GARCH Analysis in JMulTi

June 13, 2005

Helmut Herwartz, Christian Kascha

This chapter describes tools for modelling the volatility of a process that are implemented in the ARCH Analysis frame of JMulTi .

1 Univariate ARCH and GARCH models

The statistical tools described in this section are designed to model and forecast the conditional variance, or volatility, instead of the conditional mean of a variable. The analysis of the conditional variance may be useful for several reasons such as pricing an option or improving the estimation of forecast intervals.

The models described below assume that the conditional variance in time t depends on past errors and variances. They are designed to model time varying volatility, in particular volatility clustering - a feature often displayed by financial market series. The variance at time t is expected to be higher when past errors and variances were higher in the past and vice versa.

The phenomenon of time varying volatility is well known and generated a vast body of econometric literature following the seminal contributions by Engle (1982), Bollerslev (1986) and Taylor (1986) introducing the (generalized) *autoregressive conditionally heteroskedastic* ((G)ARCH) process and the stochastic volatility model, respectively. In the following we describe the basic features of these models and deal with the estimation and some extensions of the basic models that are available in JMulTi . Finally we describe how these models are implemented in JMulTi .

1.1 Basic features and theoretical properties

A simple parametric model allowing for time varying volatility is the ARCH(q) (Engle, 1982) process u_t with conditional variance σ_t^2:

$$
\begin{aligned}
u_t &= \xi_t \sigma_t,\ \xi_t \text{ iid } N(0,1), &(1)\\
\sigma_t^2 &= \omega + \gamma_1 u_{t-1}^2 + \gamma_2 u_{t-2}^2 + \ldots + \gamma_q u_{t-q}^2 &(2)\\
&= z_t' \theta &(3)
\end{aligned}
$$

An intrinsic property of the definition is that the second order moment of u_t is given conditional on an information set containing especially the history of the process. In the compact notation (3) $z_t = (1, u_{t-1}^2, \ldots, u_{t-q}^2)'$ and $\theta = (\omega, \gamma_1, \ldots, \gamma_q)'$. The $q+1$ vector θ collects the parameters of interest. The process is termed ARCH process since heteroskedasticity is parameterized conditionally in an autoregressive manner. Sufficient conditions for the conditional variances σ_t^2 to be positive are

$$\omega > 0,\ \gamma_i \geq 0,\ i = 1, \ldots, q.$$

The generalization of the ARCH process is the so-called generalized ARCH (GARCH) process (Bollerslev, 1986). A GARCH (q, p) process u_t can be written as

$$\begin{aligned} u_t &= \xi_t \sigma_t, \ \xi_t \text{ iid } N(0,1), & (4) \\ \sigma_t^2 &= \omega + \gamma_1 u_{t-1}^2 + \gamma_2 u_{t-2}^2 + \ldots + \gamma_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_p \sigma_{t-p}^2 & (5) \\ &= z_t' \theta & (6) \end{aligned}$$

Note that the case $p = 0$ in (5) covers the ARCH(q) process. Sufficient conditions for the conditional variances σ_t^2 to be positive are for the GARCH model

$$\omega > 0, \ \gamma_i, \beta_j \geq 0, \ i = 1, \ldots, q, \ j = 1, \ldots, p.$$

As provided in (5) the GARCH model is characterized by a symmetric response of current volatility to positive and negative lagged errors u_{t-1}. Since u_t is uncorrelated with its history it could be interpreted conveniently as a measure of news entering a financial market in time t.

To allow for different impacts of lagged positive and negative innovations threshold GARCH models have been introduced by Glosten, Jagannathan and Runkle (1993). The threshold GARCH(1,1) (TGARCH(1,1)) model takes the following form:

$$\sigma_t^2 = \omega + \gamma_1 u_{t-1}^2 + \gamma_1^- u_{t-1}^2 I_{(u_{t-1} < 0)} + \beta_1 \sigma_{t-1}^2. \qquad (7)$$

In (7), $I_{(\cdot)}$ denotes an indicator function that assumes the value 1 if the past innovation has been negative. The asymmetric effect is covered by the TGARCH model if $\gamma_1^- > 0$.

