No. 786

Random walk stochastic volatility models

by

Manabu Asai

July 1998

Random walk stochastic volatility models

Manabu Asai*

Doctoral Program in Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8573 Japan

May 1998

ABSTRACT

Some empirical evidences indicated the high persistence in asset return volatility. Engle and Bollerslev (1986) proposed to use the integrated GARCH (IGARCH) models for conditional volatility. Geweke (1986) and Pantula (1986) suggested the integrated log-GARCH (ILGARCH) models, which are the logarithmic extension of the IGARCH models. Harvey, Ruiz, and Shephard (1994) and Ruiz (1994) modeled as the logarithm of unobserved volatility follows a random walk process, which is called the random walk stochastic volatility (RWSV) model. As with other SV models, the likelihood function of the RWSV model is difficult to evaluate. In this paper, we briefly review the estimation methods for RWSV models. We derive an ILGARCH representation of a class of RWSV models, including linear regression models with ARMA(p,q)-RWSV errors. We propose a new QML method via the ILGARCH approach. Our Monte Carlo results indicate that QML estimator via the ILGARCH approach performs well. In the view of relative efficiency, for the parameter values found in empirical analysis, our estimators are superior to those of the Method of Moments estimator and those of the QML estimator based on the Kalman filter. We develop procedures for testing the integration of log-volatility. We present a empirical example of daily Deutsch mark/U.S. dollar exchange rates.

Key Words: Log-GARCH models, Quasi-maximum likelihood, Random walk, Stochastic Volatility,

JEL classification: C13, C22

---

*e-mail: masai@sk.tsukuba.ac.jp

The author is grateful to Professors Tsunemasa Shiba of the University of Tsukuba, Yoshihisa Baba of Soka University, and Tatsuya Kubokawa of the University of Tokyo for many useful suggestions and comments. Remaining errors are my responsibility.

1 Introduction

It is well known that asset return variance or volatility changes randomly over time. There are two types of econometric models for such effects. One is the GARCH models, developed by Engle (1982) and Bollerslev (1986). In GARCH models, such effects are captured by letting the conditional volatility be a function of squares of previous observations and past volatilities. Since the models are formulated in terms of the conditional distribution, the likelihood function may be evaluated straightforwardly. Another type is the the stochastic volatility (SV) models, in which the logarithm of an unobserved volatility is modeled as a linear stochastic process, such as an autoregression. Their statistical properties are easily obtained from the properties of the process generating the volatility component. Their main disadvantage, however, is that they are difficult to estimate by the maximum likelihood estimation method.

Some empirical evidence indicated the high persistence in asset return volatility; see, for example, Bollerslev, Chou and Kroner (1992). Engle and Bollerslev (1986) proposed to use the integrated GARCH (IGARCH) models for conditional volatility. Geweke (1986) and Pantula (1986) suggested the integrated log-GARCH (ILGARCH) models, which are the logarithmic extension of the IGARCH models. Harvey, Ruiz, and Shephard (1994) and Ruiz (1994) modeled as the logarithm of unobserved volatility follows a random walk process, which is called the random walk stochastic volatility (RWSV) model.

As with other SV models, the likelihood function of the RWSV model is difficult to evaluate. In this paper, we briefly review the estimation methods for RWSV models. We derive an ILGARCH representation of a class of RWSV models, including linear regression models with ARMA(p,q)-RWSV errors. To estimate these RWSV models, we propose a new QML method via the ILGARCH approach. Our Monte Carlo results indicate that their finite sample properties are superior to those of the Method of Moments estimator and those of the QML estimator based on the Kalman filter.

The organization of this paper is as follows. Section 2 briefly reviews the estimation methods for RWSV models, and derives an ILGARCH representation of a RWSV model. Section 3 investigates finite-sample properties of the QML estimator. Section 4 reports empirical findings for the daily Deutsch mark/U.S. dollar exchange rates. Testing procedures for integration in log-volatility are discussed in appendix C. Section 5 concludes the paper.

2 Random Walk Stochastic Volatility

A simple stationary SV model is given by

$$\begin{aligned} y_t &= \varepsilon_t \exp(h_t/2), \qquad \varepsilon_t \sim \mathrm{NID}(0,1), \\ h_t &= \gamma + \phi h_{t-1} + \sigma_\eta \eta_t, \qquad \eta_t \sim \mathrm{NID}(0,1), \end{aligned} \tag{1}$$

where ε_t is generated independently of η_t and $|\phi| < 1$. Working with exponentials ensures that $\exp(h_t)$ is always positive.

The SV model in (1) can be generalised so that h_t follows any stationary ARMA process. Altenatively, h_t can be allowed to follow a random walk. The corresponding RWSV model is then

$$\begin{aligned} y_t &= \varepsilon_t \exp(h_t/2), \\ h_t &= h_{t-1} + \sigma_\eta \eta_t. \end{aligned} \tag{2}$$

Since h_t is nonstationary, y_t is also nonstationary. The Kalman filter approach, proposed by Harvey, Ruiz, and Shephard (1994) and Ruiz (1994), is still valid if the restriction $\phi = 1$ is imposed. We consider the following state space model:

$$\begin{aligned} \ln y_t^2 &= E(\ln \varepsilon_t^2) + h_t + \zeta_t, \\ h_t &= h_{t-1} + \sigma_\eta \eta_t, \end{aligned} \tag{3}$$

where $\zeta_t = (\ln\varepsilon_t^2 - E(\ln\varepsilon_t^2))$. Since the mean and variance of $\ln\varepsilon_t^2$ are $E(\ln\varepsilon_t^2) = (\psi(1/2) - \ln(1/2)) \simeq -1.27$ and $Var(\ln\varepsilon_t^2) = \pi^2/2$ where $\psi(\cdot)$ is the Digamma function[1], ζ_t is a non-Gaussian white noise with mean zero and variance $\pi^2/2$. Estimation of σ_η^2 in model (3) can be carried out by treating ζ_t as it were i.i.d.N$(0, \pi^2/2)$. Using the results in Dunsmuir (1979), the asymptotic distribution of estimator of σ_η^2 is given by

$$\sqrt{T}(\tilde{\sigma}_\eta^2 - \sigma_\eta^2) \overset{a}{\sim} N(0, C_1(\sigma_\eta^2)),$$

where

$$C_1(\sigma_\eta^2) = \frac{2}{\sigma_\eta^2 + \pi^2}\left[(\sigma_\eta^4 + 2\sigma_\eta^2\pi^2)^{3/2} + \frac{2(\sigma_\eta^4 + 2\sigma_\eta^2\pi^2)^2}{\sigma_\eta^2 + \pi^2}\right]. \tag{4}$$

