Asia-Pacific Financial Markets **6:** 71–84, 1999.

Bayesian Estimation of ARMA-GARCH Model of Weekly Foreign Exchange Rates

TERUO NAKATSUMA[1] and HIROKI TSURUMI[2]

[1]*Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo, Japan, e-mail: cr00387@srv.cc.hit-u.ac.jp*
[2]*Department of Economics, Rutgers University, New Brunswick, N.J., U.S.A., e-mail: tsurumi@rci.rutgers.edu*

Abstract. Three Bayesian methods (Markov chain Monte Carlo, Laplace approximation and quadrature formula) are developed to estimate the parameters of the ARMA-GARCH model. The ARMA-GARCH model is applied to weekly foreign exchange rate data of five major currencies, and their stochastic volatilities are judged by the posterior probabilities of stationarity and other conditions.

Key words: GARCH, foreign exchange rate, Bayesian inference.

1. Introduction

The stylized fact about financial time series data is that they follow random walks and that their differenced series are leptokurtic and skewed. This is illustrated in Figure 1 in which the standardized kernel density of the changes in logarithm of the daily Won/U.S. dollar exchange rate from 2 January 1996 to 8 October 1998 is plotted and compared to the standardized normal density. The changes in logarithm of the Won/U.S. dollar exchange rate are skewed (with the sample skewness measure of –2.114) and highly leptokurtic (with the sample kurtosis measure of 59.29).

One way of capturing the skewed and leptokurtic distribution of financial time series is to use a simple GARCH or EGARCH model which is capable of capturing the distributional characteristics of financial time series data. Evidence in support of such a model is encouraging. Lamoureux and Lastrapes (1993), for example, show that GARCH(1,1) or EGARCH(1,1) models can improve the stock market's implicit assessment of the time varying variance. In the context of the Black-Scholes option pricing model Day and Lewis (1992) use GARCH(1,1) or EGARCH(1,1) models to examine persistence of conditional volatility inherent in the option prices, while Engle and Mustafa (1992) use ARCH models.

While many of the studies of GARCH or EGARCH models have been made in the classical or frequentist framework, there are some Bayesian analyses (Geweke, 1989a, b; Kleibergen and Van Dijk, 1993; Müller and Pole, 1995; Nakatsuma, 1997, among others). In this paper we present Bayesian estimation procedures for

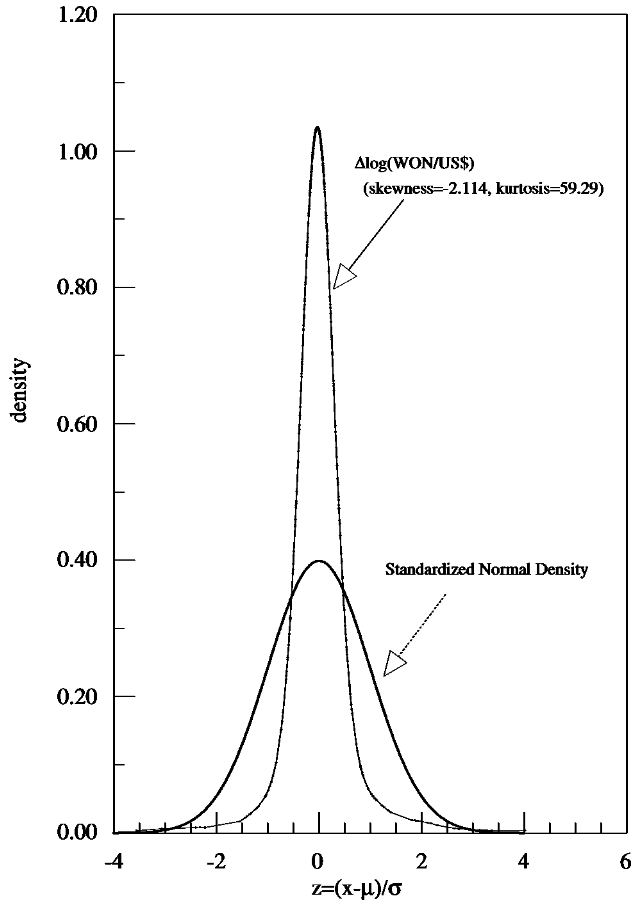

Figure 1. Kernel density of change in log(WON/US$) daily exchange rates (standardized: ($z = (x - \mu)/\sigma$)) 2 January 1996–8 October 1998

an ARMA-GARCH model (a linear regression model with an ARMA error whose conditional variance follows a GARCH process).