1.2 Estimation

Maximum Likelihood (ML) estimation of GARCH models faces the difficulty that available observations (u_t) or estimates ($\hat{u}_t$) are not independent. Therefore the specification of the joint density makes use of its representation as the product of some conditional and the corresponding marginal density. Let $\mathcal{U}_{T-1}$ denote the sequence of random variables containing $u_0, u_1, \ldots, u_{T-1}$. Assuming u_0 to be constant or drawn from a known distribution the joint distribution of a finite stretch of a GARCH-process is:

$$\begin{aligned} f(u_1, \ldots, u_T) &= f(u_T | \mathcal{U}_{T-1}) \cdot f(\mathcal{U}_{T-1}) \\ &= f(u_T | \mathcal{U}_{T-1}) f(u_{T-1} | \mathcal{U}_{T-2}) \cdots f(u_1 | \mathcal{U}_0) f(\mathcal{U}_0) & (8) \end{aligned}$$

The conditional distributions in (8) are available from the definition of the GARCH(q, p) process in (5). Then, the log-likelihood function is conditional on some initialization σ_0 given as:

$$\begin{aligned} l(\theta | u_1, \ldots, u_T) &= \sum_{t=1}^{T} l_t & (9) \\ &= \sum_{t=1}^{T} \left(-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \sigma_t^2 - \frac{1}{2} \frac{u_t^2}{\sigma_t^2} \right). & (10) \end{aligned}$$

The maximum likelihood estimator is the specific parameter vector $\hat{\theta}$ that maximizes the log-likelihood function. Compared to the common case with independent random variables the maximum of the likelihood function cannot be obtained analytically but requires iterative optimization routines.

A particular optimization routine which is often used to estimate the models in (2), (5) and (7) - and implemented in JMulTi - is the BHHH algorithm named after Berndt, Hall, Hall and Hausman (1974). According to this algorithm the $i-$th step estimate is obtained as

$$\hat{\theta}_i = \hat{\theta}_{i-1} + \phi \left(\sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \Big|_{\theta=\hat{\theta}_{i-1}} \right)^{-1} \sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta} \Big|_{\theta=\hat{\theta}_{i-1}}. \tag{11}$$

where $\phi > 0$ is used to modify the step length.

Under regularity conditions (Davidson, 2000) the ML-estimator $\hat{\theta}$ converges at rate $\sqrt{T}$ and is asymptotically normally distributed, i.e.

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, S^{-1}), \tag{12}$$

where S is the expectation of the outer product of the scores of $l_t(\theta)$,

$$S = \frac{1}{T} \sum_{t=1}^{T} E \left[\frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right].$$

The log-likelihood function in (10) is determined under the assumption of conditional normality stated in (1). Ignoring non normality of innovations ξ_t will result in a misspecification of the log likelihood function. Maximizing the misspecified Gaussian log-likelihood function is, however, justified by quasi maximum likelihood theory. For a wide variety of strictly stationary GARCH processes consistency and asymptotic normality of the QML estimator have been shown (Bollerslev and Wooldridge, 1992; Lumsdaine, 1996; Lee and Hansen, 1994). If the normality assumption is violated the covariance matrix of the QML estimator is

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, D^{-1} S D^{-1}), \tag{13}$$

where D is the negative expectation of the matrix of second order derivatives

$$D = \frac{1}{T} \sum_{t=1}^{T} -E \left[\frac{\partial^2 l_t}{\partial \theta \partial \theta'} \right]. \tag{14}$$

Analytical expressions for the derivatives necessary to implementing the BHHH algorithm or (Q)ML inference are given in Bollerslev (1986) for the general case of ARMA(p, q) processes with GARCH(q, p) error terms.

1.3 Extensions

1.3.1 Conditional Leptokurtosis

As it is often argued in empirical contributions the normal distribution specified in (1) is rarely supported by real data. If an alternative parametric distribution can reasonably be assumed exact ML methods may outperform QML estimation in terms of efficiency. On the contrary ML estimation under misspecification of the (non Gaussian) conditional distribution may yield inconsistent parameter estimates (Newey and Steigerwald, 1997).

Moreover, if the normality assumption is violated it is no longer possible to provide valid forecasting intervals for u_{t+h} given Ω_t by means of quantiles of the Gaussian distribution. To improve forecast intervals it pays to consider a leptokurtic distribution of ξ_t. Therefore GARCH models under the conditional $t-$distribution and the generalized error distribution (GED) are implemented in JMulTi . We state here for convenience the density functions of the $t-$distribution and the GED:

t-distribution A random variable u_t is $t-$distributed with v degrees of freedom, mean zero and variance σ_t^2 if it has the following density:

$$f(u_t|v) = \frac{v^{v/2}\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi}\,\Gamma(\frac{v}{2})\sqrt{\frac{(v-2)\sigma_t^2}{v}}}\left(v + \frac{v \cdot u_t^2}{(v-2)\sigma_t^2}\right)^{-(\frac{v+1}{2})}. \qquad (15)$$

