Liesenfeld, Roman; Jung, Robert C.

Working Paper

Stochastic volatility models: Conditional normality versus heavy tailed distributions

Tübinger Diskussionsbeiträge, No. 103

Provided in Cooperation with:
University of Tuebingen, Faculty of Economics and Social Sciences, School of Business and Economics

Suggested Citation: Liesenfeld, Roman; Jung, Robert C. (1997) : Stochastic volatility models: Conditional normality versus heavy tailed distributions, Tübinger Diskussionsbeiträge, No. 103, Eberhard Karls Universität Tübingen, Wirtschaftswissenschaftliche Fakultät, Tübingen

This Version is available at:
https://hdl.handle.net/10419/104916

Wirtschaftswissenschaftliche Fakultät
der Eberhard-Karls-Universität Tübingen

Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions

Roman Liesenfeld und Robert Jung

Tübinger Diskussionsbeiträge

Wirtschaftswissenschaftliche Fakultät

der Eberhard-Karls-Universität Tübingen

Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions

Roman Liesenfeld und Robert Jung

roman.liesenfeld@uni-tuebingen.de

robert.jung@uni-tuebingen.de

Diskussionsbeitrag Nr. 103
September 1997

Results of this paper are related to a project which is financially supported by the Deutsche Forschungsgemeinschaft (DFG). The authors would like to thank Walter Krämer, Martin Kukuk and Gerd Ronning for helpful comments. This paper has been presented at the Econometric Society meeting in Toulouse, August 1997.

Wirtschaftswissenschaftliches Seminar
Mohlstraße 36, D-72074 Tübingen

Abstract

Most of the empirical applications of the stochatic volatility (SV) model are based on the assumption that the conditional distribution of returns given the latent volatility process is normal. In this paper the SV model based on a conditional normal distribution is compared with SV specifications using conditional heavy-tailed distributions, especially Student's t-distribution and the generalized error distribution. To estimate the SV specifications a simulated maximum likelihood approach is applied. The results based on German stock market data reveal that the SV model with a conditional normal distribution does not adequately account for the two following empirical facts simultaneously: the leptokurtic distribution of the returns and low but slowly decaying autocorrelation function of the squared returns. It is shown that these empirical facts are more adequately captured by a SV model with a conditional heavy-tailed distribution. Finally, it turns out that the choice of the conditional distribution has systematic effects on the parameter estimates of the volatility process.

1. INTRODUCTION

The stochastic volatility (SV) model introduced by Taylor (1986) is used to account for the well documented autoregressive behavior in the volatility of financial return series. It represents an alternative to the autoregressive conditionally heteroskedastic (ARCH) model of Engle (1982) or Bollerslev's (1986) generalized ARCH (GARCH) model. The standard version of this SV model is given by

$$r_t = \mu + \exp(\lambda_t/2)u_t\ , \qquad u_t \sim i.i.d.(0,1) \tag{1a}$$

$$\lambda_t = \gamma + \delta\,\lambda_{t-1} + \nu v_t\ , \qquad v_t \sim i.i.d.\mathrm{N}(0,1)\ , \tag{1b}$$

where r_t is the return on day t and λ_t is the log volatility. The parameter μ represents the predictable part of the returns. The error processes u_t and v_t are mutually and serially independent with mean zero and unit variance. Both, u_t and v_t are unobservable. Hence λ_t, which is assumed to follow a Gaussian AR(1)-process with a persistence parameter δ, is also unobservable. For $|\delta| < 1$ the SV model is covariance stationary. The parameter ν measures the standard deviation of volatility shocks and is assumed to be greater than zero. Most of the empirical applications of the SV model are based on the additional assumption that u_t is normally distributed leading to a normal distribution for the daily returns conditional on λ_t; see for example Taylor (1986, 1994), Mahieu and Schotman (1994), Jaquier, Polson and Rossi (1994) and Kim, Shepard and Chib (1996). In the studies of Ruiz (1994) and Harvey, Ruiz and Shepard (1994) the SV model is extended to allow the conditional distribution of the returns to be more heavy-tailed than the normal distribution by using the ad-hoc assumption of a scaled Student t-distribution for u_t.

The purpose of this paper is to analyze the ability of the SV model to capture adequately the following empirical regularities of financial return series. First, the leptokurtic distribution of daily returns meaning that it is excessively peaked around zero and that it exhibits fatter tails than the corresponding normal distribution. Second, the autoregressive behavior of the volatility indicated by a typically low but very slowly decaying autocorrelation function of the squared returns. We will demonstrate that the SV model with a conditional normal distribution for the returns (SV-normal) is too restrictive to account adequately for both above mentioned empirical regularities simultaneously. Furthermore we will show that the substitution of the conditional normal distribution of the returns by a conditional heavy-tailed distribution as, for example, the Student t-distribution and the generalized error distribution (GED) can help to capture adequately both empirical regularities. Finally, it turns out that the assumption concerning the

conditional distribution of the returns affects the estimates of the parameters which governs the volatility process.

Since the latent volatility process λ_t is assumed to be serially correlated the marginal likelihood of the SV model is given by a high dimensional integral which makes the estimation by standard maximum likelihood (ML) infeasible. Hence, to estimate the SV model with different conditional return distributions we use the simulated maximum likelihood (SML) approach developed by Danielsson and Richard (1993). This estimation strategy allows to adopt the standard instruments of inference developed for the ML method.

The paper is organized as follows. Section 2 contains a brief description of the data used throughout the paper. Section 3 analyzes the theoretical predictions of the SV model with a conditional normal and conditional heavy-tailed distributions concerning the kurtosis of the returns and the autocorrelation of the squared returns. The predicted moments are compared with the corresponding empirical moments. Section 4 describes the simulated maximum likelihood (SML) estimation technique used to estimate the parameters of the SV model. Section 5 presents the results of the SML estimations of the SV model and Section 6 concludes.

2. DATA

The empirical results are based on a dataset consisting of daily closing prices for the six major German stocks listed in the DAX, the leading German stock market index: Siemens (SIE), Daimler-Benz (DAI), Volkswagen (VOW), Deutsche Bank (DBK), Bayer (BAY) and Veba (VEB). The data were obtained from the Karlsruher Kapitalmarktdatenbank (KKMDB) and are adjusted for effects of dividends and capital changes. The sample period starts on January 2, 1990 and ends on May 31, 1994 which gives a sample size of roughly 1100. The daily closing prices p_t are transformed to price changes measured in continuously compounded rates: $r_t = 100 \cdot \ln(p_t/p_{t-1})$.

