Journal of Forecasting
J. Forecast. **25**, 153–171 (2006)
Published online in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/for.982

Gamma Stochastic Volatility Models

BOVAS ABRAHAM,[1]* N. BALAKRISHNA[2] AND
RANJINI SIVAKUMAR[1]
[1] *University of Waterloo, Canada*
[2] *Cochin University of Science and Technology, India*

ABSTRACT
This paper presents gamma stochastic volatility models and investigates its distributional and time series properties. The parameter estimators obtained by the method of moments are shown analytically to be consistent and asymptotically normal. The simulation results indicate that the estimators behave well. The in-sample analysis shows that return models with gamma autoregressive stochastic volatility processes capture the leptokurtic nature of return distributions and the slowly decaying autocorrelation functions of squared stock index returns for the USA and UK. In comparison with GARCH and EGARCH models, the gamma autoregressive model picks up the persistence in volatility for the US and UK index returns but not the volatility persistence for the Canadian and Japanese index returns. The out-of-sample analysis indicates that the gamma autoregressive model has a superior volatility forecasting performance compared to GARCH and EGARCH models.

KEY WORDS stochastic volatility; GARCH; gamma sequences; moment estimation; financial time series

INTRODUCTION

Studies on financial time series reveal that changes in volatility (variance) over time occur for all classes of assets such as stocks, currency and commodities. Time series studies also indicate that the sequence of returns $\{y_t\}$ on some financial assets such as stocks often exhibit time-dependent variances and excess kurtosis in the marginal distributions. In such cases, forecasts of asset–return variance are central for financial applications such as portfolio optimization and valuation of financial derivatives. Time series models, called volatility models in the literature, have been employed to capture these salient features. As quoted by Shephard (1996), volatility models provide an excellent testing ground for the development of new non-linear and non-Gaussian time series techniques.

In a broader sense, there are two kinds of models for time-dependent variances. They are observation-driven and parameter-driven models. An example of the former is the autoregressive conditional heteroskedastic (ARCH) model introduced by Engle (1982). In this model, the variance of the

* Correspondence to: Bovas Abraham, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada. E-mail: babraham@uwaterloo.ca

series $\{y_t\}$ at time t is assumed to be a deterministic function of lagged values of the squared errors. For a review of this model and its various generalizations, such as generalized ARCH (GARCH) by Bollerslev (1986), see Shephard (1996) and Tsay (2002). In parameter-driven models, it is assumed that the time-dependent variances are random variables (r.v.'s) generated by an underlying stochastic process. For example, Taylor (1986) discussed a model for y_t defined by

$$y_t = \varepsilon_t \exp(h_t/2) \tag{1}$$

$$h_t = \gamma_0 + \gamma_1 h_{t-1} + \eta_t \tag{2}$$

where $\{\varepsilon_t\}$ and $\{\eta_t\}$ are independent Gaussian white noise processes. Note that $\{h_t\}$ is a Gaussian autoregressive (AR) process and $\exp(h_t/2)$ has a lognormal distribution. Hence (1) & (2) is referred to as a lognormal stochastic volatility model. A review of the properties of stochastic volatility (SV) models may be found in Taylor (1994), Jacquier *et al*. (1994) and Ghysels *et al*. (1996).

Most empirical applications of the SV models have assumed that the conditional distribution of returns is normal (Jacquier *et al*., 1994; Kim *et al*., 1998). Some studies (for example, Harvey *et al*., 1994; Sandmann and Koopmann, 1998; Chib *et al*., 2002; Liesenfeld and Jung, 2000) assume that the conditional distribution of returns is a Student-t distribution or a general error distribution (Box and Tiao, 1973). These studies find that the SV models with conditional heavy-tailed distributions are a better fit and capture the leptokurtic nature of the distribution better. However, the kurtosis in these models is still lower than that of typical financial time series. Andersson (2001), using a normal inverse Gaussian SV model, reports similar findings. In this paper, we focus on another class of models, the gamma SV models. The objective is to examine the ability of the gamma SV models to capture the leptokurtic nature of return distributions and the slowly decaying autocorrelation functions of squared returns.

We study the model (1) when $\{h_t\}$ is a stationary Markov sequence with exponential and gamma marginals. We note that a number of AR models are introduced for non-negative r.v.'s in the context of non-Gaussian time series. See, for example, Gaver and Lewis (1980), Sim (1986), Abraham and Balakrishna (1999) and the references cited therein. One can very well use these AR models to describe the evolution of time-dependent volatilities. Here we concentrate on a SV model (1) when $\{h_t\}$ is an AR(1) sequence with exponential and gamma marginal, and investigate its distributional and time series properties.

In the next section, we discuss the time series properties of the SV model and in the third section, we consider the distributional properties of the proposed model. In the fourth section, we present some simulation results using a range of parameter values similar to those found in stock return series and in the fifth section, we illustrate the models using daily stock returns data. A final section gives some concluding remarks.

GAMMA AUTOREGRESSIVE MODELS FOR VOLATILITY

Let y_t be the demeaned return on an asset at time t, $t = 0, \pm 1, \pm 2, \cdots$. Define

$$y_t = \varepsilon_t \sqrt{h_t} \tag{3}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (iid) standard normal r.v.'s. We assume that $\{\varepsilon_t\}$ is independent of h_t for every t. In this section, we discuss the properties of $\{y_t\}$

DOI: 10.1002/for

when $\{h_t\}$ is a Markov sequence with exponential and gamma marginals. We assume that h_t is a gamma r.v. with probability density function (pdf):

$$f(h_t; \lambda, p) = \frac{e^{-\lambda h_t} h_t^{p-1} \lambda^p}{\Gamma(p)}, \quad h_t \geq 0, \lambda > 0, p > 0 \tag{4}$$

denoted by a $G(\lambda, p)$ distribution. Then the characteristic function (cf) of y_t is given by

$$\phi(s) = E[e^{isy_t}] = \left(\frac{2\lambda}{2\lambda + s^2}\right)^p. \tag{5}$$

Thus $\{y_t\}$ is a stationary Markov sequence whose marginal distribution has the cf (5). Note that (5) is the cf of the difference of two iid gamma r.v.'s with parameters $\sqrt{2\lambda}$ and p. The distribution of such r.v.'s is referred to as a generalized Laplace distribution since at $p = 1$, the density corresponding to (5) becomes

$$f_1(y) = \sqrt{\frac{\lambda}{2}} \exp\{-|y|\sqrt{2\lambda}\}, \quad -\infty < y < \infty, \lambda > 0 \tag{6}$$

which is the Laplace pdf.

The properties of generalized Laplace distributions and their applications are discussed by Mathai (1993). The odd moments of y_t are zero and its even moments are given by

$$E[y_t^{2r}] = (2r-1)(2r-3)\ldots 3 \cdot 1 \cdot \frac{\Gamma(p+r)}{\Gamma(p)} \lambda^{-r}, \quad r = 1, 2, \cdots. \tag{7}$$

Then $V(y_t) = p\theta$ where $\theta = 1/\lambda$, and the kurtosis of y_t becomes

$$K = \frac{E(y_t^4) - [E(y_t)]^4}{[\mathrm{Var}(y_t)]^2} = 3 + \frac{3}{p} > 3. \tag{8}$$

Note that if $p = 1$, then $K = 6$ which is the kurtosis of a Laplace distribution and as $p \to \infty$, $K \to 3$, that corresponding to a normal distribution. By choosing a smaller p, one can get a distribution with larger kurtosis. So y_t has a leptokurtic marginal distribution. The literature on financial time series indicates that the return series shows the tendency to follow leptokurtic distributions (see, for example, Hsieh, 1991). Hence the generalized Laplace distribution is a good candidate for modelling such data. Now we describe some of the models for generating the volatility sequence $\{h_t\}$.

