Realized Volatility and Modeling Stock Returns as a Mixture of Normal Random Variables: the GARCH-Skew-t Model

Young Il Kim
The Ohio State University

Econometric Society
2008 North American Summer Meeting
GARCH-skew-t modeling

Main issue:

– Equity returns show three stylized facts: volatility clustering, negative skewness and leptokurtosis

– This paper provides a new empirical guidance for modeling a skewed and fat-tailed error distribution along with GARCH effects for equity returns based on empirical findings on Realized Volatility
GARCH-skew-t modeling

Stylized facts for equity returns

- Volatility clustering; 2nd moment
- Negative skewness; 3rd moment
- Excess kurtosis (thick tails); 4th moment
GARCH-skew-t modeling

ARCH/GARCH class of models
- Successful at capturing volatility clustering
  Eg) GARCH(1,1) with Gaussian error

GARCH models with Gaussian error
- Unable to account for asymmetric mass in tail parts of distribution for equity returns
GARCH-skew-t modeling

GARCH(1,1) standardized residuals for monthly NYSE index returns
GARCH-skew-t modeling

Overcoming the limitation of traditional GARCH modeling
- Explore alternative error distributions underlying GARCH models
  - Symmetric Stable distribution by McCulloch (1985)
  - Student t by Bollerslev (1987)
  - (Many others)

This line of approach
- Mainly empirical and pragmatic (Forsberg & Bollerslev, 2002)
- But where does the (asymmetric) non-Gaussian distribution come from?
GARCH-skew-t modeling

New approach in this essay:
The Mixture-of-Distributions Hypothesis (MDH); Clark (1973)
+ Stochastic nature of volatility (volatility risk)
(with variance-mean mixture framework for the asymmetric distribution)
\[ \text{The distribution of returns can be derived as a mixture of the two distributions: Normal + the distribution of volatility} \]

Note
- MDH: Distribution of returns conditional on the rate of information arrival to the market is Gaussian
- Normality of returns conditional on volatility
- The Realized Volatility (RV) is employed as a proxy for the stochastic variance (Schwert 1989, Campbell, Lettau, Malkiel & Xu 2001)

\[ RV_t = \sum_{i=1}^{N_t} (r_{it} - \bar{r}_t)^2 \]
GARCH-skew-t modeling

Asymmetric non-Gaussian distribution of returns is derived as a mixture of the following two distributions:

- Distribution of returns conditional on $RV$: Normal
  \[ r_t \mid RV_t \sim N(\mu + \beta RV_t, RV_t) \]

- Distribution of $RV$: Inverted Chi-square
  \[ RV_t^{-1} \sim \frac{1}{h^2 \nu} \chi_\nu^2 \]

\[ \Rightarrow \text{Distribution of returns:} \]
\[ r \sim f_{\tilde{r}}(r) = \int f_{\tilde{r} \mid RV}(r \mid RV)f_{RV}(RV)dRV \]
\[ r_t \mid I_{t-1} \sim f_{\tilde{r}}(r_t \mid I_{t-1}) = \int f_{\tilde{r} \mid RV}(r_t \mid RV_t, I_{t-1})f_{RV}(RV_t \mid I_{t-1})dRV \]
GARCH-skew-t modeling

The unconditional mixture distribution is derived by
Aas & Haff (2006): as a special limiting case of the GH distribution
Kim & McCulloch (2007): as a normal inverted-chi-square mixture distribution

With skewness ($\beta \neq 0$)

$$f_{\tilde{r}}(r_t \mid \mu, h, \nu, \beta) = \frac{h^\nu \sqrt{\nu/2} |\beta|^{(\nu+1)/2} K_{(\nu+1)/2} (\sqrt{r_t - \mu}) \sqrt{\pi \Gamma(\nu/2)} 2^{(\nu+1)/2-1} \left(\sqrt{(r_t - \mu)^2 + h^2 \nu}\right)}{\sqrt{(r_t - \mu)^2 + h^2 \nu}} \exp(\beta (r_t - \mu))$$

Without skewness ($\beta = 0$)

$$f_{\tilde{r}}(r_t \mid \mu, h, \nu) = \frac{1}{h} t_{\nu} \left(\frac{r_t - \mu}{h}\right)$$

where $K_{\nu}(\cdot)$ is the modified Bessel function of the second type of order $\nu$
GARCH-skew-t modeling

