Estimation of Stochastic Volatility Models by Simulated Maximum Likelihood Method

by

Ji Eun Choi

A research paper presented to the
University of Waterloo
in partial fulfillment of the requirements for the degree of
Master of Mathematics
in
Actuarial Science

Waterloo, Ontario, Canada, 2004

Abstract

The Stochastic Volatility (SV) model, introduced by Taylor (1986) is used for capturing the empirical properties of financial time series. However, most of the applications are based on the assumption that the conditional distribution of the returns given the log volatilities is normal. This paper overviews those properties and compares the SV model with the heavy-tailed error distribution(Student t-distribution) and the SV model with the normal error distribution. The Simulated Maximum Likelihood (SML) method is applied to estimate the parameters and the latent volatility. Furthermore an empirical analysis with several return series shows that the SV-t specification adequately account for the well-known properties of the financial series: a high kurtosis of the returns and low but slowly decaying autocorrelation or the squared returns.

1 Introduction

It is generally acknowledged that the volatility of many financial returns series is not constant over time. Over the past two decades two main classes of models have been developed that capture the time-varying autocorrelated volatility process: the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and the Stochastic Volatility (SV) model. GARCH models define the time-varying variance as a deterministic function of past squared innovations and lagged conditional variances whereas SV model defines volatility as a logarithmic first order autoregressive process. Although SV models are more sophisticated than GARCH models, their empirical application has been limited under the assumption that the conditional distribution of returns is normal, given the latent volatility process. This SV-normal model is not able to capture the empirical regularities of financial return series: first, volatility clustering is often observed. That is, large changes tend to be followed by large changes and small changes tent to be followed by small changes; second, financial time series often exhibit leptokurtosis, meaning that their distribution is symmetrical in shape, similar to a normal distribution, but the center peak is much higher, so it has a fat tail; third, squared returns exhibit serial correlation whereas little or no serial dependence can be detected in the return series itself. In the papers of Ruiz (1994), Harvey, Ruix and Shephard (1994), Sandmann and Koopman (1998) and Chib, Nardari and Shephard (1998) the SV model is extended to allow the conditional distribution of the returns to be more heavy-tailed distribution than the normal distribution by using Student t-distribution for the standardized residual.

The purpose of this paper is to examine the ability of the SV model to capture the properties of financial return series mentioned earlier under the assumptions based on a conditional normal distribution for log volatility and a conditional Student t-distribution for the returns. It is finally shown that such assumptions regarding the conditional distribution of the returns systematically affects the estimates of the parameters in the latent volatility process.

Another reason for the limited empirical applications of the SV model is the difficulty to evaluate the likelihood function directly. The marginal likelihood function of the SV model is given by a high dimensional integral , which cannot be calculated by standard maximum likelihood method (ML). In this paper, the simulated maximum likelihood (SML) approach developed by Danielsson and Richard (1993) is employed to calculate this integral. This estimation method allows us to demonstrate the statistical inference with the standard instruments of inference for the ML method.

This paper is organized as follows: Section 2 describes volatility models and more details regarding SV model. Section 3 contains the data which have been used for this study and some statistics of the returns and the squared returns. Section 4 describes the SML and the Accelerated Gaussian Importance Sampling (AGIS) method and Section 5 provides the results of parameter estimation and some diagnostic checking for our model. Finally, section 6 gives a summary.

2 Volatility Models

A series is called conditionally heteroscedastic if the conditional variance depends on time while the unconditional variance is constant. Throughout the paper, volatility means the conditional variance of an asset return. Models of such a volatility are referred to as the conditional heteroscedastic models. Since volatility evolves over time, modelling the volatility plays an important role in both option trading and risk management. Also, it can improve the efficiency in parameter estimation and the accuracy in interval forecast.

Three main univariate models are discussed in this section: the autoregressive conditional heteroscedastic (ARCH) model of Engle (1982), Bollerslev's (1986) the generalized ARCH (GARCH), and the stochastic volatility (SV) model introduced by Taylor (1986).

2.1 The ARCH and GARCH Models

Engle's paper (1982) introduced the ARCH model to express the conditional variance of today's return as a function of previous observations. The basic idea of the ARCH model is that the mean-corrected asset returns are serially uncorrelated, but dependent of past observations. The ARCH(q) model is defined by

$$\begin{aligned} r_t &= \sqrt{\lambda_t} u_t, \quad u_t \backsim iid(0,1), \\ \lambda_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i}^2, \end{aligned}$$

where r_t is the return on day t and u_t is a white noise process. $\sum_{i=1}^{q} \alpha_i < 1$ is a necessary and sufficient condition for r_t to be a weakly stationary series. The order q of the process determines the volatility persistence, which increases with the value of q.

The Generalized ARCH (GARCH) model is an extension of Engel's work by Bollerslev (1986) that allows the conditional variance to depend on the previous conditional variance and the squares of previous returns. The possibility that estimated parameters in ARCH models do not satisfy the stationarity condition increases with lag. Thus GARCH model is an alternative to ARCH model. The GARCH(p,q) model is defined by

$$r_t = \sqrt{\lambda_t} u_t, \qquad u_t \backsim iid(0,1),$$
$$\lambda_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i}^2 + \sum_{j=1}^{p} \beta_j \lambda_{t-j},$$

where λ_t is the conditional variance of r_t given $R_{t-1} = (r_{t-1}, r_{t-2}, \ldots)$ and the parameters $(\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)$ are restricted such that $\lambda_t > 0$ for all t.

2.2 The SV model

The standard version of the Stochastic Volatility (SV) model is given by

$$r_t = exp\{\lambda_t/2\} u_t, \tag{1}$$

$$\lambda_t = \alpha + \beta \lambda_{t-1} + \gamma v_t, \tag{2}$$

where r_t is the return on day t and λ_t is the log volatility, that is $\lambda_t = 2 \ln \sigma_t$ where σ_t is the return volatility. Both u_t and v_t are identically and independently distributed random variables with zero mean and unit variance, but v_t is normally distributed. These error processes are stochastically independent and unobservable. The unobservable volatility process λ_t is assumed as a Gaussian AR(1) process with a persistence parameter β. For $\mid \beta \mid < 1$, the latent volatility process is stationary. The parameter γ^2 represents the variance of volatility shocks and is assumed to be positive. In most cases, the error u_t is assumed to be normal, but it is observed that a heavy-tailed and leptokurtic distribution for u_t better captures the empirical regularity of the financial time series. In the next section we give some statistical properties of two SV models: the SV model with a normally distributed error u_t (SV-Normal) and the SV model with a heavy-tailed error u_t(SV-t).

2.2.1 Properties of the SV-Normal model

For $\mid \beta \mid < 1$, the latent volatility process $\lambda_t \sim N(\mu, \sigma^2)$ with $\mu = \frac{\alpha}{1-\beta}$ and $\sigma^2 = \frac{\gamma^2}{1-\beta^2}$ denote the unconditional mean and variance of λ_t, respectively. Letting $\sigma_t^2 = exp\{\lambda_t\}$ and assuming that $E(u_t^4) < \infty$, the moments $E(r_t^2)$ and $E(r_t^4)$ are expressed as follows

$$E(r_t^2) = E(\sigma_t^2) = exp\{\mu + \sigma^2/2\}, \tag{3}$$

$$E(r_t^4) = E(\sigma_t^2)^2 exp\{\sigma^2\}E(u_t^4) = exp\{2\mu + 2\sigma^2\}E(u_t^4) \tag{4}$$

Putting equations (3) and (4) into the definition of the kurtosis, the kurtosis for the unconditional distribution of the returns is given by

$$\kappa = E(r_t^4)/E(r_t^2)^2 = E(u_t^4)exp\{\sigma^2\}. \tag{5}$$

This expression of the predicted kurtosis in equation (5) has two components: the first one is the baseline-kurtosis due to the term $E(u_t^4)$ that represents the kurtosis of the standardized residuals and the second one is the kurtosis due to the variation in the volatility process λ_t. Under the SV model with a conditional normal distribution for the returns(SV-normal), the baseline-kurtosis $E(u_t^4)$ is equal to three and so an unconditional kurtosis of the returns is greater than three.