Appendix A gives the derivation of this asymptotic variance. Note that the asymptotic variance in Ruiz (1994, p.304),

$$\frac{2}{\sigma_\eta^2}\left[(\sigma_\eta^4 + 2\sigma_\eta^2\pi^2)^{3/2} + \frac{2\pi^4\sigma_\eta^4}{\sigma_\eta^2 + 2\pi^2}\right], \tag{5}$$

contains typogarphic errors and thus it should not be used[2].

A method of moments estimator is also given in Ruiz (1994). Consider the state space form in (3). The stationary form of $\ln y_t^2$ is given by

$$\Delta\ln y_t^2 = \sigma_\eta\eta_t + \Delta\zeta_t.$$

Given that η_t and ζ_t are mutually uncorrelated and $\sigma_\zeta^2 = \pi^2/2$, thet variance of η_t is given by

$$\sigma_\eta^2 = \sigma_{\Delta\ln y_t^2}^2 - \pi^2.$$

[1]See Abramovits and Stegun (1970, p.943).

[2]Instead of typogarphic errors, since Table 6 of Ruiz (1994, p.298) which provides asymptotic standard deviations of QML estimator via the KF coincides the our results in equation (4), Ruiz (1994) do not have to change her conclusion that the efficiency of the method of moments estimator compared with the QML estimator via Kalman filtering methods is exceptionally low for RWSV models.

A method of moments estimator of σ_η^2 is then given by

$$\check{\sigma}_\eta^2 = \check{\sigma}^2_{\Delta \ln y_t^2} - \pi^2,$$

where $\check{\sigma}^2_{\Delta \ln y_t^2}$ is the sample variance of $\Delta \ln y_t^2$. If $\sigma_\eta^2 > 0$, then $\sqrt{T}(\check{\sigma}_\eta^2 - \sigma_\eta^2)$ has an asymptotic normal distribution with zero mean and variance

$$C_2(\sigma_\eta^2) = 2[(\sigma_\eta^2 + \pi^2)^2 + \pi^4]. \tag{6}$$

Jacquier, Polson, and Rossi (1994) suggested a Bayesian inference and used the Markov chain Monte Carlo (MCMC). This method can be applied both stationary and nonstationay SV models. The MCMC is, however, computationally intensive, and thus we do not deal with it in this paper.

Asai (1998) derived a log-GARCH representation of a class of SV models, including linear regression models with ARMA(p,q)-SV errors. To estimate these SV models, he proposed a QML method via the log-GARCH approach based on either a Gaussian or a standardized t distribution. His Monte Carlo results indicate that their finite sample properties are superior to those of the Genelalized Method of Moments estimator and those of the QML estimator based on the Kalman filter; and close to those of the nonlinear filtering maximum likelihood estimator, which is a conputationally intensive method.

Asai's (1998) basic idea is that a SV model in y_t is interpreted as an ARMA process in $\ln y_t^2$, and a log-GARCH model in x_t as an ARMA process in $\ln x_t^2$. Thus, there is a link between a SV model and a log-GARCH model. Following the spirit of Asai (1998), we derive a ILGARCH representation of RWSV models. Concentrating out h_t in model (3), we obtain

$$\ln y_t^2 = \ln y_{t-1}^2 + \sigma_\eta \eta_t + \zeta_t - \zeta_{t-1}. \tag{7}$$

Let us note that a Gaussian process variable combined with an MA(1) variable reduces to an MA(1) process; see Hamilton (1994, chapter 4, pp.102–106). Using this fact, the last three terms of equation (7) becomes an MA(1) process,

$$\sigma_\eta \eta_t + (\zeta_t - \zeta_{t-1}) = v_t - \theta v_{t-1}, \quad v_t \sim \mathrm{WN}(0, \sigma^2) \tag{8}$$

where $\mathrm{WN}(0, \sigma^2)$ denotes a white noise process with mean zero and variance σ^2, and

$$\sigma^2 = \frac{\pi^2}{2\theta}, \quad \theta = \left(1 + \frac{\sigma_\eta^2}{\pi^2}\right) - \sqrt{\left(1 + \frac{\sigma_\eta^2}{\pi^2}\right)^2 - 1} \tag{9}$$

The other solution of the quadratic equation for θ does not satisfy the invertibility condition. Note that equation (9) implies $0 < \theta < 1$. As a result, the model reduces to an integrated MA(1) process in $\ln y_t^2$:

$$\ln\, y_t^2 = \ln\, y_{t-1}^2 + v_t - \theta v_{t-1}, \quad v_t \sim \mathrm{WN}(0, \sigma^2). \tag{10}$$

where v_t is a skewed and leptokurtic white noise.