The statistical properties of the data are summarized in Table 1. Notice that the kurtosis of the returns is well above three, the value associated with the normal distribution. In order to detect a serial correlation in the returns the Ljung-Box statistic for 20 lags $LB_r(20)$ is calculated. Under the null hypotheses of no autocorrelation $LB_r(20)$ is $\chi^2_{(20)}$ distributed. The results indicate that the returns are serially uncorrelated. Table 2 shows the autocorrelation coefficients of the squared returns at different lags. The autocorrelation starts at a low level indicated by a

first-order autocorrelation smaller than 0.15 and declines with increasing lags very slowly. Even for lag 40 the autocorrelation for all six stocks is greater than the asymptotic standard error given by 0.03. The Ljung-Box statistic for the squared data $LB_{r^2}(30)$, which is $\chi^2_{(30)}$ distributed under the null hypotheses of no serial correlation, demonstrates that the positive autocorrelation in the squared returns is statistically significant for all stocks.

3. DISTRIBUTIONAL AND TIME SERIES PROPERTIES OF THE SV MODEL

In this section the theoretical predictions of the SV model (1a) and (1b) with regard to the kurtosis of the returns and the autocorrelation of the squared returns are confronted with the empirical findings of Section 2.

Defining $e_t = r_t - \mu$ as the return residuals, the kurtosis κ of the unconditional distribution of the returns resulting from the SV model (1a) and (1b) is given by (for the derivation see Appendix A)

$$\kappa = \mathrm{E}(e_t^4)/\mathrm{E}(e_t^2)^2 = \mathrm{E}(u_t^4)\exp\{\sigma_\lambda^2\} \qquad \text{where} \qquad \sigma_\lambda^2 = \nu^2/(1-\delta^2)\,. \tag{2}$$

The parameter σ_λ^2 represents the unconditional variance of λ_t. Equation (2) clearly shows that the kurtosis of the unconditional return distribution consists of two components: the *baseline-kurtosis* which is due to the kurtosis of the standardized errors in the return equation $\mathrm{E}(u_t^4)$ and the kurtosis which is due to the variation in the volatility process λ_t. In the case of the conditional normality assumption for the returns, the baseline-kurtosis is given by $\mathrm{E}(u_t^4) = 3$ leading to an unconditional kurtosis of the returns of more than three. This is consistent with the observed leptokurtosis of the empirical distribution of financial return series. Notice that as $\nu \to 0$ the kurtosis of the returns predicted by the SV model converges to the baseline-kurtosis $\mathrm{E}(u_t^4)$. For $\nu = 0$ the SV model is no longer identified.

The autocorrelation function (ACF) of the squared return residuals $e_t^2 = (r_t - \mu)^2$ resulting from the SV model (1a), (1a) can be shown to be of the following form (for the derivation see Appendix A):

$$\rho(\tau) = \frac{\exp\{\sigma_\lambda^2\delta^\tau\} - 1}{\mathrm{E}(u_t^4)\exp\{\sigma_\lambda^2\} - 1}\,, \qquad \tau = 1, 2, \ldots\,. \tag{3}$$

Hence for a normally distributed error u_t and for $\delta > 0$ the SV model predicts a positive autocorrelation in the squared return residuals, which is in accordance with the observed persistence

in the volatility of returns. Furthermore, the ACF is characterized by an exponetially decaying rate determined by the parameter δ. Thus the persistence of shocks in the volatility process of the returns depends on δ.

To evaluate the ability of the SV model to account adequately for the observed high kurtosis of the returns and the autocorrelation of the squared returns starting with a low first-order autocorrelation and decaying very slowly, the theoretical ACF can be expressed as a function of the theoretical kurtosis, as proposed by Teräsvirta (1996). According to Equation (2) and (3), for the SV model this functional relationship between $\rho(\tau)$ and κ is given by

$$\rho(\tau) = \frac{\left[\kappa/\mathrm{E}(u_t^4)\right]^{\delta^\tau} - 1}{\kappa - 1} . \tag{4}$$

For a given lag τ and given values of $\mathrm{E}(u_t^4)$ and δ this function provides the theoretically attainable $\rho(\tau)/\kappa$-combinations of the SV model which can be compared with the empirical autocorrelation/kurtosis combinations for the return data. Taking $\tau = 1$ and assuming that u_t is normally distributed Equation (4) reduces to

$$\rho(1) = \frac{(\kappa/3)^\delta - 1}{\kappa - 1} . \tag{5}$$

Figure 1 shows this function for different values of δ together with the empirical first-order autocorrelation/kurtosis combinations of the six DAX stocks. Since typical estimates of δ for return series with a slowly decaying ACF of the squared data show values which are greater than 0.9, we choose $\delta = 0.9$, $\delta = 0.95$ and $\delta = 0.99$. It is obvious from Figure 1 that the empirical autocorrelation/kurtosis combinations are not attainable by the SV model with a conditional normal distribution. The empirical autocorrelation/kurtosis combinations lie in a region which is limited by $\rho(1) < 0.15$ and $\kappa > 7$. In contrast to that, the SV model with normal errors and with $\delta \geq 0.9$ predicts for a kurtosis $\kappa \geq 7$ a first-order autocorrelation $\rho(1) \geq 0.19$. On the other hand, for a first-order autocorrelation $\rho(1) \leq 0.15$ the predicted values of the kurtosis are restricted by $\kappa \leq 5.21$ if $\delta \geq 0.9$. Hence one can conclude that in the SV-normal model a low but persistent autocorrelation in the squared return residuals (i.e. $\rho(1) \leq 0.15$ and $\delta \geq 0.9$) is not compatible with a high kurtosis in the return distribution (i.e. $\kappa > 7$).

Next we consider leptokurtic distributions for the error u_t which are characterized by a kurtosis $\mathrm{E}(u_t^4) > 3$. As can be inferred from Equation (2) and (3), relative to the normal assumption for u_t, a leptokurtic distribution increases ceteris paribus the kurtosis of the returns and leads ceteris paribus to a decline in the level of the ACF. A prominent leptokurtic distribution used

in the class of GARCH models is the t-distribution (see for example Bollerslev (1987)). The density function of a t-distributed random variable u_t with mean zero and variance normalized to one is given by

$$f(u_t) = [\pi(\omega-2)]^{-1/2} \frac{\Gamma((\omega+1)/2)}{\Gamma(\omega/2)} \left[1 + \frac{u_t^2}{\omega-2}\right]^{-(\omega+1)/2}, \qquad \omega > 2\,, \tag{6}$$

where the parameter ω represents the degrees-of-freedom. As long as $\omega > 4$ the kurtosis of the t-distribution is $\mathrm{E}(u_t^4) = 3(\omega-2)/(\omega-4)$ which is greater than three if $\omega < \infty$. For $\omega \to \infty$ the t-distribution approaches a normal distribution.