Organization of the paper is as follows. In Section 2 we present the ARMA-GARCH model and three Bayesian estimation methods: quadrature, Laplace approximation and MCMC algorithms. In Section 3 we discuss the conditions for stationarity and other properties of the GARCH(1,1) process and show how to test these properties. In Section 4 we estimate ARMA-GARCH models of weekly foreign exchange rates and test these properties of the GARCH(1,1)process. Concluding remarks are given in Section 5.

2. ARMA-GARCH Model

Let us consider the linear regression model with an ARMA-GARCH error process

$$
\begin{aligned}
y_t &= x_t\gamma + u_t, \quad t = 1, \cdots, T, \\
u_t &= \sum_{j=1}^{p} \phi_j u_{t-j} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}\,, \qquad \epsilon_t|\Im_{t-1} \sim \mathrm{N}(0,\ \sigma_t^2), \\
\sigma_t^2 &= \omega + \sum_{j=1}^{r} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2\,,
\end{aligned}
\tag{1}
$$

where y_t is a scalar regressand and x_t is a $1 \times k$ vector of regressors; γ is a $k \times 1$ vector of regression coefficients, and $\Im_{t-1}$ is an increasing sequence of σ-fields generated by $\{y_{t-1}, y_{t-2}, \cdots\}$; ϕ_i's are the coefficients of the autoregressive process, and θ_j's are the coefficients of the moving average process; $\omega,\ \alpha_j$'s and β_j's are the coefficients of the GARCH process. Later we shall introduce specific constraints on the parameters of the ARMA-GARCH process. Let $\phi = (\phi_1, \cdots, \phi_p)^{'},\ \theta = (\theta_1, \cdots, \theta_q)^{'},\ \alpha = (\alpha_1, \cdots, \alpha_r)^{'},\ \beta = (\beta_1, \cdots, \beta_s)^{'},\ \delta = (\gamma^{'}, \phi^{'}, \theta^{'}, \omega, \alpha^{'}, \beta^{'})^{'},\ Y = (y_1, \cdots, y_T)^{'},\ X = (x_1^{'}, \cdots, X_T^{'})^{'}$ and

$$
\Phi(L) = \sum_{j=1}^{p} \phi_j L^j, \qquad \Theta(L) = \sum_{j=1}^{q} \theta_j L^j,
$$

$$
A(L) = \sum_{j=1}^{r} \alpha_j L^j, \qquad B(L) = \sum_{j=1}^{s} \beta_j L^j\,,
$$

where L is the lag operator.

The joint posterior density is given by

$$
p(\delta|Y, X) = \frac{\ell(\delta|Y, X)p(\delta)}{\int \ell(\delta|Y, X)p(\delta)\mathrm{d}\delta}\,, \tag{2}
$$

where $\ell(\cdot)$ is the likelihood function given by

$$
\ell(\delta|\text{data}) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma_t} \exp\left\{-\frac{1}{2\sigma_t^2}\,\epsilon_t^2\right\}\,, \tag{3}
$$

where

$$
\epsilon_t = y_t - x_t\gamma - \Phi(L)(y_t - x_t\gamma) - \Theta(L)\epsilon_t\,.
$$

Given the joint posterior density, we derive the marginal posterior density of a parameter of interest, say δ_1, the first element in δ by integrating out all other elements in δ:

$$
p(\delta_1|Y, X) = \int p(\delta|Y, X)\mathrm{d}\,\delta_{-1}\,, \tag{4}
$$

where $\delta_{-1} = (\delta_2, \cdots \delta_{k+p+q+s+r+1})^{'}$.

To derive the marginal posterior probability density there are several numerical integration procedures: a quadrature formula, Laplace (or modal) approximation, importance sampling method, and Markov chain Monte Carlo (MCMC) method. In this paper we shall use a quadrature formula, Laplace approximation and MCMC method, since an application of the importance sampling method to a GARCH model is already made in Kleibergen and Van Dijk (1993). A quadrature formula does not need general explanation. In the appendix we explain how to handle an MA(1) portion of the ARMA(1,1)-GARCH(1,1) model, and suggest a generalization to an ARMA(p, q) model. Here, we shall explain the Laplace approximation and MCMC procedures.