We next show RWSV process in y_t is equivalent to an ILGARCH(1,1) process in y_t with a symmetric non-Gaussian noise. Define a standardized process z_t and a positive predetermined variable σ_t by

$$z_t \equiv \frac{1}{c_z} \left(\varepsilon_t \prod_{i=1}^{\infty} |\varepsilon_{t-i}|^{(\theta-1)\theta^{i-1}} \right) \exp\left[\frac{\sigma_\eta}{2} \eta_t + \frac{\sigma_\eta}{2} \sum_{i=1}^{\infty} \theta^i \eta_{t-i} \right], \tag{11}$$

$$\ln \sigma_t^2 \equiv \ln c_z^2 + (1-\theta)(1-\theta L)^{-1} \ln y_{t-1}^2, \tag{12}$$

where

$$\begin{aligned} c_z \;=\; \exp & \left[\frac{\sigma_\eta^2}{4(1-\theta^2)} - \frac{c_\varepsilon}{2} + \frac{\theta}{2} \psi\left(\frac{1}{2}\right) \right. \\ & \left. + \frac{1}{2} \sum_{i=1}^{\infty} \left[\ln\Gamma\left((\theta-1)\theta^i + \frac{1}{2} \right) - \ln\Gamma\left(\frac{1}{2}\right) \right] \right], \end{aligned}$$

$c_\varepsilon \equiv E(\ln \varepsilon_t^2)$, and L denotes the lag operator. Since $c_z < \infty$ if $\sigma_\eta > 0$, z_t is a weak stationary process and its moments up to fourth order are[3]:

$$E(z_t) = 0, \quad E(z_t^2) = 1, \quad E(z_t^3) = 0,$$

[3]The derivation of the moments of z_t in this paper is available upon request.

$$E(z_t^4) = \frac{3}{c_z^4} \exp\left[\frac{2\sigma_\eta^2}{1-\theta^2} - 2c_\varepsilon + 2\theta\psi\left(\frac{1}{2}\right)\right.$$
$$\left. + \sum_{i=1}^{\infty}\left[\ln\Gamma\left(2(\theta-1)\theta^i + \frac{1}{2}\right) - \ln\Gamma\left(\frac{1}{2}\right)\right]\right],$$
$$E(z_t z_{t-j}) = 0, \qquad \text{for } j > 1.$$

All odd-moments are equal to zero. σ_t is measurable with respect to the time $t-1$ information set.

Rewriting ε_t in terms of z_t or σ_t as

$$\varepsilon_t = \ln z_t^2 - \ln c_z^2 = \ln y_t^2 - \ln\sigma_t^2 - \ln c_z^2 \tag{13}$$

and substituting equation (13) to (10), we obtain an equivalent representation of the RWSV process in y_t,

$$\begin{aligned} y_t &= \sigma_t z_t, \quad z_t \sim \mathrm{WN}(0,1), \\ \ln\sigma_t^2 &= (1-\theta)\ln c_z^2 + (1-\theta)\ln y_{t-1}^2 + \theta\ln\sigma_{t-1}^2. \end{aligned} \tag{14}$$

Note that z_t is serially uncorrelated with mean zero and variance one, and that $\ln c_z^2$ is the function of θ. Therefore the RWSV process in y_t can be interpreted as an ILGARCH(1,1) process in y_t which has a heavy-tailed and symmetric conditional distribution. It should be noted that theorem 2.1 by Nelson (1991) indicates y_t in (14) is nonstationary.

The procesure in this section is also applied to linear regression models with ARMA(p,q)-RWSV errors; see appendix B.

3 A New QML Estimation Method

In the previous section, we derived a ILGARCH representation of a RWSV model, To estimate RWSV models, we propose a new QML method via the ILGARCH approach based on the Gaussian distribution. We also analyze the finite sample properties of QML estimator of σ_η^2 using Monte Carlo experiments. Using response surface methodology, we investigate finite sample biases and compare estimated asymptotic standard errors to sample standard deviations. We examine their relative efficiencies with respect to other estimators.

Under a set of mild regularity conditions, the quasi-maximum likelihood estimator proposed by White (1982) is consistent and asymptotically normal. Lee and Hansen (1994) established the consistency and asymptotic normality properties of the quasi-maximum likelihood estimator of the GARCH(1,1) and IGARCH(1,1) models assuming that the standardized variable z_t is stationary and ergodic with a bounded fourth conditional moment. Unfortunately, as with other ARCH models including EGARCH models, a satisfactory asymptotic theory for the log-GARCH/ILGARCH is as yet unavailable. In the remainder of this paper, we assume that the QML estimator is consistent and asymptotically normal[4]. The asymptotic distribution for the QML estimator of the parameter θ takes the form

$$\sqrt{T}(\hat{\theta}_{QMLE} - \theta_0) \overset{a}{\sim} N(0, b_0/a_0^2) \tag{15}$$

where a_0 is the expected value of the Hessian evaluated at the true parameter, and b_0 is the expected value of the outer product of the gradients evaluated at the true parameter. Hence, the asymptotic distribution for the QML estimator of $f(\theta) \equiv .5\pi^2(1-\theta)^2/\theta = \sigma_\eta^2$ takes the form

$$\sqrt{T}(f(\hat{\theta}_{QMLE}) - f(\theta_0)) \overset{a}{\sim} N(0, C), \tag{16}$$

where

$$C = (\partial f(\theta_0)/\partial\theta)^2 b_0/a_0^2. \tag{17}$$

As in Ruiz (1994), we can obtain smoothed estimates of log-volatility, h_t, by applying the extended Kalman filter, described in Anderson and Moore (1979).

[4]This is the usual practice in papers that use ARCH models. See, e.g., Bollerslev, Engle, and Nelson (1995).

