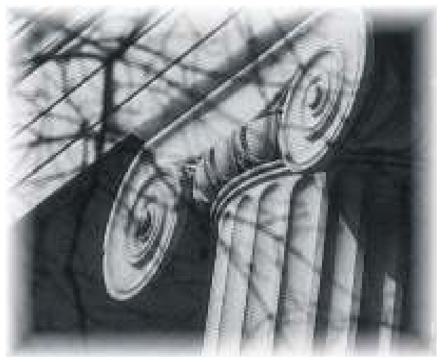
THE UNIVERSITY OF ALABAMA

ECONOMICS, FINANCE AND LEGAL STUDIES

WORKING PAPER SERIES



Working Paper No. 07-09-06

A Closed-Form Asymptotic Variance-Covariance Matrix for the Quasi Maximum Likelihood Estimator of the GARCH(1,1) Model

Jun Ma

Culverhouse College of Commerce and Business Administration
The University of Alabama

A Closed-Form Asymptotic Variance-Covariance Matrix for the Quasi-Maximum Likelihood Estimator of the GARCH(1,1) Model

Jun Ma
Department of Economics, Finance and Legal Studies
University of Alabama

First draft March 2, 2006
This draft May 1, 2008

JEL *classification:* C12, C22

Abstract: This paper presents a closed-form asymptotic variance-covariance matrix of the Quasi-Maximum Likelihood Estimator (QMLE) for the GARCH(1,1) model. The robust 'sandwich' asymptotic variance matrix is shown to be a product of the function of higher moments of innovation and the inverse of negative expected Hessian, whose closed-form in terms of only model parameters is then derived via a local approximation. Taking inverse of it, the variance-covariance matrix is readily obtained. A Monte Carlo simulation experiment demonstrates that this analytical formula works well for both normal and non-normal innovations in admissible parameter regions.

Key words: GARCH, Quasi-Maximum Likelihood Estimator, asymptotic variance-covariance matrix

The author gratefully acknowledges financial support from Buechel Fellowship at the University of Washington. I am grateful to Tim Bollerslev and Dennis Kristensen for the help at the early stage of this work. I appreciate helpful comments from Charles R. Nelson, Richard Startz, Eric Zivot and Zhen-Qing Chen on the manuscript. But I am entirely responsible for the content of this paper.

Correspondence:
Department of Economics, Finance and Legal Studies
Box 870224
University of Alabama
Tuscaloosa, AL 35487-0224
Email Addresses: jma@cba.ua.edu
Phone: (205)348-8985

1. INTRODUCTION

The GARCH(1,1) model has become a benchmark in modeling time-varying volatility since its introduction by Bollerslev (1986). Despite its wide applications in both academia and industry, there has not yet been a closed-form asymptotic variance for its estimator in the stationary region[1]. This paper works out a closed-form asymptotic variance matrix for the GARCH(1,1) QMLE in terms of only model parameters in stationary region.

The general consistency and asymptotic normality of the GARCH(1,1) QMLE have been well established by Bollerslev and Wooldridge (1992), Lee and Hansen (1994), and Lumsdaine (1996), which will be briefly discussed in Section 2. In section 3, it is shown that the robust asymptotic variance-covariance matrix for the GARCH QMLE is simply the product of a function of higher moments of innovation and the inverse of negative expected Hessian. A local approximation is taken in the expected Hessian matrix and consequently, the derivation breaks down to those of the auto-covariance and cross-covariance structures for squared variables, which are then derived. This results in a closed-form expected Hessian, taking the inverse of which the asymptotic variance-covariance matrix is readily obtained. In section 4, a Monte Carlo experiment is conducted to show that this formula works very well for admissible parameters. Section 5 concludes.

