Parametric Forecasting of Value at Risk Using Heavy Tailed Distribution

Josip Arnerić

University of Split, Faculty of Economics, Croatia

Elza Jurun

University of Split, Faculty of Economics Split, Croatia

Snježana Pivac

University of Split, Faculty of Economics Split, Croatia

Abstract

This paper deals with modeling volatility of returns of Pliva stocks on Zagreb Stock Exchange for Value at Risk forecasting. Volatility reaction and volatility persistence are measured using asymmetric GARCH process. Croatian capital market characteristic is absence of intensive reaction on "good" information. But it is confirmed that Pliva stocks volatility on Croatian capital market are under dominant influence of "bad" information. If the data are heavy tailed, the VaR calculated using Normal assumption differs significantly from Student's t-distribution. The fact that kurtosis and degrees of freedom from Student's distribution are closely related is used in estimation procedure of GARCH model. The complete procedure of Value at Risk forecasting for Croatia is established with assumption that returns follows extreme value distribution, precisely Student's t-distribution with non-integer degrees of freedom. The optimization problem is solved by FinMetrics module of S-Plus package.

Key words: *Value at Risk, asymmetric GARCH process, leverage effect, volatility persistence, heavy tailed distribution, non-integer degrees of freedom*

1. Introduction

Value at Risk (VaR) has become the most common measure that financial analysts use to quantify market risk. Even so VaR is proposed, by Basel Committee on Banking Supervision in 1996, as the basis for calculation of capital requirements, within establishing banks internal risk models. VaR is defined as the maximum potential loss of financial instrument with a given probability (usually 1% or 5%) over a certain time period.

There are many methodologies for calculating VaR, but for simplicity they can be classified into parametric and nonparametric models. In category of parametric models the most are used GARCH(p,q) models and within nonparametric Monte Carlo simulation is most popular.

However, it isn't easy to estimate VaR when stochastic process which generates distribution of returns is not known. Unfortunately the assumption that the returns are independently and identically normally distributed is not satisfied. Furthermore, empirical research about financial markets reveals following facts:

- financial return distributions are leptokurtic, i.e. they have heavy tails and a higher peak than a normal distribution,
- equity returns are typically negatively skewed and
- squared return series shows significant autocorrelation, i.e. volatilities tends to cluster.

According to first two facts it is important to examine which probability density function capture heavy tails and asymmetry the best. According to the third fact it is important to correctly specify conditional mean and conditional variance equations from GARCH family models. So, it is well-known that returns from financial instruments such as exchange rates, equity prices and interest rates measured over short time intervals, i.e. daily or weekly, are characterized by high kurtosis. It is important to note that kurtosis is both a measure of peakdness and fat tails of the distribution.

2. Real assumption of extreme value distribution

If the distribution of returns heavy tailed, the VaR calculated using normal assumption differs significantly from Student's t-distribution. As it is known Student's distribution belongs to family of extreme value distributions. In case of volatility modeling and VaR estimating of Pliva stock returns on Zagreb Stock Exchange it is found that kurtosis and degrees of freedom from Student's t-distribution are closely related.

To identify outliers and another extreme values Box and Whisker plot has been used. Statistical significance of existing heavy tailed distribution has been shown by Q-Q plot and tested using Jarque-Bera test.

In practice, the kurtosis is often larger than six, leading to estimate of non-integer degrees of freedom between four and five. Thus, degrees of freedom can easily be estimated using the method of moments.

Generally, there are three parameters that define a probability density function (pdf):

- location parameter,
- scale parameter and
- shape parameter.

The most common measure of location parameter is the mean. The scale parameter measure variability of pdf, and the most commonly used is variance or standard deviation.

The shape parameter (skewness and/or kurtosis) determines how the variation is distributed about the location parameter.

3. Non-integer degrees of freedom estimation

The density of a non-central Student t-distribution has the following form:

$$f(x)=\frac{\Gamma\left(\frac{df+1}{2}\right)}{\Gamma\left(\frac{df}{2}\right)\sqrt{\pi\cdot\beta\cdot df}}\left(1+\frac{(x-\mu)^2}{\beta\cdot df}\right)^{-\frac{1+df}{2}}, \tag{1}$$

where μ is location parameter, β is scale parameter and *df* is a shape parameter, or degrees of freedom and $\Gamma(\cdot)$ is gamma function. Standard t-distribution assumes $\mu=0$, $\beta=1$, with integer *df*. However, there are no mathematical reasons why the degrees of freedom should be an integer. Even so, the degrees of freedom can be estimated using method of moments, which means that kurtosis and degrees of freedom are closely related:

$$k=\frac{6}{df-4}+3 \quad for \quad every \quad df>4\,. \tag{2}$$

So, when empirical distribution is leptokurtic, then Student's t-distribution with parameter $4<df\leq 30$ should be used to allow heavy tails of high kurtosis distribution.