The autocorrelation function (ACF) of the squared returns is defined by

$$\rho(h) = Cov(r_t^2, r_{t-h}^2)/Var(r_t^2), \qquad h = 1, 2, \ldots \tag{6}$$

Provided $\mid \beta \mid < 1$, the autocovariance of the squared returns is give by

$$\begin{aligned} Cov(r_t^2, r_{t-h}^2) &= Cov(\sigma_t^2, \sigma_{(t-h)}^2) \qquad \textit{by independence of } u_t \\ &= (exp\{\sigma^2\beta^h\} - 1)E(\sigma_t^2)^2 \end{aligned} \tag{7}$$

Under the additional assumption $E(u_t^4) < \infty$, its variance is

$$\begin{aligned} Var(r_t^2) &= E(\{\sigma_t^2\}^2 u_t^4) - E(h_t u_t^2)^2 \\ &= E(\{\sigma_t^2\}^2)E(u_t^4) - E(\sigma_t^2)^2 \qquad \textit{by independence} \\ &= (Var(\sigma_t^2) + E(\sigma_t^2)^2)E(u_t^4) - E(\sigma_t^2)^2 \\ &= E(\sigma_t^2)^2 \left\{ \left(\frac{Var(\sigma_t^2)}{E(\sigma_t^2)^2} + 1 \right) E(u_t^4) - 1 \right\} \\ &= E(h_t)^2[exp\{\sigma^2\}E(u_t^4) - 1]. \end{aligned} \tag{8}$$

Inserting the equations (7) and (8) into the equation (6), the ACF of the squared return r_t^2 is given by

$$\rho(h) = \frac{exp\{\sigma^2\beta^h\} - 1}{E(u_t^4)exp\{\sigma^2\} - 1}, \qquad h = 1, 2, \ldots. \tag{9}$$

Thus, the SV-normal model predicts a positive autocorrelation in the squared returns which is exponentially decaying out. The rate is determined by the parameter β and hence the persistence of volatility shocks depends on the value of β.

2.2.2 Properties of the SV-t model

The heavy-tailed distributions for the error u_t means that the fourth moment of u_t is greater than three. This characteristic implies that the kurtosis of a conditional heavy-tailed distribution for the returns is larger than that of conditional normal returns and that the level of ACF of the squared returns declines. It can be verified by the equations (5) and (9). A well-known leptokurtic distribution is the student t-distribution. The density function of a t-distributed random variable u_t with mean zero and unit variance is given by

$$f(u_t) = [\pi(\omega-2)]^{-1/2}\frac{\Gamma((\omega+1)/2)}{\Gamma(\omega/2)}\left[1 + \frac{u_t^2}{(\omega-2)}\right]^{-\frac{(\omega+1)}{2}} \qquad \omega > 2 \tag{10}$$

where $\Gamma(\cdot)$ is the gamma function and the parameter ω denotes the degree of freedom.

The kurtosis of the t-distribution is give by, as long as $\omega > 4$,

$$\begin{aligned}\kappa &= exp\{\sigma^2\}E(u_t^4) \\ &= exp\{\sigma^2\}\left[3\frac{(\omega-2)}{(\omega-4)}\right].\end{aligned}$$

If $\omega < \infty$, the kurtosis is greater three and does not exist if $\omega = 4$. Also, the t-distribution approaches a normal distribution as ω goes to infinity. Note that $\rho(1) = \frac{exp\{\sigma^2\beta\}-1}{E(u_t^4)exp\{\sigma^2\}-1}$ when $h = 1$ and so κ can be considered as a function of $\rho(1)$. Liesenfeld and Jung (2000) shows using an empirical study that the SV-t model can capture the relationship between κ and $\rho(1)$ better than the SV-normal model. Therefore, it can be concluded in some financial

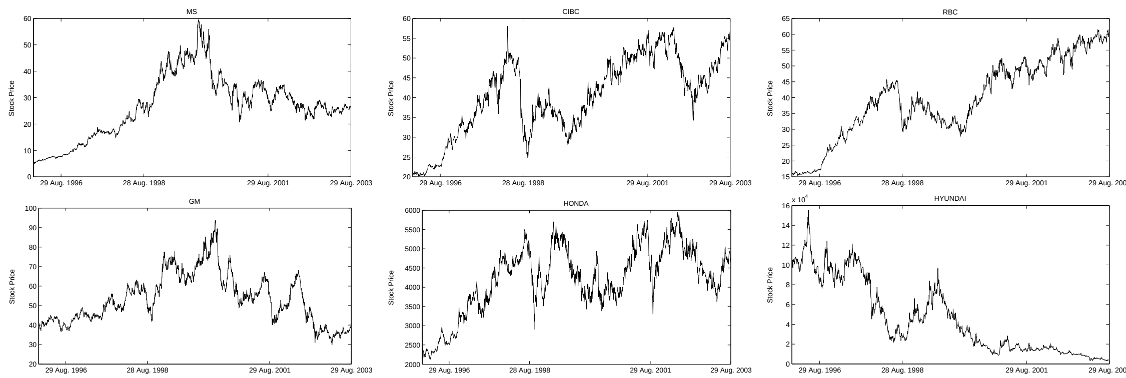

Figure 1: Plots of the stock prices over time

return series that a leptokurtic error distribution such as a Student t-distribution helps to capture the empirical properties of the financial series: a high kurtosis of the returns and a low but slowly decaying ACF of the squared returns.

3 DATA

Daily return series are used in this paper: stock prices of Microsoft, CIBC, Royal Bank(RBC), General Motors(GM), Honda, and Hyundai. All the prices have the same period from January 1, 1996 to August 29, 2003, so the total number of observations is 2000. The data are provided from the Department of Accounting at the University of Waterloo. Figure 1. illustrates the stock prices versus time.

Denoting the daily price by p_t, the transformed log return is give by

$$r_t = \ln \left\{ \frac{p_t}{p_{t-1}} \right\}.$$

The summary statistics of the return series are given in Table 1. The kurtosis of the returns for all series is above three, which is the kurtosis of a normal distribution. These values imply that the distribution of the returns is leptokurtic.

From the Autocorrelation(ACF) of the returns in Figure 2, it is observed that the log return series are serially uncorrelated, but not independent as usual. However, in CIBC and RBC series the first lag autocorrelation appears to be large. This is because the limits are very narrow due to the large number of observations. Figure 3 gives some evidence for the well-known properties of the financial return series that the ACF of the squared returns is low, but declines with increasing lags very slowly. This is an indication that the conditional variance depends on time.

Table 1. Summary statistics of data

Statistics	MS	RBC	CIBC	GM	HONDA	HYUNDAI
Sample size	2000	2000	2000	2000	2000	2000
Mean	0.00079	0.00067	0.00051	-0.00000	0.00040	-0.0016
Std. dev.	0.0249	0.0151	0.0186	0.0214	0.0236	0.0485
Kurtosis	7.3005	5.5355	8.7963	5.5507	7.0661	4.9994
Minimum	-0.1697	-0.0801	-0.1704	-0.1454	-0.1502	-0.1618
Maximum	0.1786	0.0795	0.1	0.0984	0.1318	0.2256

4 Estimation Method

4.1 Parameter estimation

The likelihood function associated with the known observations $R = \{r_t\}_{t=1}^T$ and the vector of the latent variables $\Lambda = \{\lambda_t\}_{t=1}^T$ is given by

$$f(R|\theta) = \int_{\mathbb{R}^T} f(R, \Lambda|\theta) d\Lambda, \tag{11}$$

where $\theta = (\alpha, \beta, \gamma)$ denotes the vector of parameters to be estimated and T is the total number of observations. The latent process λ_t in the SV model makes the direct calculation of the integral in (11) difficult. In this paper, the Simulated Maximum Likelihood (SML) approach is employed, which was introduced by Danielsson and Richard (1993) to estimate the parameters in the model. This method depends on Monte Carlo (MC) integration in evaluating the likelihood (11).