Another distribution which also allows $\mathrm{E}(u_t^4)$ to be greater than three is the GED distribution described by Box and Tiao (1973). The density of a GED distributed random variable u_t with mean zero and variance one is given by

$$f(u_t) = \frac{\vartheta \exp\{-1/2|u_t/\beta|^\vartheta\}}{\beta\Gamma(1/\vartheta)2^{1+1/\vartheta}}, \qquad 0 < \vartheta < \infty\,, \tag{7}$$

with

$$\beta = \left[2^{-2/\vartheta}\frac{\Gamma(1/\vartheta)}{\Gamma(3/\vartheta)}\right]^{1/2}.$$

For $\vartheta = 2$ the GED collapses to a Normal distribution and for $\vartheta < 2$ the kurtosis given by $\mathrm{E}(u_t^4) = \Gamma(1/\vartheta)\Gamma(5/\vartheta)/[\Gamma(3/\vartheta)]^2$ is greater than three (see Johnson and Kotz (1970)).

Figure 2 shows the first-order autocorrelation/kurtosis combinations which are according to Equation (4) predicted by a SV model with a leptokurtic error distribution. In the left-hand panel a baseline-kurtosis of $\mathrm{E}(u_t^4) = 4$ is used resulting from $\omega = 10$ in the t-distribution and $\vartheta \approx 1.4$ in the GED distribution. In the right-hand panel the baseline-kurtosis is six representing either a t-distribution with $\omega = 6$ or a GED distribution with $\vartheta \approx 1$. It can be seen that in contrast to a SV model with a normal error distribution, the predicted $\rho(1)/\kappa$ combinations for a persistence parameter $\delta \geq 0.9$ and a leptokurtic error distribution are now able to cover the empirical points. An increase in the baseline-kurtosis $\mathrm{E}(u_t^4)$ moves up the predicted $\rho(1)/\kappa$ combinations towards the empirical combinations. Therefore we conclude that in order to make the SV model compatible with a high kurtosis of the returns and a low but slowly decaying ACF of the squared returns it seems to be useful to allow for a leptokurtic error distribution.

In the following sections, the SV model is estimated under the assumption of a normal error distribution as well as under the assumption of a t-distribution (SV-t) and a GED distribution (SV-GED). These different specifications are compared concerning their ability to capture the observed distributional and dynamic patterns of the return series.

4. ESTIMATION METHOD

In the SV model (1a) and (1b) the latent variable λ_t is autocorrelated. Thus in order to estimate the SV model one has to solve the problem that the likelihood function is defined by an integral whose dimension is given by the number of observations:

$$f(R \mid \theta) = \int_{\Re^T} f(R, \Lambda \mid \theta)\, d\Lambda \,, \tag{8}$$

where $f(R, \Lambda \mid \theta)$ denotes the joint probability function of the vectors $R = \{r_t\}_{t=1}^T$ and $\Lambda = \{\lambda_t\}_{t=1}^T$. The vector θ contains the parameters to be estimated. For this high dimensional integral no closed form solution exists and standard numerical integration methods can not be applied. Hence to estimate the SV model we use the SML estimator developed by Danielsson and Richard (1993). This estimation strategy relies on Monte Carlo (MC) techniques to estimate the integral (8) and can be motivated as follows. If the number of MC replications is large enough SML allows to adopt the standard instruments for inference developed for maximum likelihood. Furthermore, the application of SML for the SV model with different assumptions concerning the error distribution is straightforward, as long as the corresponding density function can be expressed in an explicit form.

In the SML approach the integral (8) is estimated using an importance sampling technique and subsequently maximized with respect to θ. Therefore the joint density $f(R, \Lambda \mid \theta)$ is factorized in an importance sampling function (IF) $\psi(\Lambda \mid R)$ and a remainder function (RF) $\phi(\Lambda, R)$ such that the equation

$$f(R, \Lambda \mid \theta) = \phi(\Lambda, R)\psi(\Lambda \mid R) \tag{9}$$

holds. The expected value of the RF evaluated over the distribution defined by the IF is given by

$$\mathrm{E}_\psi[\phi(\Lambda, R)] = \int_{\Re^T} \phi(\Lambda, R)\, \psi(\Lambda \mid R)\, d\Lambda = \int_{\Re^T} f(R, \Lambda \mid \theta)\, d\Lambda = f(R \mid \theta) \,. \tag{10}$$

Since the integral $f(R \mid \theta)$ can be expressed as the expectation $\mathrm{E}_\psi[\phi(\Lambda, R)]$ a natural estimator for $f(R \mid \theta)$ is the following sample average:

$$\hat{f}_N(R \mid \theta) = \frac{1}{N} \sum_{n=1}^{N} \phi(\Lambda_n, R) \,, \tag{11}$$

where $\{\Lambda_n\}_{n=1}^N$ represents a simulated random sample of size N drawn from the probability distribution $\psi(\Lambda \mid R)$. The precision of this estimator of the integral can arbitrarily be increased

by raising the simulation sample size N. An initial factorization of $f(R, \Lambda \mid \theta)$ according to the conditions (9) and (10) is obtained as follows

$$\psi_0(\Lambda \mid R) \equiv \prod_{t=1}^{T} f(\lambda_t \mid \lambda_{t-1}) \tag{12}$$

$$\phi_0(\Lambda, R) \equiv \prod_{t=1}^{T} f(r_t \mid \lambda_t) \,, \tag{13}$$

where $f(\lambda_t \mid \lambda_{t-1})$ is the density of the conditional distribution of λ_t given λ_{t-1}, which is according to Equation (1b) a Gaussian distribution. $f(r_t \mid \lambda_t)$ is the density of the return in period t conditional on the value of λ_t. For a SV-normal model $f(r_t \mid \lambda_t)$ is given by the density of a normal distribution with mean μ and variance $\exp\{\lambda_t\}$. In a SV-t specification $f(r_t \mid \lambda_t)$ has the form

$$f(r_t \mid \lambda_t) = [\pi(\omega-2)\exp\{\lambda_t\}]^{-1/2} \frac{\Gamma((\omega+1)/2)}{\Gamma(\omega/2)} \left[1 + \frac{(r_t-\mu)^2}{\exp\{\lambda_t\}(\omega-2)}\right]^{-(\omega+1)/2} \tag{14}$$

and in a SV-GED model

$$f(r_t \mid \lambda_t) = \frac{\vartheta \exp\{-1/2|(r_t-\mu)/(\exp\{\lambda_t/2\}\beta)|^{\vartheta}\}}{\exp\{\lambda_t/2\}\beta\Gamma(1/\vartheta)2^{1+1/\vartheta}} \,, \tag{15}$$

where β is defined in Equation (7). The initial IF (12) and the corresponding RF (13) can be used to construct a naive MC-estimate of the integral $f(R \mid \theta)$ for a given value of the parameter vector θ, as outlined in Equation (11).