2. THE ASYMPTOTICS OF GARCH(1,1) MLE

An archetype GARCH(1,1) model may be written as:

$$\varepsilon_t = \sqrt{h_t} \cdot \xi_t \tag{1}$$

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} \tag{2}$$

Here, the mean of equation (1) is set to be zero without loss of generality since the information matrix is block-diagonal (see Bollerslev (1986)). ξ_t is independently and identically distributed with zero mean and unit variance, i.e., *i.i.d.(*0,1), with finite higher moments (see Lumsdaine (1996) for details).

Write up the log-likelihood function:

$$L_T(\theta) = T^{-1} \sum_{t=1}^{T} l_t(\theta) \tag{3}$$

$$l_t(\theta) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log h_t - \frac{1}{2}\frac{\varepsilon_t^2}{h_t} \tag{4}$$

Where, $\theta = (\omega, \alpha, \beta)'$ and $\hat{\theta}_T$ maximizes the quasi log-likelihood function for a given sample data $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T\}$ and therefore is the QMLE. Note if ξ_t is normally distributed, $\hat{\theta}_T$ becomes MLE. In practice, the evaluation of (3) and (4) conditions upon an initial assignment of h_0. The difference of various choices of

[1] Kristensen and Linton (2006) provide a closed-form estimator alternative to the (Q)MLE. Jensen and Rahbek (2004) obtain a 'nearly' analytical variance for non-stationary GARCH since their result still involves expectations of a nonlinear form of innovation, which gives a lower bound for the variance of stationary GARCH.

h_0, however, is negligible asymptotically given a stationary and ergodic underlying process.

The Gradient at each time t of the log-likelihood function is:

$$s_t(\theta) = \frac{\partial h_t}{\partial \theta}\frac{1}{2h_t}(\frac{\varepsilon_t^{\,2}}{h_t} - 1) \tag{5}$$

By law of iterated expectation we have a standard result (see Bollerslev and Wooldridge (1992)):

$$E[s_t(\theta)] = 0 \tag{6}$$

The Hessian at each time t is given by:

$$H_t(\theta) = \frac{\partial^2 l_t}{\partial\theta\partial\theta'} = (\frac{\varepsilon_t^{\,2}}{h_t} - 1)\frac{\partial}{\partial\theta'}[\frac{1}{2h_t}\frac{\partial h_t}{\partial\theta}] - \frac{1}{2h_t^{\,2}}\frac{\partial h_t}{\partial\theta}\frac{\partial h_t}{\partial\theta'}\frac{\varepsilon_t^{\,2}}{h_t} \tag{7}$$

Again by law of iterated expectation, we have (see Engle (1982)):

$$E[H_t(\theta)] = -E[\frac{1}{2h_t^{\,2}}\frac{\partial h_t}{\partial\theta}\frac{\partial h_t}{\partial\theta'}] \tag{8}$$

Lumsdaine (1996) proves the consistency and asymptotic normality of the QMLE by assuming a compact and convex parameter space along with a strict stationarity and ergodicity condition for the GARCH(1,1) model, derived in Nelson (1990):

$$E[\ln(\beta + \alpha\cdot\xi_t^2)] < 0 \tag{9}$$

To have a well-defined finite unconditional volatility, I further impose a stronger restriction as in Bollerslev (1986):

$$\alpha + \beta < 1 \tag{10}$$

One can easily verify that condition (10) along with the assumption that ξ_t has unit variance is sufficient to derive condition (9) via Jensen's inequality.

The asymptotic result for the QMLE is given by:

$$T^{1/2}(\hat{\theta}_T - \theta_0)\overset{A}{\sim} N(0,(I_0 C_0^{-1} I_0)^{-1}) \tag{11}$$

Where θ_0 is true parameter value; I_0 is negative expected Hessian evaluated at θ_0:

$$I_0 = -E[\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}]\Big|_{\theta_0} = E[\frac{1}{2h_t^{\,2}}\frac{\partial h_t}{\partial\theta}\frac{\partial h_t}{\partial\theta'}]\Big|_{\theta_0} \tag{12}$$

And

$$C_0 = E[\frac{\partial l_t(\theta)}{\partial \theta} \cdot \frac{\partial l_t(\theta)}{\partial \theta'}]\Big|_{\theta_0} \tag{13}$$

Plug (5) into (13) we can easily derive:

$$C_0 = E[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'}]\Big|_{\theta_0} \cdot \kappa = I_0 \cdot \kappa \tag{14}$$

Here, $\kappa = {E(\xi_t^2 - 1)^2}/{2} < \infty$. Note if ξ_t is normally distributed, then $\kappa = 1$. For fat-tailed distributions, κ is typically greater than 1.