The conditional mixture distribution is derived by Kim & McCulloch (2007): as a mixture of the following two distributions

\[ r_t \mid RV_t, I_{t-1} \sim N(\mu + \beta RV_t, RV_t) \]

\[ RV_t^{-1} \mid I_{t-1} \sim \frac{1}{h_t^2} \chi^2_{\nu} \]

\[ h_t^2 = h(I_{t-1}; \varphi)^2 = c + bh_{t-1}^2 + ae_{t-1}^2 \]

The conditional skew-t model is easily cast into the common form of GARCH(1,1) model as below (Kim & McCulloch 2007).

\[ r_t = \mu + h_t \varepsilon_t \]

\[ h_t^2 = h(I_{t-1}; \varphi) = c + bh_{t-1}^2 + ae_{t-1}^2 \]

\[ \varepsilon_t \sim \text{Skewt} \left( \mu^*, h^*, \nu, \bar{\beta} \right) \]

\[ E[\varepsilon_t] = 0 \quad \text{Var}[\varepsilon_t] = 1 \]
GARCH-skew-t modeling

Parameter estimates with standard errors in parentheses

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.00469</td>
<td>0.0140</td>
<td>0.0183</td>
</tr>
<tr>
<td></td>
<td>(0.00173)</td>
<td>(0.00105)</td>
<td>(0.00124)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-</td>
<td>0</td>
<td>-6.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.02)</td>
</tr>
<tr>
<td>$h$</td>
<td>0.0537</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.00129)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model 1 (naïve): $r_t = \mu + hZ_t$, $Z_t \sim iidN(0,1)$

Model 2: $r_t = \mu + RV_{t}^{1/2}Z_t$, $Z_t \sim iidN(0,1)$

Model 3: $r_t = \mu + \beta RV_t + RV_t^{1/2}Z_t$, $Z_t \sim iidN(0,1)$
### GARCH-skew-t modeling

Summary statistics of residuals and Jarque-Bera test statistics

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0</td>
<td>-0.125</td>
<td>-0.0327</td>
</tr>
<tr>
<td>Std</td>
<td>1</td>
<td>1.23</td>
<td>1.21</td>
</tr>
<tr>
<td>Skw</td>
<td>-0.517</td>
<td>0.137</td>
<td>0.0428</td>
</tr>
<tr>
<td>Kts</td>
<td>10.5</td>
<td>2.83</td>
<td>2.88</td>
</tr>
<tr>
<td>JB</td>
<td>2257.6</td>
<td>4.29</td>
<td>0.974</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.117)</td>
<td>(0.615)</td>
</tr>
</tbody>
</table>
GARCH-skew-t modeling

Gaussian PP-plots of normalized returns for each model
GARCH-skew-t modeling

The substantially negative skewness is reflected in a significantly negative correlation between returns and $RV$

Regression of residuals on the realized standard deviations

<table>
<thead>
<tr>
<th></th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(r_t - \mu)/\sqrt{RV_t}$</td>
<td>0.122</td>
<td>-0.102</td>
</tr>
<tr>
<td></td>
<td>(0.694)</td>
<td>(0.0690)</td>
</tr>
<tr>
<td>$\sqrt{RV_t}$</td>
<td>-6.42</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>(1.49)</td>
<td>(1.48)</td>
</tr>
</tbody>
</table>
### GARCH-skew-t modeling

Ljung-Box portmanteau test statistics for residuals and squared residuals

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q(1)$</td>
<td>8.25</td>
<td>0.143</td>
<td>0.155</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.00408)</td>
<td>(0.705)</td>
<td>(0.694)</td>
</tr>
<tr>
<td>$Q(10)$</td>
<td>33.3</td>
<td>16.0</td>
<td>23.1</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000240)</td>
<td>(0.0994)</td>
<td>(0.0105)</td>
</tr>
<tr>
<td>$Q^2(1)$</td>
<td>49.9</td>
<td>5.24</td>
<td>2.23</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.0221)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>$Q^2(10)$</td>
<td>431.8</td>
<td>30.3</td>
<td>25.4</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000761)</td>
<td>(0.00457)</td>
</tr>
</tbody>
</table>
GARCH-skew-t modeling

Portmanteau test statistics indicates that normality of returns conditional on $RV$ is confirmed even when conditioned on past information set – i.e. set of returns and squared returns $I_{t-1}$.