Exponential autoregressive model (EAR)

The simplest one is the exponential autoregressive model of order one [EAR(1)] introduced by Gaver and Lewis (1980), given by

$$h_t = \phi h_{t-1} + I_t E_t, \quad 0 \leq \phi < 1 \tag{9}$$

where $\{I_t\}$ and $\{E_t\}$ are two mutually independent iid sequences such that E_t is an exponential r.v. with pdf

DOI: 10.1002/for

$$g(E_t;\lambda)=\lambda e^{-\lambda E_t},\quad E_t\geq 0,\lambda>0, \text{where } \lambda=\frac{1}{\theta} \tag{10}$$

and

$$P[I_t=0]=1-P[I_t=1]=\phi. \tag{11}$$

Then $\{h_t\}$ defines a stationary Markov sequence with exponential marginals. Assume that I_tE_t is independent of ε_t for every t. The value of ϕ in the range (0,1) ensures stationarity. Under this setup, $\{y_t\}$ is a stationary sequence of Laplacian r.v.'s with marginal pdf (6). Similarly one can take $\{h_t\}$ as the exponential sequences of Lawrance and Lewis (1981) and Sim (1990).

It can be shown that $E(h_{t+1}|y_t, y_{t-1}\ldots) = \phi h_t + (1-\phi)\theta$, $E(h_t)=\theta$, $V(h_t)=\theta^2$, $E[h_t^2]=2\theta^2$ and $\gamma_k(h) = \text{Cov}(h_t, h_{t-k}) = \theta^2\phi^k$. Hence, $\rho_k(h) = \text{Corr}(h_t, h_{t-k}) = \phi^k$, which goes to zero as $k\to\infty$, and this autocorrelation function (acf) is very similar to that of the usual Gaussian AR(1) process.

In addition, $\gamma_k[y^2] = \text{Cov}[y_t^2, y_{t-k}^2] = \theta^2\phi^k$ and $V[y_t^2] = 5\theta^2$. Hence the lag k autocorrelation

$$\rho_k[y^2]=\text{Corr}[y_t^2, y_{t-k}^2]=\frac{\phi^k}{5}. \tag{12}$$

Thus if $k = 1$, $\rho_1[y^2] = \phi/5$. In practice, the persistence parameter $\phi\approx 0.9$, which implies that $\rho_1[y^2] \approx 0.18$, a value usually observed in stock return series (Liesenfeld and Jung, 2000).

Gamma autoregressive model (GAR)

We also study the properties of a SV model when $\{h_t\}$ is a first order gamma autoregressive [GAR(1)] sequence of Gaver and Lewis (1980). In this case,

$$h_t=\phi h_{t-1}+\eta_t,\quad 0\leq\phi<1, t=1,2,\cdots, \tag{13}$$

where $\{\eta_t\}$ is an iid sequence and h_t has a $G(\lambda, p)$ distribution independent of η_t. The distribution of η_t is specified as

$$\eta_t=\sum_{j=l}^{N}\phi^{U_j}E_j \tag{14}$$

where $\{U_l\}$ is a sequence of iid uniform (0,1) r.v.'s, the E_j's are iid exponentials as described earlier, and N is Poisson with mean $p \log(1/\phi)$. The r.v.'s U_j, E_j and N are mutually independent for every j. If $N\equiv 0$, then we take $\eta_j\equiv 0$. The sequence $\{h_t\}$ defined by (13) is stationary and has $G(\lambda, p)$ as the marginal distribution.

We can show that

$$E(h_{t+1}|y_t, y_{t-1}\ldots)=\phi h_t+(1-\phi)p\theta$$
$$E(h_t)=p\theta,\quad E[h_t^2]=p(p+1)\theta^2,\quad V(h_t)=p\theta^2,\quad \text{and}\quad \gamma_k(h)=p\theta^2\phi^k.$$

Hence, $\rho_k(h) = \text{Corr}(h_t, h_{t-k}) = \phi^k$, which goes to zero as $k\to\infty$ and this acf is very similar to that of the usual Gaussian AR(1) process.

DOI: 10.1002/for

In addition, $V[y_t^2] = p\theta^2(3 + 2p)$ and $\gamma_k[y^2] = p\theta^2\phi^k$. Hence the lag k autocorrelation is

$$\rho_k[y_t^2] = \mathrm{Corr}[y_t^2, y_{t-k}^2] = \frac{\phi^k}{3+2p}. \tag{15}$$

Note:

(i) If $p = 1$, the acf simplifies to $\rho_1[y^2] = \phi^k/5$, and $\rho_1[y^2] = \phi/5$ is the acf of an EAR(1) SV model.
(ii) When $k = 1$, $\rho_1[y^2] = \phi/(3 + 2p)$ and also if $p = 0.5$ (this corresponds to a kurtosis $K = 9$, which is observed in practice), $\rho_1[y^2] = \phi/4$. In addition, if the persistence parameter $\phi \approx 0.9$, then $\rho_1[y^2] \approx 0.22$, again a reasonable value seen in practice (Liesenfeld and Jung, 2000).

When the kurtosis is large (i.e. p is small), $\rho_1[y^2]$ will be larger. For instance, if $p = 1/3$ then kurtosis $K = 12$ and $\rho_1[y^2] = \phi/3.67$.

From (8) and (15), the acf and the kurtosis are related as follows:

$$\rho_k[y^2] = \frac{\phi^k}{3+2p} = \frac{(K-3)}{3(K-1)}\phi^k \tag{16}$$

This implies that $\rho_1[y^2] = \frac{(K-3)}{3(K-1)}\phi$. For $K = 9$ observed in practice, $\rho_1[y^2] = 0.25\phi$. Similarly, we can show that the kurtosis $K = \frac{3[\rho_k(y^2) - \phi^k]}{3\rho_k(y^2) - \phi^k}$.

When $k = 1$,

$$K = \frac{3[\rho_1(y^2) - \phi]}{3\rho_1(y^2) - \phi} = \frac{3[\phi - \rho_1(y^2)]}{\phi - 3\rho_1(y^2)} \tag{17}$$

(For the persistence parameter, $\phi \approx 0.9$ and acf, $\rho_1[y^2] \approx 0.2$ that are observed in practice, $K \approx 7$.)

This implies that for the model, the persistence parameter ϕ has to be more than three times larger than $\rho_1[y^2]$. Figure 1 shows the relation between the kurtosis, K and $\rho_1[y^2]$ as implied by the GAR(1) model, for values of ϕ ranging from 0.35 to 0.95. Note that the persistence parameter ϕ is restricted to the range (0, 1) to ensure stationarity. Figure 1 indicates that the GAR specification covers the empirical points that relate to all the stock index returns except for Canada. It should be noted that the data points for Canada and Japan are somewhat away from those of the USA and UK. In comparison with Andersson (2001), we find that the GAR(1) model seems to respond better to the co-existence of a high kurtosis and low acf.