Laplace approximation: The marginal posterior density of δ_1 by the Laplace approximation is given by

$$p(\delta_1|Y, X) \propto |\Sigma^*|\, p(\delta_1, \widehat{\delta}_2|Y, X)\,, \tag{5}$$

where

$$\Sigma^* = \left[-\frac{\partial^2 \ln\, p(\delta_1, \delta_2|Y, X)}{\partial\delta_1 \partial\delta_2}\bigg|_{\widehat{\delta}_2} \right]^{-\frac{1}{2}}$$

and $\widehat{\delta}_2$ maximizes the joint posterior pdf given δ_1. Laplace's method is based on the normal distribution, and to use it we need to see that Σ^* is positive definite. In obtaining $\widehat{\delta}_2$ we need to use a constrained maximization algorithm since some of the parameters follow inequality constraints.

MCMC method: To construct an MCMC for the ARMA-GARCH model, we note that the ARMA-GARCH model can be represented by the following auxiliary ARMA models:

AM1: regression model with an ARMA(p,q) error

$$y_t = X_t\gamma + \sum_{j=1}^{p} \phi_j(y_{t-j} - X_{t-j}\gamma) + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}, \tag{6}$$

where $\epsilon_t \sim \mathrm{N}(0,\ \sigma_t^2)$.

AM2: ARMA(ℓ,s) model of the squared errors, ϵ_t^2

$$\epsilon_t^2 = \omega + \sum_{j=1}^{\ell} (\alpha_j - \beta_j)\epsilon_{t-j}^2 + w_t - \sum_{j=1}^{s} \beta_j w_{t-j}, \tag{7}$$

where $w_t \sim \mathrm{N}(0,\ 2\sigma_t^4)$, and $\ell = \max\{r, s\}$, $\alpha_j = 0$ for $j > r$, and $\beta_j = 0$ for $j > s$. Given *AM1* and *AM2* models, we conduct an MCMC scheme as follows:

(a) Generate (γ, ϕ, θ) from *AM1* given $\{\sigma_t^2\}$ and the rest of parameters.

(b) Generate (ω, α, β) from *AM2* given $\{\epsilon_t^2\}$, $\{\sigma_t^2\}$, and the rest of the parameters.

(c) Apply the Metropolis–Hastings (MH) algorithm after each parameter is generated. The MH algorithm is

(i) generate a candidate of δ from the proposal distribution given $\delta^{(j)}$, where $\delta^{(j)}$ is the value of δ at the j-th iteration.

(ii) Accept or reject this candidate by

$$\delta^{(j+1)} = \begin{cases} \delta & \text{with probability } a \\ \delta^{(j)} & \text{with probability } 1 - a, \end{cases} \tag{8}$$

where

$$a = \min\left\{\frac{\pi(\delta)}{\pi(\delta^{(j)})} \cdot \frac{g(\delta, \delta^{(j)})}{g(\delta^{(j)}, \delta)},\ 1\right\} \tag{9}$$

and $\pi(\delta)$ is the density of the target distribution and $g(\delta^{(j)}, \delta)$ is the density of the proposal distribution.

(d) Repeat (a)–(c) until the sequences become stable.

In the MCMC method, we update $\{\epsilon_t^2\}$ and $\{\sigma_t^2\}$ every time the parameters δ are updated. See Chib and Greenberg (1994), Müller and Pole (1995) and Nakatsuma (1966) for more details.

Let us illustrate the quadrature, Laplace, and MCMC methods by using simulated data that are obtained from the following data generating process:

$$\begin{aligned} y_t &= \gamma_0 + u_t, \qquad u_t = \phi_1 u_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}, \\ \sigma_t^2 &= \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{aligned} \tag{10}$$

We set $\gamma_0 = 1,\ \theta = .1,\ \phi = .8,\ \omega = 1,\ \alpha_1 = .15,\ \beta_1 = .85$ and set the sample size T at 200. This is an IGARCH(1,1) model since $\alpha_1 + \beta_1 = 1$. Table I presents the posterior means and standard deviations of the parameters. From Table I we observe that except for γ_0, the posterior means and standard deviations produced by the three methods are similar to each other. The posterior pdf's of α_1 and of β_1 in Figures 1(a) and 1(b) show that the three methods yield comparable results.