The Monte Carlo experiment considers finite sample properties across two dimensions. First, we investigate the effect of small to medium sized samples using sample sizes of 250, 500, 1000, and 2000 observations. Second, we consider the effect of widely differing parameter values given by $\sigma_\eta^2 = \{0.0009, 0.0025, 0.0049, 0.0081, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10\}$. This gives 14 design points and a total of 56 experiments. Parameter values are chosen to cover the region of $(0.0034, 0.0194)$, which is empirical results for daily exchange rates reported in Harvey, Ruiz, and Shephard (1994). The model given by equation (2) is estimated via the ILAGRCH representaion in (14) using Berndt, Hall, Hall, Hausman (1994) algorithm with analytical derivatives. Pre-sample values of the logarithm of conditional variance are set to zero. This is analogous to the IGARCH analysis; see Lumsdaine (1995). The true parameter values transformed by equation (9) are used as starting values for the parameters in the estimation algorithm. To approximate the infinite sum which appears in c_z, $\frac{\partial}{\partial\theta}\ln c_z^2$, and $\frac{\partial^2}{\partial\theta^2}\ln c_z^2$, we use the number large enough so that the difference between approximate sum and the infinite sum is less than 10^{-5}. These numbers are reported in Table 1. For each design point and sample size, we use 200 replications. All conputations were carried out on Pentium PC using GAUSS version 3.2.30.

Response surface methodology facilitates understanding of experimental evidence because large amounts of experimental data can be summarized using simple function forms. It also provides applied researchers a simple tool for computing outcomes at points in the parameter space that are not included in the experimental study. Another advantage, especially for computationally intensive processes, is that a large number of replications is not required. Increasing the number of replications, however, improve the precision of the estimates and makes it less likely that outliers influence the regression results. See Maasoumi and Phillips (1982), Hendry (1982), and Davidson and MacKinnon (1993) for detailed discussions of the merits of response surface methodology.

The response surface regression for bias is

$$\frac{\ln\hat{\sigma}_\eta^2 - \ln\sigma_\eta^2}{v_1} = \frac{\Psi_1(T,\sigma_\eta^2)}{v_1} + u_1 \tag{18}$$

where σ_η^2 is the true parameter value in the DGP, T is the sample size, and u_1 is the regression error term. v_1 denotes the logarithms of the root mean squared errors of parameter estimates (over replications) and implies the heteroskedasticity transform. Although the true functional form of $\Psi_1(T,\sigma_\eta^2)$ is unknown, it may be approximated using polynomial approximations to the true functional forms as:

$$\hat{\Psi}_1(T,\sigma_\eta^2) = \underset{(20.1)}{63.5}\,T^{-1} - \underset{(236.8)}{903.9}\,\sigma_\eta T^{-1} + \underset{(617.2)}{1806.8}\,\sigma_\eta^2 T^{-1} \tag{19}$$

$$\mathrm{R}^2 = 0.53, \quad \mathrm{SE} = 0.017, \quad \mathrm{W} = 0.00, \quad \mathrm{JB} = 13.5,$$

where standard errors are in parenthesis. W denotes White's (1980) test for heteroskedasticity and/or functional form misspecification. W is distributed as a $\chi^2(8)$ under the null hypothesis of homoskedasticity. JB denotes Jarque and Bera's (1987) test for normality. Under the null bypothesis of normality, JB is destributed as $\chi^2(2)$. W is not rejected at the five percent significance level. This indicates the validity of the response surface specifications[5]. All parameters are siginificant at the five percent level. In this case, the bias will be minimized when $\sigma_\eta^2 \simeq 0.0625$. The bias may disapper as the sample size increases.

The next question of interest is how the asymptotic variances, C, computed in each experiment using analytical first and second derivatives compare with T times sample variances of the estimated parameters obtained from the simulations. Specifically, one might expects the asymptotic variance ($AVAR$) should approximate the corresponding T times sample vari-

[5]For the Gaussian QMLE, there is no apriori reason to believe that the residuals of response surfaces will be normally distributed and have unit variance.

ances ($SVAR$). The response surface regression for $AVAR$ is

$$\frac{\ln(AVAR) - \ln(T \cdot SVAR)}{v_2} = \frac{\Psi_2(T, \sigma_\eta^2)}{v_2} + u_2 \tag{20}$$

where u_2 is the regression error term and $v_2 = \ln$(sample standard deviation of $AVAR$ for the experiment) is a heteroskedasticity transformation. We specified the functional form of $\Psi_2(T, \sigma_\eta^2)$ as:

$$\hat{\Psi}_2(T, \sigma_\eta^2) = -\underset{(1069.6)}{1838.7}\, T^{-1} - \underset{(12805.5)}{1035.4}\, \sigma_\eta T^{-1} + \underset{(33749.3)}{1170.7}\, \sigma_\eta^2 T^{-1} \tag{21}$$

$$\mathrm{R}^2 = .13, \quad \mathrm{SE} = .51, \quad \mathrm{W} = 0.00, \quad \mathrm{JB} = 6.01,$$

where standard errors are in parenthesis. W is not rejected at the five percent level. All parameters are not siginificant at the five percent level. The coefficent of T^{-1} is significant at the ten percent level. This is a desirable property since it suggests that the difference between $AVAR$ and $T \cdot SVAR$ is not a function of model parameter, though it is a function of the sample size.

Remainder of this section, we exmine the relative efficiency of QML esimator via the ILGARCH approach to Method of moment estimator and/or to QML esimator based on the Kalman filtering procedure, which are discussed in section 2. We do not consider the Bayesian Markov chain Monte Carlo (MCMC) method of Jacquier, Polson, and Rossi (1994) in our simulation for practical computational reasons. To obtain more precise estimator of C in (17), We conducted another Monte Caro simulation. With sample size 5000 and σ_η^2 = { 0.0009, 0.0049, 0.01, 0.09 }, we calculated the estimate of $\sqrt{C}$ using the sample gradients and sample Hessian evaluated at the true parameter, and obtained its sample mean by 1000 replications. Table 2 shows the asymptotic standard error for three estimators. ILGARCH indicates the estimates of asymptotic standard error, $\sqrt{C}$, of QML estimator via the ILGARCH approach. KF and MM denote analytical asymptotic standard error of QML estimator via the Kalman filtering procedure and method of moments estimator, $\sqrt{C_1(\sigma_\eta^2)}$ and $\sqrt{C_2(\sigma_\eta^2)}$, respectively. Squre root of relative efficiency is in parenthesis. It is possible to observe that, for the parameter values between .0009 and .09, the efficiency of the method of moments estimator and the QML estimator via Kalman filtering methods compared with QML estimator via the ILGARCH approach is low.