Therefore, we have:

$$T^{1/2}(\hat{\theta}_T - \theta_0) \overset{A}{\sim} N(0, \kappa \cdot I_0^{-1}) \tag{15}$$

Again, if ξ_t is normally distributed, the asymptotic variance is simply the inverse of information matrix. For non-normal distributions, however, the variance needs to be adjusted by a scale κ.

By recursion and assuming the process extends infinitely far into the past, we have the analytical result:

$$\frac{\partial h_t}{\partial \theta} = \begin{bmatrix} \sum_{i=1}^{\infty} \beta^{i-1} \\ \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}{}^2 \\ \sum_{i=1}^{\infty} \beta^{i-1} h_{t-i} \end{bmatrix} \tag{16}$$

Note Fiorentini, Calzolari and Panattoni (1996) derive the first and second derivatives in GARCH models, which include this specific result. Combine results (12) and (16) to get the symmetric and positive definitive matrix *I*:

$$I = \frac{1}{2} \begin{bmatrix} E\left[\frac{(\sum_{i=1}^{\infty} \beta^{i-1})^2}{h_t^2}\right] & E\left[\frac{\sum_{i=1}^{\infty} \beta^{i-1} \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}{}^2}{h_t{}^2}\right] & E\left[\frac{\sum_{i=1}^{\infty} \beta^{i-1} \sum_{i=1}^{\infty} \beta^{i-1} h_{t-i}}{h_t{}^2}\right] \\ -- & E\left[\frac{(\sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}{}^2)^2}{h_t^2}\right] & E\left[\frac{\sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}{}^2 \sum_{i=1}^{\infty} \beta^{i-1} h_{t-i}}{h_t{}^2}\right] \\ -- & -- & E\left[\frac{(\sum_{i=1}^{\infty} \beta^{i-1} h_{t-i})^2}{h_t^2}\right] \end{bmatrix} \tag{17}$$

Next, I take a local approximation for each element in the neighborhood of small α to avoid taking the expectation of a nonlinear form. After all, most empirical work adopting GARCH(1,1) estimation end up with an estimated $\hat{\alpha}$ ranging roughly from 0.05 to 0.10; see, e.g., Zivot (2008) and reference therein. Take the element *I*(1,1) for illustration purpose:

$$E\left[\frac{(\sum_{i=1}^{\infty}\beta^{i-1})^2}{h_t^2}\right] \approx \frac{(1-\alpha-\beta)^2}{\omega^2}\cdot\left(E\left[\left(\sum_{i=1}^{\infty}\beta^{i-1}\right)^2\right]\right) \tag{18}$$

Note here the approximation is valid given each element is bounded on the parameter space. In this way, we can deal with only the linear part on the numerators. It is easy to derive the analytical expressions for the numerators of *I*(1,1), *I*(1,2) and *I*(1,3). Next, I show how to derive those of *I*(2,2), *I*(3,3) and *I*(2,3).