As shown by Danielsson and Richard (1993) this initial factorization of $f(R, \Lambda \mid \theta)$ in the IF (12) and the RF (13) is inefficient in the sense that the resulting MC sampling variance of the estimator for the integral increases dramatically with the dimension of the integral T. To solve this inefficiency problem Danielsson and Richard (1993) proposed an acceleration method, called Accelerated Gaussian Importance Sampling (AGIS). This AGIS method search for an IF which minimizes the MC sampling variance of the corresponding RF given by

$$\mathrm{Var}_{\psi}[\phi(\Lambda, R)] = \int_{\Re^T} [\phi(\Lambda, R) - f(R \mid \theta)]^2 \psi(\Lambda \mid R)\, d\Lambda \,, \tag{16}$$

while preserving the conditions (9) and (10). The numerical solution to this minimizing problem is provided by the AGIS procedure (see Appendix B for a description of the AGIS procedure). The experiences of Danielsson and Richard (1993) showed that the application of the AGIS method to estimate a SV model for $T = 2000$ observations reduces the necessary MC sample size to $N = 50$ or less. Here the SV specifications are estimated using a simulation sample size of $N = 500$ and four iterations for the AGIS algorithm described in Appendix B.

5. EMPIRICAL RESULTS

The SML estimation results for the SV-normal model are summarized in Table 3. The estimates of the persistence parameter δ are always highly significant and lie between 0.91 (VOW) and 0.96 (DBK). This result is in accordance to those reported in the studies of, for example, Taylor (1994) and Jacquier, Polson and Rossi (1994) and indicates a high persistence of volatility shocks. The fact that the estimates of δ are greater than 0.9 implies according to Figure 1 that the combinations of the kurtosis κ and the first-order autocorrelation $\rho(1)$ predicted by the SV-normal model are not compatibel with the combinations of the data. Thus even if the implied estimates of κ are for all stocks greater than three, in nearly all cases the implied estimate is not high enough to adequately capture the observed kurtosis given in Table 1. In the case of SIE, for example, the explained kurtosis is only about a half of the observed kurtosis. On the other hand the implied estimates of $\rho(1)$ are for all stocks substantially greater than their sample counterparts given in Table 2. For example, for SIE the predicted first-order autocorrelation is 0.18 where as the empirical autocorrelation is only 0.13. Hence one can conclude that the SV-normal specification does not adequately capture the observed high kurtosis of the returns and the low first-order autocorrelation of the squared data simultaneously.

The estimation results for the SV-t model are displayed in Table 4. To test the SV-normal against the SV-t specification the likelihood-ratio statistics $LR_{1/\omega=0}$ of the Null hypotheses $1/\omega = 0$ is used. Since the value of $1/\omega$ under the Null hypotheses is on the boundary of the admissible parameter space the appropriate distribution of $LR_{1/\omega=0}$ under the Null is given by

$$LR_{1/\omega=0} \sim \frac{1}{2}\chi^2_{(0)} + \frac{1}{2}\chi^2_{(1)} ,$$

where $\chi^2_{(0)}$ is a degenerate distribution with all its mass at the origin, see Harvey (1989) and Ruiz (1994). Hence the appropriate critical value for a significance level α is the 2α quantil of the $\chi^2_{(1)}$-distribution. As can be infered from the values of $LR_{1/\omega=0}$ the SV model with normal errors is rejected against the specification with t-distributed errors at the 5% level for VOW and BAY and at the 1% level for the remaining stocks. Thus, in contrast to our results which show that the SV-t model represents a significant improvement compared to the SV-normal specification, Ruiz (1994) found only a very weak evidence in favor of the SV-t model using a quasi maximum likelihood approach to estimate the SV model for the yen/dollar exchange rate

.

In Table 5 the results for the SV-GED specification are summarized. The estimates of $1/\vartheta$ indicate that the implied kurtosis of the error u_t is greater than three for all stocks. But the

LR statistics of the hypotheses $1/\vartheta = 0.5$ shows that the SV-normal model is rejected at the 5% level against a GED distribution with $1/\vartheta \neq 0.5$ only for DAI and SIE. Since the SV-t and the SV-GED model are non-nested the classical testing procedures cannot be applied to compare these specifications. Hence we base the choice between the two specifications upon the maximized log-likelihood values $\ln L$. This criterion indicates that the SV-t specification has a better fit than the SV-GED model for all stocks. In view of these results only the SV-t model will be subject of further comparisons with the SV-normal specification.

As for the SV-normal model the estimated persistence parameters in the SV-t specification are greater than 0.9 but smaller than one. But the estimates resulting from the SV-t specification are for all stocks greater than those from the SV-normal model. Furthermore, the estimates of the variance parameter ν from the SV-t model are in all cases smaller than those from the SV-normal model. We shall return to these differences in the parameter estimates shortly. The estimated degrees-of-freedom ω vary between 6.4 (DBK) and 11.5 (VOW) and the corresponding baseline-kurtosis $\mathrm{E}(u_t^4)$ lie in a range between 5.5 (DBK) and 3.8 (VOW). Thus according to Figure 2 the increase of the baseline-kurtosis, relative to the SV-normal model, moves the predicted $\rho(1)/\kappa$ combinations towards the empirical ones. The comparison of Tables 3 and 4 shows that for all six stocks the specification with a t-distributed error leads to an increase in the implied estimates of the kurtosis κ accompanied with a decrease in the estimate for the first-order autocorrelation $\rho(1)$ relative to the SV-normal model. Except for DAI and BAY the implied estimates of both κ and $\rho(1)$ are closer to the corresponding empirical moments in the SV-t specification as compared to the SV-normal model. In the case of DAI and BAY only the estimates of $\rho(1)$ are closer to their empirical counterparts. Hence one can conclude that for four stocks substituting the normal distribution by the leptokurtic t-distribution helps to capture simultaneously the observed low first-order autocorrelation and the high kurtosis more adequately.

As noted above the SML estimates of the persistence parameter of the volatility process δ from the SV-t model are greater than those from the SV-normal model. The differences between the estimates of δ measured in standard errors from the SV-t model are 5.1 (SIE), 2.1 (DAI), 2.6 (VOW), 3.8 (DBK), 1.7 (BAY) and 2.9 (VEB). Hence except for BAY the differences can be regarded as significant. On the other hand the substitution of the conditional normal distribution by a conditional t-distribution leads to a decrease in the SML estimates of the variance parameter of the volatility ν which is in general significant. The differences are 5.4 (SIE), 2.5 (DAI), 3.3 (VOW), 4.6 (DBK), 1.8 (BAY) and 3.3 (VEB) standard errors of the SV-t model, respectively.

