Analysis Skewness in GARCH model

Ziqi Zhang

Analysis skewness in GARCH model

Ziqi Zhang

Supervisor: Chang-li He

School of Economics and Social Sciences, Högskolan Dalarna,

SE-781 88, Borlänge, Sweden

D-Level Thesis

June, 2009

Abstract

In this thesis, the model is built under the skewed-type GARCH models. The assumption of distribution in GARCH model are Normal Distribution, Student-t Distribution, Generalized Error Distribution and there skewed version. Compare these models and find a best one to analysis the data in this case. The data is from DAX index (Germany stock market), which not obey the Normal Distribution and skewness exist. In order to capture the feature of data, the conditional error distribution of GARCH model has been changed, from Standard Normal Distribution to Skewed Distribution.

Key words: GARCH model, Skewed Distribution

1 Introduction

Skewness, non-symmetry in distribution, becomes important in economic variables such as stock index return and exchange rate. The phenomenon of skewness has been recognized in the empirical financial literature for several years and it is well-know the stock return distributions exhibit negative skweness and kurtosis. Many people do the research about this departure from normality of return distribution. Black (1976) and Christie (1982) introduce the leverage effect to stochastic bubbles. Campbell and Hentschel (1992) explain the ranging from the so-called volatility feedback effect. Campbell R. Harvey and Akhtar Siddique (1999) published the Autoregressive Conditional Skewness on Journal of Financial and Quantitative Analysis.

The Generalized Autoregressive Conditional Heteroscedasticity model (GARCH model), introduced by Engle (1982) and Bollerslev (1986), and allowed for time-varying volatility and much more flexible lag structure compare with ARCH model. In our knowledge, the conditional error distribution is Gaussian distribution; however, there are three basic assumptions of the conditional error distribution: the Gaussian distribution; Student-t Distribution; Generalized Error Distribution. Until now, GARCH model is one of the powerful tools to analysis the financial time series data.

In this thesis, there is an introduction about GARCH model and two extra assumptions of conditional error distribution: Student-t Distribution and Generalized Error Distribution.

In the data, I choose the Germany stock market data (DAX index) from 2 January 2006 to 27 April 2009. Analysis the basic information of data first, the data have stylized fact after we print the log-return. Compare with the density of Normal Distribution, the result is the data which from DAX index is not obey the Normal Distribution and it dependent with it own past. I built the conditional mean equation with AR(1) model first, then, use GARCH model to analysis the conditional variance, use the value of AIC and log likelihood to select the model which will describe the DAX index best.

2 Method

2.1 The GARCH (p, q) process

One of the traditional and classical assumptions of conventional time series and econometric models is constant variance. The ARCH (Autoregressive Conditional Heteroskedastic) process introduced by Engle (1982) changed this assumption, is the revolution which reveal the connection between conditional and the unconditional second order moment. The ARCH process allowed the conditional variance to change over time as a function of past errors leaving the unconditional variance constant; it is the first model that provides a systematic form for volatility modeling.

After Engle propose ARCH model, this type of models has a widely used in modeling economic phenomena and financial time series. However, after ARCH process is appropriated, people found out there are lots of weakness part in this new type of models, like the long lag length, a large number of parameters and it not easy to control the exist of negative variance, in order to solve this, Bollerslev (1986) proposed the generalized ARCH, GARCH (Generalized Autoregressive Conditional Heteroscedasticity) model.

Let denote a stochastic process which is a real-value and discrete ε_t, the GARCH (p, q) process is

$$\varepsilon_t = z_t \sigma_t, \quad z_t \sim D_\nu(0,1) \tag{2.1.1}$$

$$\sigma_t^2 = \varpi + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.1.2}$$

Where,

$$p \geq 0, q > 0$$

$$\alpha_0 > 0, \alpha_i \geq 0, i = 1, \ldots q$$

$$\beta_i \geq 0, i = 1, \ldots p$$

In above equation, $D_\nu(0,1)$ is the probability density function of the residual with zero mean and unit variance, ν are additional distribution parameters to describe the skew and shape parameters of the distribution. When p=0, the process becomes to the ARCH (q) process, in this process, the conditional variance as linear function of it own past only. While, when p=q=0, ε_t is white noise. We generate the residual from mean equation and we call above process as variance equation.