3. THE DERIVATION OF A CLOSED FORM INFORMATION MATRIX

To derive analytical expressions for the numerators of *I*(2,2), *I*(2,3) and *I*(3,3), we need to work out the auto-covariance and cross-covariance structures for $\{\varepsilon_t^2\},\{h_t\}$. To see why, express out the terms:

$$E\left(\sum_{i=1}^{\infty}\beta^{i-1}\varepsilon_{t-i}{}^2\right)^2 = E(\sum_{i=1}^{\infty}\beta^{2(i-1)}\varepsilon_{t-i}^4 + 2\sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\beta^{i+j-2}\varepsilon_{t-i}^2\varepsilon_{t-j}^2) \tag{19}$$

$$E\left(\sum_{i=1}^{\infty}\beta^{\,i-1}h_{t-i}\right)^2 = E(\sum_{i=1}^{\infty}\beta^{2(i-1)}h_{t-i}^2 + 2\sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\beta^{i+j-2}h_{t-i}h_{t-j}) \tag{20}$$

$$E(\sum_{i=1}^{\infty}\beta^{\,i-1}\varepsilon_{t-i}{}^2\sum_{i=1}^{\infty}\beta^{\,i-1}h_{t-i}) = E[\sum_{i=1}^{\infty}\beta^{2(i-1)}\varepsilon_{t-i}^2h_{t-i} + \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\beta^{i+j-2}(\varepsilon_{t-i}^2h_{t-j} + h_{t-i}\varepsilon_{t-j}^2)] \tag{21}$$

The derivation of the auto-covariance structures for $\{\varepsilon_t^2\},\{h_t\}$ starts from the ARMA(1,1) representation for the GARCH(1,1) model:

$$\varepsilon_t{}^2 = \omega + (\alpha+\beta)\cdot\varepsilon_{t-1}{}^2 + w_t - \beta w_{t-1} \tag{22}$$

Where innovation $w_t = h_t(\xi_t^2 - 1)$ is a Martingale Difference Sequence (MDS):

$$E[w_t \mid I_{t-1}] = 0 \tag{23}$$

Here $\{I_t\}$ denotes the information filtration and w_t is adapted to I_t.

Furthermore, notice that $\{h_t\}$ has an AR(1) representation:

$$h_t = \omega + (\alpha+\beta)\cdot h_{t-1} + \alpha w_{t-1} \tag{24}$$

Note that h_t is adapted to I_{t-1}.

The assumption (10) implies a finite unconditional variance for ε_t:

$$E[\varepsilon_t^2] = E[h_t] = \frac{\omega}{1-\alpha-\beta} < \infty\,, \tag{25}$$

However, for the existence of a finite fourth moment, we need to impose one more restriction for parameter space as derived in Bollerslev (1988):

$$3\alpha^2 + 2\alpha\beta + \beta^2 < 1 \tag{26}$$

Under this restriction, we have:

$$E[\varepsilon_t^4] = 3E[h_t^2] = \frac{3\omega^2(1+\alpha+\beta)}{(1-3\alpha^2-2\alpha\beta-\beta^2)(1-\alpha-\beta)} \tag{27}$$

And the following autocorrelations for both $\{\varepsilon_t^2\}$ and $\{h_t\}$:

$$\rho_1^{\varepsilon 2} = \frac{\alpha(1-\alpha\beta-\beta^2)}{1-2\alpha\beta-\beta^2}\text{, and } \rho_i^{\varepsilon 2} = (\alpha+\beta)^{i-1}\rho_1, i = 2,3,\cdots \tag{28}$$

$$\rho_i^h = (\alpha+\beta)^i, i = 1,2,\cdots \tag{29}$$

These autocorrelations have also been independently derived in Bollerslev (1988) and Kristensen and Linton (2006). One can derive them following Harvey (1993, Chapter 1). He and Terasvirta (1999) also work out the general fourth moment structure of a squared GARCH process.