Hence for five stocks the choice of the conditional distribution has a significant effect on the SML estimates of the parameters in the volatility equation δ and ν which seems to be systematically. To analyze this more formally we estimated the SV-t model with predetermined values of $1/\omega$. Starting at $1/\omega = 0$ which characterizes the normal distribution, $1/\omega$ is gradually raised by one standard deviation of $1/\omega$ given in Table 4 until the value of the unrestricted estimate of this parameter is reached. Figure 3 shows the resulting combinations of the predetermined values of $1/\omega$ and the corresponding SML estimates of δ and ν. It can be seen that for all stocks the SML estimates of δ are monotonically increasing and the SML estimates of ν are monotonically decreasing in $1/\omega$, respectively. These results demonstrate that the effects of the choice between a normal and a heavy-tailed t-distribution on the SML estimates of the volatility parameters δ and ν are systematic. A possible interpretation of these effects is as follows. The use of t-distributed errors in the mean equation of the returns filters return data in the extreme tail areas which are captured in the SV-normal model by a stronger variation in the volatility process leading to a higher predicted variance parameter of the error in the volatility process ν. This predicted higher variance of the volatility in the SV-normal model dampens the autogeressive structure in the volatility process which leads to a low persistence parameter δ relative to the SV-t specification.

6. SUMMARY AND CONCLUSION

This paper analyzes the SV model with different assumptions concerning the conditional distribution of the daily returns. Three distributional assumptions are compared: The normal distribution (SV-normal), the t-distribution (SV-t) and the GED distribution (SV-GED). To estimate the SV specifications in which the unobservable volatility process is serially correlated the simulated maximum likelihood (SML) approach developed by Danielsson and Richard (1993) is applied.

The results using data from the six major German stocks can be summarized as follows. First, the SV-normal model does not adequately capture the observed high kurtosis of the returns and the low first-order autocorrelation of the squared returns simultaneously. Second, according to the likelihood-ratio test the SV-t model represents for all six stocks a significant improvement compared to the SV-normal model, whereas the SV-GED model is significantly better than the SV-normal model only for two stocks. The log-likelihood values shows that the fit of the SV-model with a t-distribution is better than with a GED-distribution. Third, for four stocks

the substitution of the normal distribution by a t-distribution helps to capture simultaneously the high kurtosis of the returns and the low first-order autocorrelation in the squared returns. Fourth, the comparison between the SV-normal and the SV-t specification shows that the choice of the conditional distribution has a significant and systematic effect on the SML estimates of the parameters which direct the volatility process. In fact, the estimates of the persistence parameter of the volatility process is systematically higher and the estimates of the variance parameter of the volatility shocks are systematically lower in the SV-t than in the SV-normal model, respectively.

APPENDIX A: MOMEMTS OF THE STOCHASTIC VOLATILITY MODEL

The specification for the return residuals $e_t = r_t - \mu$ is

$$e_t = \exp(\lambda_t/2)u_t\ , \qquad u_t \sim i.i.d.(0,1)$$

$$\lambda_t = \gamma + \delta\,\lambda_{t-1} + \nu v_t\ , \qquad v_t \sim i.i.d.\mathrm{N}(0,1)\ ,$$

where u_t and v_t are stochastically independent. For $|\delta| < 1$ the unconditional distribution of λ_t is given by $\lambda_t \sim \mathrm{N}(\mu_\lambda, \sigma_\lambda^2)$ with $\mu_\lambda = \gamma/(1-\delta)$ and $\sigma_\lambda^2 = \nu^2/(1-\delta^2)$.

The kurtosis for the unconditional distribution of the returns is

$$\kappa = \mathrm{E}[(r_t-\mu)^4]/\mathrm{E}[(r_t-\mu)^2]^2 = \mathrm{E}(e_t^4)/\mathrm{E}(e_t^2)^2\ . \tag{A.1}$$

Defining $h_t = \exp\{\lambda_t\}$, the moments $\mathrm{E}(e_t^4)$ and $\mathrm{E}(e_t^2)$ are given by

$$\mathrm{E}(e_t^4) \;=\; \mathrm{E}(h_t^2u_t^4) = \mathrm{E}(h_t^2)\mathrm{E}(u_t^4) \tag{A.2}$$

$$\mathrm{E}(e_t^2) \;=\; \mathrm{E}(h_tu_t^2) = \mathrm{E}(h_t)\mathrm{E}(u_t^2) = \mathrm{E}(h_t) \tag{A.3}$$

and the kurtosis can be written as

$$\kappa = \frac{\mathrm{E}(u_t^4)\mathrm{E}(u_t^2)}{\mathrm{E}(h_t)^2} = \mathrm{E}(u_t^4)\left(\frac{\mathrm{Var}(h_t)}{\mathrm{E}(h_t)^2}+1\right)\ . \tag{A.4}$$

The squared variation coefficient of h_t is given by $\mathrm{Var}(h_t)/\mathrm{E}(h_t)^2 = \exp\{\sigma_\lambda^2\} - 1$. Hence the kurtosis of the returns is

$$\kappa = \mathrm{E}(u_t^4)\exp(\sigma_\lambda^2)\ , \tag{A.5}$$

which corresponds to Equation (2).

The autocorrelation function of the squared residuals e_t^2 is defined by

$$\rho(\tau) = \mathrm{Cov}(e_t^2, e_{t-\tau}^2)/\mathrm{Var}(e_t^2) \ , \qquad \tau = 1, 2, \ldots . \tag{A.6}$$

The autocovariance $\mathrm{Cov}(e_t^2, e_{t-\tau}^2)$ which is independent of the distributional assumption concerning u_t is given by (see for example Jacquier, Polson and Rossi (1994), Appendix, p. 387)

$$\mathrm{Cov}(e_t^2, e_{t-\tau}^2) = \mathrm{Cov}(h_t, h_{t-\tau}) = \mathrm{E}(h_t)^2(\exp\{\sigma_\lambda^2 \delta^\tau\} - 1) \ . \tag{A.7}$$

The variance of the squared residuals is

$$\begin{aligned} \mathrm{Var}(e_t^2) &= \mathrm{E}(h_t^2 u_t^4) - \mathrm{E}(h_t u_t^2)^2 = \mathrm{E}(h_t^2)\mathrm{E}(u_t^4) - \mathrm{E}(h_t)^2 \\ &= \mathrm{E}(h_t)^2[\mathrm{E}(u_t^4)\Big(\frac{\mathrm{Var}(h_t)}{\mathrm{E}(h_t)^2} + 1\Big) - 1] \\ &= \mathrm{E}(h_t)^2[\mathrm{E}(u_t^4)\exp\{\sigma_\lambda^2\} - 1] \ . \end{aligned} \tag{A.8}$$

If the Equations (A.7) and (A.8) are inserted in Equation (A.6) the autocorrelation function is given by

$$\rho(\tau) = \frac{\exp\{\sigma_\lambda^2 \delta^\tau\} - 1}{\mathrm{E}(u_t^4)\exp\{\sigma_\lambda^2\} - 1} \ , \tag{A.9}$$

which corresponds to Equation (3).