Thus, $r_t \mid RV_t, I_{t-1} \sim N(\mu + \beta RV_t, RV_t)$

→ Confirming the basic tenet of the MDH
**GARCH-skew-t modeling**

The significant GARCH effects for returns are captured by the Realized Volatility (RV) series; most of autocorrelation for RV is represented by the simple GARCH(1,1) variance.

<table>
<thead>
<tr>
<th>Ljung-Box test statistics for RV and the scaled RV</th>
<th>$RV_t$</th>
<th>$RV_t / h_t^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q(1) (p-value)</td>
<td>268.4</td>
<td>11.2</td>
</tr>
<tr>
<td>(0.000)</td>
<td></td>
<td>(0.000837)</td>
</tr>
<tr>
<td>Q(5) (p-value)</td>
<td>713.8</td>
<td>13.4</td>
</tr>
<tr>
<td>(0.000)</td>
<td></td>
<td>(0.0197)</td>
</tr>
<tr>
<td>Q(10) (p-value)</td>
<td>1231.4</td>
<td>15.8</td>
</tr>
<tr>
<td>(0.000)</td>
<td></td>
<td>(0.106)</td>
</tr>
</tbody>
</table>
GARCH-skew-t modeling

The distribution of $RV$ conditional on past returns is examined with the inverted chi-square.

$$RV_t^{-1} | I_{t-1} \sim \frac{1}{(h_t h_t)^2} \chi^2_v$$

$$RV_t | I_{t-1} \sim f_{RV}(RV_t | h_t, \bar{h}, \nu, I_{t-1}) = \frac{(h_t \bar{h})^\nu (\nu / 2)^{\nu/2}}{\Gamma(\nu / 2)} \frac{1}{RV_t^{\nu/2+1}} \exp(-\frac{(h_t \bar{h})^2 \nu}{2RV_t})$$

$$RV_t / h_t^2 \sim f_{RV}(RV_t | \bar{h}, \nu) = \frac{\bar{h}^\nu (\nu / 2)^{\nu/2}}{\Gamma(\nu / 2)} \frac{1}{RV_t^{\nu/2+1}} \exp(-\frac{\bar{h}^2 \nu}{2RV_t})$$
GARCH-skew-t modeling

$RV$ scaled by the GARCH(1,1) variance is fit to the scale-inverse-chi-square density function.

Histogram and the inverted Chi-square PP-plot for the scale-adjusted realized volatilities from Jan. 1926 to Dec. 2005
GARCH-skew-t modeling

Histogram and the inverted Chi-square PP-plot for the scale-adjusted realized volatilities from Jan. 1966 to Dec. 2005
GARCH-skew-t modeling

Histogram and the inverted Chi-square PP-plot for the scale-adjusted realized volatilities from Jan. 1976 to Dec. 2005
Inference on the distribution of returns conditional on past returns according to the mixing rule

→ The conditional skew-t, or the GARCH-skew-t model

\[
r_t \mid I_{t-1} \sim f_{r_t \mid I_{t-1}}(r_t \mid I_{t-1}) = \int_0^\infty f_{r_t \mid RV_t, I_{t-1}}(r_t \mid RV_t, I_{t-1}) f_{RV_t, I_{t-1}}(RV_t \mid I_{t-1}) dRV
\]
GARCH-skew-t modeling

The implied GARCH-skew-t model is examined and tested by monthly U.S. stock market returns from Jan. 1926 to Dec. 2005.