Manipulating by the standard formulas for expectation, covariance and variance, summing up the geometric series, and plugging in the expressions for the second and fourth moments, I obtained the closed forms of (19) and (20) (See Appendix for details):

$$E\left(\sum_{i=1}^{\infty}\beta^{i-1}\varepsilon_{t-i}^{2}\right)^2 = \frac{\omega^2}{(1-2\alpha\beta-\beta^2)(1-\alpha-\beta)}\left[\frac{3(1+\alpha+\beta)}{1-3\alpha^2-2\alpha\beta-\beta^2} + \frac{2\beta}{(1-\beta)^2}\right] \tag{30}$$

$$E\left(\sum_{i=1}^{\infty}\beta^{i-1}h_t\right)^2 = \frac{\omega^2}{(1-\beta^2)(1-\alpha\beta-\beta^2)(1-\alpha-\beta)}\left[\frac{(1+\alpha\beta+\beta^2)(1+\alpha+\beta)}{1-3\alpha^2-2\alpha\beta-\beta^2} + \frac{2\beta}{1-\beta}\right] \tag{31}$$

Lastly, to derive the analytical expression for (21), I transform the cross-covariance between $\{\varepsilon_t^2\}$ and $\{h_t\}$ to known auto-covariance of $\{\varepsilon_t^2\}$ and $\{h_t\}$ by taking advantage of the MDS property of w_t (See Appendix for details). These results, along with the previous work, allow us to derive the following closed form expression for the numerator of $I(2, 3)$:

$$E(\sum_{i=1}^{\infty}\beta^{i-1}\varepsilon_{t-i}{}^{2}\sum_{i=1}^{\infty}\beta^{i-1}h_{t-i})$$
$$=\frac{\omega^2(1+\alpha+\beta)}{(1-\alpha-\beta)(1-3\alpha^2-2\alpha\beta-\beta^2)(1-\beta^2)}(\frac{1}{1-\alpha\beta-\beta^2}+\frac{3\alpha\beta}{1-2\alpha\beta-\beta^2}) \quad (32)$$
$$+\frac{\omega^2\beta}{(1-\alpha-\beta)^2(1-\beta^2)}(\frac{2}{1-\beta}-\frac{\alpha+\beta}{1-\alpha\beta-\beta^2}-\frac{\alpha}{1-2\alpha\beta-\beta^2})$$

To finish this section, I list the results for numerators of $I(1,1)$, $I(1,2)$ and $I(1,3)$:

$$E[(\sum_{i=1}^{\infty}\beta^{i-1})^2]=\frac{1}{(1-\beta)^2} \quad (33)$$
$$E[\sum_{i=1}^{\infty}\beta^{i-1}\sum_{i=1}^{\infty}\beta^{i-1}\varepsilon_{t-i}{}^{2}]=\frac{1}{(1-\beta)^2}\cdot\frac{\omega}{1-\alpha-\beta} \quad (34)$$
$$E[\sum_{i=1}^{\infty}\beta^{i-1}\sum_{i=1}^{\infty}\beta^{i-1}h_{t-i}]=\frac{1}{(1-\beta)^2}\cdot\frac{\omega}{1-\alpha-\beta} \quad (35)$$

The above derivations result in a closed-form negative expected Hessian in terms of only model parameters, taking the inverse of which, one can obtain the asymptotic variance-covariance matrix after re-scaling it with constant κ. The constant κ can either be pre-specified if *a priori* information is available or easily estimated from $\hat{\xi}_t$.

4. MONTE CARLO SIMULATION EXPERIMENTS

To evaluate how well this analytical formula works, I carry out two set of Monte Carlo simulation experiments. In the first experiment ξ_t is drawn from a normal distribution and so the estimator is MLE. In the second experiment, ξ_t is drawn from a standardized student-*t* distribution with degree of freedom 10 and this corresponds to evaluating the variance of QMLE. The sample size is fixed at T = 1000 and 6 sets of parameter values are chosen for each experiment:

$$\begin{pmatrix}\omega\\ \alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1\\ 0.05\\ 0\end{pmatrix},\begin{pmatrix}1\\ 0.10\\ 0\end{pmatrix},\begin{pmatrix}1\\ 0.05\\ 0.5\end{pmatrix},\begin{pmatrix}1\\ 0.10\\ 0.5\end{pmatrix},\begin{pmatrix}1\\ 0.05\\ 0.8\end{pmatrix},\begin{pmatrix}1\\ 0.10\\ 0.8\end{pmatrix}$$