4 Empirical Example: Daily Exchange Rates

The data consist of daily observations on the Deutsch mark/U.S. dollar exchange rate over Nobember 24, 1989 through December 31, 1996 period, for a total of 1852[6]. A broad consensus has emerged that nominal exchange rates over the free float period are best described as non-stationay, or $I(1)$, tyope processes; see, *e.g.* Baillie and Bollerslev (1989). We shall, therefore, concentrate on the nominal returns; *i.e.* $r_t = \ln S_t - \ln S_{t-1}$, where S_t denotes the spot Deutsch mark/U.S. dollar exchange rate at day t. Table 3 shows Box-Ljung Q-statistics for several transformations of r_t minus its sample mean. The chi-square 5 percent critical value for ten degrees of freedom is 18.3. There appears to be little serial dependence in the levels of the returns, whereas strong serial dependence in the squared returns and their logarithms. Note that in the presence of heteroskedasticity, Ljung-Box Q-statistics will tend to over-reject.

We consider the following MA(1)-SV model:

$$
\begin{aligned}
r_t &= \delta + e_t - b e_{t-1} \\
e_t &= \varepsilon_t \exp(h_t/2) \qquad \varepsilon_t \sim \text{NID}(0,1), \\
h_t &= \gamma + \phi h_{t-1} + \sigma_\eta \eta_t, \qquad \eta_t \sim \text{NID}(0,1),
\end{aligned}
$$

The MA(1) term is included to take account of the weak serial dependence in the mean. Following Asai (1998), the above MA(1)-SV model in r_t may

[6]The data were obtained through Datastream.

be interpreted as an MA(1)-log-GARCH(1,1) model in r_t

$$r_t = \delta + e_t - be_{t-1}$$
$$e_t = \sigma_t z_t,$$
$$\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln e_{t-1}^2 + \beta \ln\sigma_{t-1}^2,$$

where the definition of transformed parameters, α_0, α_1 and β_1, are given in appendix C.

The first column of table 4 reports the QML estimates based on the above representation. The robust QML covariance estimators proposed in a very general framework by White (1982) are used to compute the standard errors. Testing procedures for unit root in variance are discussed in appendix C. The t-values for the null of $\phi = \alpha_1 + \beta_1 = 1$ are calucalted under the MA(1)-log-GARCH(1,1) representation. The t-value for $\phi = 1$ is -1.1579 and cannot be rejected the null of unitroot in variance at five percent siginificant level. Since many paremeters are insignificant at five percent level, the model may be misspecified.

The second column of table 4 indicates the QML estimates of SV model, excluding the MA term. The t-value for $\phi = 1$ is -1.1604 and cannot be rejected the null of unitroot in variance. To calculate KPSS test statistic, we have to determine the lag truncation parameter, l. According to the results of Andrews (1991, pp.834-835), l is automatically selected as 336. The critical value for five percent level and ten percent level are 0.463 and 0.347 respectively; see Kwiatkowski *at al.* (1992). The KPSS statistic is 0.4285 and is rejected at ten percent siginificance level. We conclude that log-volatility of r_t has a unitroot.

The QML estimates of MA(1)-RWSV and RWSV models are given in third and fourth columns of table 4, respectively. The t-value of the coefficent of MA(1), b, is significant, in this case. Therefore the MA(1)-RWSV model is preffered.

5 Concluding Remarks

In this paper, we briefly review the estimation methods for RWSV models. We derive a ILGARCH representation of a class of RWSV models, including linear regression models with ARMA(p,q)-RWSV errors. To estimate these RWSV models. we propose a new QML method via the ILGARCH approach. Our Monte Carlo results indicates that QML estimator via the ILGARCH approach performs well. In the view of relative efficiency, for the parameter velues found in empirical analysis our estimators are superior to those of the Method of Moments estimator and those of the QML estimator based on the Kalman filter. We developed procedures for testing the integration of log-volatility. We presented a empirical example of daily Deutsch mark/U.S. dollar exchange rates.

Appendices

A The limiting variance

Using the results in Dunsmuir (1979), the asymptotic variance of the QML estimator via the KF approach is given by

$$C_1(\sigma_\eta^2) = 2A^{-1} + A^{-1}BA^{-1},$$

with

$$A = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\partial \ln g(e^{i\lambda})}{\partial \sigma_\eta^2}\right)^2 d\lambda,$$

where $g(e^{i\lambda})$ is the spectral generating function of the stationary form of $\ln y_t^2$, and

$$B = \kappa\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial \ln g(e^{i\lambda})}{\partial \sigma_\eta^2}d\lambda\right]^2,$$

where κ is the measure of excess kurtosis of $\ln\varepsilon_t^2$ and $\kappa = 4$; see Abramovits and Stegun (1970, p.943).