3. Testing the Properties of the GARCH(1,1) Process

The GARCH(1,1) process has attracted attentions in the literature because of its relative easiness of estimation (either in Bayesian or maximum likelihood procedures) and also because of the fact that its theoretical properties have been worked out by many, notably by Nelson (1990) and by Hansen (1991). Within the sampling

Table I. Posterior means and standard deviations of the parameters of the ARMA(1,1)-GARCH(1,1) Model: Generated data, $T = 200$

	Quadrature Method	Laplace Method	MCMC Method
γ_0	0.3734	0.1002	0.1016
	(1.8982)	(1.3775)	(3.5287)
ϕ_1	0.7948	0.8005	0.8080
	(0.0515)	(0.0505)	(0.0504)
θ_1	0.1311	0.1571	0.1446
	(0.0805)	(0.0803)	(0.0815)
ω	1.3943	1.1108	1.1583
	(0.6639)	(0.9169)	(0.7925)
α_1	0.1455	0.1603	0.1709
	(0.0700)	(0.0771)	(0.0755)
β_1	0.7998	0.7828	0.8296
	(0.0647)	(0.0893)	(0.0622)

Notes: (1) Figures in parentheses are posterior standard deviations; (2) Data are generated by Equation (10) for the sample of $T = 200$.

theory framework testing of an IGARCH or L^2-near epoch dependent (L^2-NED) is done by, sometimes called, ‘classical’ tests such as the likelihood ratio, Wald, and Lagrange tests. Lumsdaine (1995) evaluates the finite sample properties of the classical tests of the IGARCH(1,1) model. Testing of the properties of the IGARCH(1,1) model that involve the expected value of nonlinear random variates, however, eludes an easy access. Within the Bayesian framework such a testing is relatively easy, although computationally it may be more involved than a classical testing procedure. Kleibergen and Van Dijk (1993) use an importance sampling method to derive the posterior probabilities of the stationarity condition of the IGARH(1,1) model from the bivariate posterior distributions.

Let us obtain the marginal posterior densities of the certain conditions of the GARCH(1,1) model using the MCMC procedure. First, let us enumerate the conditions:

(i) L^2-NED process: Hansen (1991) shows that the GARCH(1,1) process is L^2-NED if

$$\beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2 < 1\,, \tag{11}$$

when ϵ_t/σ_t ($\equiv z_t$) is *i.i.d.* normal.

(ii) Finiteness of the unconditional variance: Nelson (1990) shows that the unconditional variance of the GARCH(1,1) model is finite if

$$\mathrm{E}(\beta_1 + \alpha_1 z_t^2 - 1) < 0, \quad \text{or} \quad \beta_1 + \alpha_1 - 1 < 0, \tag{12}$$

if z_t is *i.i.d.* normal.

(iii) Finiteness of the unconditional standard deviation: If

$$\mathrm{E}(\beta_1 + \alpha_1 z_t^2)^{\frac{1}{2}} - 1 < 0\,, \tag{13}$$

then the unconditional standard deviation of the GARCH(1,1) model is finite (see Nelson (1990)).

(iv) Stationarity and ergodicity: If

$$\mathrm{E}[\ln(\beta_1 + \alpha_1 z_t^2)] < 0 \tag{14}$$

then the conditional variance of the GARCH(1,1) model is strictly stationary and ergodic.

The regions of $(\alpha_1,\ \beta_1)$ satisfying (11)–(14) are derived in Nelson (1990) for $z_t \sim \mathrm{NID}(0,\ 1)$.

Let us first derive the posterior probability densities of the left hand side quantities in (11)–(14) by the MCMC method, and once the posterior pdf's are obtained then we can obtain the posterior probabilities that satisfy inequalities in (11)–(14). For the NED condition (11) we calculate

$$\begin{aligned} &\mathrm{Prob}\ \left\{\beta_1^2 + 2\,\alpha_1\beta_1 + 3\,\alpha_1^2 < 1\right\} \\ &\quad \doteq \frac{1}{M}\sum_{j=1}^{M} \mathrm{I}\left[-\frac{(\beta_1^{(j)})^2 + 2\alpha_1^{(j)}\beta_1^{(j)} + 3(\alpha_1^{(j)})^2 - 1}{h_j}\right], \end{aligned} \tag{15}$$