The mean equation of an univariate time series x_t can be described as:

$$x_t = E(x_t \,|\, \Omega_{t-1}) + \varepsilon_t \tag{2.1.3}$$

Where $E(\cdot|\cdot)$ denote the conditional expectation, Ω_{t-1} is the information set at time t-1, ε_t is residuals of the time series. In return series, if the mean is small and the return is uncorrelated, we can use a simple, which return is residual to be the mean equation. The most common way to model the mean equation is use ARMA mean equation.

2.2 Skewed Distribution

Fernandez and Steel proposed Skewed Distribution in 1998, which allows skewness in any continuous and symmetric distribution by changing the scale at each side of the model

$$f\left(z|\xi\right) = \frac{2}{\xi + \frac{1}{\xi}}[f(\xi z)H(-z) + f(\frac{z}{\xi})H(z)] \tag{2.2.1}$$

In above model, ξ is a shape parameter which is positive and describes the degree of asymmetry. $\xi = 1$ is the symmetric distribution with $f(z|\xi = 1) = f(z)$. $H(z) = (1 + sign(z))/2$ is the Heaviside unit step function. The mean and variance of $f(z|\xi)$ depend on ξ are given by

$$\mu_\xi = M_1(\xi - \frac{1}{\xi})$$

$$\sigma_\xi^2 = (M_2 - M_1^2)(\xi^2 + \frac{1}{\xi^2}) + 2M_1^2 - M_2$$

$$M_r = 2\int_0^\infty x^r f(x)dx$$

M_r is the r-th moment of $f(x)$ on the positive real line. If a skewed distribution functions with zero mean and unit variance, we call it Standardized Skewed Distribution.

The probability function $f(z|\xi)$ of a Standardized Skewed Distribution can be wrote as

$$f(z|\xi\theta) = \frac{2\sigma}{\xi + \frac{1}{\xi}} f(z_{u_\xi \sigma_\xi}|\theta) \qquad (2.2.2)$$

Where $z_{\mu_\xi \sigma_\xi} = \xi^{sign(\sigma_\xi z + u\xi)}(\sigma_\xi z + u_\xi)$, $f()$ is any standardized symmetric distribution function, like the Standard Student-t Distribution, the Standard Generalized Error Distribution or Standard Normal Distribution. In above equation, μ is location parameter, σ is the standard deviation parameter, ξ is a shape parameter which model the skewness, θ is an optional set of shape parameter that models higher moments of even order like ν in GED and Student-t Distribution.

The probability density function of the Skewed-Normal Distribution with parameter α is given by

$$f(x) = 2\phi(x)\Phi(\xi x), \qquad (2.2.3)$$

Where ξ is a fixed parameter,

$$\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$$

$$\Phi(\xi x) = \int_{-\infty}^{\xi x} \phi(t)dt$$

The fixed parameter ξ is called the shape parameter, it control the different shape of density function.

Base on the Standard Skewed Distribution, the Skewed-Student-t Distribution can be writing down in above way:

$$f(z_t|\xi,\nu) = \begin{cases} \frac{2\sigma}{\xi + \xi^{-1}} g[\xi(sz_t + \mu)|\nu] \cdots\cdots z_t < -\frac{\mu}{\sigma} \\ \frac{2\sigma}{\xi + \xi^{-1}} g[(sz_t + \mu)/\xi|\nu] \cdots\cdots z_t \geq -\frac{\mu}{\sigma} \end{cases} \qquad (2.2.4)$$

$z_t \sim SKST(0,1,\xi,\nu)$, $g(\cdot|\nu)$ is the density function of Standard Student-t Distribution. ξ is non-symmetry parameter, ν test the fat tail, μ and

σ^2 is the mean and variance of non-Standard Student-t Distribution.

$$\mu = \frac{\Gamma(\frac{v-1}{2})\sqrt{v-2}}{\sqrt{\pi}\Gamma(\frac{v}{2})}(\xi - \frac{1}{\xi})$$

$$\sigma^2 = (\xi^2 + \frac{1}{\xi^2} - 1) - \mu^2$$

The density function of the Standardized Skewed-Generalized Error Distribution is

$$f(z_t \,|\, v, \xi) = v(2\theta \cdot \Gamma(1/v))^{-1} \cdot \exp(-\frac{|z_t - \delta|^v}{[1 - sign(z_t - \delta)\xi]^v \theta^v}) \qquad (2.2.5)$$

Where,

$$\theta = \Gamma(1/v)^{0.5}\Gamma(3/v)^{-0.5}S(\xi)^{-1}$$

$$\delta = 2\xi \cdot \mathrm{A}S(\xi)^{-1}$$

$$S(\xi) = \sqrt{1 + 3\xi^2 - 4\mathrm{A}^2\xi^2}$$

$$\mathrm{A} = \Gamma(2/v)\Gamma(1/v)^{-0.5}\Gamma(3/v)^{-0.5}$$

Where the shape parameter v control the height and fat-tail of the density function with constraint $v > 0$, while ξ is a skewness parameter of the density with $-1 < \xi < 1$. In the case of positive skewness, the density function skews toward to the right, vice versa. Sign is the sign function. When $v = 2$, λ =0, the SGED distribution turns out to be the standard normal distribution.