ESTIMATION OF PARAMETERS

The likelihood-based inference in SV models is quite complicated as the likelihood function of $(y_1, y_2, \ldots, y_T)$ is a T-dimensional integral. A number of methods are proposed for estimating the parameters and a comprehensive survey may be seen in Shephard (1996) and Tsay (2002). The GMM approach to estimation has been used, among others, by Melino and Turnbull (1990), Duffie and

DOI: 10.1002/for

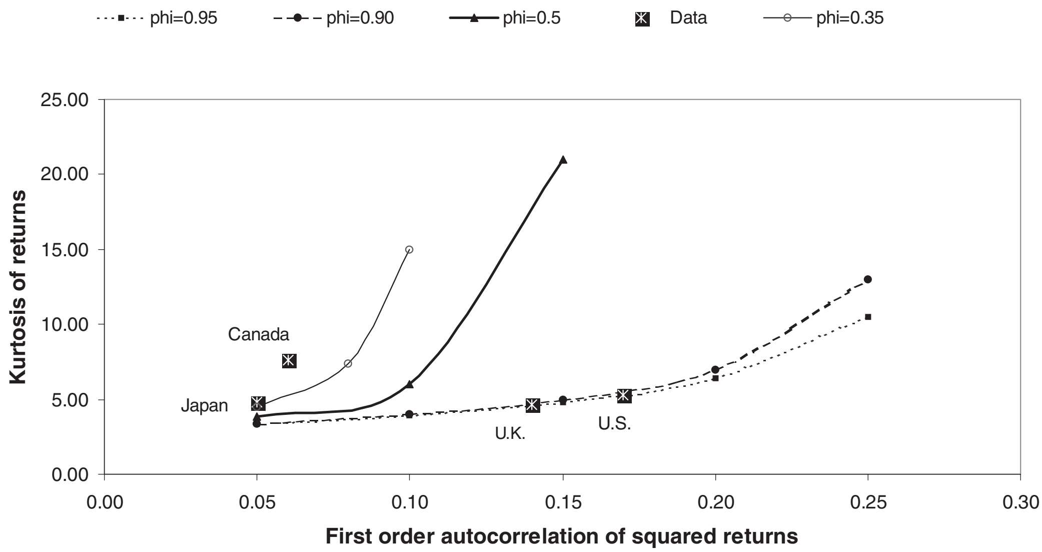

Figure 1. GAR(1) model–combination of first-order autocorrelation of squared returns and kurtosis

Singleton (1993), and Andersen and Sorensen (1995). Jacquier *et al*. (1994) analysed the model (1) using a Bayesian approach based on Markov chain Monte Carlo (MCMC) techniques. Kim *et al*. (1998) use the MCMC approach as well. Harvey *et al*. (1994) use the quasi maximum likelihood (QML) approach, while Gallant *et al*. (1997) and Gallant and Tauchen (1998) adopt the efficient method of moments (EMM) approach. Fridman and Harris (1998) discuss a maximum likelihood approach for SV models in which the authors use a recursive numerical integration procedure to calculate the likelihood. Liesenfeld and Jung (2000) use the simulated maximum likelihood (SML) approach and suggest that MCMC and SML have the best small sample properties.

For the SV models described above, we adopt the method of moments to estimate the parameters. We are aware that its efficiency does not compare favourably with more complicated approaches indicated above. However, we use this method because its estimation simplicity and analytical tractability are the basic premises of the study.

Let β be the vector of unknown parameters and $\hat{\beta}_T$ be its estimator based on a sample of size T. We apply the theory developed by Hansen (1982) to show that the moment estimators are consistent and asymptotically normal (CAN). Hansen (1982) proved the results under the following assumptions:

A_1 Let $\{y_t, -\infty < t < \infty\}$ be a stationary and ergodic sequence whose finite-dimensional distribution depends on a q-dimensional parameter vector.

A_2 The parameter space S is an open subset of R^q that contains the true parameter β_0.

A_3 Let f be a function f: $R^q \times S \to R^r$, $r \geq q$.
Asume that $f(\cdot, \beta)$ and $\partial f/\partial\beta$ are Borel measurable functions for $\beta \in S$ and $\partial f(x, \cdot)/\partial\beta$ is continuous on S for each $x \in R^q$.

DOI: 10.1002/for

A_4 Let $f_t(\cdot, \beta) = f(y_t^{(\cdot)}, \beta)$ and assume that $\partial f_1/\partial\beta$ is first moment continuous at β_0 and $D = E\left[\frac{\partial}{\partial\beta} f(y_t, \beta_0)\right]$ exists, is finite and has full rank.

A_5 Let $\omega_t = f(x_t, \beta_0)$, $-\infty < t < \infty$ and $\vartheta_j = E(\omega_0|\omega_{-j}, \omega_{-j-1}, \cdots) - E(\omega_0|\omega_{-j-1}, \omega_{-j-2}, \cdots)$, $j \geq 0$.

The assumptions are that $E(\omega_0\omega_0')$ exists and is finite, $E(\omega_0|\omega_{-j}, \omega_{-j-1}, \ldots)$ converges in mean square to zero and $\sum_{j=0}^{\infty} E[\vartheta_j'\vartheta_j]^{1/2} < \infty$.

Now we state the following result from Hansen (1982).

Result Suppose that the sequence $\{y_t, -\infty < t < \infty\}$ satisfies the assumptions $A_1 - A_5$. Then $\{\sqrt{T}(\hat{\beta}_T - \beta_0), T \geq 1\}$ converges in distribution to a normal random vector with mean 0 and dispersion matrix $[DS^{-1}D']^{-1}$. D is as in A_4 and

$$S = \sum_{k=-\infty}^{\infty} \Gamma_{(k)}, \quad \Gamma_{(k)} = E(\omega_t\omega_{t-k}^1) \tag{18}$$

We have verified that the assumptions $A_1 - A_5$ hold good for all the SV models given earlier. From the general theory of AR(1) sequences it follows that $\{h_t\}$ defined by (9) and (13) is stationary and ergodic. Hence the corresponding y_t sequences also have these properties. Then by the pointwise ergodic theorem, the moment estimators are consistent for their expectations.

In the following discussion, the suffixes 1 and 2 of the matrices D, S and Σ denote that they represent the matrices associated with the EAR(1) and GAR(1) models, respectively.

EAR(1) process

Now we estimate the parameters of an SV model when $\{h_t\}$ is an EAR(1) process defined by (9). In this case, the parameter vector $\beta = (\theta, \phi)^1$, $\theta = 1/\lambda$. We have $E(y_t^2) = \theta$ and $E(y_t^2y_{t-1}^2) = (1 + \phi)\theta^2$.

Let

$$f(y_t, \beta_0) = \begin{pmatrix} y_t^2 - \theta \\ y_t^2 y_{t-1}^2 - (1+\phi)\theta^2 \end{pmatrix} \quad \text{and} \quad \beta_0 = \begin{pmatrix} \theta_0 \\ \phi_0 \end{pmatrix}. \tag{19}$$

The moment estimator of β can be obtained by solving $f(y_t, \beta_0) = 0$.[1]

The estimator $\hat{\beta}_T = (\hat{\theta}_T, \hat{\phi}_T)$ is given by

$$\hat{\theta}_T = \overline{Y}_2 \quad \text{and} \quad \hat{\phi}_T = \frac{\overline{Y}_{22} - \overline{Y}_2^2}{\overline{Y}_2^2} \tag{20}$$

[1] In general, there is no restriction for choosing the number of moment conditions. A larger number of conditions may improve the precision but will lead to computational complications. We have chosen a just-identified system because it leads to simple expressions for the estimators and explicit evaluation of the dispersion matrices, since simplicity was a basic premise of the study. If we increase the number of conditions, then the analytics will be much more complicated. There appears to be a tradeoff between the gains in efficiency and the analytical tractability when moving beyond a just-identified system. We thank the referee for pointing out the inefficiencies in using a just-identified framework cited by Andersen and Sorensen (1995).