The stationary from of $\ln y_t^2$ in equation (7) is given by,

$$\Delta \ln y_t^2 = \sigma_\eta \eta_t + (1-L)\zeta_t.$$

The spectral generating function of $\Delta \ln y_t^2$ is given by

$$g(e^{i\lambda}) = \sigma_\eta^2 + |1-e^{i\lambda}|^2\frac{\pi^2}{2} = \sigma_\eta^2 + \pi^2(1-\cos\lambda),$$

and thus

$$\frac{\partial \ln g(e^{i\lambda})}{\partial \sigma_\eta^2} = \frac{1}{\sigma_\eta^2 + \pi^2(1-\cos\lambda)}.$$

Using the results in Abramovits and Stegun (1970, p.78),

$$\int_{-\pi}^{\pi}\frac{\partial \ln g(e^{i\lambda})}{\partial \sigma_\eta^2}d\lambda$$

$$
\begin{aligned}
&= \frac{1}{\pi^2}\int_{-\pi}^{\pi}\frac{1}{(1+\sigma_\eta^2/\pi^2)-\cos\lambda}d\lambda \\
&= \frac{1}{\pi^2}\left[\frac{2}{\sqrt{(1+\sigma_\eta^2/\pi^2)^2-1}}\arctan\frac{((1+\sigma_\eta^2/\pi^2)+1)\tan\frac{\lambda}{2}}{\sqrt{(1+\sigma_\eta^2/\pi^2)^2-1}}\right]_{-\pi}^{\pi} \\
&= \frac{2\pi}{\sqrt{\sigma_\eta^4+2\sigma_\eta^2\pi^2}}.
\end{aligned}
$$

Similarily, we obtain from the results in Prudnikov *at al.* (1986, subsection 1.5.9),

$$
\begin{aligned}
&\int_{-\pi}^{\pi}\left(\frac{\partial \ln g(e^{i\lambda})}{\partial\sigma_\eta^2}\right)^2 d\lambda \\
&= \frac{1}{\pi^4}\int_{-\pi}^{\pi}\frac{1}{[(1+\sigma_\eta^2/\pi^2)-\cos\lambda]^2}d\lambda \\
&= \frac{1}{\pi^4}\left[\frac{\sin\lambda}{[(1+\sigma_\eta^2/\pi^2)^2-1][(1+\sigma_\eta^2/\pi^2)-\cos\lambda]}\right]_{-\pi}^{\pi} \\
&\quad +\frac{(1+\sigma_\eta^2/\pi^2)}{\pi^4[(1+\sigma_\eta^2/\pi^2)^2-1]}\int_{-\pi}^{\pi}\frac{1}{(1+\sigma_\eta^2/\pi^2)-\cos\lambda}d\lambda \\
&= 2\pi\frac{(\sigma_\eta^2+\pi^2)}{(\sigma_\eta^4+2\sigma_\eta^2\pi^2)^{3/2}}.
\end{aligned}
$$

Therefore the analytical expression of A, B and C_1 are given by

$$
A=\frac{\sigma_\eta^2+\pi^2}{(\sigma_\eta^4+2\sigma_\eta^2\pi^2)^{3/2}},\qquad B=\frac{4}{\sigma_\eta^4+2\sigma_\eta^2\pi^2},
$$

$$
C_1(\sigma_\eta^2)=\frac{2}{\sigma_\eta^2+\pi^2}\left[(\sigma_\eta^4+2\sigma_\eta^2\pi^2)^{3/2}+\frac{2(\sigma_\eta^4+2\sigma_\eta^2\pi^2)^2}{\sigma_\eta^2+\pi^2}\right].
$$

B ARMA-RWSV Models

We now consider a linear regression model with ARMA(p,q)-RWSV errors, or simply ARMA-RWSV model:

$$
y_t = X_t\delta+u_t, \tag{22}
$$

$$
\begin{aligned}
A(L)u_t &= B(L)e_t, & (23)\\
e_t &= \varepsilon_t \exp(h_t/2), \quad \varepsilon_t \sim \text{NID}(0,1), & (24)\\
h_t &= h_{t-1} + \sigma_\eta \eta_t, \quad \eta_t \sim \text{NID}(0,1), & (25)
\end{aligned}
$$

where ε_t is generated independently of η_t, X_t is a $1 \times k$ vector, δ is a $k \times 1$ parameter vector,

$$
A(L) = 1 - a_1 L - \cdots - a_p L^p,
$$
$$
B(L) = 1 - b_1 L - \cdots - b_q L^q,
$$

and L is the lag operator. The Kalman filtering method is hard to apply estimating ARMA-RWSV models.

The ARMA-RWSV model of (22)-(25) can be interpreted as an ARMA-ILGARCH(1,1) model in a similar fashion to the simple RWSV model (2):

$$
\begin{aligned}
y_t &= X_t\delta + u_t,\\
A(L)u_t &= B(L)e_t,\\
e_t &= \sigma_t z_t, \quad z_t \sim \text{WN}(0,1), & (26)\\
\ln\sigma_t^2 &= (1-\theta)\ln c_z^2 + (1-\theta)\ln e_{t-1}^2 + \theta \ln\sigma_{t-1}^2, & (27)
\end{aligned}
$$

where the definition of the transformed parameter θ is the same as in the RWSV case. z_t has the heavy-tailed and symmetric conditional distribution.

C Test for Integration in Log-Volatility

The appropriate procedure for testing for integration in variance is not yet clear. In this section, we discuss three possible way.

One possible way is to apply augmented Dickey-Fuller test and/or Phillps-Perron test to $\ln y_t^2$. However, the reliability of such unit root tests in this situation is questionable. According to Asai (1998) and Harvey, Ruiz, and Shephard (1994), simple SV process (1) in y_t can be interpreted as ARMA(1,1) process in $\ln y_t^2$:

$$\ln y_t^2 = \gamma^* + \phi \ln y_{t-1}^2 + \zeta_t - \theta^* \zeta_{t-1}, \tag{28}$$

where ζ_t is white noise and $\gamma^* = \gamma - 1.27(1 - \phi)$. Since the variance of ζ_t typically dominates the variance of $\sigma_\eta \eta_t$, the parameter θ^* will be close to unity for values of ϕ close to one. For example, when $\phi = .98$ and $\sigma_\eta = .166$, which was used in Monte Carlo experiments of Jacquier, Polson, and Rossi (1994), θ^* is .92. As shown in Pantula (1991) and Schwert (1989), when moving-average parameter is very close to one, unit root tests reject the null hyposthesis of a unit root too often since the model is difficult to distinguish from white noise.

