossible to expect that accounting for non-normal stable innovations might improve the fit of the stochastic volatility model.

References

Akrigay, V., and G.G. Booth, 1988, The stable-law model of stock returns, Journal of Business and Economic Statistics 6, 51-57.

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.

Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics 31, 307-327.

Bollerslev, T., R. Chou, and K.F. Kroner, 1992, ARCH modeling in finance, Journal of Econometrics 52, 5-59.

Csorgo, S., 1987, Testing for stability, in Revesz, P. et al (eds), Goodness of Fit, Colloquia Mathematica Societatis Janos Bolyai, 45, 101-132, North Holland.

Danielsson, J., 1994, Stochastic volatility in asset prices: Estimation with simulated maximum likelihood, Journal of Econometrics 64, 375-400.

Danielsson, J., and J.-F. Richard, 1993, Accelerated Gaussian importance sampler with application to dynamic latent variable models, Journal of Applied Econometrics 8, S153-S173.

Geweke, J., 1996, Simulation-based Bayesian inference for economic time series, working paper 570, Research Department, Federal Reserve Bank of Minneapolis.

Ghose D., and K.F. Kroner, 1995, The relationship between GARCH and symmetric stable processes: Finding the source of fat tails in financial data, Journal of Empirical Finance 2, 225-251.

Groenendijk, P.A., A. Lucas, and C.G. de Vries, 1995, A note on the relationship between GARCH and symmetric stable processes, Journal of Empirical Finance 2, 253-264.

Hall, J.A., B.W. Brorsen, and S.H. Irwin, 1989, The distribution of futures prices: A test of the stable Paretian and mixture of normals hypotheses, Journal of Financial and Quantitative Analysis 24, 105-116.

Jacquier. E., N.G. Polson, and P.E. Rossi, 1994, Bayesian analysis of stochastic volatility models, Journal of Business and Economic Statistics 12, 371-417.

Jansen, D.W., and C.G. de Vries, 1991, On the frequency of large stock returns: Putting booms and busts into perspective, Review of Economics and Statistics 73, 18-24.

Koutrouvelis, I.A., and S.G. Meintanis, 1999, Testing for stability based on the empirical characteristic function with applications to financial data, Journal of Statistical Computation and Simulation 64, 275-300.

Lau, H.S., D.W. Gribbin, and R. Harris, 1992, How prevalent are stable-Paretian distributed financial variables?, Decision Sciences 23, 1240-1250.

Lau H.S., and A.H.L. Lau, 1994, The reliability of the stability-under-addition test for the stable-Paretian hypothesis, Journal of Statistical Computation and Simulation 48, 67-80.

Liu, S.M., and B.W. Brorsen, 1995, Maximum likelihood estimation of a GARCH-stable model, Journal of Applied Econometrics 10, 273-285.

Mandelbrot, B., 1963, The variation of certain speculative prices, Journal of Business 36, 394-419.

McCulloch J.H., 1994, Numerical approximation of the symmetric stable distribution and density, working paper, Department of Economics, Ohio State University.

McCulloch J.H., 1997, Measuring tail thickness to estimate the stable index α: A critique, Journal of Business and Economic Statistics 15, 74-81.

Schotman P., and H. K. van Dijk, 1991, A Bayesian analysis of the unit root in real exchange rates, Journal of Econometrics 49, 195-238.

Tierney L. and J.B. Kadane 1986 Accurate approximations for posterior moments and marginal densities, Journal of the American Statistical Association 81, 82-86.

Tierney L., R.E. Kass and J.B. Kadane 1989, Fully exponential Laplace approximations to expectations and variances of nonpositive functions, Journal of the American Statistical Association 84, 710-716.

Tsionas, E.G., 1999, Monte Carlo inference in econometric models with symmetric stable disturbances, Journal of Econometrics, 88, 365-401.

Tucker, A.L., 1992, A reexamination of finite- and infinite-variance distributions as models of daily stock returns, Journal of Business and Economic Statistics 10, 73-81.

Table 1. Empirical results for GARCH model

	μ	ρ	ψ_0	ψ_1	ψ_2
CD	-.0062 (.0054)	.017 (.022)	.0008 (.0003)	.949 (.008)	.043 (.007)
FF	-.0085 (.018)	-.009 (.028)	.0035 (.0014)	.964 (.008)	.028 (.006)
DM	-.0102 (.015)	-.017 (.026)	.0043 (.002)	.961 (.008)	.029 (.006)
IL	-.0078 (.0091)	**-.054** (.022)	.0028 (.001)	.950 (.008)	.045 (.007)
SF	-.0108 (.021)	-.003 (.019)	.0081 (.003)	.955 (.009)	.030 (.006)
BP	.0053 (.012)	-.015 (.025)	.0021 (.0008)	.962 (.006)	.032 (.0051)
JY	.0012 (.013)	-.011 (.075)	.011 (.004)	.932 (.015)	.051 (.011)

Notes: Boldface figures denote that zero is excluded from the asymptotic highest posterior density interval.