In (15) $\Gamma(.)$ denotes the Gamma function. Recall that for $v \to \infty$ the density in (15) coincides with the Gaussian density.

generalized error distribution According to this distribution with shape parameter v, a zero mean random variable u_t with variance σ_t^2 has the following density:

$$f(u_t|\theta, v) = v \exp\left(-\frac{1}{2}\left|\frac{u_t}{\lambda \cdot \sigma_t}\right|^v\right)\left[2^{\frac{v+1}{v}}\Gamma\left(\frac{1}{v}\right)\lambda \cdot \sigma_t\right]^{-1}, \qquad (16)$$

where λ is defined as

$$\lambda = \left[\frac{\Gamma(\frac{1}{v})}{2^{\frac{2}{v}}\Gamma(\frac{3}{v})}\right]^{0.5}. \qquad (17)$$

In case $v = 2$ the density in (16) is equal to the $N(0, \sigma_t^2)$ density, the distribution becomes leptokurtic if $v < 2$. For $v = 1$ ($v \to \infty$) the GED coincides with the (approximates the) double exponential (rectangular) distribution (Harvey, 1990).

1.4 Implementation and Specification in JMulTi

In JMulTi the basic ARCH(q), GARCH(q, p) and TGARCH(q, p) models given in (2), (5) and (7) can be estimated up to orders $q = 5$ and $q = p = 2$, respectively. Maximum likelihood estimation can be performed under the assumptions of conditional normality (1), a conditional $t-$distribution (15) or a conditional GED (16). Consequently one has to specify the basic model, the lag-lengths and the assumed conditional distribution.

The output consists of the parameter values (where `gamma` denotes the parameters of lagged errors and `beta` denotes the parameters of lagged variances), the respective t-values, the variance-covariance matrix and the value of the log likelihood function. Additionally the estimated residuals $\hat{\xi}_t$ can be analyzed.

The empirical mean of the outer product of the log-likelihood scores (12) is used to estimate the covariance matrix of the ML estimates (Davidson, 2000). The BHHH in (11) is implemented using analytical derivatives of the corresponding likelihood functions. Derivatives of the Gamma function, $\Gamma'(.)$, are evaluated numerically. Depending on the order of the process ARCH-parameters (γ_1, γ_1^-) are initialized with small values whereas initial GARCH parameters are larger without violating the condition for covariance stationarity. For a given initialization of these parameters the initial deterministic variance component is obtained using its link to the unconditional expectation $E[u_t^2]$. At each iteration the BHHH algorithm compares alternative parameters given by the step lengths. Different step lengths, ϕ, are employed to economize on computing time.

2 Multivariate GARCH(1,1)

Multivariate GARCH models are a conceptually straighforward generalization of univariate models. Problems stem from the fact that a very large parameter space is involved, posing analytical and computational problems. The representation being used in JMulTi is the BEKK form ((Baba, Engle, Kraft and Kroner, 1990)) in its simplest form with $N = p = q = 1$. The estimation is implemented for a GARCH$(1, 1)$ model, see ((Herwartz, 2004)) for a more detailled description.