The other possible way is to use KPSS test proposed by Kwiatkowski *at al.* (1992). The KPSS test is basically different from other unit root tests since the null hypothesis is that an observed series is stationary. We consider the data-generating-process (DGP) of $\ln y_t^2$:

$$\ln y_t^2 = \alpha + \xi_t, \quad t = 1, 2, \ldots, T$$

Kwiatkowski *at al.* (1992) assume the components representation $\xi_t = R_t + V_t$, where r_t is a random walk,

$$R_t = R_{t-1} + \nu_t, \;\; r_0 = 0, \quad \nu_t \sim \text{iid}(0, \sigma_\nu^2),$$

and V_t is ARMA process that satisfies Assumption 2.1 of Phillips (1987, p.280). They test the stationarity hypothesis $\text{H}_0 : \sigma_\nu^2 = 0$, which implies that $\ln y_t^2 = V_t$ is ARMA process. We can test, therefore, whether the DGP of $\ln y_t^2$ is ARMA process in (28) or random walk process in (3)

Let $\hat{\xi}_t$ be the residuals from a regression of $\ln y_t$ on intercept, and S_t be the partial sum process of the $\hat{\xi}_t$: $S_t = \sum_{j=1}^{t} \hat{\xi}_j$, $t = 1, \ldots, T$. Let σ^2 be the long-run variance of the errors v_t, and consider the Newey-West (1987) estimator of σ^2:

$$s^2(l) = T^{-1} \sum_{t-1}^{T} \hat{\xi}_t^2 + 2T^{-1} \sum_{s=1}^{l} w(s, l) \sum_{t=s+1}^{T} \hat{\xi}_t \hat{\xi}_{t-s},$$

where $w(s,l) = 1 - s/(l+1)$, which guarantees the nonnegativity of $s^2(l)$. For consistency of $s^2(l)$ under the null hypothesis it is necessary that the lag truncation parameter $l \to \infty$ and $T \to \infty$. l may be automatically selected by the resluts of Andrews (1991). The KPSS statistic for testing the null of stationarity can then be expressed as follows:

$$\text{KPSS} = T^{-2}\sum_{t=1}^{T} S_t^2/s^2(l).$$

Critical values for this statistic are provided in Table 1 of Kwiatkowski *et al.* (1992).

The third way depends on the results of Asai (1998). He showed the simple SV process in y_t of equation (1) can be interpreted as a log-GARCH(1,1) model in y_t

$$y_t = \sigma_t z_t,$$
$$\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln e_{t-1}^2 + \beta \ln\sigma_{t-1}^2,$$

where

$$\begin{aligned}
\alpha_0 &\equiv \gamma + (1-\phi)c_\eta + (1-\theta^*)\ \ln c_z^{*2}, \\
\alpha_1 &\equiv \phi - \theta^*, \\
\beta_1 &\equiv \theta^*,
\end{aligned}$$

and

$$\begin{aligned}
\theta^* &= \frac{1}{2\phi}\left[1 + \phi^2 + \frac{2\sigma_\eta^2}{\pi^2} - \sqrt{\left(1 - \phi^2 + \frac{2\sigma_\eta^2}{\pi^2}\right)^2 + \frac{8\phi^2\sigma_\eta^2}{\pi^2}}\right], \\
c_z^* &= \exp\left[\frac{\sigma_\nu^2}{4(1-\theta^2)} - \frac{c_\eta}{2} - \frac{(\theta-\phi)\theta}{2(1-\theta)}\psi\left(\frac{1}{2}\right)\right. \\
&\qquad \left. + \frac{1}{2}\sum_{i=1}^{\infty}\left[\ln\Gamma\left((\theta-\phi)\theta^i + \frac{1}{2}\right) - \ln\Gamma\left(\frac{1}{2}\right)\right]\right].
\end{aligned}$$

Since $\phi = \alpha_1 + \beta_1$, the null hypothesis for $\phi = 1$ can be tested by t statistics for the null hypothesis that $\alpha_1 + \beta_1 = 1$. As is the test for IGARCH(1,1) model, this test can be easily conducted.

References

[1] Abramovitz, M. and N. Stegun, 1970, Handbook of Mathematical Functions (Dover Publications, NY).

[2] Andersen, T.G., and B.E. Sørensen, 1996, GMM estimation of a stochastic volatility model: a Monte Carlo study, Journal of Business and Economic Statistics 14, 328–353.

[3] Andrews, D.W.K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, 817–858.

[4] Asai, M., 1998, A new method to estimate stochastic volatility models: a log-GARCH approach, forthcoming in: Journal of the Japan Statistical Society 28.

[5] Baillie, R.T. and T. Bollerslev, 1989, The message in daily exchange rates: a conditional-variance tail, Journal of Business and Economic Statistics 7, 297–305.

[6] Berndt, E.K., B.H. Hall, R.E. Hall and J.A. Hausman, 1974, Estimation and inference in nonlinear structural models Annals of Economic and Social Measurement 4, 653–665.

[7] Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307–327.

[8] Bollerslev, T., R.Y. Chou and K.F. Kroner, 1992, ARCH modeling in finance: a review of the theory and empirical evidence, Journal of Econometrics 52, 5–59.

[9] Bollerslev, T., R.F. Engle and D.B. Nelson, 1995, ARCH models, in: R.F. Engle and D. McFadden, eds., The Handbook of Econometrics Vol.4 (North-Holland, Amsterdam) 2959–3038.

[10] Davidson, R. and J.G. MacKinnon, 1993, Estimation and Inference in Econometrics (Oxford University Press, Oxford).

[11] Dunsmuir, W., 1979, A central limit theorem for parameter estimation in stationary times series and its applications to models for a signal observed white noise, Annals of Statistics 7, 490–506.

[12] Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica 50, 987–1008.