DOI: 10.1002/for

where $\bar{Y}_2 = (1/T)\sum_{t=1}^{T} y_t^2$ and $\bar{Y}_{22} = (1/T)\sum_{t=1}^{T} y_t^2 y_{t-1}^2$.

The matrix S_1 is given by equation (18) as

$$S_1 = \Gamma_{(0)} + 2\sum_{k=1}^{\infty} \Gamma_{(k)}.$$

The sequence $\{\Gamma_{(k)}\}$ of matrices is shown in the Appendix. Hence the matrix S_1 associated with the EAR(1) model is given by

$$S_1 = \left(\frac{\theta^2}{1-\phi}\right)\begin{bmatrix} 5-3\phi & \theta(5+2\phi+3\phi^2-2\phi^3) \\ \left(\frac{\theta}{1+\phi}\right)\begin{pmatrix}15+29\phi-5\phi^2 \\ -21\phi^3-6\phi^4\end{pmatrix} & \left(\frac{\theta^2}{1+\phi}\right)\begin{pmatrix}45+102\phi+98\phi^2-68\phi^3 \\ -91\phi^4-14\phi^5-16\phi^6\end{pmatrix}\end{bmatrix}. \tag{21}$$

For our choice of f in (19), we have the matrix $D_1 = E\left[\frac{d}{d\beta} f(y_t, \beta)\right]$ given by

$$D_1 = \begin{bmatrix} -1 & -2(1+\phi)\theta \\ 0 & -\theta^2 \end{bmatrix} \tag{22}$$

Hence, by our previous result, the asymptotic dispersion matrix of $\hat{\beta}_T$ for the exponential SV model is given by Σ_1/T, where

$$\Sigma_1 = [D_1 S_1^{-1} D_1']^{-1} \tag{23}$$

Next we obtain the asymptotic dispersion matrix of $\hat{\beta}_T$ for a GAR(1) model.

GAR(1) process

When $\{h_t\}$ is a GAR(1) process defined by (13) & (14), the parameter vector is given by $\beta = (\theta, p, \phi)^1$, where $\theta = 1/\lambda$. Here we choose

$$f(y_t, y_{t-1}, \beta) = \begin{pmatrix} y_t^2 - p\theta \\ y_t^4 - 3p(p+1)\theta^2 \\ y_t^2 y_{t-1}^2 - (\phi p + p^2)\theta^2 \end{pmatrix}, \quad t = 1, 2, \ldots, T. \tag{24}$$

The moment estimator of β obtained by solving $f(y_t, y_{t-1}, \beta_0) = 0$ is given by

$$\hat{\beta}_T = (\hat{\theta}_T, \hat{p}_T, \hat{\phi}_T)^1 \tag{25}$$

where

$$\hat{\theta}_T = \frac{\left\{\frac{\bar{Y}_4}{3} - (\bar{Y}_2)^2\right\}}{\bar{Y}_2}, \quad \hat{p}_T = \frac{\bar{Y}_2}{\hat{\theta}_T}, \quad \text{and} \quad \hat{\phi}_T = \frac{\bar{Y}_{22} - \bar{Y}_2^2}{\frac{\bar{Y}_4}{3} - (\bar{Y}_2)^2}.$$

DOI: 10.1002/for

The notations $\overline{Y}_2$ and $\overline{Y}_{22}$ are as defined in (19) and $\overline{Y}_4 = \sum_{t=1}^{T} y_t^4 \Big/ T$.

To apply our result, we need to compute the matrices $\Gamma_{(k)}$, $k = 0, \pm1, \pm2, \cdots$.

Let

$$\Gamma_{(k)} = \begin{bmatrix} \gamma_{11}^{(k)} & \gamma_{12}^{(k)} & \gamma_{13}^{(k)} \\ \gamma_{21}^{(k)} & \gamma_{22}^{(k)} & \gamma_{23}^{(k)} \\ \gamma_{31}^{(k)} & \gamma_{32}^{(k)} & \gamma_{33}^{(k)} \end{bmatrix}, \quad k = 0, \pm1, \pm2, \cdots \tag{26}$$

where $\Gamma_{(k)} = \Gamma_{(-k)}$, $k = 1, 2, \cdots$.

The matrix S_2 for the GAR(1) SV model is given by $S_2 = \Gamma_{(0)} + 2\sum_{k=1}^{\infty} \Gamma_{(k)}$.

We denote it by

$$S_2 = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \tag{27}$$

The elements for $\Gamma_{(k)}$ and S_2 are shown in the Appendix.

The matrix D in this case is denoted by D_2, which is given by $D_2 = E\left[\frac{d}{d\beta} f(y_t, \beta)\right]$.

$$D_2 = \begin{bmatrix} -p & -6p(p+1)\theta & -2(p\phi + p^2)\theta \\ -\theta & -3(2p+1)\theta^2 & -(\phi + 2p)\theta^2 \\ 0 & 0 & -p\theta^2 \end{bmatrix}. \tag{28}$$

Hence the asymptotic dispersion matrix becomes Σ_2/T, where

$$\Sigma_2 = [D_2 S_2^{-1} D_2']^{-1}. \tag{29}$$

SIMULATION

We carry out a simulation study to understand the performance of the estimators with $T = 2000$ observations since several empirical applications use sample sizes around $T = 2000$ (Andersson, 2001; Liesenfeld and Jung, 2000). The parameter estimates were obtained for each simulated series. This was repeated 1000 times. The averages and standard deviations of the estimates over the 1000 repetitions are reported. In addition, we also report the estimates of the asymptotic standard deviation obtained from (23) for the EAR(1) model and from (29) for the GAR(1) model.

In Table I, we present the results from the EAR (1) model for $\theta = 1.00$ and 2.00 and $\phi = 0.25$, 0.50, 0.75 and 0.90. For the EAR(1) model, the implicit value of p is 1 and hence the kurtosis is 6.

DOI: 10.1002/for

Table I. Averages of $\hat{\theta}$ and $\hat{\phi}$ from 1000 simulations based on the EAR model for specified values of θ and ϕ

(θ, ϕ)	$\hat{\theta}$			$\hat{\phi}$		
	Mean	Standard deviation	Estimates of asymptotic standard deviations	Mean	Standard deviation	Estimates of asymptotic standard deviations
1.00, 0.25	0.9987	0.0547	0.0532	0.2410	0.1365	0.1400
1.00, 0.50	1.0013	0.0582	0.0592	0.4846	0.1717	0.1787
1.00, 0.75	1.0036	0.0731	0.0742	0.7144	0.1819	0.2288
1.00, 0.90	0.9992	0.1071	0.1072	0.8135	0.1741	0.2785
2.00, 0.25	1.9966	0.1041	0.1065	0.2438	0.1361	0.1400
2.00, 0.50	2.0004	0.1182	0.1183	0.4999	0.1733	0.1787
2.00, 0.75	2.0008	0.1490	0.1483	0.7181	0.1791	0.2288
2.00, 0.90	2.0091	0.2067	0.2145	0.8090	0.1705	0.2785

Note: The number of observations in the time series in each simulation is 2000. The estimates of the asymptotic standard deviation are obtained from equation (23).