In finite sample space, the SML method performs almost identical to MCMC(See Danielsson, 1994). The standard instruments for inference in ML estimation can be used even

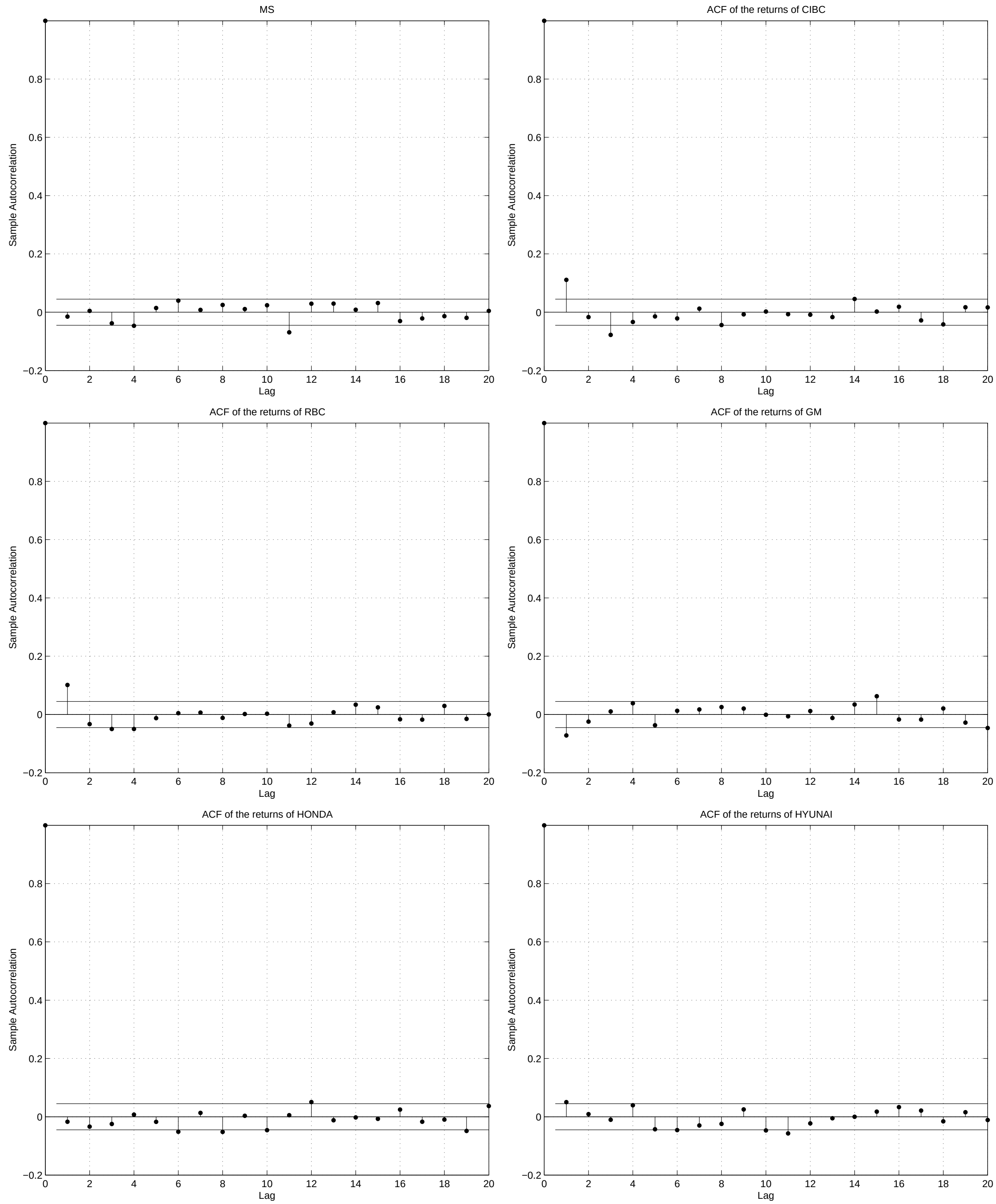

Figure 2: The ACF of the returns on stock return series

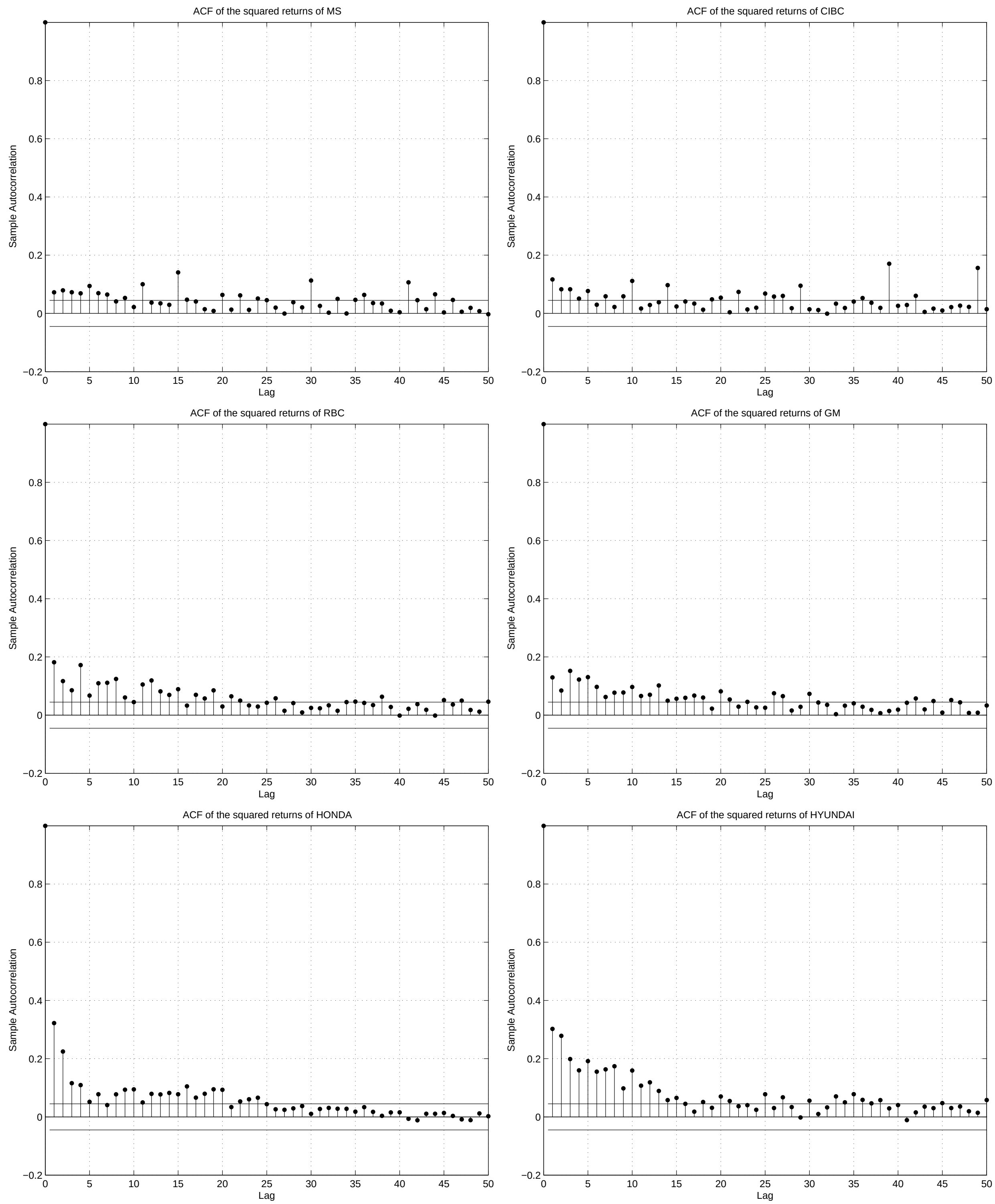

Figure 3: The ACF of the squared returns on stock return series

if the number of MC iterations are very large. Moreover, the applications to SV models can not only be obtained easily once SML process is implemented but extended to the multivariate case with the latent process λ_t.

Danielsson and Richard (1993) introduced the SML approach using an important sampling method and an Accelerated Gaussian Importance Sampling (AGIS). Next, we discuss the SML method with these two sampling techniques.