where $\alpha_1^{(j)}$ and $\beta_1^{(j)}$ $(j = 1, \cdots, M)$ are realized values in the MCMC simulation; $\mathrm{I}(\cdot)$ is an indicator function. We choose the Gaussian kernel as K, and h_i is decided by a local smoothing method. For (ii) the condition of the finite unconditional variance we also use the same method. However, the conditions in (iii) and (iv) involve the expected values of nonlinear functions. We obtain the expected values by a Monte Carlo method. For example, the expectation of $\ln(\beta_1^{(j)} + \alpha_1^{(j)} z^2)$ is evaluated by

$$\mathrm{E}[\ln(\beta_1^{(j)} + \alpha_1^{(j)} z^2)] \doteq \frac{1}{N}\sum_{h=1}^{N} \ln(\beta_1^{(j)} + \alpha_1^{(j)}(z^{(h)})^2, \tag{16}$$

where $z^{(h)}$ is drawn from $N(0, 1)$. Once we obtain the expectation for each pair of $(\alpha_1^{(j)},\ \beta_1^{(j)})$, we can apply the same kernel estimation method as in (15) to estimate the posterior probabilities of the finite mean and strict stationarity and ergodicity.

4. ARMA(1,1)-GARCH(1,1) Model for Weekly Foreign Exchange Rates

In the literature GARCH models have been applied to time series on stock prices, returns on assets, and foreign exchange rates to examine volatility. Engle and Bollerslev (1986) and Lumsdaine (1995) are notable examples of GARCH models as applied to data on foreign exchange rates. Let us estimate ARMA(1,1)-GARCH(1,1) models of weekly foreign exchange rates for five currencies: British pound, Canadian dollar, Deutsche mark, Japanese yen, and Swiss franc using the U.S. dollar as the base currency. The ARMA(1,1)-GARCH(1,1) model is a simple but basic specification to examine the existence of a unit root and of noninvertibility as well as the properties of the GARCH(1,1) model given in the previous section.

The sample period is from the week of 4 June 1974 to the week of 18 May 1987. The logarithm of the closing rates on Wednesday (Thursday if Wednesday is a holiday) are used. The sample size is 729. Since the constant term, γ_0 in (3) becomes unidentifiable when ϕ_1 is unity and data on foreign exchange rates often follow the pattern of a random walk with drift, we subtract the first observation from the current values.

We report in here the estimation results using the MCMC method since the estimated results for the parameters of the model by the Laplace and quadrature methods are similar to those by the MCMC method and since the MCMC method is the easiest procedure to obtain the posterior pdf's of the conditions (iii) and (iv) (involving the computation of the expected values) of the properties on the GARCH(1,1) model in the previous section. The number of iteration of the Markov chain sampling is 11,000, and we discard the first 1,000 as burn-in. Thus, the size of the Monte Carlo sample is 10,000. Posterior means and standard deviations are given in Table II, while the posterior probabilities of the properties of the GARCH(1,1) model are presented in Table III. The expected values in (13) and (14) are computed by 10,000 drawings of $z \sim N(0, 1)$.

The estimated results show that ϕ_1 of all of the five foreign exchange rates are tightly distributed around the value of unity, indicating that they follow random walks. The sizes of θ_1 judged by the posterior pdf's or quickly by the two point estimates, the posterior means and standard deviations, indicate that the MA(1) processes are significant but well within the invertible regions. The constant term of the conditional variance process, ω is extremely small for all of the five exchange rates.

Let us examine the properties of the GARCH(1,1) process, since they are important in evaluating the volatility of the exchange rate data. For the Canadian dollar, all the posterior probabilities in Table III are more than 99%. The British pound and Japanese yen show patterns similar to the Canadian dollar, but their posterior

Table II. Posterior means and standard deviations of the parameters of the ARMA(1,1)-GARCH(1,1) model