3 Data Analysis

The data used in this thesis is DAX index (Germany stock market), which is daily closing values of the index. P_t as the closing value from the period 2 January 2006 to 27 April 2009.

Usually, people do not analysis P_t directly, use return instead of daily close value. Campbell, Lo and Mackinlay give two mean reasons for using returns. First, for average investors, return of an asset is a complete and scale-free summary of the investment opportunity. Second, return series are easier to handle than price series because the former have more attractive statistic properties.

Figure 1 is the line graph of daily data, from that figure; we can only know the basic information of data, for instead, there are two peaks during this period.

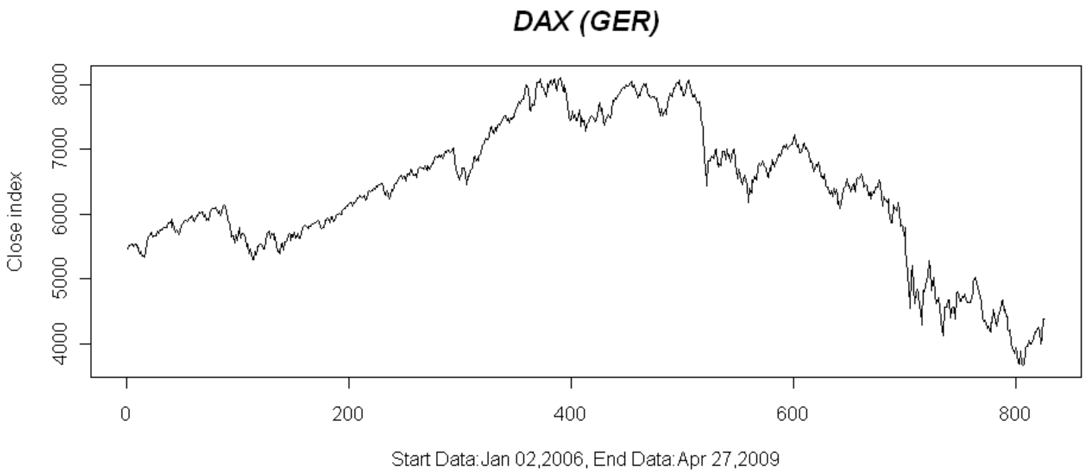

Figure 1.The daily closing value of DAX

Instead of that it is common to focus on log return in order to capture the feature of the data, R_t as the return of closing value. Use log transform of the ratio of return and it own past.

$$R_t = \log(P_t / P_{t-1}) = \log P_t - \log P_{t-1}$$

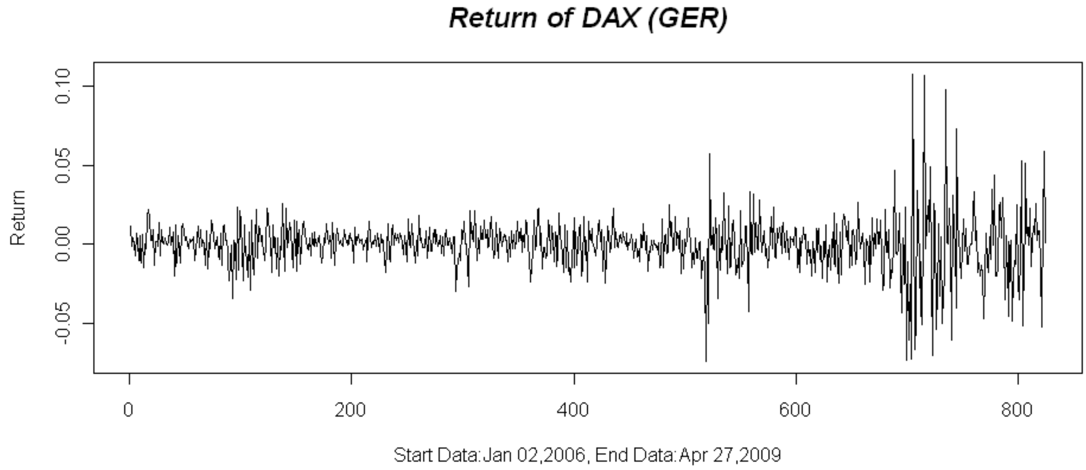

Figure 2. Return of DAX

As Mandelbrot (1963) wrote: a large changes tends to be followed by large changes, of either sign, and small changes tend to be followed by small changes. It is clear from visual inspection of Figure 2, the returns are not i.i.d through time, volatility clustering phenomenon is immediately apparent.