Table 2. Empirical results for stochastic volatility model

	μ	ρ	δ	ϕ	σ
CD	-.007 (.005)	.008 (.019)	-1.624 (.911)	.427 (.321)	.853 (.149)
FF	-.007 (.010)	**-.040** (.018)	-.081 (.067)	.938 (.050)	.358 (.144)
DM	-.004 (.014)	**-.038** (.018)	-.013 (.007)	.989 (.006)	.146 (.035)
IL	-.004 (.009)	**-.075** (.017)	-.583 (.259)	.547 (.200)	.886 (.144)
SF	-.014 (.014)	**-.039** (.017)	-.982 (.044)	-.005 (.055)	.933 (.046)
BP	-.0012 (.009)	**-.057** (.017)	-2.017 (.395)	-.327 (.256)	1.067 (.112)
JY	-.015 (.014)	**-.046** (.019)	-.510 (.314)	.534 (.285)	.888 (.197)

Notes: Boldface figures denote that zero is excluded from the asymptotic highest posterior density interval.

Table 3. Empirical results for symmetric stable model

	α	μ	ρ	σ
CD	1.686 (.029)	-.0093 (.0064)	.0039 (.065)	.172 (.0037)
FF	1.668 (.029)	-.0061 (.047)	-.034 (.047)	.373 (.0081)
DM	1.695 (.012)	-.0042 (.012)	-.037 (.020)	.398 (.0097)
IL	1.648 (.039)	-.0047 (.008)	**-.078** (.019)	.378 (.009)
SF	1.717 (.021)	-.014 (.052)	-.035 (.041)	.443 (.008)
BP	1.576 (.036)	.0026 (.007)	**-.055** (.017)	.328 (.017)
JY	1.663 (.033)	-.017 (.019)	**-.035** (.017)	.413 (.009)

Notes: Boldface figures denote that zero is excluded from the asymptotic highest posterior density interval.

Table 4. Empirical results for stable GARCH model

	α	μ	ρ	ψ_1	ψ_2
CD	1.791 (.033)	-.0033 (.006)	.027 (.04)	.946 (.0094)	.020 (.003)
FF	1.723 (.037)	-.0095 (.008)	-.03 (.02)	.958 (.0098)	.012 (.003)
DM	1.744 (.036)	-.011 (.011)	-.**038** (.018)	.961 (.009)	.012 (.003)
IL	1.729 (.016)	-.008 (.065)	-.**075** (.011)	.953 (.0007)	.016 (.002)
SF	1.744 (.041)	-.015 (.025)	-.**032** (.018)	.965 (.009)	.010 (.002)
BP	1.692 (.038)	.0077 (.007)	-.**039** (.018)	.959 (.009)	.013 (.003)
JY	1.710 (.036)	-.0181 (.013)	-.**042** (.021)	.955 (.009)	.011 (.002)

Notes: Estimates for ψ_0 are omitted. Boldface figures denote that zero is excluded from the asymptotic highest posterior density interval.

Table 5. Empirical results for stable stochastic volatility model

	α	μ	ρ	δ	ϕ	σ_ε
CD	1.706 (.03)	-.009 (.005)	.004 (.017)	-1.41 (.05)	.599 (.02)	.147 (.77)
FF	1.968 (.01)	-.008 (.011)	-.**043** (.01)	-.741 (.05)	.635 (.25)	.799 (.25)
DM	1.986 (.01)	-.004 (.014)	-.**040** (.01)	-.697 (.06)	.629 (.03)	.735 (1.7)
IL	1.974 (.02)	-.004 (.009)	-.**075** (.01)	-.722 (.06)	.635 (.03)	.790 (1.8)
SF	1.984 (.01)	-.011 (.01)	-.034 (.02)	-.689 (.06)	.589 (.03)	.745 (1.7)
BP	1.954 (.03)	.0015 (.009)	-.**061** (.02)	-.814 (.04)	.633 (.02)	.842 (.02)
JY	1.976 (.03)	-.015 (.013)	-.**046** (.02)	-.653 (.74)	.634 (.41)	.776 (.41)

Notes: Boldface figures denote that zero is excluded from the asymptotic highest posterior density interval.

Table 6. Log-marginal likelihood values

	GARCH	SV	Stable	Stable GARCH	Stable SV
CD	-397.59	-405.93	-431.27	**-345.9**	-445.86
FF	-2301.	-2231.9	-2263.5	**-2223.4**	-2243.9
DM	-2383.	-2351.7	-2379.8	**-2345.**	-2361.5
IL	-2302.	-2291.7	-2324.	**-2227.7**	-2306.7
SF	-2633.5	**-2580.4**	-2609.2	-2585.7	-2594.4
BP	-2085.7	-2059.	-2096.	**-2011.**	-2077.9
JY	-2631.6	-2507.4	-2525.1	**-2471.5**	-2517.9

Note: Results for stable, stable GARCH and stable SV are given for $\underline{h}_\alpha = 10^{-4}$. Boldface values denote the model with the largest log-marginal likelihood value.

Figure 1. Prior distributions of characteristic exponent