Parameter	British Pound	Canadian Dollar	Deutsche Mark	Japanese Yen	Swiss Franc
ϕ_1	1.001	1.001	1.001	1.003	1.001
	(0.00132)	(0.00105)	(0.00232)	(0.00269)	(0.00145)
θ_1	0.05827	0.08501	0.08702	0.08806	0.07520
	(0.0395)	(0.0405)	(0.0367)	(0.0383)	(0.0378)
ω	1.065×10^{-5}	5.769×10^{-6}	5.797×10^{-6}	5.012×10^{-6}	5.139×10^{-6}
	(0.378×10^{-6})	(2.20×10^{-6})	(3.54×10^{-6})	(1.41×10^{-6})	(2.93×10^{-6})
α_1	0.1279	0.2290	0.1322	0.0879	0.1371
	(0.0314)	(0.0447)	(0.0329)	(0.0177)	(0.0246)
β_1	0.8231	0.5887	0.8455	0.8914	0.8518
	(0.0406)	(0.0104)	(0.0418)	(0.0211)	(0.0269)

Note: The figures in parentheses are posterior standard deviations.

Table III. Posterior probabilities of the properties of the GARCH(1,1) model

Condition	British Pound	Canadian Dollar	Deutsche Mark	Japanese Yen	Swiss Franc
L^2-NED					
$\beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2 < 1$	90.1%	99.0%	57.9%	89.3%	31.0%
Finite variance					
$\beta_1 + \alpha_1 < 1$	95.8%	99.7%	86.1%	94.4%	74.0%
Finite Std. Dev.					
$E(\beta_1 + \alpha_1 z^2)^{\frac{1}{2}} < 1$	95.3%	99.4%	85.8%	91.4%	79.9%
Stationarity					
$E[\ln(\beta_1 + \alpha_1 z^2)] < 0$	97.9%	99.8%	94.0%	97.1%	92.0%

probabilities are a few percentage points less. In particular the probabilities of L^2-NED for the British pound and yen are about 90%. The Deutsche mark and Swiss franc exhibit patterns different from the other three exchange rates: the posterior probabilities of L^2-NED are low, and those of the finite variance and standard deviation are less than 90%, indicating that the probability that the GARCH(1,1) process does not have finite moments cannot be ignored as far as the Deutsche mark and Swiss franc are concerned.

The posterior means of α_1 and of β_1 for the four currencies other than the Canadian dollar in Table II are close to each other but as for the Deutsche mark and Swiss franc their posterior probabilities of L^2-NED, finite variance, and of finite standard deviation are different from those of the British pound and Japanese yen. The reason why they are different is that the joint posterior pdf's of α_1 and of β_1 for the Deutsche mark and Swiss franc are concentrated on the region of $\alpha_1 \in (.08, .15)$ and $\beta_1 \in (.8, 1.0)$, and this is the region where conditions (11)–(14) converge. On the other hand, the joint posterior pdf's of α_1 and of β_1 of the British pound and Japanese yen are more spread out, thus leading to the higher probabilities for these conditions.

5. Conclusions

In this paper we presented three Bayesian estimation procedures of the ARMA-GARCH regression model: a quadrature formula, Laplace approximation and MCMC procedures. Among the three procedures the MCMC procedure we can easily obtain the posterior probability densities of such properties of the GARCH(1,1) process as near epoch dependence (NED), finite unconditional variance, strict stationarity and ergodicity. Using the weekly data on the five foreign exchange rates, we showed that these posterior probabilities of the properties of the GARCH(1,1) process can differ widely even if the point estimates such as posterior means and standard deviations are similar. Although we limited most of our attention to the ARMA(1,1)-GARCH(1,1) process, the Bayesian estimation procedures such as the MCMC and the Laplace approximation can be easily applied to a higher order ARMA(p, q)-GARCH(r, s) model or to a GARCH-in-the-mean or other variates of the GARCH models.

Appendix: Quadrature Procedure for an ARMA(1,1)-GARCH(1,1) Model

The model is given by

$$\begin{aligned} y_t &= x_t\beta + u_t \,, \\ u_t &= \phi\, u_{t-1} + \epsilon_t + \theta\,\epsilon_{t-1} \,, \\ h_t &= \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1 h_{t-1} \,. \end{aligned} \tag{17}$$

.