[13] Engle, R.F. and T. Bollerslev, 1986, Modelling the persistence of conditional variances, Econometric Reviews 5, 1–50.

[14] Geweke, J., 1986, Modeling the persistence of conditional variances: a comment, Econometric Reviews 5, 57–61.

[15] Hamilton, J.D., 1994, Time Series Analysis (Princeton University Press, NJ).

[16] Harvey, A.C., E. Ruiz and N. Shephard, 1994, Multivariate stochastic variance models, Review of Economic Studies 61, 247–264.

[17] Hendry, D.F., 1982, Monte Carlo experimantation in econometrics, in: Z. Griliches and M.D. Intriligator, eds., The Handbook of Econometrics Vol.2 (North-Holland, Amsterdam) 937–976.

[18] Jacquier, E., N.G. Polson and P.E. Rossi, 1994, Bayesian analysis of stochastic volatilities, Journal of Business and Economic Statistics 12, 371–389.

[19] Jarque, C.M. and A.K. Bera, 1987, A test for normality of observations and regression residuals, International Statistical Review 55, 163–172.

[20] Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, 1992, Testing the null hypothesis of stationary against the alternative of a unit root: How sure are we that economic time series have a unit root?, Journal of Econometrics 54, 159–178.

[21] Lee, S.W. and B.E. Hansen, 1994, Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, Econometric Theory 10, 29-52.

[22] Lumsdaine, R.L., 1995, Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: a Monte Carlo investigation, Journal of Business and Economics Statistics 13, 1–10.

[23] Maasoumi, E. and P.C.B. Phillips, 1982, On the behayior of inconsistent instrumental variable estimators, Journal of Econometrics 19, 183–201.

[24] Nelson, D.B., 1991, Conditional heteroskedasticity in asset returns: a new approach, Econometrica 59, 347–370.

[25] Newey, W.K. and K.D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.

[26] Pantula, S.G., 1986, Modeling the persistence of conditional variances: a comment, Econometric Reviews 5, 79–97.

[27] Pantula, S.G., 1991, Asymptotic distributions of unit-root tests when the process is nearly stationary, Jouranl of Business and Economic Statistics 9, 63–71.

[28] Phillips, P.C.B., 1987, Time series regression with a unit root, Econometrica 55, 277–301.

[29] Prudnikov, A.P., Yu.A. Brychkov and O.I. Marichev, 1986, Integrals and Series Vol.1 (Gordon and Breach Science Publishers, NY).

[30] Ruiz, E., 1994, Quasi-maximum likelihood estimation of stochastic volatility models, Jouranl of Econometrics 63, 289–306.

[31] Schwert, G.W., 1989, Tests for unit roots: a Monte Carlo investigation, Jouranl of Business and Economic Statistics 7, 147–159.

[32] White, H., 1980, A heteroskedasticity-consistent covariances matrix estimator and a direct test for heteroskedasticity, Econometrica 48, 817–838.

[33] White, H., 1982, Maximum likelihood estimation of misspecified models, Econometrica 50, 1–25.

Table 1: The Number N for approximating the infinite summations in c_z, $\frac{\partial}{\partial\theta}\ln c_z^2$ and $\frac{\partial^2}{\partial\theta^2}\ln c_z^2$

σ_η^2	0.0009	0.0025	0.0049	0.0081	
N	2000	1500	1100	800	
σ_η^2	0.01	0.02	0.03	0.04	0.05
N	700	600	500	400	350
σ_η^2	0.06	0.07	0.08	0.09	0.1
N	350	350	350	300	300

Table 2: Asymptotic standard error of QML estimator

σ_η^2	ILGARCH	KF	MM
0.09	0.71255	0.77922 (0.9144)	19.8294 (0.0359)
0.05	0.40598	0.48810 (0.8318)	19.7893 (0.0205)
0.01	0.10106	0.13916 (0.7262)	19.7492 (0.0051)
0.0049	0.05845	0.08049 (0.7262)	19.7441 (0.0030)
0.0009	0.01799	0.02220 (0.8103)	19.7401 (0.0009)

Notes: ILGARCH indicates the estimates of asymptotic standard error, $\sqrt{C}$, of QML estimator via the ILGARCH approach (1000 reprications). KF and MM denote analytical asymptotic standard errors of QML estimator via the Kalman filtering method and method of moments estimator, $\sqrt{C_1(\sigma_\eta^2)}$ and $\sqrt{C_2(\sigma_\eta^2)}$, respectively. Squre root of relative efficiency is in parenthesis.

Table 3: Box-Ljung Q-statistics based on ten lags

	r_t	r_t^2	$\ln r_t^2$
Q(10)	18.4	132.8	58.2

Sample period: 1 January 1990 to 31 December 1996.
The critical value is given by $\chi^2_{.05}(10) = 18.3$

Table 4: Quasi-maximum likelihood estimates

	MA(1)-SV	SV	MA(1)-RWSV	RWSV
δ_0	-0.0001136	-0.0001099	-0.0001802	-0.0001908
	(0.0001686)	(0.0001522)	(0.9267 $\times 10^{-7}$)	(0.8729 $\times 10^{-7}$)
b	-0.04515	—	-0.06218	—
	(0.03032)		(0.0001565)	
γ	-0.1236	-0.06848	—	—
	(0.1110)	(0.06470)		
ϕ	0.9868	0.9923	—	—
	(0.01136)	(0.006660)		
σ_η^2	0.004174	0.003237	0.01085	0.009930
	(0.003576)	(0.002204)	(0.003519)	(0.003146)
quasi-log-likelihood	3.5880	3.5915	3.5076	3.4950
H_0:$\phi = 1$ t-value	-1.1579	-1.1604	—	—

Notes: Sample period: 1 January 1990 to 31 December 1996.
Robust standard errors are in parentheses.
t-values for the null of $\phi = \alpha_1 + \beta_1 = 1$ are calucalted under the log-GARCH(1,1) representation.