APPENDIX B: THE ACCELERATED GAUSSIAN IMPORTANCE SAMPLER

The AGIS method uses a numerical and iterative procedure to solve the minimizing problem

$$\min_{\psi} \ \mathrm{Var}_\psi[\phi(\Lambda, R)] \tag{B.1}$$

subject to the constraints

$$f(R, \Lambda \mid \theta) = \phi(\Lambda, R)\psi(\Lambda \mid R) \qquad \text{and} \qquad \mathrm{E}_\psi[\phi(\Lambda, R)] = f(R \mid \theta) \ ,$$

where $\mathrm{Var}_\psi[\phi(\Lambda, R)]$ is given by Equation (16). Therefore a variance reduction function $\xi(\Lambda, Q)$ is defined where Q is a matrix of parameters to be determined. This variance reduction function

is used to construct a new pair of a IF and a RF by transforming the initial pair given in Equations (12) and (13) as follows:

$$\psi(\Lambda \mid R) = [\psi_0(\Lambda \mid R)\, \xi(\Lambda, Q)]/k(Q) \tag{B.2}$$

$$\phi(\Lambda, R) = [\phi_0(\Lambda, R)\, k(Q)]/\xi(\Lambda, Q)\,. \tag{B.3}$$

$k(Q)$ represents the integration constant which ensures that the new IF is a proper probability function and is given by

$$k(Q) = \int_{\Re^T} \psi_0(\Lambda \mid R)\, \xi(\Lambda, Q)\, d\Lambda\,. \tag{B.4}$$

The transformations (B.2) and (B.3) retains the constraints of the minimizing problem given above while changing the variance of $\phi(\Lambda, R)$. For the variance reduction function the AGIS method uses the following functional form

$$\xi(\Lambda, Q) = \prod_{t=1}^{T} \xi(\lambda_t, Q_t) \tag{B.5}$$

with

$$\xi(\lambda_t, Q_t) = \exp\{-\frac{1}{2}\eta_t' Q_t \eta_t\} \qquad \text{and} \qquad \eta_t' = (\lambda_t, \lambda_{t-1}, 1)\,.$$

Even if this choice for $\xi(\Lambda, Q)$ is dictated by analytical and computational reasons, it can be shown that it works exceptionally well. To determine $Q = \{Q_t\}_{t=1}^T$ and hence $\xi(\Lambda, Q)$ the following iterative procedure with j steps is applied. Starting with the initial IF $\psi_0(\Lambda \mid R)$ an initial simulated sample $\{\Lambda_{0,n}\}_{n=1}^N$ is drawn and used to run the following regression for every time period t=1, ..., T :

$$\ln \psi_0(\lambda_{0,n,t}) = a_{1,t} + b_{1,t}\lambda_{0,n,t} + c_{1,t}\lambda_{0,n,t}^2 + \text{residual} \qquad n = 1, ..., N\,. \tag{B.6}$$

Then the OLS-estimates of the coefficients are used to construct for every time period t the matrix

$$\hat{Q}_{1,t} = \begin{pmatrix} -2\hat{c}_{1,t} & 0 & -\hat{b}_{1,t} \\ 0 & 0 & 0 \\ -\hat{b}_{1,t} & 0 & -2\hat{a}_{1,t} \end{pmatrix} \qquad t = 1, ..., T\,. \tag{B.7}$$

With $\hat{Q}_1$=$\{\hat{Q}_{1,t}\}_{t=1}^T$ a first new IF is given by $\psi_1(\Lambda \mid R)$=$\psi_0(\Lambda \mid R)\xi(\Lambda, \hat{Q}_1)/k(\hat{Q}_1)$. A second step IF $\psi_2(\Lambda \mid R)$ is constructed in the same fashion by drawing a random sample from $\psi_1(\Lambda \mid R)$ and regressing $\ln \psi_0(\lambda_{1,n,t})$ on a constant, $\lambda_{1,n,t}$ and $\lambda_{1,n,t}^2$. With the resulting sequence of matrixes $\hat{Q}_2$ one can determine $\psi_2(\Lambda \mid R)$. This procedure is repeated until $\hat{Q}_j$ is sufficiently close to the

one-step-ahead matrixes $\hat{Q}_{j-1}$. Danielsson and Richard (1993) showed that the convergence is reached very fast, typically after less than 5 iterations. Finally the simulated sample $\{\Lambda_{j,n}\}_{n=1}^{N}$ from $\psi_j(\Lambda \mid R)$ is used to calculate the jth step AGIS estimate of the integral:

$$\hat{f}_{N,j}(R \mid \theta) = \frac{1}{N} \sum_{n=1}^{N} \frac{\phi_0(\Lambda_{j,n}, R)\, k(\hat{Q}_j)}{\xi(\Lambda_{j,n}, \hat{Q}_j)} \tag{B.8}$$

To obtain the SML estimate of θ, this jth step AGIS estimate of the likelihood function is maximized with respect to θ.

REFERENCES

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

Bollerslev, T. (1987), "A Conditionally Heteroskedastic Time Series Model for Security Prices and Rates of Return Data," *The Review of Economics and Statistics 59*, 542-547.

Box, G.E. and , P.E. Tiao (1973), *Bayesian Inference in Statistical Analysis,* Addison-Wesley, Reading, Massachusetts.

Danielsson, J., and J.F. Richard (1993), "Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models," *Journal of Applied Econometrics 8*, 153-173.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50, 987-1007.

Harvey, A. (1989), *Forecasting, Structural Time Series and Kalman Filter,* Cambridge University Press, Cambridge.

Harvey, A., Ruiz, E., and N. Shephard (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies* 61, 247-264.

Jacquier, E., Polson, N. G., and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics* 12, 371-389.

Johnson, N.L. and , S. Kotz (1973), *Distributions in Statistics: Continuous Univariate Distributions-2,* John Wiley and Sons, New York.

Kim, S., Shepard, N., and S. Chib. (1996), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models ," *mimeo, Nuffield College, Oxford University.*

Mahieu, R.J., and P.C. Schotman (1994), "Stochastic Volatility and the Distribution of Exchange Rate News," *Working paper 94-22, Limburg Institute of Financial Economics, Maastricht.*

Ruiz, E. (1994), "Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models," *Journal of Econometrics* 63, 289-306.

Teräsvirta, T. (1996), "Two Stylized Facts and the GARCH(1,1) Model," *Working paper, No.96, Stockholm School of Economics.*

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

Taylor, S.J. (1994), "Modeling Stochastic Volatility: A Review and Comparative Study," *Mathematical Finance* 4, 183-204.

Table 1. Summary statistics of the daily returns.