Estimation by the QMLE
- Two step estimation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.00646</td>
<td>(2.00*10^{-6})</td>
</tr>
<tr>
<td>$c$</td>
<td>5.99*10^{-5}</td>
<td>(8.47*10^{-9})</td>
</tr>
<tr>
<td>$b$</td>
<td>0.120</td>
<td>(0.000199)</td>
</tr>
<tr>
<td>$a$</td>
<td>0.862</td>
<td>(0.000662)</td>
</tr>
</tbody>
</table>
GARCH-skew-t modeling

QQ-plot of GARCH(1,1) standardized residuals

QQ-plot of GARCH-standardized residuals
GARCH-skew-t modeling

Estimation of symmetric t and skew-t for GARCH(1,1) standardized residuals

<table>
<thead>
<tr>
<th></th>
<th>Gaussian</th>
<th>Symmetric t</th>
<th>Skew t</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^*$</td>
<td>0</td>
<td>0</td>
<td>0.741</td>
</tr>
<tr>
<td>$h^*$</td>
<td>1</td>
<td>0.846</td>
<td>0.815</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$\infty$</td>
<td>7.04</td>
<td>10.3</td>
</tr>
<tr>
<td>$\bar{\beta}$</td>
<td>0</td>
<td>0</td>
<td>-0.732</td>
</tr>
</tbody>
</table>

Log Lik. -1360.8 -1336.9 -1322.9

AIC 2725.6 2679.8 2653.8

BIC 2735.4 2694.4 2673.3
GARCH-skew-t modeling

PP-plots of the estimated Gaussian and symmetric t
GARCH-skew-t modeling

PP-plot of the estimated skew-t distribution
GARCH-skew-t modeling

Estimates of the location parameter is sensitive to distributional assumptions

Eg) GARCH(1,1)-normal vs. GARCH(1,1)-Student t
   - Estimates of the conditional mean are different by 2.5% at an annual rate while the GARCH(1,1)-t fits the sample much better.
   - Nevertheless, the conditional standard deviations are not much different.

→ Motivation for the robust GLS; Best Linear Unbiased Estimator

\[ r_t / h_t = \mu / h_t + \varepsilon_t \]

with \( E[\varepsilon_t] = 0 \quad Var[\varepsilon_t] = 1 \)
GARCH-skew-t modeling

GARCH(1,1) standard deviations with Gaussian and Student t errors

Abs. demeaned returns
Std. dev. with Student t
Std. dev. with Gaussian
GARCH-skew-t modeling

QQ-plot of the GLS-standardized returns
### GARCH-skew-t modeling

Estimation of the symmetric t and the skew-t distribution for the GLS-standardized residuals

<table>
<thead>
<tr>
<th>Distributions</th>
<th>Gaussian</th>
<th>Symmetric t</th>
<th>Skew t</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^*$</td>
<td>0</td>
<td>0</td>
<td>0.779</td>
</tr>
<tr>
<td>$h^*$</td>
<td>1</td>
<td>0.852</td>
<td>0.814</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$\infty$</td>
<td>7.29</td>
<td>10.6</td>
</tr>
<tr>
<td>$\bar{\beta}$</td>
<td>0</td>
<td>0</td>
<td>-0.777</td>
</tr>
</tbody>
</table>

Log Lik.  
-1360.5  
-1337.9  
-1322.6  

AIC        
2725      
2681.8    
2653.2

BIC        
2734.7    
2696.4    
2672.7
GARCH-skew-t modeling

PP-plots for Gaussian and Student t distributions
GARCH-skew-t modeling

PP-plot of the estimated skew-t distribution
# GARCH-skew-t modeling

## Mean, standard deviation, skewness and kurtosis of GLS-residuals

<table>
<thead>
<tr>
<th></th>
<th>Residuals</th>
<th>Gaussian</th>
<th>Symmetric t</th>
<th>Skew t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.00580</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Std</td>
<td>0.999</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Skw</td>
<td>-0.791</td>
<td>0</td>
<td>0</td>
<td>-0.825</td>
</tr>
<tr>
<td>Kts</td>
<td>5.27</td>
<td>3</td>
<td>4.82</td>
<td>6.21</td>
</tr>
</tbody>
</table>
GARCH-skew-t modeling

CDF for GLS residuals

- Skew t
- Symmetric t
- Gaussian
- Empirical
GARCH-skew-t modeling

1-CDF for GLS residuals

- Skew t
- Symmetric t
- Gaussian
- Empirical
GARCH-skew-t modeling

Conclusion:

• New empirical guidance for modeling a skewed & fat-tailed error distribution along with GARCH effects for equity returns

- Distribution of returns conditional on $RV$ and past returns is normal.
- Distribution of $RV$ conditional on past returns is the inverted-chi-square.

→ The distribution of returns conditional on past returns should be well approximated by the GARCH-skew-t model; confirmed by U.S. stock market returns.