4.1.1 Important Sampling (IS)

To obtain the MC estimate of $f(R|\theta)$, the joint density function of $f(R,\Lambda|\theta)$ is factorized into an importance sampling function (IF) $\psi(\Lambda|R)$ and a remainder function (RF) $\phi(\Lambda, R)$ such that

$$f(R,\Lambda|\theta) = \phi(\Lambda, R)\psi(\Lambda|R). \tag{12}$$

An initial factorization of $f(R,\Lambda|\theta)$ is derived from (12) as follows:

$$\psi_0(\Lambda|R) = \prod_{t=1}^{T} f(\lambda_t|\lambda_{t-1}) \tag{13}$$

$$\phi_0(\Lambda, R) = \prod_{t=1}^{T} f(r_t|\lambda_t) \tag{14}$$

where $f(\lambda_t|\lambda_{t-1})$ is the conditional density of λ_t given λ_{t-1} satisfying the equation (2), which is a normal distribution and $f(r_t|\lambda_t)$ is the density of the t-th day return conditional on λ_t. In a SV-normal model, $f(r_t|\lambda_t)$ is given by

$$f(r_t|\lambda_t) = \frac{1}{\sqrt{2\pi exp\{\lambda_t\}}} exp\left\{-\frac{r_t^2}{2exp\{\lambda_t\}}\right\},$$

and in a SV-t distribution model, $f(r_t|\lambda_t)$ has a form

$$f(r_t|\lambda_t) = \frac{1}{\sqrt{\pi(\omega-2)exp\{\lambda_t\}}} \frac{\Gamma((\omega+1)/2)}{\Gamma(\omega/2)} \left[1 + \frac{r_t^2}{exp\{\lambda_t\}(\omega-2)}\right]^{-\frac{(\omega+1)}{2}}.$$

Since the expected value of the RF is given by

$$\begin{aligned} E_\psi[\phi(\Lambda, R)] &= \int_{\mathbb{R}^T} \phi(\Lambda, R)\psi(\Lambda|R)\, d\Lambda \\ &= \int_{\mathbb{R}^T} f(R, \Lambda|\theta)\, d\Lambda \\ &= f(R|\theta), \end{aligned} \tag{15}$$

the MC sampling mean is given by

$$\hat{f}_N(R|\theta) = \frac{1}{N}\sum_{n=1}^{N} \phi_0(\Lambda_n, R) \tag{16}$$

where $\{\Lambda_n\}_{n=1}^N$ denotes N numbers of the simulated sample from the probability distribution $\psi(\Lambda|R)$. Thus the ML estimate of θ is obtained by maximizing $\ln[\hat{f}(R|\theta)]$.

4.1.2 Accelerated Gaussian Important Sampling (AGIS)

It often occurs that the initial IF (13) and RF (14) in IS technique leads to inefficiency in the integral calculation. That is, the MC sampling variance remarkably increases with the dimension of the integral T. The Accelerated Gaussian Important Sampling (AGIS) method, proposed by Danielsson and Richard (1993) can solve this inefficiency problem. The AGIS method is based on minimizing the variance of the remainder function $\phi(\Lambda, R)$ while the conditions (12) and (15) hold. That is, the process used in AGIS solves a minimizing problem

$$\min_{\psi} Var_\psi[\phi(\Lambda, R)] \tag{17}$$

subject to the constraints

$$f(R, \Lambda|\theta) = \phi(\Lambda, R)\psi(\Lambda|R) \quad and \quad E_\psi[\phi(\Lambda, R)] = f(R|\theta) \tag{18}$$

where the variance of $\phi(\Lambda, R)$ evaluated over an importance function $\psi(\Lambda|R)$ is given by

$$Var_\psi[\phi(\Lambda, R)] = \int_{\mathbb{R}^T} [\phi(\Lambda, R) - f(R|\theta)]^2 \psi(\Lambda|R)\, d\Lambda.$$

Let us define a variance reduction function $\xi(\Lambda, \mathcal{Q})$ such that a new IF and a new RF are given by

$$\psi_1(\Lambda|R) = [\psi_0(\Lambda|R)\xi(\Lambda, \mathcal{Q})]/k(\mathcal{Q}) \tag{19}$$

$$\phi_1(\Lambda, R) = [\phi_0(\Lambda, R)k(\mathcal{Q})]/\xi(\Lambda, Q), \tag{20}$$

where $k(\mathcal{Q})$ is the constant which makes the new IF $\psi_1(\Lambda|R)$ a probability density function and is given by

$$k(\mathcal{Q}) = \int_{\mathbb{R}^T} \psi_0(\Lambda|R)\xi(\Lambda, \mathcal{Q})\, d\Lambda. \tag{21}$$

These transformations for the new RF and IF are suggested to change the variance of the remainder function while keeping the constraints (18) and $\xi(\Lambda, \mathcal{Q})$ is defined by

$$\xi(\Lambda, \mathcal{Q}) = \prod_{t=1}^{T} \xi(\lambda_t, \mathcal{Q}_t) \tag{22}$$

with

$$\xi(\lambda_t, \mathcal{Q}_t) = exp\{-\frac{1}{2}\eta_t' \mathcal{Q}_t \eta_t\} \quad and \quad \eta_t' = (\lambda_t, \lambda_{t-1}, 1).$$

This form reduces the computational burden for calculating a new IF and RF. To obtain $\mathcal{Q} = \{\mathcal{Q}_t\}_{t=1}^T$ and hence $\xi(\Lambda, \mathcal{Q})$, the following iterations are needed.

- **Step 0 (Initialization)**

  1. Generate a set of N independent random vectors $\{\lambda_{0,t}\}$ for $t = 1, 2, \ldots, T$, satisfying the following conditions: When $t = 1$, $\lambda_{0,t}|\, \theta \sim \mathbf{N}(\alpha_0, \gamma_0^2)$, where α_0 and γ_0 are initial values of α and γ respectively. When t is greater than one, $\lambda_{0,t}|\, \lambda_{t-1}, \theta \sim \mathbf{N}(\alpha + \beta\lambda_{t-1}, \gamma^2)$.
  2. Construct the initial IF $\psi_0(\Lambda|R)$ in (13) by calculating the normal density of $\lambda_{0,t}$ conditional on $\lambda_{0,t-1}$ and the initial RF $\phi_0(\Lambda, R)$ in (14).

- **Step 1**

  1. Run the following linear regression for $t = 1, 2, \ldots, T$ and $n = 1, 2, \ldots, N$:

  $$\ln \phi_0(\lambda_{0,n,t}) = a_{1,t}\lambda_{0,n,t}^2 + b_{1,t}\lambda_{0,n,t} + c_{1,t} + \varepsilon, \tag{23}$$

  where ε is a residual and $\phi_0(\cdot)$ is obtained from the previous step.

  2. Construct the matrix $\mathcal{Q}_1 = \{\mathcal{Q}_{1,t}\}_{t=1}^T$ with the OLS estimates of the coefficients, which is given by,

  $$\hat{\mathcal{Q}}_{1,t} = \begin{pmatrix} -2\hat{a}_{1,t} & 0 & -\hat{b}_{1,t} \\ 0 & 0 & 0 \\ -\hat{b}_{1,t} & 0 & -2\hat{c}_{1,t} \end{pmatrix}$$

  for $t = 1, 2, \ldots, T$.

  3. Construct a first new IF $\psi_1(\Lambda|R) = \psi_0(\Lambda|R)\xi(\hat{\mathcal{Q}}_1)/k(\hat{\mathcal{Q}}_1)$ with $\hat{\mathcal{Q}}_1 = \{\hat{\mathcal{Q}}_{1,t}\}_{t=1}^T$. This implies that a new random variable $\lambda_{1,t}$ has a distribution, given by

  $$\mathbf{N}\left(\frac{\alpha_0 + \beta_0\lambda_{0,t-1} + (\sqrt{2\hat{a}_{1,t}})^{-1}\gamma_0\hat{b}_{1,t}}{(1-\sqrt{2\hat{a}_{1,t}})\gamma_0},\ \gamma_0^2(1-\gamma_0\sqrt{2\hat{a}_{1,t}})^2\right). \tag{24}$$

- **Step i**

  1. Construct the $i-th$ matrix $\mathcal{Q}_i = \{\mathcal{Q}_{i,t}\}_{t=1}^T$ with the coefficients obtained by regressing $\ln \phi_0(\lambda_{i-1,n,t})$ on $\lambda_{i-1,n,t}$ and $\lambda_{i-1,n,t}^2$

  2. Determine a $i-th$ step IF $\psi_i(\Lambda|R) = \psi_{i-1}(\Lambda|R)\xi(\hat{\mathcal{Q}}_{i-1})/k(\hat{\mathcal{Q}}_{i-1})$.

This iteration algorithm is repeated until $\hat{\mathcal{Q}}_i$ is sufficiently close to $\hat{\mathcal{Q}}_{i-1}$. Usually, the number of iterations is less than five. Then the MC sample mean is given by

$$\hat{f}_N(R|\,\theta) = \frac{1}{N}\sum_{n=1}^{N}\frac{\phi_0(\Lambda_{i,n}, R)k(\hat{\mathcal{Q}}_i)}{\xi(\Lambda_{i,n}, \hat{\mathcal{Q}}_i)} \tag{25}$$

The SML estimate of θ is hence obtained by maximizing the AGIS estimate of the likelihood function (25) with respect to θ.