The negative expected Hessian matrix based upon the simulations is computed by averaging the realized matrices across the simulated data paths for each set of parameter values. The variance matrix based on numerical procedure is obtained by inverting this average of realized negative Hessian. The number of simulations is set to be 100,000. Table 1 and 2 give the performance comparisons of the variance matrix for both normal and student-*t* innovations. Overall, the analytical formula works fairly well, especially when the innovation comes from a normal distribution. There is, however, a tendency of deterioration as α gets larger. Note, also, although the analytical result (15) states the fact that the innovation drawn from a fat-tailed distribution ought to give a wider confidence interval than those from a normal distribution, numerical evaluation based on the ‘sandwich’ formula somehow does not seem to produce this correction. This may be due to the fact that the finite sample difference between the inverse of negative expected Hessian (I_0^{-1}) and the expected inner product of score (C_0) is more than just the scalar κ.

5. CONCLUSION

In this paper, I derive an analytical asymptotic variance-covariance matrix for the GARCH(1,1) QMLE. The derivation relies on a local approximation and the Monte Carlo simulation experiments demonstrate that this formula works fairly well in the admissible parameters region.

References

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, Econometrica 31, 307-327.

Bollersleve, T., 1988, On the correlation structure for the generalized autoregressive conditional heteroskedastic process, Journal of Time Series Analysis 9, 121-131

Bollerslev, T. and J. Wooldridge, 1992, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, Econometric Reviews 11, 143-172.

Brooks, C., S. Burke, and G. Persand, 2001, Benchmarks and the accuracy of GARCH model estimation, International Journal of Forecasting 17, 45-56.

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

Fiorentini, G., G. Calzolari, and L. Panattoni, 1996, Analytical derivatives and the computation of GARCH estimates, Journal of Applied Econometrics 11, 399-417.

Hamilton, J., 1994, Time Series Analysis (Princeton University Press).

Harvey, A. C., 1993, Time Series Models, 2nd Edition (London, Harvester Wheatsheaf).

He, C. and T. Terasvirta, 1999, Fourth moment structure of the GARCH(*p,q*) process, Econometric Theory 15, 824-846.

Jensen, S. T., and Rahbek, A. 2004), Asymptotic inference for nonstationary GARCH, Econometric Theory, 20, 1203-1226.

Kristensen, D. and O. Linton, 2006, A closed-form estimator for the GARCH(1,1)-model, Econometric Theory 22, 323-337.

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

Lumsdaine, R. L., 1996, Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, Econometrica 64, 575-596.

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

Zivot, Eric, 2008, Practical Issues in the Analysis of GARCH Models, forthcoming in the *Handbook of Financial Time Series*, Springer-Verlag.

Appendix

1. *The Derivation of (30) and (31):*
   By standard formulas, we have:

$$\begin{aligned} E[\varepsilon_t^2 \varepsilon_{t-i}^2] &= Cov(\varepsilon_t^2, \varepsilon_{t-i}^2) + E[\varepsilon_t^2]E[\varepsilon_{t-i}^2] \\ &= \rho_i Var(\varepsilon_t^2) + (E[\varepsilon_t^2])^2 \\ &= \rho_i (E[\varepsilon_t^4] - (E[\varepsilon_t^2])^2) + (E[\varepsilon_t^2])^2 \end{aligned} \tag{A.1}$$