Statistic	SIE	DAI	VOW	DBK	BAY	VEB
Mean	.004	.005	−.001	.000	.031	.036
Std. Dev.	1.198	1.531	1.798	1.245	1.348	1.306
Skewness	.024	−.026	−.084	−.444	.231	−.266
Kurtosis	10.597	7.250	8.402	9.551	7.052	10.323
Minimum	−9.381	−7.984	−13.200	−10.167	−5.964	−9.475
Maximum	8.519	8.739	12.497	6.520	8.659	8.630
$LB_r(20)$	22.159	17.396	7.769	22.546	24.641	10.795
	(.332)	(.627)	(.993)	(.312)	(.215)	(.951)

NOTE: Sample size is 1095 for SIE, DAI, VOW and 1096 for DBK, BAY, VEB.
Marginal significance level are in parentheses.

Table 2. Autocorrelation of the squared returns

Lag	SIE	DAI	VOW	DBK	BAY	VEB
			Autocorrelation coefficients			
1	.137	.104	.022	.053	.124	.096
2	.093	.217	.088	.124	.182	.126
3	.152	.147	.118	.157	.111	.144
4	.053	.158	.050	.093	.134	.139
5	.084	.059	.056	.044	.037	.054
6	.048	.100	.044	.053	.077	.080
7	.026	.070	.021	.045	.039	.042
8	.053	.110	.045	.069	.071	.077
9	.055	.086	.013	.060	.114	.085
10	.027	.095	.031	.039	.051	.034
20	.005	.066	.009	.006	.064	.032
30	.089	.044	.031	.061	.042	.053
40	.049	.081	.043	.093	.144	.080
50	−.005	.021	.010	−.014	.041	.037
			Ljung-Box statistic			
$LB_{r^2}(30)$	140.3	378.5	56.6	118.9	200.0	190.5
	(.000)	(.000)	(.002)	(.000)	(.000)	(.000)

NOTE: Marginal significance level are in parentheses.

Table 3. SML estimation of the SV-normal model.

Parameter	SIE	DAI	VOW	DBK	BAY	VEB
μ	−.0010	.0034	.0236	.0235	.0356	.0592
	(.0250)	(.0347)	(.0385)	(.0249)	(.0240)	(.0264)
γ	−.0014	.0211	.0841	.0006	.0111	−.0007
	(.0079)	(.0103)	(.0263)	(.0074)	(.0110)	(.0097)
δ	.9526	.9535	.9055	.9629	.9322	.9515
	(.0165)	(.0142)	(.0285)	(.0127)	(.0193)	(.0162)
ν	.2488	.2715	.3031	.2408	.3439	.3091
	(.0408)	(.0391)	(.0492)	(.0359)	(.0468)	(.0480)
κ	5.856	6.754	4.997	6.652	7.399	8.232
$\rho(1)$	.184	.203	.147	.204	.206	.223
$\ln L$	−1612.93	−1877.46	−2117.89	−1641.72	−1745.56	−1643.33

NOTE: Standard errors are in parentheses. The estimation are based on a simulation sample size $N = 500$ and four AGIS iterations.

Table 4. SML estimation of the SV-t model.

Parameter	SIE	DAI	VOW	DBK	BAY	VEB
μ	.0032	.0001	.0274	.0227	.0365	.0589
	(.0248)	(.0237)	(.0224)	(.0252)	(.0293)	(.0249)
γ	.0006	.0133	.0502	.0014	.0093	.0016
	(.0032)	(.0075)	(.0160)	(.0039)	(.0078)	(.0057)
δ	.9881	.9753	.9475	.9877	.9621	.9794
	(.0070)	(.0101)	(.0157)	(.0065)	(.0175)	(.0096)
ν	.1058	.1795	.1941	.1229	.2341	.1842
	(.0266)	(.0354)	(.0331)	(.0255)	(.0593)	(.0376)
$1/\omega$	.1213	.1145	.0871	.1560	.0927	.1102
	(.0260)	(.0316)	(.0311)	(.0259)	(.0372)	(.0265)
κ	7.084	8.260	5.496	10.182	8.116	9.612
$\rho(1)$	.098	.125	.093	.092	.145	.146
$\ln L$	−1605.06	−1872.78	−2115.80	−1628.46	−1743.58	−1638.12
$LR_{1/\omega=0}$	15.74	9.36	4.18	26.52	3.98	10.42
	[.000]	[.001]	[.021]	[.000]	[.023]	[.000]

NOTE: Standard errors are in parentheses and p-values are in brackets. The estimation are based on a simulation sample size $N = 500$ and four AGIS iterations.

Table 5. SML estimation of the SV-GED model.

Parameter	SIE	DAI	VOW	DBK	BAY	VEB
μ	−.0014	−.0026	.0220	.0155	.0355	.0566
	(.0249)	(.0253)	(.0310)	(.0250)	(.0240)	(.0250)
γ	−.0005	.0164	.0761	.0017	.0110	.0009
	(.0069)	(.0081)	(.0230)	(.0050)	(.0104)	(.0076)
δ	.9603	.9673	.9160	.9804	.9357	.9652
	(.0168)	(.0112)	(.0250)	(.0090)	(.0209)	(.0148)
ν	.2219	.2142	.2766	.1615	.3320	.2533
	(.0485)	(.0397)	(.0505)	(.0327)	(.0607)	(.0521)
$1/\vartheta$	.5410	.6145	.5361	.7053	.5181	.5788
	(.0507)	(.0459)	(.0436)	(.0489)	(.0485)	(.0520)
κ	5.798	6.644	4.916	6.802	7.378	8.188
$\rho(1)$	.168	.162	.134	.138	.200	.196
$\ln L$	−1612.54	−1874.90	−2117.82	−1633.03	−1745.45	−1642.34
$LR_{1/\vartheta=.5}$	.78	5.12	.14	17.38	.22	1.98
	[.377]	[.024]	[.708]	[.000]	[.639]	[.159]

NOTE: Standard errors are in parentheses and p-values are in brackets. The estimation are based on a simulation sample size $N = 500$ and four AGIS iterations.

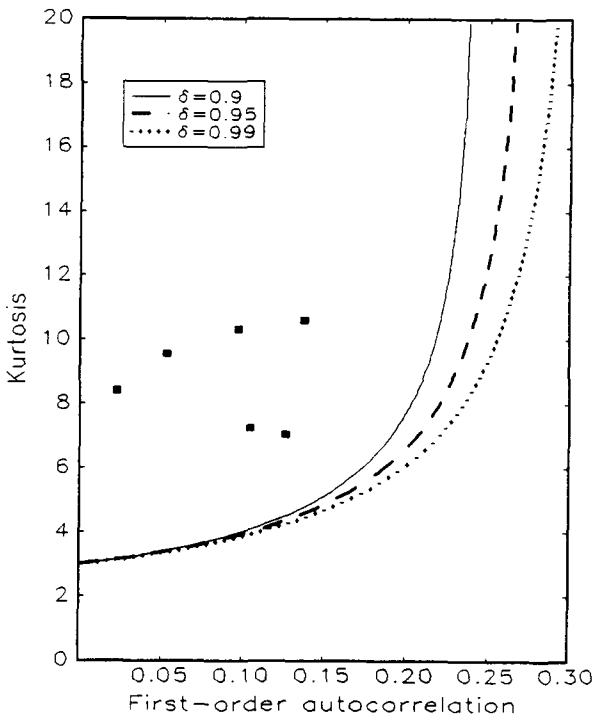

Figure 1. Combinations of the kurtosis of the returns and the first-order autocorrelation of the squared returns for the data, and the predicted combinations of the SV model with normal errors for various values of δ.