We find that the estimates of θ and ϕ are close to the true values. For instance, at $\theta = 1.00$ and $\phi = 0.25$, the average $\hat{\theta}$ is 0.9987 and the average $\hat{\phi}$ is 0.2410. The standard deviations of the estimates and the estimates of the asymptotic standard deviation are also close. For example, at $\theta = 1.00$ and $\phi = 0.25$, the standard deviation for $\hat{\theta}$ from the simulations is 0.0547 and the asymptotic standard deviation of $\hat{\theta}$ is 0.0532. For $\phi = 0.75$ and 0.90, in simulations where $\hat{\phi}$ was greater than one, the value was truncated to 0.99. Hence, the corresponding standard deviations of the estimates from the simulations appear to be smaller than the asymptotic standard deviations.

In Table II, we present the simulation results for the GAR(1) model. The procedure adopted is similar to that described above. For this model, the values of kurtosis for $p = 0.30$, 0.50 and 1.50 are 13, 9 and 5, respectively. We provide the estimates for series with 2000 observations. As before, we find that the estimates of θ and ϕ are close to the true values. The standard deviations of $\hat{p}$ and $\hat{\theta}$ increase as ϕ increases. The standard deviations of the estimates and the estimates of the asymptotic standard deviation are close. As in the case of the EAR(1) model, in the GAR(1) model also, for $\phi = 0.75$ and 0.90 where $\hat{\phi}$ was greater than one, the value was truncated to 0.99.

Overall, the simulation results suggest that the moment estimators behave reasonably well for the EAR(1) and GAR(1) models for large samples. In unreported results, we repeated the simulations for series with 1000 and 5000 observations and find that the results for 1000 observations are qualitatively similar to the results for series with 2000 observations. The results for 5000 observations are much better than those for 2000 observations.

We next illustrate the model using stock index returns data for Canada, Japan, the UK and the USA for the period 1997–2002, which is comparable in size to our simulation illustration ($T = 2000$).

APPLICATION TO STOCK RETURN DATA

We examine daily stock price index returns for the period 1997–2002. We use the closing index data on Canada (TSE300), Japan (TOPIX), the UK (FTSE100) and the USA (S&P500 composite). The returns are defined as the first difference of the logarithm of prices and are demeaned. In Table III,

DOI: 10.1002/for

Table II. Averages of $\hat{p}$, $\hat{\theta}$ and $\hat{\phi}$ from 1000 simulations based on the GAR model for specified values of p, θ and ϕ

(p, θ)	ϕ	$\hat{p}$			$\hat{\theta}$			$\hat{\phi}$		
		Mean	Standard deviation	Estimates of asymptotic standard deviations	Mean	Standard deviation	Estimates of asymptotic standard deviations	Mean	Standard deviation	Estimates of asymptotic standard deviations
0.30, 1.00	0.25	0.3316	0.0794	0.0977	0.9722	0.3473	0.3587	0.2654	0.1188	0.1344
	0.50	0.3291	0.0834	0.0987	0.9833	0.3420	0.3612	0.5124	0.1674	0.2050
	0.75	0.3309	0.0868	0.1024	0.9765	0.3357	0.3721	0.7548	0.1863	0.2674
	0.90	0.3384	0.0966	0.1132	0.9619	0.3607	0.4051	0.8640	0.1621	0.2955
0.30, 2.00	0.25	0.3253	0.0778	0.0977	1.9798	0.6726	0.7173	0.2615	0.1214	0.1346
	0.50	0.3325	0.0783	0.0987	1.9316	0.6290	0.7225	0.5215	0.1717	0.2080
	0.75	0.3302	0.0865	0.1024	1.9788	0.7702	0.7441	0.7505	0.1915	0.2968
	0.90	0.3333	0.0967	0.1132	1.9844	0.7627	0.8102	0.8601	0.1609	0.4342
0.50, 1.00	0.25	0.5442	0.1244	0.1418	0.9780	0.2997	0.3098	0.2626	0.1236	0.1314
	0.50	0.5381	0.1218	0.1432	0.9909	0.3055	0.3120	0.5210	0.1646	0.1830
	0.75	0.5438	0.1264	0.1481	0.9847	0.3307	0.3207	0.7524	0.1793	0.2263
	0.90	0.5515	0.1438	0.1625	0.9728	0.3619	0.3473	0.8676	0.1509	0.2432
0.50, 2.00	0.25	0.5368	0.1165	0.1418	1.9777	0.5660	0.6197	0.2608	0.1249	0.1316
	0.50	0.5429	0.1195	0.1432	1.9492	0.5577	0.6240	0.5199	0.1627	0.1863
	0.75	0.5455	0.1335	0.1481	1.9629	0.6183	0.6416	0.7609	0.1841	0.2566
	0.90	0.5557	0.1502	0.1625	1.9375	0.6574	0.6946	0.8602	0.1578	0.3801
1.50, 2.00	0.25	1.6089	0.3831	0.3879	1.9882	0.5805	0.5493	0.2604	0.1643	0.1560
	0.50	1.6197	0.3884	0.3905	1.9741	0.5494	0.5522	0.5237	0.2011	0.1672
	0.75	1.5944	0.3863	0.3996	2.0046	0.5438	0.5633	0.7508	0.1946	0.1748
	0.90	1.6452	0.4398	0.4266	1.9450	0.5566	0.5974	0.8430	0.1725	0.2866

Note: The number of observations in the time series in each simulation is 2000. The estimates of the asymptotic standard deviation are obtained from equation (29).

Table III. Summary statistics of daily stock index returns (1997–2002) 1565 observations

	Mean (%)	Standard deviation	Skewness	Kurtosis	Ljung–Box statistics Q_{12}^{2} #
Canada TSE300	0.0130	1.1676	–0.6116	7.5906	146.26***
Japan TOPIX	–0.0227	1.3279	–0.0057	4.7415	124.29***
UK FTSE100	0.0072	1.2876	–0.1846	4.6780	826.07***
USA S&P500	0.0164	1.3130	–0.0656	5.3280	215.99***

Note: #Ljung–Box statistics for up to 12th-order serial correlation in squared returns. *** Denotes the Ljung–Box statistic is significant at the 1% level.

we present the descriptive statistics for the returns on the indexes. The kurtosis of returns ranges from 4.7 to 7.6 and is much higher than would be expected if returns were normally distributed.

Since the EAR(1) model has the limitation that its implied unconditional kurtosis is always six, we omit the EAR(1) model in this part of the analysis and focus on the GAR(1) model. We compare the GAR(1) model with the basic GARCH(1, 1) model and the EGARCH(1, 1) model. The comparison with the EGARCH model would be especially useful since the GAR(1) model does not allow for skewness or for asymmetric responses of volatility to news.

In Table IV panel A, we report the parameter estimates for each of the return series for the GARCH(1, 1), EGARCH(1, 1) and GAR(1) models. The values of the persistence parameter, $\hat{\beta}$ in the GARCH(1, 1) and EGARCH(1, 1) models suggest significant persistence of volatility. For most of the series, the values of $\hat{\phi}$, the persistence parameter in the SV models are smaller than that in the EGARCH(1, 1) and GARCH(1, 1) models, consistent with Shephard's (1996) findings. Notice that for Canada and Japan, the persistence parameter estimates, $\hat{\phi}$ are much lower in the GAR(1) model. As shown in Figure 1, the behaviour of the returns for Canada and Japan is somewhat different from that of the USA and UK in terms of the first lag autocorrelation of the squared returns and the kurtosis of the returns. Thus GAR(1) is not picking up the persistence in volatility as well as GARCH(1, 1) or EGARCH(1, 1). The values of $\hat{\theta}$ lie between 0.9 and 2.1 and the values of $\hat{p}$ range from 0.6 to 1.80. In panel B, we present the Ljung–Box statistics as diagnostics of the autocorrelations of standardized and squared standardized residuals. The diagnostics indicate that the GAR(1) model specification is better than the GARCH(1, 1) and EGARCH(1, 1) models for the USA and UK returns. However, the GAR(1) model does worse than the other two models for the Canadian and Japanese returns, consistent with the findings of low persistence in volatility for these two returns in panel A. Further model testing with post-sample data will be considered later.