4.2 Volatility Estimation

Once the estimates of parameters are obtained, the unobservable volatility σ_t is estimated as the conditional expectation $E(\sigma_t|\mathrm{R},\ \hat{\theta})$ where $\sigma_t = exp\{\lambda/2\}$. $\hat{\theta}$ is the SML estimate of the parameter θ and R is the vector of the returns. The conditional expectation is given by

$$E(\sigma_t|\mathrm{R},\ \hat{\theta}) = \frac{\int_{\mathbb{R}^T} exp\{\lambda_t/2\} f(\mathrm{R},\Lambda|\hat{\theta}) d\Lambda}{\int_{\mathbb{R}^T} f(\mathrm{R},\Lambda|\hat{\theta}) d\Lambda} \tag{26}$$

$$= \frac{E_\psi[exp\{\lambda_t/2\}\phi(\Lambda, R)]}{E_\psi[\phi(\Lambda, R)]} \tag{27}$$

The equation (27) can be derived by factorizing the joint density $f(\mathrm{R},\Lambda|\hat{\theta})$ by (12) and applying then (13) to both integrals in (26). In order to evaluate the conditional expectation $E(\sigma_t|\mathrm{R},\hat{\theta})$, determine the IF $\psi(\Lambda|R)$ by the AGIS algorithm and compute then the estimate of the expectations in (27) by using MC integration. Finally, the MC estimates of $E(\sigma_t|\mathrm{R},\hat{\theta})$ are obtained as a by-product of the likelihood evaluation given $\hat{\theta}$ and the IF $\psi(\Lambda|R)$.

5 Empirical Results

In our empirical example, the AGIS method is used with the number of observations $T = 2000$, a MC sample size of $N = 50$, and AGIS iterations of $J = 4$. Moreover, the MC standard deviations are calculated by repeating the estimation thirty times.

5.1 Parameter Estimation Results

The SV-normal estimation results are summarized in Table 2 along with the statistical standard error of the parameter estimates and the MC sampling standard deviation of the parameter estimates. The small MC standard deviations and the standard errors of the estimates indicate that the SML estimates are quite precise. The estimated β are highly significant in all cases. However, MS and RBC data, the $\hat{\beta}'s$ are slightly lower than the rest.

Table 2. SML estimation of the SV-normal model

Parameter	MS	RBC	CIBC	GM	HONDA	HYUNDAI
α	0.0034	-0.03426	-0.00897	0.00188	-0.0212	0.00498
MC std. dev.	0.0045	0.0005	0.0004	0.0024	0.0013	0.0027
Std. error	0.0122	0.0024	0.0428	0.0423	0.0488	0.0198
β	0.8998	0.91243	0.97236	0.97822	0.95478	0.95132
MC std. dev.	0.0065	0.0002	0.0005	0.0004	0.0011	0.0042
Std. error	0.0124	0.0498	0.0591	0.0379	0.0102	0.0628
γ	0.14352	0.10982	0.15421	0.28423	0.10845	0.1245
MC std. dev.	0.0009	0.0012	0.0014	0.0007	0.0007	0.0041
Std. error	0.0599	0.0544	0.0521	0.0547	0.0573	0.0106

The estimation results for the returns with a Student t distribution with $\omega = 10$ are displayed in Table 3. It is easily observed that the MC standard deviations and the standard errors for all the parameter estimates are quite smaller than the ones from the SV-normal model. This result may be already expected from the fact that the unconditional variance of the latent process is equal to $\sigma^2 = \gamma^2/(1-\beta^2)$, which happens to be smaller for the SV-t than the SV-normal in all the series. The estimates of β are all around 0.95, which are reasonable. Furthermore, the standard error of the parameter estimates in most cases are quite small, but the estimates of β and γ for GM are somewhat high. The estimated values of kurtosis in most cases are quite close to the actual value in Table. 1 except for MS data.

5.2 Volatility Estimation Result

In Figure 4, 5, and 6, the MC estimates of the sequence of volatilities $E(\sigma_t|\, R, \hat{\theta})$ resulting from the SV-normal(dash line) and the SV-t models(thick line) are presented along with the log returns. Note that all the log returns seem to be stationary compared to the actual stock prices in Figure 1. For MS data, the volatility estimates from SV-normal does not seem to reflect the movement of the return series compared to the one from SV-t model. In the estimated volatility plot for GM shows that the estimated volatility from SV-normal model is slightly more stable than from SV-t model and both estimates do not reflect the movement of the return series. In the rest of cases, the estimated volatilities from the SV-t model exhibit smoother movements than the ones from the SV-normal model.

Table 3. SML estimation of the SV-t distribution model with $\omega = 10$

Parameter	MS	RBC	CIBC	GM	HONDA	HYUNDAI
α	0.00234	-0.02896	-0.00761	0.00082	-0.01925	0.00344
MC std. dev.	0.0001	0.0003	0.0003	0.0005	0.0006	0.0004
Std. error	0.0073	0.0089	0.0072	0.0323	0.0347	0.0082
β	0.92213	0.96356	0.98876	0.98245	0.99241	0.98211
MC std. dev.	0.00002	0.0001	0.0004	0.0001	0.0002	0.00005
Std. error	0.0088	0.0453	0.0291	0.0841	0.0377	0.0379
γ	0.09318	0.08997	0.12653	0.11423	0.08793	0.08436
MC std. dev.	0.0006	0.0003	0.0002	0.0002	0.0004	0.0007
Std. error	0.0512	0.0291	0.0376	0.0772	0.0312	0.0536
$\hat{\kappa}$	4.2389	4.4791	8.1866	5.8202	6.6697	4.8890

5.3 Model Diagnostics

Now some model validation is needed for justifying our model. Let us define the standardized error from equation (1):

$$\hat{u}_t \equiv \frac{r_t}{\hat{\sigma}_t} \tag{28}$$

where $\hat{\sigma}_t = exp\{\hat{\lambda}_t/2\}$ and $T = 1001, 1002, \ldots, 2000$. For convenience of notations, we label the time periods as $t = 1, 2, \ldots, 1000$. The $\hat{\sigma}_t$'s are the estimated volatility by the estimation method presented in Section 4 and r_t's are the observed returns. If the model is appropriate for our data, then there should be no autocorrelations of the standardized error u_t and the squared standardized error u_t^2. Figures 7 and 8 exhibit the ACF of $\hat{u}_t$ and $\hat{u}_t^2$. Most plots show no significant serial autocorrelations, but some plots show somewhat large autocorrelations at certain lags. For example, RBC and CIBC seem to have large

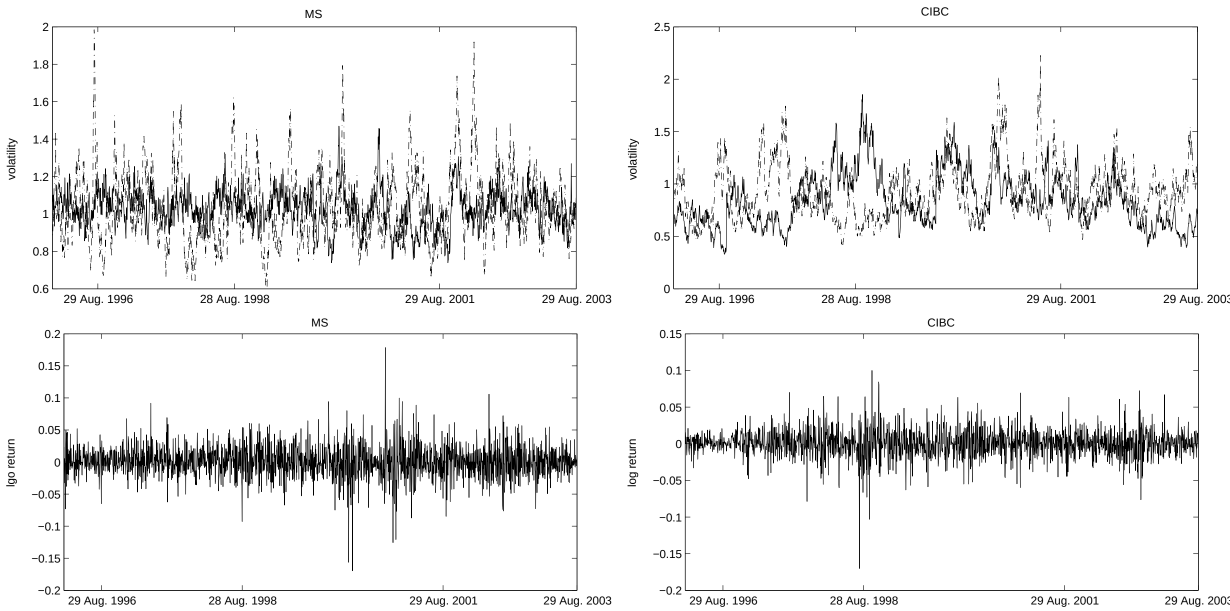

Figure 4: Top panels: MC estimates of volatilities and bottom panels: corresponding log return series for MS and CIBC