Plug (A.1) into (19) and Sum up the infinite geometric series to get:

$$\begin{aligned} E\left(\sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^{2}\right)^2 &= \frac{E[\varepsilon_t^4]}{1-\beta^2} + \frac{2\beta}{1-\beta^2}(E[\varepsilon_t^2 \varepsilon_{t-1}^2] + \beta E[\varepsilon_t^2 \varepsilon_{t-2}^2] + \cdots) \\ &= \frac{E[\varepsilon_t^4]}{1-\beta^2} + \frac{2\beta}{1-\beta^2} \sum_{i=1}^{\infty} \beta^{i-1} E[\varepsilon_t^2 \varepsilon_{t-i}^2] \\ &= \frac{E[\varepsilon_t^4]}{1-\beta^2} + \frac{2\beta}{1-\beta^2} \{\sum_{i=1}^{\infty} \beta^{i-1}((\alpha+\beta)^{i-1} \rho_1^{\varepsilon 2}(E[\varepsilon_t^4] - (E[\varepsilon_t^2])^2) + (E[\varepsilon_t^2])^2)\} \\ &= \frac{E[\varepsilon_t^4]}{1-\beta^2} + \frac{2\beta}{1-\beta^2} \{E[\varepsilon_t^4] \frac{\rho_1^{\varepsilon 2}}{1-(\alpha+\beta)\beta} - (E[\varepsilon_t^2])^2 \frac{\rho_1^{\varepsilon 2}}{1-(\alpha+\beta)\beta} + (E[\varepsilon_t^2])^2 \frac{1}{1-\beta}\} \end{aligned} \tag{A.2}$$

Plugging in the expressions for $E[\varepsilon_t^4]$, $\rho_1^{\varepsilon 2}$, and $E[\varepsilon_t^2]$, one can obtain (30). Likewise, (31) is derived.

2. *The Derivation of (31):*
   By MDS property of w_t, we have:

$$E[w_{t-1} \cdot h_{t-1}] = 0 \Rightarrow E[(\varepsilon_{t-1}^2 - h_{t-1})h_{t-1}] = 0 \Rightarrow E[\varepsilon_{t-1}^2 h_{t-1}] = E[h_{t-1}^2] \tag{A.3}$$

Here, notice that h_{t-1} is adapted to I_{t-2}.
Applying the law of iterative expectation, we have the following result:

$$E[w_{t-1} \cdot h_{t-i}] = 0 \Rightarrow E[(\varepsilon_{t-1}^2 - h_{t-1})h_{t-i}] = 0 \Rightarrow E[\varepsilon_{t-1}^2 h_{t-i}] = E[h_{t-1} h_{t-i}], i = 2,3\cdots \tag{A.4}$$

In the same way, we can derive:

$$E[w_{t-1}\varepsilon_{t-i}^2] = 0 \Rightarrow E[(\varepsilon_{t-1}^2 - h_{t-1})\varepsilon_{t-i}^2] = 0 \Rightarrow E[h_{t-1}\varepsilon_{t-i}^2] = E[\varepsilon_{t-1}^2 \varepsilon_{t-i}^2], i = 2,3,\cdots \tag{A.5}$$

Table 1. Comparison of the Asymptotic Variance-Covariance Matrix When Innovation is Normally Distributed