In Table V, we find that the kurtosis captured by the GAR(1) model is fairly close to the kurtosis obtained from the stock return data for all four return series. This is broadly in agreement with our findings in Figure 1, except for Canadian index returns. It may be noted that the GAR(1) model is better able to capture the kurtosis for the series compared to that documented by Andersson (2001) using the normal inverse Gaussian SV model.

In Table VI, we present the results of an out-of-sample analysis with the squared return series. The parameters were estimated for the period 1997–2002 and the out-of-sample analysis is performed for the first 100 trading days of 2003.

For the GAR(1) model, we note that $E(h_{T+1}|y_T, y_{T-1} \ldots) = \phi h_T + (1 - \phi)p\theta$. Hence, we can estimate the one-step prediction of h_{T+j} given $y_{T+j-1}, y_{T+j-2}, \ldots$ as $h_{T+j} = E(h_{T+j}|y_{T+j-1}, y_{T+j-2} \ldots)$, $j = 1, 2, \ldots, 100$, by estimating ϕ, p and θ at every point in the post-sample period. From this, we obtain

DOI: 10.1002/for

Table IV. Parameter estimates and diagnostics of the volatility models of stock index returns (1997–2002: 1565 observations). The models are specified as follows: Returns: $y_t = \varepsilon_t \sqrt{h_t}$; GARCH(1,1): $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$; EGARCH(1, 1): $\ln(h_t) = \alpha_0 + \alpha_1 \left[\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \left(\frac{2}{\pi}\right)^{0.5} - \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right] + \beta h_{t-1}$; GAR(1): $h_t = \phi h_{t-1} + \eta_t$ where η_t is defined by equation (14)

Panel A: Parameter estimates

	GARCH(1,1)			EGARCH(1,1)				GAR(1)		
	α_0	α_1	β	α_0	α_1	β	γ	p	θ	ϕ
Canada TSE300	0.0367*** (0.0050)	0.0939*** (0.0104)	0.8830*** (0.0095)	0.0096*** (0.0031)	0.1491*** (0.0170)	0.9682*** (0.0038)	0.6520*** (0.1105)	0.6529*** (0.1988)	2.0848*** (0.6899)	0.2534 (0.1515)
Japan TOPIX	0.0747*** (0.0223)	0.0687*** (0.0128)	0.8899*** (0.0218)	0.0240*** (0.0061)	0.1129*** (0.0198)	0.9637*** (0.0095)	0.5883*** (0.1236)	1.6948*** (0.5020)	1.0647*** (0.3334)	0.3423* (0.1892)
UK FTSE100	0.0405*** (0.0111)	0.0951*** (0.0140)	0.8804*** (0.0180)	0.0099*** (0.0037)	0.1053*** (0.0221)	0.9755*** (0.0054)	0.9105*** (0.2548)	1.7962*** (0.6004)	1.1*** (0.3837)	0.9241*** (0.1172)
USA S&P500	0.0810*** (0.0178)	0.0888*** (0.0123)	0.8650*** (0.0187)	0.0204*** (0.0050)	0.0758*** (0.0178)	0.9551*** (0.0062)	2.0805*** (0.4887)	1.2909*** (0.3953)	1.3346*** (0.4417)	0.9398*** (0.2606)

Note: Standard errors are denoted in parentheses. ***, **, and * denote significance at the 1%, 5% and 10% levels.

Panel B: Diagnostics of standardized and squared standardized residuals—Ljung–Box statistics

	GARCH(1,1)		EGARCH(1,1)		GAR(1)	
	Standardized residuals	Squared standardized residuals	Standardized residuals	Squared standardized residuals	Standardized residuals	Squared standardized Residuals
	Q_{12}	Q_{12}^2	Q_{12}	Q_{12}^2	Q_{12}	Q_{12}^2
Canada TSE300	42.59***	2.25	57.00***	13.80	15.58	73.13***
Japan TOPIX	25.44**	19.88*	24.63**	30.63***	20.61*	51.68***
UK FTSE100	21.45**	5.55	37.18***	21.81**	14.12	0.47
USA S&P500	13.58	9.16	18.41	58.30***	7.98	0.09

Note: ***, **, and * denote significance at the 1%, 5% and 10% levels.

Table V. Kurtosis of stock returns implied by the estimated models (1997–2002: 1565 observations)

	Data	GARCH(1,1)	EGARCH(1,1)	GAR(1)
Canada–TSE300	7.5906	4.8872	6.5368	7.5949
Japan–TOPIX	4.7415	3.3952	4.4633	4.7701
UK–FTSE100	4.6780	4.7902	4.4903	4.6702
US–S&P500	5.3280	3.6351	5.3376	5.3240

Table VI. Out-of-sample, one-day volatility forecast evaluation for the first 100 days of trading in 2003

	Mean absolute errors of prediction			Ljung–Box statistics for GAR(1) model	
	GARCH(1,1)	EGARCH(1,1)	GAR(1)	Standardized residuals Q_{12}	Squared standardized residuals Q^2_{12}
Canada–TSE300	1.0272	1.0532	0.9330	13.14	20.10*
Japan–TOPIX	2.6893	2.7093	1.6072	15.73	18.14
UK–FTSE100	4.6068	4.1394	3.2201	0.91	0.13
USA–S&P500	3.3147	3.4275	2.5136	16.45	16.23

Note: ***, **, and * denote significance at the 1%, 5% and 10% levels.

the 'prediction' errors $d_j = y^2_{T+j} - \hat{h}_{T+j}$ as in Andersson (2001) and obtain the mean absolute errors,

$$MAE = \sum_{j=1}^{100} |d_j| / 100.$$

Table VI presents MAE for GARCH(1, 1), EGARCH(1, 1) and GAR(1) models and we observe that the GAR(1) model has the smallest mean absolute error of prediction in all four cases.

For further model evaluations with post-sample data, we compute the standardized 'errors', $\hat{\varepsilon}_{T+j} = y_{T+j}/\hat{h}_{T+j}$, $j = 1, 2, \ldots, 100$. The diagnostics in Table VI indicate that the autocorrelations of standardized and squared standardized residuals are negligible for all the series, except possibly for the Canadian returns series.

Recall that Figure 1 indicates that the GAR(1) model seems to respond well to the co-existence of a high kurtosis and low acf. Figure 1 also shows that the GAR specifications cover the empirical points denoted as data that relate to stock index returns for Japan, the USA and UK and suggests that the GAR(1) model captures the leptokurtic nature of return distributions and the slowly decaying autocorrelation functions of squared returns for three return series. Overall, we find that the GAR(1) model fits the stylized features of stock returns better than the GARCH(1, 1) and EGARCH(1, 1) models for the USA and UK stock index returns. It may be noted that Andersson (2001) reported that the GARCH(1, 1) model outperformed the normal inverse Gaussian SV model for Canada and the UK but did not for the USA and Japan for the period 1991–1998.