The summary statistics are presented in Table 1, the value of mean is -0.0002659, the skewness exist, the kurtosis is larger than 3, base on above, the result is clearly showing the data have non-normality and asymmetry's feature, the value of Jarque Brea test is 2380.356, p-value is small, reject the null hypothesis of normal distribution.

Table 1. summary statistics for log return on DAX index: standard deviation, mean, median, skewness, kurtosis and JB test (the null hypothesis of JB test is normality)

Quantity	**Value**
Mean	-0.0002659
Std Dev	0.01673975
Skewness	0.2643346
Kurtosis	8.2772
JarqueBera	2380.356
(P-value)	(0.0000)

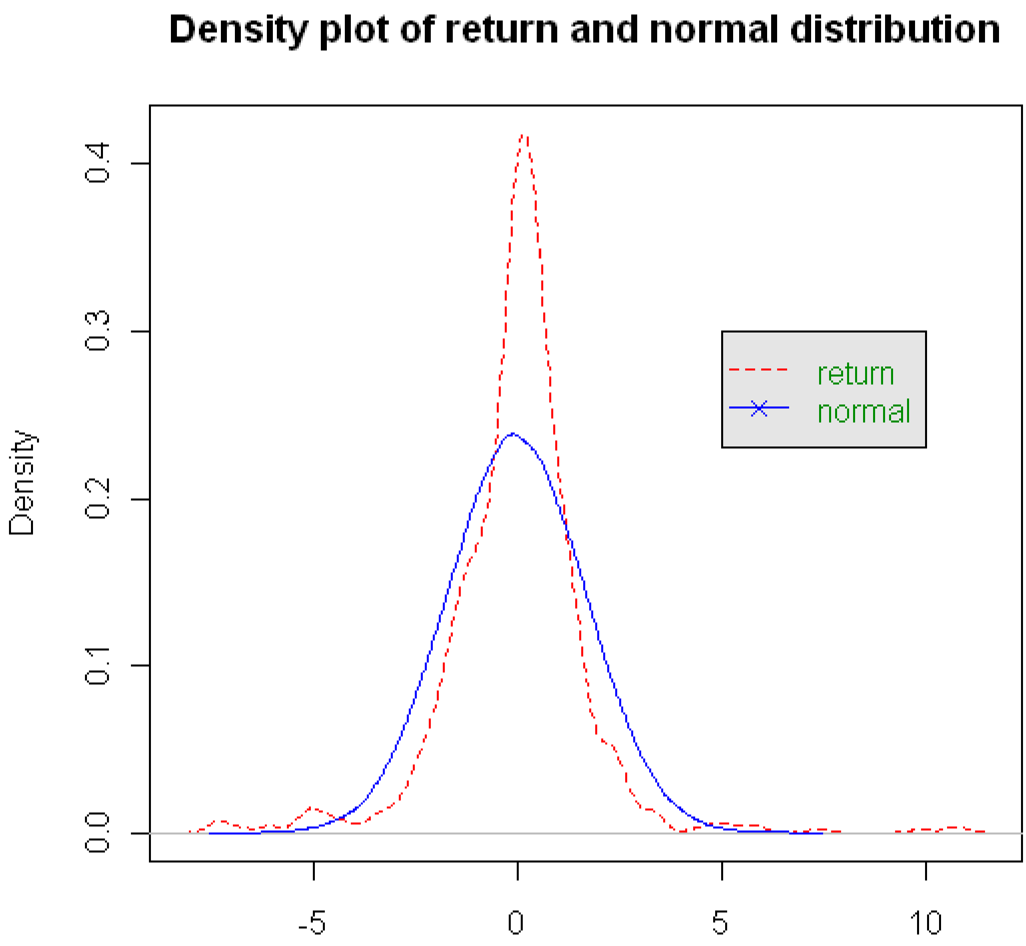

Figure 3. Compare the density of return and normal distribution

(times 100 to the return)

From the Figure 3, there is evidently different between return and the date from normal distribution, these two densities have different shape, tails and peak. Return is thinner, has a high peak and non-symmetry compare with the density which from the normal distribution; obviously, it is not the normal distribution.