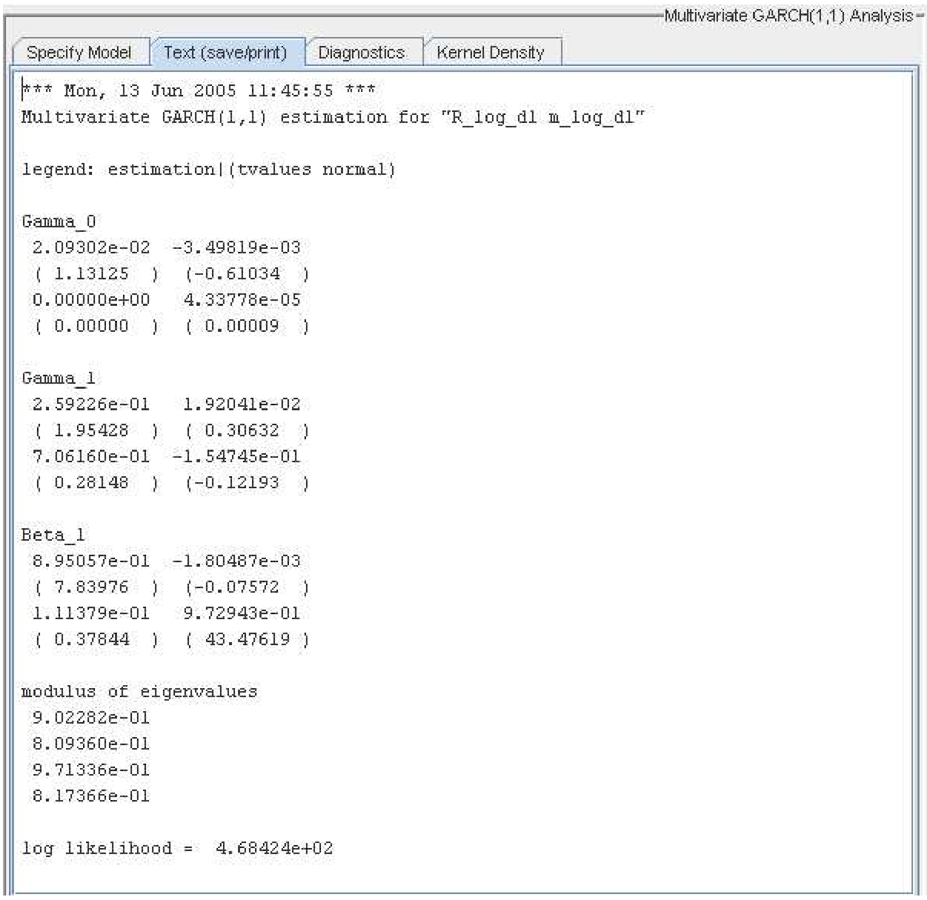

Figure 1: Screenshot of output for multivariate GARCH$(1, 1)$ estimation

Multivariate GARCH$(1, 1)$ models can be specified for dimensions of 2, 3, and 4 variables. The model is estimated with a quasi maximum likelihood (QML) estimator under normality assumption. As an option it is possible to compute the exact QML t-ratios which require to evaluate the 1st and 2nd order derivatives of the Gaussian likelihood function analytically. The latter is left as an option to the user because the computation might be quite time consuming.

- **Input** - The procedure is defined in terms of the second moments of a serially uncorrelated but conditionally heteroskedastic K-dimensional vector $u_t = (u_{1,t}, u_{2,t}, ..., u_{K,t})$ that is normally distributed $u_t|\Omega_{t-1} \sim N(0, \Sigma_t)$. Input variables are taken for realizations of u_t and should satisfy those assumptions.

- **Estimation** - The output (Figure 1) consists of the parameter values (where `gamma` denotes the parameter matrices of lagged errors and `beta` denotes the parameter matrix of lagged variances), the respective t-values, and the value of the log likelihood function. Furthermore, the modulus of the eigenvalues of the polynomial describing the unconditional mean of the covariance process are given to check for covariance stationarity.

- **Diagnostics** - To check whether the residuals meet the required assumptions, a number of multivariate diagnostic tests can be applied. Available are the Portmanteau test, the ARCH-LM test, plots of the AC and PAC functions of the residuals, Jarque-Bera tests for nonnormality, as well as plots of the estimated standard deviation process. It is also possible to plot estimated univariate GARCH$(1, 1)$ processes together with the multivariate variance processes to analyse the differences.

- **Kernel Density Estimation** - Kernel density estimates can be done for each of the estimated residual series $\hat{\xi}_t$.

References

Baba, Y., Engle, R., Kraft, D. and Kroner, K. (1990). Multivariate simultaneous generalized ARCH, *mimeo*, UCSD.

Berndt, E., Hall, B., Hall, R. and Hausman, J. (1974). Estimation and inference in nonlinear structural models, *Annals of Economic and Social Measurement* **3/4**: 653–665.

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

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

Davidson, J. (2000). *Econometric theory*, Blackwell publishers, Oxford.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdoms inflations, *Econometrica* **50**: 987–1007.

Glosten, L., Jagannathan, R. and Runkle, D. (1993). Relationship between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**: 1779–1801.

Harvey, A. C. (1990). *The econometric analysis of time series*, 2nd edn, Philip Allan, Hemel Hempstead.

Herwartz, H. (2004). Conditional heteroskedasticity, *in* H. Lütkepohl and M. Krätzig (eds), *Applied Time Series Econometrics*, Cambridge University Press.

Lee, S. and Hansen, B. (1994). Asymtotic theory for the GARCH(1,1) quasi maximum likelihood estimator, *Econometric Theory* **10**: 29–52.

Lumsdaine, R. (1996). Consistency and asymptotic normality of the quasi maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, *Econometrics* **64**: 575–596.

Newey, W. and Steigerwald, D. (1997). Asymptotic bias for quasi maximum likelihood estimators in conditional heteroskedasticity models, *Econometrica* **65**: 587–599.

Taylor, S. J. (1986). *Modelling financial time series*, John Wiley & Sons, Chichester.