DOI: 10.1002/for

CONCLUSION

In this paper, we propose models for returns in which return volatilities evolve according to stationary gamma and exponential autoregressive specifications. We obtain moment estimators for the parameters of the models and show analytically that these estimators are consistent and asymptotically normal. In the in-sample analysis with stock index return data we find that return models with gamma stochastic volatility processes capture the leptokurtic nature of return distributions and the slowly decaying autocorrelation functions of squared returns for the USA and UK index return series. However, the GAR(1) model does not pick up the persistence in volatility for the Canadian and Japanese index returns as well as the GARCH(1, 1) and EGARCH(1, 1) models. It should also be noted that the GARCH(1, 1) and EGARCH(1, 1) models are not fitting the Japanese index return series well. In the out-of-sample analysis we show that the gamma autoregressive stochastic volatility has a superior forecasting performance compared to the commonly used GARCH and EGARCH models.

APPENDIX

EAR(1) process

The sequence $\{\Gamma_{(k)}\}$ of matrices is,

$$\Gamma_{(0)} = \begin{bmatrix} 5\theta^2 & \theta^3(5+5\phi+6\phi^2) \\ \theta^3(5+5\phi+6\phi^2) & \theta^4\begin{pmatrix} 36(3\phi^2+2\phi+1) \\ -(1+\phi\rho)^2 \end{pmatrix} \end{bmatrix}, \quad \text{for } k=0$$

$$\Gamma_{(1)} = \begin{bmatrix} \phi\theta^2 & \phi\theta^3(1+\phi+2\phi^2) \\ \theta^3(11\phi+5) & \theta^4\begin{pmatrix} 5+16\phi+17\phi^2 \\ +12\phi^3+18\phi^4 \end{pmatrix} \end{bmatrix}, \quad \text{for } k=1$$

$$\Gamma_{(k)} = \begin{bmatrix} \phi^k\theta^2 & \rho^k\theta^3(1+\phi+2\phi^2) \\ \phi^{k-1}\theta^3(1+\phi+2\phi^k) & \rho^k\theta^4\begin{pmatrix} 1+4\phi+3\phi^2 \\ +2\phi^{k+1}+10\phi^{k+2} \end{pmatrix} \end{bmatrix}, \quad \text{for } k=2,3,\cdots$$

and

$$\Gamma_{(-k)} = \Gamma_{(k)} \quad \text{for } k=1,2,\cdots.$$

GAR(1) process

Let

$$\Gamma_{(k)} = \begin{bmatrix} \gamma_{11}^{(k)} & \gamma_{12}^{(k)} & \gamma_{13}^{(k)} \\ \gamma_{21}^{(k)} & \gamma_{22}^{(k)} & \gamma_{23}^{(k)} \\ \gamma_{31}^{(k)} & \gamma_{32}^{(k)} & \gamma_{33}^{(k)} \end{bmatrix}, \quad k=0,\pm1,\pm2,\cdots,$$

DOI: 10.1002/for

where $\Gamma_{(k)} = \Gamma_{(-k)}$, $k = 1, 2, \cdots$.

The elements of $\Gamma_{(0)}$ are given by

$$\gamma_{11}^{(0)} = (2p^2 + 3p)\theta^2$$
$$\gamma_{12}^{(0)} = \gamma_{21}^{(0)} = 6p(p+1)(2p+5)\theta^3$$
$$\gamma_{13}^{(0)} = \gamma_{31}^{(0)} = p(6\phi^2 + 6\phi p + 2p + 3p^3 - \phi p^2)\theta^3$$
$$\gamma_{22}^{(0)} = 3p(p+1)(32p^2 + 172p + 207)\theta^4$$
$$\gamma_{23}^{(0)} = \gamma_{32}^{(0)} = 15p\theta^4\{p^3(1+6\phi^2) + p^2(6\phi^2 + 3\phi + 3) + p(6\phi^2 + 3\phi + 2) + 6\phi^3\} - 3p^2(p+1)(p+\phi)\theta^4$$
$$\gamma_{33}^{(0)} = 9\theta^4 p(p+1)(6\phi^2 + p + p^2 + 4p\phi) - \theta^4 p^2(\phi + p)^2$$

Similarly, the following are the elements of $\Gamma_{(k)}$:

$$\gamma_{11}^{(k)} = p\theta^2\phi^k, \quad k = 1, 2, \ldots$$
$$\gamma_{12}^{(k)} = 6p(p+1)\theta^3\phi^k, \quad k = 1, 2, \ldots$$
$$\gamma_{13}^{(k)} = p(p + p\phi + 2\phi^2)\theta^3\phi^k, \quad k = 1, 2, \ldots$$
$$\gamma_{21}^{(k)} = 6p(p + \phi^k)\theta^3\phi^k, \quad k = 1, 2, \ldots$$
$$\gamma_{22}^{(k)} = 18p(p+1)(2p + 3\phi^k)\theta^4\phi^k, \quad k = 1, 2, \ldots$$
$$\gamma_{23}^{(k)} = 6\theta^4\{p(p + 2p\phi + 3\phi)\phi^{k+1} + p^2(p + p\phi + 2\phi)\}\phi^{k+1}, k = 1, 2, \ldots$$
$$\gamma_{31}^{(1)} = p(5p\phi + 2p^2 + 6\phi + 3p)\theta^3$$
$$\gamma_{31}^{(k)} = p(p + p\phi + 2\phi^k)\theta^3\phi^{k-1}, \quad k = 2, 3, \ldots$$
$$\gamma_{32}^{(1)} = 6p(p+1)\theta^4(7p\phi + 2p^2 + 5p + 15\phi)$$
$$\gamma_{32}^{(k)} = 6p(p+1)\theta^4\{p(1-\phi) + 2p(1-\phi^{k-1}) + (2p+3)\phi^k\}\phi^{k-1}, \quad k = 2, 3, \ldots$$
$$\gamma_{33}^{(1)} = 3\theta^4\begin{bmatrix} p(p+1)\{6\phi^4 + 5p\phi^2 + p^2\phi^3 + (1-\phi)p\phi(6p^2 + 2\phi + 3p + 2)\} \\ +(1-\phi)p^3\{(1-\phi)(1+\phi+p-p\phi) - 6\phi^3\} \end{bmatrix} - p^2\theta^4(p+\phi)^2$$
$$\gamma_{33}^{(k)} = p\theta^4\left\{2(p+1)\phi^2(2p + 3\phi^k) + 2p\phi(1-\phi)(1 + 2p + \phi^k) + p^2(1-\phi)^2\right\}\phi^k, \quad k = 2, 3, \ldots$$

The matrix S for the GAR(1) SV model is given by $S = \Gamma_{(0)} + 2\sum_{k=1}^{\infty}\Gamma_{(k)}$.