From on above analysis, in sample period, the returns have volatility clustering and non-symmetry. We use ARMA mean equation to fit the return. After compare with the significant of parameters and AIC criterion, the AR (1) which removes the intercept model has better ability to fit the return.

Analysis the residual from above model, the residual is independent, which by Box-Ljung test at lag equal to one, we accept the null hypothesis, the series is independent, and the square of residual is dependent, it is possible to use GARCH model. However, the distribution of return is non-symmetry; the residual which generated from AR (1) is different with normal distribution, which we can see from Figure 4. The skewness and kurtosis exist in residual series; we can not use standard GARCH to be the variance equation.

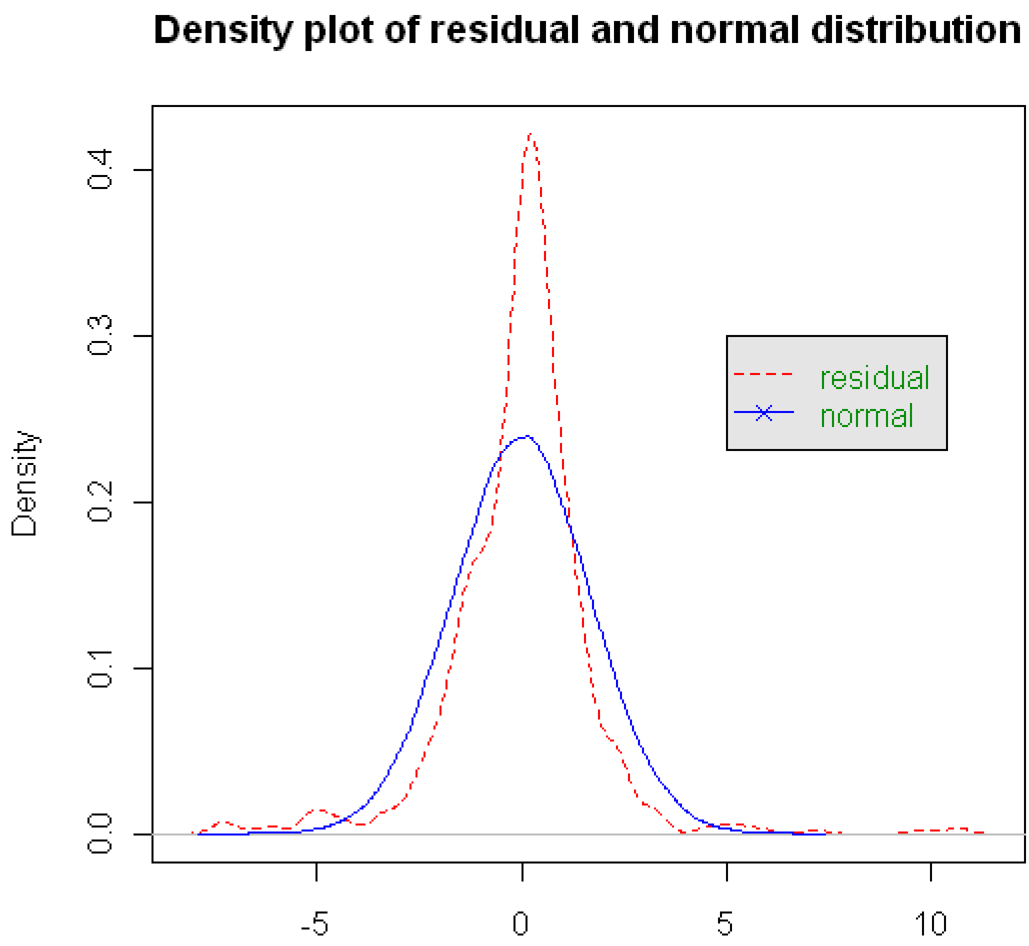

Figure 4. Density plot of residual from AR (1) and normal distribution

In order to capture skewness, asymmetry in distribution, one way is modelling series with GARCH-type models which having a skewed marginal distribution; another way is apply a class of GARCH models allowing asymmetries or nonlinearities in the first and the second conditional moment, whereas the marginal distribution is symmetry. The method which used in this thesis is using a skewed marginal distribution in GARCH-type models.

As far as we know, we usually use standard GARCH, which the distribution is Normal Distribution, there are another two conditional distribution are quite often used, which are Student -t-Distribution and GED. In order to capture the skewness, there is skewed versions distribution. In this thesis, the model is built under the skew-type GARCH models.