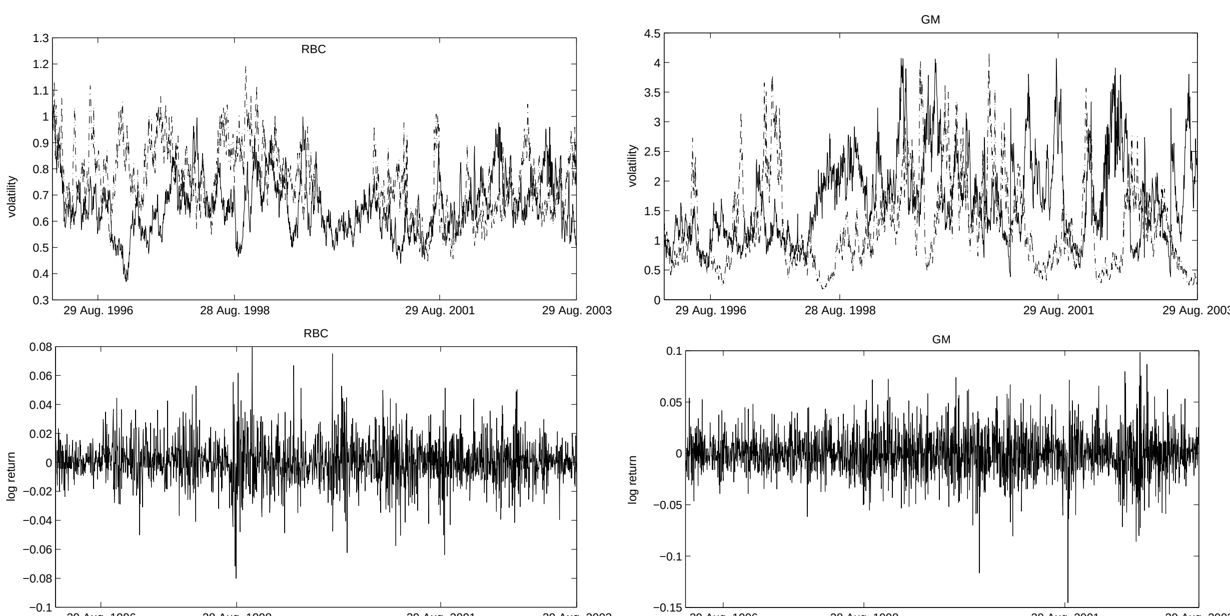

Figure 5: Top panels: MC estimates of volatilities and bottom panels: corresponding log return series for RBC and GM

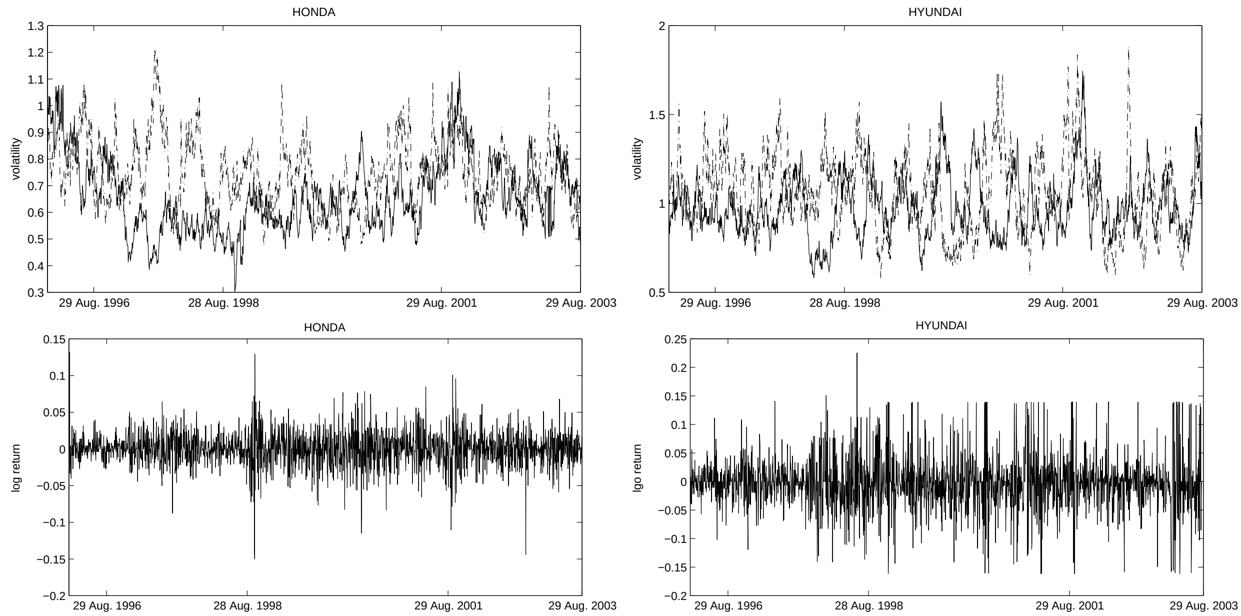

Figure 6: Top panels: MC estimates of volatilities and bottom panels: corresponding log return series for Honda and Hyundai

autocorrelations at lag one, but this may be due to the narrow limits resulting from a large sample size. The standard deviations of $\hat{u}_t$ are summarized in Table 4. This result is expected from the original assumption $u_t \sim t(0, 1)$.

Table 4. The standard deviation of the error processes u_t and v_t

Parameter	MS	RBC	CIBC	GM	HONDA	HYUNDAI
SD of u_t	0.7835	0.6235	0.6737	0.7347	0.5293	0.6821
SD of v_t	0.9452	0.9867	0.9847	0.9213	1.0382	0.9663

For further model checking, we examine the error process $\{v_t^*\}$, where $v_t^* = \gamma v_t$ in (2) with same criterion. The $\hat{v}_t$'s are generated from

$$\hat{\lambda}_t = \hat{\alpha} + \hat{\beta}\hat{\lambda}_{t-1} + \hat{\gamma}\hat{v}_t, \qquad t = 1, 2, \ldots, 1000$$

where $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are the SML parameter estimates and $\hat{\lambda}_t$ and $\hat{\lambda}_{t-1}$ are estimated volatilities. Figure 9 shows the ACF of the estimated error $\hat{v}_t^*$. The plots show little autocorrelation in $\hat{v}_t^*$. The standard deviations of $\hat{v}_t$ in Table 4 are very close to one, which is the standard deviation of the error process v_t. The normal probability plots in Figure 10 checks the