We denote it by

$$S_2 = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

where

DOI: 10.1002/for

$$S_{11}=\left(\frac{p\theta^2}{1-\phi}\right)(2p(1-\phi)+3-\phi)$$

$$S_{12}=\left(\frac{6p(p+1)\theta^3}{1-\phi}\right)(5+2p-3\phi-2p\phi)$$

$$S_{13}=\left(\frac{p\theta^3}{1-\phi}\right)(6\phi^2+6p\phi+2p+3p^2-4p^2\phi-2\phi^3-4p\phi^2+p^2\phi^2)$$

$$S_{21}=\left(\frac{6p\theta^3}{1-\phi^2}\right)[(p+1)(2p+5)(1-\phi^2)+2\phi\{p(1+\phi)+\phi\}]$$

$$S_{22}=\left(\frac{3p(p+1)\theta^4}{1-\phi^2}\right)[(1-\phi^2)(32p^2+172p+210)+12\phi\{2p(1+\phi)+3\phi\}]$$

$$\begin{aligned}S_{23}&=15p\theta^4[p^3(1+6\phi^2)+p^2(6\phi^2+3\phi+3)+p(6\phi^2+3\phi+2)+6\phi^3]\\&\quad-3p^2(p+1)(p+\phi)\theta^4+\frac{12p\phi^2}{(1-\phi^2)}\{\phi^2(p+2p\phi+3\phi)+p(1+\phi)(p+p\phi+2\phi)\}\theta^4\end{aligned}$$

$$S_{31}=p\theta^3\left[(6\phi^2+16p\phi+8p+7p^2+12\phi-p^2\phi)+\left(\frac{2\phi}{1-\phi^2}\right)\left\{p(1+\phi)^2+2\phi^2\right\}\right]$$

$$\begin{aligned}S_{32}&=15p\theta^4[p^3(1+6\phi^2)+p^2(6\phi^2+3\phi+3)+p(6\phi^2+3\phi+2)+6\phi^3]-3p^2(p+1)(p+\phi)\theta^4\\&\quad+12p(p+1)\theta^4\left\{7p\phi+2p^2+5p+15\phi+\left(\frac{\phi}{1-\phi^2}\right)[p(3-\phi)(1+\phi)-2p\phi+(2p+3)\phi^2]\right\}\end{aligned}$$

$$\begin{aligned}S_{33}&=\theta^4\left\{9p(p+1)(6\phi^2+p+p^2+4p\phi)-p^2(p+\phi)^2\right\}\\&\quad+6\theta^4\begin{bmatrix}p(p+1)\{6\phi^4+5p\phi^2+p^2\phi^3+(1-\phi)p\phi(6\phi^2+2\phi+3p+2)\}\\+(1-\phi)p^3\{(1-\phi)(1+\phi+p-p\phi)-6\phi^3\}\end{bmatrix}-2p^2(p+\phi)^2\theta^4\\&\quad+2p\theta^4\begin{bmatrix}\left(\frac{\phi^2}{(1-\phi)}\right)\left(4p(p+1)\phi^2+2p\phi(1-\phi)(1+2p)+p^2(1-\phi)^2\right)\\+\left(\frac{\phi^4}{1-\phi^2}\right)\{6(p+1)\phi^2+2p\phi(1-\phi)\}\end{bmatrix}.\end{aligned}$$

ACKNOWLEDGEMENTS

We thank the editor and two anonymous referees for helpful suggestions that have improved the paper. The paper benefited from presentations at the 54th Session of the International Statistical Institute, 2003, Berlin; International Symposium on Forecasting, 2003, Merida, Mexico; and the Financial Econometrics Conference, 2002, University of Waterloo. B. Abraham was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Abraham B, Balakrishna N. 1999. Inverse Gaussian autoregressive models. *Journal of Time Series Analysis* **20**: 605–618.

DOI: 10.1002/for

Andersen TG, Sorensen BE. 1995. GMM estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Business and Economic Statistics* **14**: 328–352.
Andersson J. 2001. On the normal inverse Gaussian stochastic volatility model. *Journal of Business and Economic Statistics* **19**: 44–54.
Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**: 307–327.
Box GEP, Tiao GC. 1973. *Bayesian Inference in Statistical Analysis*. John Wiley: New York.
Chib S, Nardari F, Shephard N. 2002. Markov chain Monte Carlo methods for generalized stochastic volatility models. *Journal of Econometrics* **108**: 281–316.
Duffie D, Singleton KJ. 1993. Simulated moments estimation of Markov models of asset process. *Econometrica* **61**: 929–952.
Engle RF. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**: 987–1007.
Fridman M, Harris L. 1998. Maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business and Economic Statistics* **16**: 284–291.
Gallant AR, Tauchen G. 1998. Reprojection of partially observed systems with applications to interest rate diffusions. *Journal of the American Statistical Association* **93**: 10–24.
Gallant AR, Hsieh D, Tauchen G. 1997. Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics* **81**: 159–192.
Gaver DP, Lewis PAW. 1980. First order autoregressive gamma sequences and point processes. A*dvances in Applied Probability* **12**: 727–745.
Ghysels E, Harvey AC, Renault E. 1996. Stochastic volatility. In *Statistical Methods in Finance*, Rao CR, Maddala GS (eds). North-Holland: Amsterdam; 119–191.
Hansen LP. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* **50**: 1029–1054.
Harvey AC, Ruiz E, Shephard N. 1994. Multivariate stochastic variance models. *Review of Economics Studies* **61**: 247–264.
Hsieh DA. 1991. Chaos and non-linear dynamics: applications to financial markets. *Journal of Finance* **46**: 1839–1877.
Jacquier E, Polson NG, Rossi PE. 1994. Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics* **12**: 371–417.
Kim S, Shephard N, Chib S. 1998. Stochastic volatility, likelihood inference and comparison with GARCH models. *Review of Economic Studies* **65**: 361–393.
Lawrance AJ, Lewis PAW. 1981. A new autoregressive time series model in exponential variables (NEAR(1)). A*dvances in Applied Probability* **13**: 826–845.
Liesenfeld R, Jung RC. 2000. Stochastic volatility models: conditional normality versus heavy-tailed distributions. *Journal of Applied Econometrics* **15**: 137–160.
Mathai AM. 1993. Generalized Laplace distribution with application. *Journal of Applied Statistical Science* **1**: 169–178.
Melino A, Turnbull SM. 1990. Pricing foreign currency options with stochastic volatility. *Journal of Econometrics* **45**: 239–265.
Sandmann G, Koopman SJ. 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics* **87**: 271–301.
Shephard N. 1996. Statistical aspects of ARCH and stochastic volatility. In *Time Series Models in Econometrics, Finance and Other Fields*, Cox DR, Hinkley DV, Barndorff-Nielsen OE (eds). Chapman & Hall: London; 1–67.
Sim CH. 1986. Simulation of Weibull and gamma autoregressive stationary processes. *Communication Stat. Simul-Computations* **15**: 1141–1146.
Sim CH. 1990. First order autoregressive model for gamma and exponential processes. *Journal of Applied Probability* **27**: 325–332.
Taylor SJ. 1986. *Modelling Financial Time Series.* John Wiley: Chicester.
Taylor SJ. 1994. Modelling Stochastic Volatility. *Mathematical Finance* **4**: 183–204.
Tsay RS. 2002. *Analysis of Financial Time Series.* Wiley Interscience: New York.

DOI: 10.1002/for

Authors' biographies:
Bovas Abraham is Professor of Statistics in the Department of Statistics and Actuarial Science at the University of Waterloo, Canada. His research interests are in time-series analysis and the application of statistics in business and industry.

N. Balakrishna is Associate Professor of Statistics in the Department of Statistics at the Cochin University of Science and Technology, Cochin, India. His research interests include statistical analysis of time series, inference for stochastic processes and chaotic communications.

Ranjini Sivakumar is Associate Professor of Finance in the Centre for Advanced Studies in Finance, School of Accountancy, University of Waterloo, Canada. Her research interests are in financial econometrics and corporate finance.

Authors' addresses:
Bovas Abraham, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada.

N. Balakrishna, Department of Statistics, Cochin University of Science and Technology, Cochin 682022, India.

Ranjini Sivakumar, Centre for Advanced Studies in Finance, School of Accountancy, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada.

DOI: 10.1002/for