The equations above may be rewritten as

$$y_t(\phi) = x_t(\phi)\beta + v_t \,, \tag{18}$$

where $y_t(\phi) = y_t - \phi\, y_{t-1},\ x_t(\phi) = x_t - \phi\, x_{t-1}$, and $v_t = \epsilon_t + \theta\,\epsilon_{t-1}$. The variance of v_t becomes

$$\mathrm{var}(v_t) = h_t + \theta^2 h_{t-1} \,, \qquad t = 1, \cdots, n$$

or in matrix form

$$
\begin{aligned}
\mathrm{Var}(v) &= \begin{bmatrix} h_1+\theta^2h_0 & \theta h_1 & 0 & 0 & \cdots & 0 & 0 \\ \theta h_1 & h_2+\theta^2h_1 & \theta h_2 & 0 & \cdots & 0 & 0 \\ 0 & \theta h_2 & h_3+\theta^2h_2 & \theta h_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & & & \vdots \\ & & & & & & \theta h_{n-1} \\ 0 & \cdots & \cdots & & & \theta h_{n-1} & h_n^2+\theta^2h_{n-1} \end{bmatrix} \\
&= C\,H\,C^{'},
\end{aligned}
$$

where

$$
C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \theta & 1 & 0 & \cdots & & 0 \\ 0 & \theta & 1 & 0 & \cdots & \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \cdots & 0 & \theta & 1 \end{bmatrix}
$$

and $H = \mathrm{Diag}(h_1, h_2, \cdots, h_n)$. The inverse of C is given by, except for the first element,

$$
P = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ c & 1 & 0 & 0 & 0 \\ c^2 & c & 1 & & 0 \\ \vdots & \vdots & & & \vdots \\ c^{n-1} & c^{n-2} & c^{n-3} & \cdots\; c & 1 \end{bmatrix}.
$$

The likelihood function may be expressed as

$$
\begin{aligned}
\ell(\delta|\mathrm{data}) &\propto |H|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[y(\phi)-X(\phi)\gamma\right]^{'}C^{'-1}H^{-1}C^{-1}\left[y(\phi)-X(\phi)\gamma\right]\right\} \\
&= |H|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[y(\phi,\theta)-X(\phi,\theta)\gamma\right]^{'}H^{-1}\left[y(\phi,\theta)-X(\phi,\theta)\gamma\right]\right\}, \qquad (19)
\end{aligned}
$$

where

$$
y(\phi,\theta) = P\,y(\phi), \quad y(\phi) = \begin{bmatrix} y_1(\phi) \\ y_2(\phi) \\ \vdots \\ y_n(\phi) \end{bmatrix},
$$

$$
X(\phi,\theta) = P\,X(\phi), \quad \text{and} \quad X(\phi) = \begin{bmatrix} x_1(\phi) \\ x_2(\phi) \\ \vdots \\ x_n(\phi) \end{bmatrix}.
$$

Let the prior pdf be given by

$$p(\delta) \propto \text{ constant}\,, \tag{20}$$

then the joint posterior pdf is given by

$$p(\delta|\text{data}) \propto \ell(\delta|\text{data}) \tag{21}$$

and we obtain the bivariate posterior pdf's and the marginal posterior pdf of parameters of interest by incorporating the assumptions A1–A5:

A1: $-1 < \theta < 1\,,$
A2: $\alpha_0 > 0\,,$
A3: $\alpha_1 > 0\,,$
A4: $\beta_1 > 0\,,$
A5: $\alpha_1 + \beta_1 < 1\,.$

For example, the bivariate posterior pdf of α_1 and γ is derived by

$$p(\alpha_1, \beta_1|\text{data}) = \int_{R^k} \int_{-1}^{1} \int_{R_\phi} \int_0^\infty p(\gamma, \theta, \phi, \alpha_0|\text{data}) \mathrm{d}\alpha_0 \, \mathrm{d}\phi \, \mathrm{d}\theta \, \mathrm{d}\gamma\,, \tag{22}$$

where R^k are k-dimensional regions for γ and R_ϕ is the region for ϕ and we include the unit root and nonstationary region. Then the marginal posterior pdf of α_1 and that of γ are given by

$$p(\alpha_1|\text{data}) = \int_0^{1-\beta_1} p(\alpha_1, \beta_1|\text{data}) \mathrm{d}\beta_1\,, \tag{23}$$

$$p(\beta_1|\text{data}) = \int_0^{1-\alpha_1} p(\alpha_1, \beta_1|\text{data}) \mathrm{d}\alpha_1\,. \tag{24}$$

We use either a Simpson's rule or a Gaussian quadrature procedure to carry out numerical integration.