Parameter Values		Closed-Form Expression Variance Matrix			**S.E.**	Numerical Evaluation Variance Matrix			**S.E.**	Difference of S.E. in %
ω	1	0.4421	0.0000	-0.4179	**0.6649**	0.4418	-0.0003	-0.4177	**0.6647**	-0.03%
α	0.05	--	0.0010	-0.0010	**0.0315**	--	0.0014	-0.0009	**0.0370**	17.15%
β	0	--	--	0.3980	**0.6309**	--	--	0.3977	**0.6307**	-0.03%
ω	1	01222	0.0000	-0.1078	**0.3496**	0.1224	-0.0004	-0.1080	**0.3499**	0.08%
α	0.1	--	0.0010	-0.0010	**0.0313**	--	0.0017	-0.0009	**0.0412**	31.50%
β	0	--	--	0.0980	**0.3130**	--	--	0.0981	**0.3132**	0.07%
ω	1	0.7215	0.0112	-0.3347	**0.8494**	0.8180	0.0146	-0.3818	**0.9044**	6.48%
α	0.05	--	0.0009	-0.0059	**0.0300**	--	0.0013	-0.0077	**0.0354**	17.93%
β	0.5	--	--	0.1566	**0.3957**	--	--	0.1795	**0.4237**	7.06%
ω	1	0.1930	0.0054	-0.0814	**0.4393**	0.2403	0.0085	-0.1038	**0.4902**	11.57%
α	0.1	--	0.0009	-0.0031	**0.0300**	--	0.0015	-0.0047	**0.0391**	30.26%
β	0.5	--	--	0.0356	**0.1887**	--	--	0.0462	**0.2149**	13.91%
ω	1	0.4996	0.0093	-0.0837	**0.7068**	0.5741	0.0118	-0.0976	**0.7577**	7.19%
α	0.05	--	0.0005	-0.0019	**0.0230**	--	0.0007	-0.0024	**0.0266**	15.93%
β	0.8	--	--	0.0145	**0.1203**	--	--	0.0171	**0.1307**	8.59%
ω	1	0.1413	0.0038	-0.0171	**0.3759**	0.1795	0.0060	-0.0239	**0.4236**	12.71%
α	0.1	--	0.0004	-0.0008	**0.0208**	--	0.0008	-0.0013	**0.0280**	34.85%
β	0.8	--	--	0.0025	**0.0503**	--	--	0.0037	**0.0606**	20.65%

Table 2. Comparison of the Asymptotic Variance-Covariance Matrix
When Innovation is Drawn from Student-*t* Distribution with Degree of Freedom 10.

Parameter Values		Closed-Form Expression Variance Matrix			**S.E.**	Numerical Evaluation Variance Matrix			**S.E.**	Difference of S.E. in %
ω	1	0.5985	0.0000	-0.5658	**0.7737**	0.3494	-0.0004	-0.3277	**0.5911**	-23.60%
α	0.05	--	0.0014	-0.0014	**0.0367**	--	0.0014	-0.0008	**0.0374**	1.91%
β	0	--	--	0.5388	**0.7340**	--	--	0.3111	**0.5578**	-24.01%
ω	1	0.1485	0.0000	-0.1310	**0.3853**	0.0993	-0.0007	-0.0857	**0.3152**	-18.19%
α	0.1	--	0.0012	-0.0012	**0.0345**	--	0.0020	-0.0008	**0.0445**	28.99%
β	0	--	--	0.1191	**0.3450**	--	--	0.0774	**0.2782**	-19.36%
ω	1	1.0823	0.0168	-0.5021	**1.0403**	0.7902	0.0148	-0.3690	**0.8889**	-14.55%
α	0.05	--	0.0014	-0.0089	**0.0367**	--	0.0013	-0.0079	**0.0363**	-1.09%
β	0.5	--	--	0.2350	**0.4846**	--	--	0.1738	**0.4169**	-13.97%
ω	1	0.2895	0.0081	-0.1221	**0.5380**	0.2471	0.0094	-0.1070	**0.4971**	-7.60%
α	0.1	--	0.0014	-0.0047	**0.0367**	--	0.0018	-0.0052	**0.0422**	14.99%
β	0.5	--	--	0.0534	**0.2311**	--	--	0.0479	**0.2189**	-5.28%
ω	1	0.7494	0.0140	-0.1256	**0.8657**	0.5652	0.0121	-0.0964	**0.7518**	-13.16%
α	0.05	--	0.0008	-0.0029	**0.0282**	--	0.0008	-0.0025	**0.0276**	-2.13%
β	0.8	--	--	0.0218	**0.1473**	--	--	0.0170	**0.1303**	-11.54%
ω	1	0.2120	0.0057	-0.0257	**0.4604**	0.1878	0.0066	-0.0253	**0.4334**	-5.86%
α	0.1	--	0.0006	-0.0012	**0.0255**	--	0.0009	-0.0015	**0.0304**	19.22%
β	0.8	--	--	0.0038	**0.0616**	--	--	0.0040	**0.0630**	2.27%