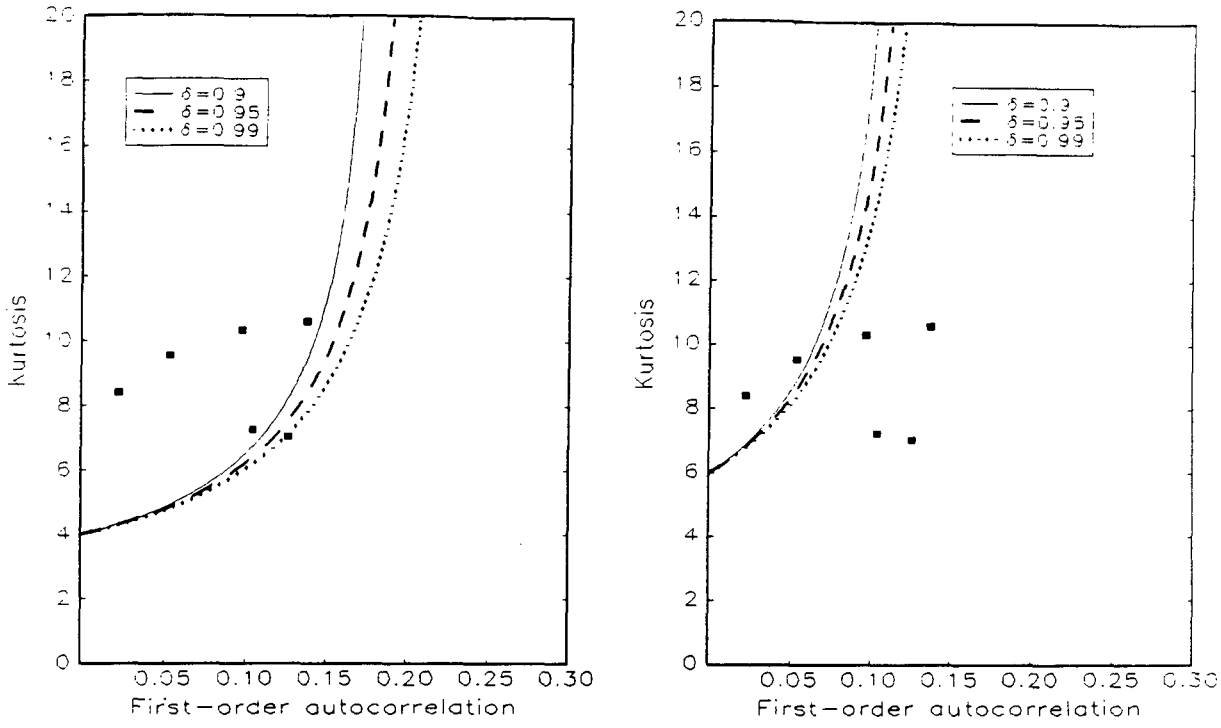

Figure 2. Combinations of the kurtosis of the returns and the first-order autocorrelation of the squared returns for the data, and the predicted combinations of the SV model with a leptokurtic error distribution for various values of δ. In the left-hand panel the baseline-kurtosis is $E(u_t^4) = 4$ and in the right-hand panel $E(u_t^4) = 6$.

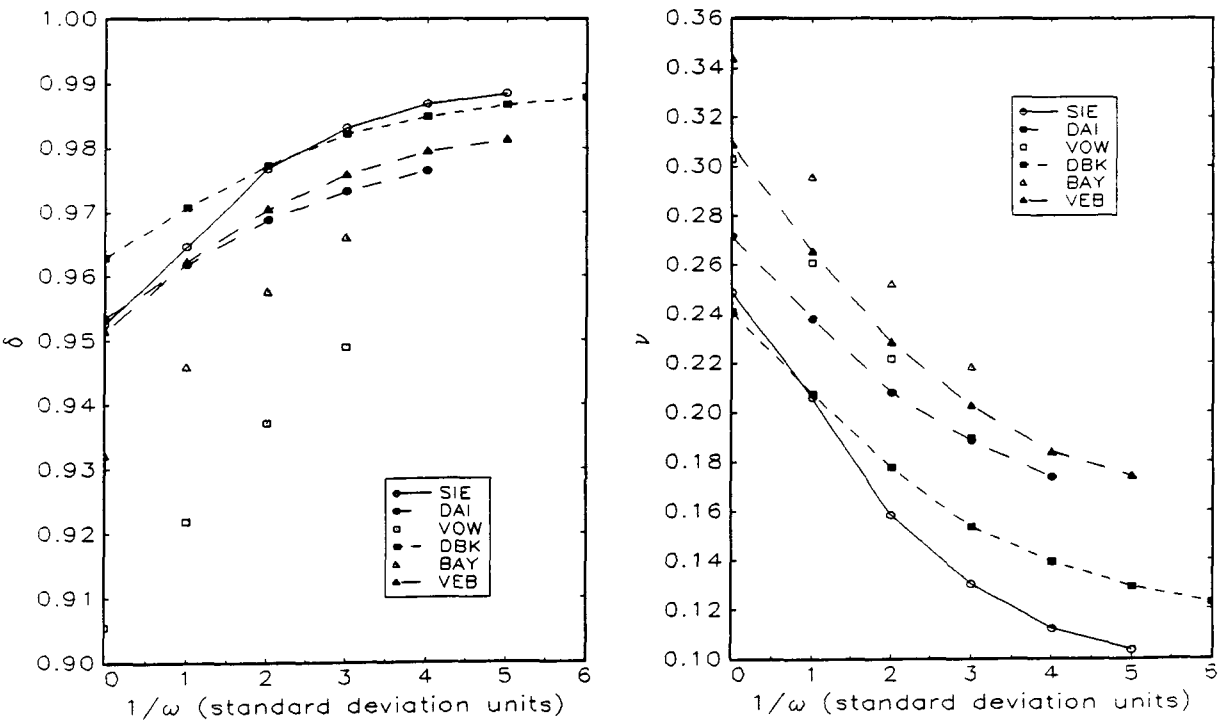

Figure 3. Combinations of predetermined values of $1/\omega$ in the SV-t model measured in standard deviations given in Table 4 and the resulting SML estimates of δ and ν. The left-hand panel contains the combinations of $1/\omega$ and δ, and the right-hand panel those of $1/\omega$ and ν.