The order of GARCH model is p=q=1, GARCH (1, 1) model, the conditional distribution of GARCH model is skewed version of Normal Distribution, Student-t Distribution, GED. We select the residual from AR (1), make the innovation of residual with GARCH model. The result of comparison is below:

Table 2: The parameters from three models and the significant level is 1% level

Ljung-Box* is the test for the square of residual

	SGED	P-value	SNORM	P-value	SSTD	P-value
mu	7.520e-04	0.03209	6.483e-04	0.0668	7.341e-04	0.0444
ar 1	-7.134e-02	0.00468	-8.806e-02	0.0161	-6.998e-02	0.0516
omega	2.474e-06	0.05158	2.607e-06	0.0138	2.134e-06	0.0753
alpha	1.176e-01	1.07e-05	1.270e-01	0	1.157e-01	0
beta	8.773e-01	0	8.680e-01	0	8.818e-01	0
skew	8.825e-01	0	8.339e-01	0	8.750e-01	0
shape	1.303e+00	0	-----	-----	7.129e+00	0
Log likelihood	2444.845	-----	2418.282	-----	2441.025	------
AIC	-5.9099	-----	-5.8504	-----	-5.9006	------
ARCH-LM Test	13.2525	0.3509	13.4094	0.3399	13.1742	0.3565
Ljung-Box Q(10)	12.1734	0.2736	11.3307	0.3323	12.0687	0.2805
Ljung-Box Q(20)	20.7081	0.4145	19.6541	0.4797	20.6656	0.4170
Ljung-Box Q(10)*	7.1754	0.7088	7.1259	0.7135	7.20566	0.7059

Form the p-value of these models, the parameters are significant at 10% level, all the ARCH-LM tests accept the null hypothesis, which mean the heteroskedasticity is removed and these models are fitted with data. The value of Box-Ljung test is located in the accept region, which is the data is independent, the residual which from GARCH model is independent and the Box-Ljung test about the square of residual is not significant either, this phenomenon show the effect of GARCH model, make a good innovation of residual.

Log likelihood and AIC are the two criterions to select which model will be the final one to use, the biggest log likelihood value of these is 2444.845 and the smallest AIC is -5.9099, both of them from Skewed Generalized Error Distribution, can fit the data better than other two models. However, there is not a big different between the Skewed Student-t Distribution and Skewed Generalized Error Distribution, one reason of that is both of them have shape and skew parameters.

In all, the result of comparison is the conditional variance function which will be used in GARCH (1, 1) is Skewed–Generalized Error Distribution.

4 Conclusions

This thesis is basic on the theory that in the assumption of GARCH model, there has different conditional error distribution, Student-t Distribution, the Normal Distribution and Generalized Error Distribution and their own skewed version.

After analysis the data, it is clearly to see from density plot that the DAX is not like normal distribution, it has scenes and high kurtosis, after plot the daily return, the data is not independent, I use ARMA model to generate the residual which is independent, stationary and the square of residual is dependent, the ARCH-LM test prove the GARCH effect exist in the residual, however, if I use standard GARCH to analysis the data which is not submit the normal distribution, the information will wrong if we do the further analysis.

I built three GARCH models which have different conditional distribution: Skewed Normal Distribution, Skewed Student-t Distribution, Skewed Generalized Error Distribution. After compare with the three different conditional error distributions and base on the log likelihood, AIC criterion and the predict ability, the model which is the best for DAX in this thesis is Skewed Generalized Error Distribution.

Reference

Bollerslev T. 1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31: 307-327

Engle R. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica.50:987-1008

Anders Wilhelmsson. 2006. Garch Forecasting Performance under Different Distribution Assumption. Journal of Forecasting 25: 561-578

Campbell R.Harvey and Akhtar Siddique. Autoregressive Conditional Skewness. Journal of Financial and Quantitatice Analysis Dec 1999 465-487

Ruey S. Tsay Analysis of Financial Time Series Second Edition

James D.Hamilton Time Series Analysis

Nelson D.B., (1991); Conditional Heteroscedasticity in Asset Returns: A New Approach, Econometrica, 59, 347–370.

Lambert P. Laurent S (2001); Modelling Financial Time Series Using GARCH-Type Model and a Skewed Student Density, Mimeo University de Liege.

Högskolan Dalarna
791 88 Falun
Tel 023-77 80 00

Rapport 200x:nr
ISBN
ISSN