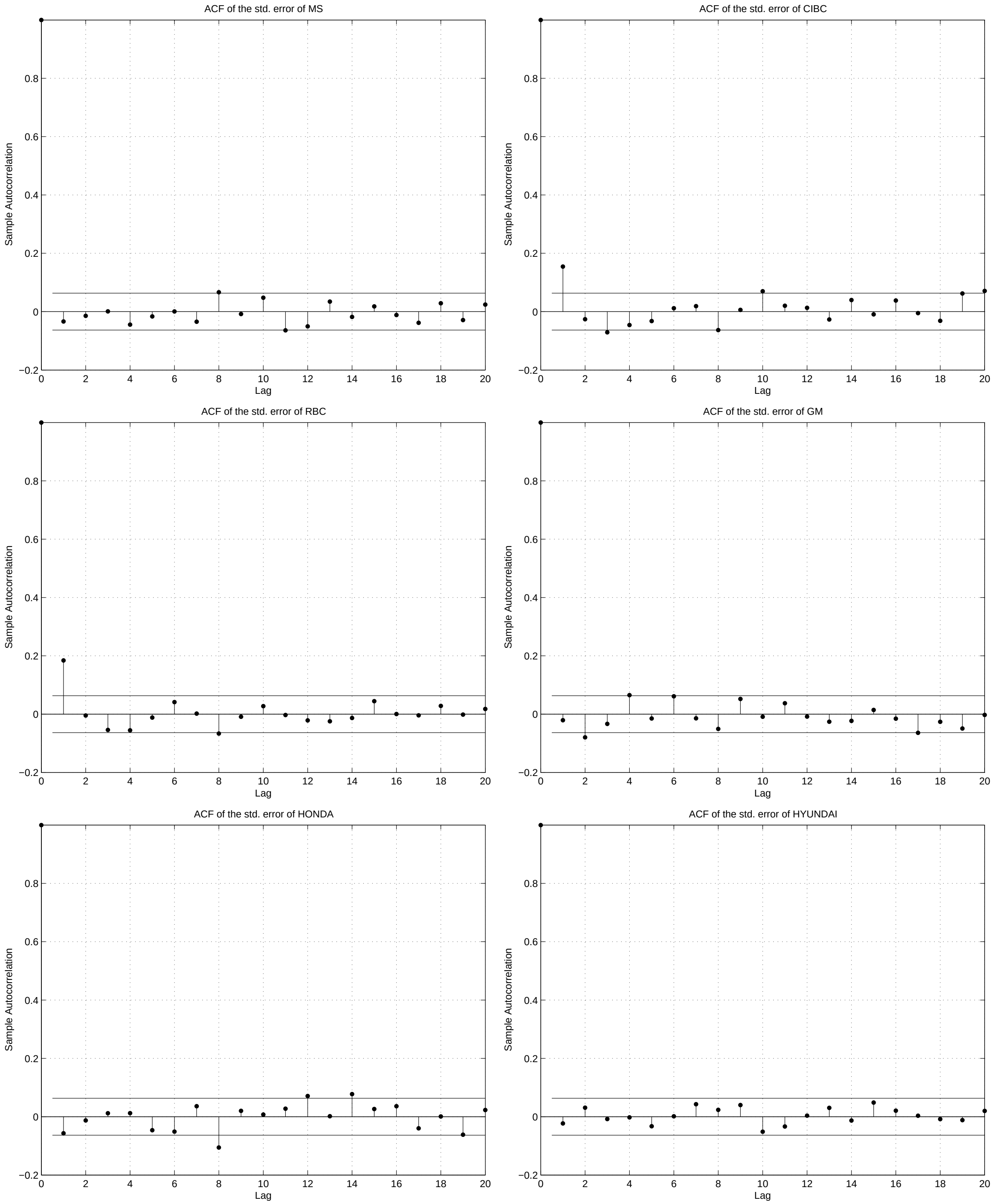

Figure 7: The ACF of the standard residual $\hat{u}_t$

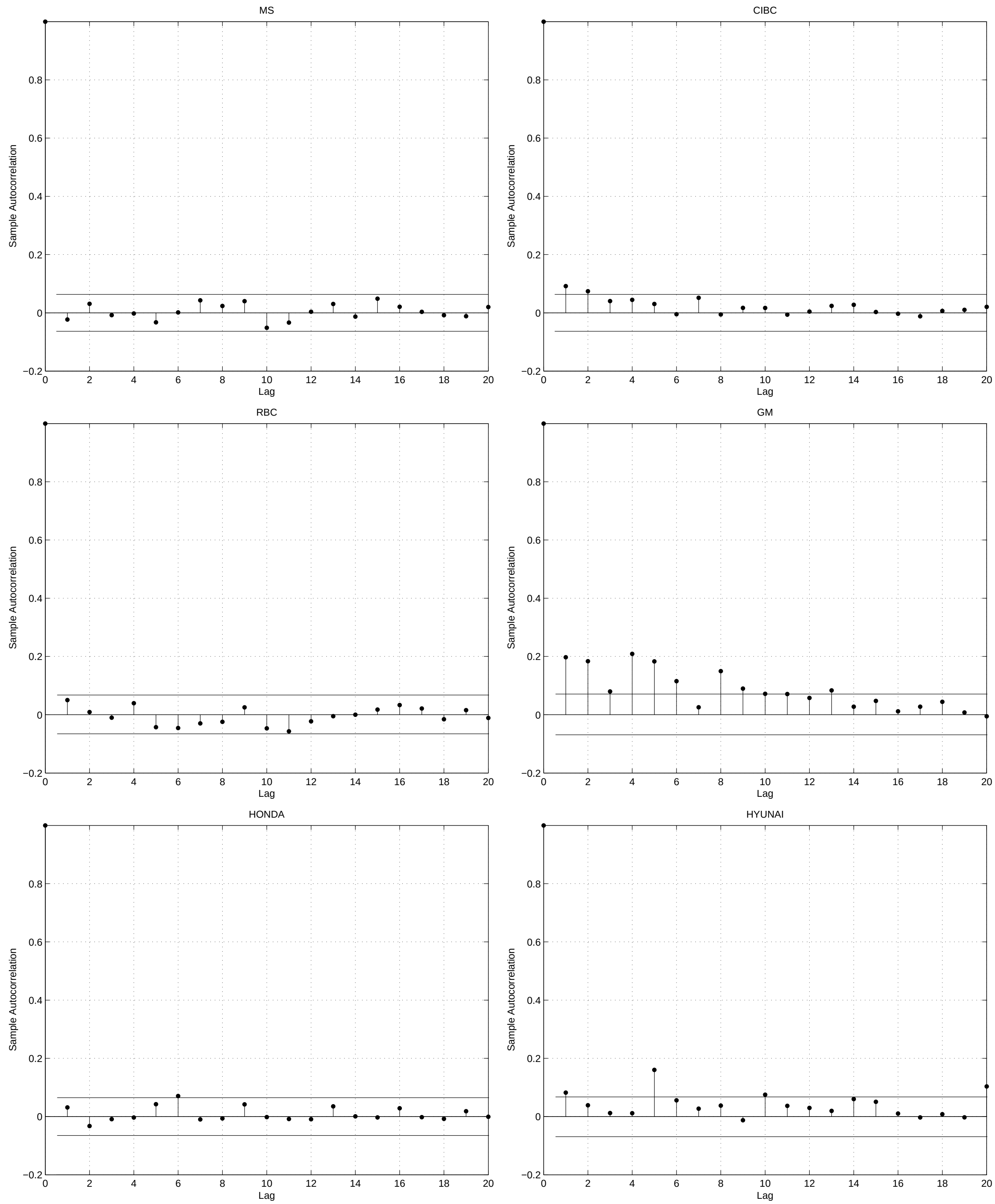

Figure 8: The ACF of the squared standard residual $\hat{u}_t^2$

normality of $\hat{v}_t^*$. In general, the normal probability plots are close to straight lines, so the $\hat{v}_t^*$ seems to be normal. However, the plots for CIBC, GM, and Honda exhibit some non-normality.

6 Conclusion

This paper analyzes the SV model with a normal error distribution and a leptokurtic error distribution (t-distribution). The Simulated Maximum Likelihood (SML) method, proposed by Danielsson and Richard (1993) is applied.

The results from the empirical example can be summarized as follows. First, the SML approach with the Accelerated Importance Sampling (AGIS) technique produces a high accuracy of parameter estimation. Second, we find that the SV-t model captures the properties of high kurtosis and slowly decaying ACF of squared returns usually seen in the financial time series. Finally, the SV-t model is an reasonable choice for the return series in our example.