The quadrature procedure above can be extended to any higher order ARMA(p, q)-GARCH(r, s) model in theory, but when the dimension of parameters go over 7, the computational burden increases enormously, and thus we may resort to a Laplace approximation procedure. Let us explain how to implement it to obtain the marginal posterior pdf of a parameter of interest. An ARMA(p, q) error process

$$u_t = \sum_{j=1}^{p} \phi_j u_{t-j} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

may be written as

$$u_t = \frac{\Theta(B)}{\Phi(B)}\,, \tag{25}$$

where $\Phi(B) = 1-\phi_1 B - \cdots - \phi_p B^p$ and $\Theta(B) = 1+\theta_1 B + \cdots + \theta_q B^q$. Suppose that we are interested in deriving the posterior pdf of a unit root parameter ρ. Then $\Phi(B)$ may be rewritten as

$$\Phi(B) = (1-\phi)\Phi_1(B)\,,$$

where $\Phi_1(B)$ is the $(p-1)$-th order polynomial in lag operator B. Let us assume that all the roots of $\Phi_1(B)$ and $\Theta(B)$ lie outside the unit circle. Then we may expand $\Phi_1(B)/\Theta(B)$ as

$$\frac{\Phi_1(B)}{\Theta(B)} = 1 - \pi_1 B - \pi_2 B^2 - \cdots - \pi_\ell B^\ell - \cdots$$

Then $y_t = x_t\gamma + u_t$ may be rewritten as

$$y_t(\phi) = x_t(\phi)\gamma + \sum_{i=1}^{\ell} \pi_i \mathrm{e}_{t-i}(\phi) + \epsilon_t\,, \tag{26}$$

when we truncate the lags of e_{t-i} at $i = \ell$, where

$$\mathrm{e}_t(\phi) = y_t(\phi) - x_t(\phi)\gamma, \quad y_t(\phi) = y_t - \phi\, y_{t-1}\,, \quad x_t(\phi) = x_t - \phi\, x_{t-1}\,.$$

Forming the likelihood function using $\epsilon_t \sim \mathrm{N}(0, \sigma_t^2)$, and with a prior pdf we obtain the marginal posterior density of ϕ as in Equation (5) in the text by setting $\delta_1 = \phi$. All the inequality constraints on the parameters of the GARCH(r, s) are incorporated in the maximization procedure.

References

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity, *J. Econometrics* **31**, 307–327.

Chib, S. and Greenberg, E. (1994) Bayes inference in regression models with ARMA(p, q) errors, *J. Econometrics* **64**, 183–206.

Day, T. E. and Lewis, C. M. (1992) Stock market volatility and the information content of stock index options, *J. Econometrics* **52**, 267–287.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation, *Econometrica* **50**, 987–1008.

Engle, R. F. and Bollerslev, T. (1986) Modeling the persistence of conditional variances, *Econometric Reviews* **5**, 1–50, 81–87.

Engle, R. F. and Mustafa, C. (1992) Implied ARCH models from options prices, *J. Econometrics* **52**, 289–311.

Geweke, J. (1989a) Bayesian inference in econometric models using Monte Carlo integration, *Econometrica* **57**, 1317–1339.

Geweke, J. (1989b) Exact predictive densities for linear models with ARCH disturbances, *J. Econometrics* **40**, 63–86.

Hansen, B. E. (1991) GARCH(1,1) processes are near epoch dependent, *Economic Letters* **36**, 181–186.

Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57**, 97–109.

Kleibergen, F. and Van Dijk, H. K. (1993) Non-stationarity in GARCH models: A Bayesian analysis, *J. Appl. Econometrics* **8**, 341–361.

Lamoureux, C. G. and Lastrapes, W. D. (1993) Forecasting stock return variance: toward an understanding of stochastic implied volatilities, *Rev. Financ. Stud.* **5**, 293–326.

Lumdaine, R. L. (1995) Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: A Monte Carlo investigation, *J. Busin. Econom. Stat.* **13**, 1–10.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953) Equations of state calculations by fast computing machines, *J. Chem. Phys.* **21**, 1087–1092.

Müller, P. and Pole, A. (1995) Monte Carlo posterior integration in GARCH models, mimeograph, Duke University.

Nakatsuma, T. (1996) Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach, mimeograph, Rutgers University.

Nelson, D. B. (1990) Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory* **6**, 318–334.