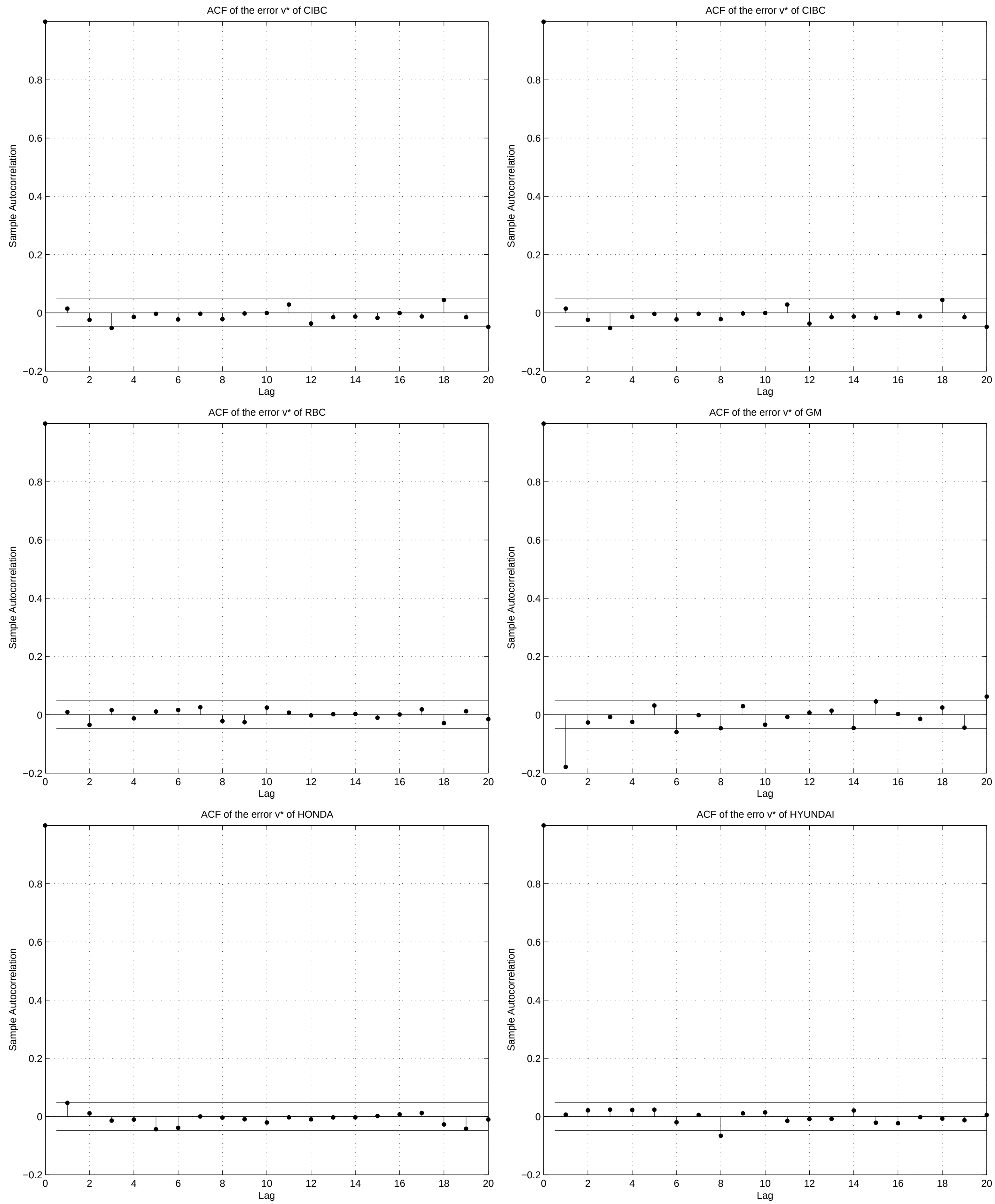

Figure 9: The ACF of $\hat{v}_t^*$

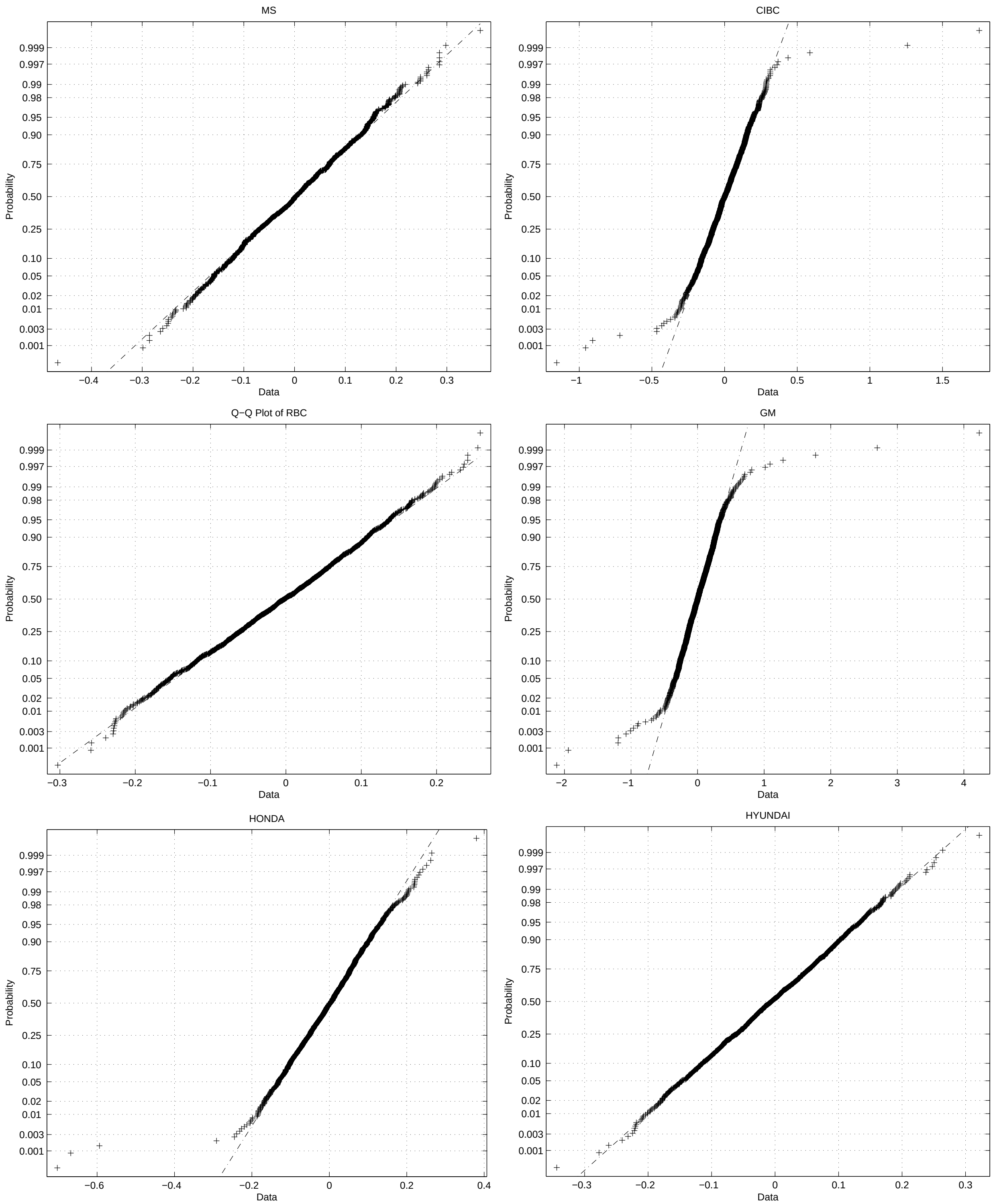

Figure 10: Q-Q plots of $\hat{v}_t^*$

References

[1] Liesenfeld, R. and Jung, R. C. (2000), "Stochastic volatility models: conditional normality versus heavy-tailed distributions", *Journal of Appiled Econometrics*, 15, 137-160

[2] Danielsson, J. and J. F. Richard (1993), "Accelerated Gaissian importance sampler with application to dynamic latent variable models", *Journal of Applied Econometrics*, 8, S153-S173

[3] Abraham, B. and Ledolter, J., "Statistical methods for forecasting", John Wiley & Sons, Inc.

[4] Poon, S. H. and Granger C. (2001), "Forecasting financial market volatility, A Review"

[5] Tsay, R. S., "Analysis of financial time series", John Wiley & Sons, Inc.

[6] Bai, X., Russell, G. and Tiao, C. (2003), "Kurtosis of GARCH and Stochastic Volatility models with non-mormal innovations", *Journal of Econemetrics*, 114, 349-360

[7] Jackel, P.,"Monte Carlo methods in finance", John Wilet & Sons, Ltd.

[8] Casella, G. and Berger, R. L., "Statistical inference", 2nd Ed., Duxbury.

[9] Hull, J. C.,"Options, futures, and other derivatives ", 5th Ed, Prentice Hall, Inc.