First two central moments are given as:

$$\begin{aligned}\mu_2 &= E\left[(x-\mu)^2\right]=\frac{\beta\cdot df}{df-2}\\ \mu_4 &= E\left[(x-\mu)^4\right]=\frac{3\beta^2 df^2}{(df-2)(df-4)}\end{aligned}\,, \tag{3}$$

with excess kurtosis (greater then 3):

$$k^*=\frac{\mu_4}{\mu_2^2}-3=\frac{6}{df-4}\,. \tag{4}$$

Hence, we may apply method of moments to get consistent estimators:

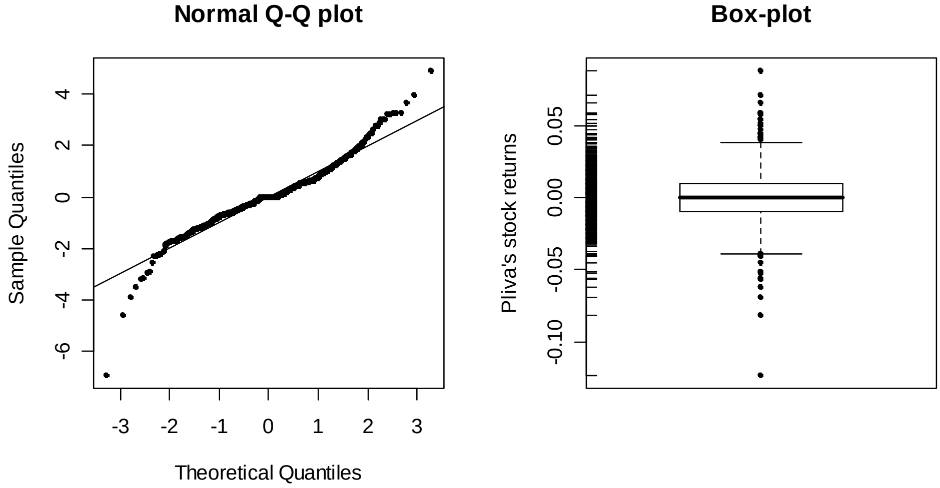
$$\begin{aligned}\hat{\mu}_2 &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}\\ \hat{\mu}_4 &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})^4}{n}\\ \hat{k}^* &= \frac{\hat{\mu}_4}{\hat{\mu}_2^2}-3\\ \hat{df} &= 4+\frac{6}{\hat{k}^*}\\ \hat{\beta} &= \left(\frac{3+\hat{k}^*}{3+2\cdot\hat{k}^*}\right)\cdot\hat{\sigma}^2\end{aligned}\,, \tag{5}$$

where variance from sample $\hat{\sigma}^2$ is biased estimator of scale parameter β.

4. Heavy tails diagnostics

There are various analytical and graphical methods to detect heavy tails from observed distribution. The most common used are Jarque-Bera and Kolmogorov-Smirnov test, while Box-plot and Q-Q plot graphically determines fat tails. From Figure 1. it is obvious that outliers and extreme values causes fat tails, which are most interesting for risk managers.

Figure 1.

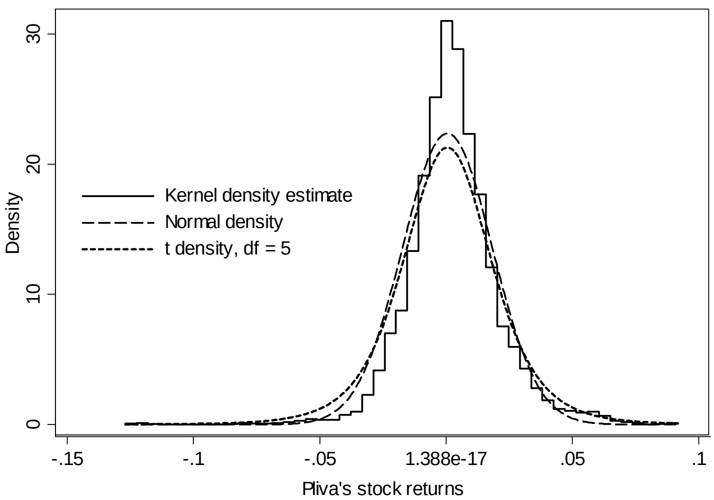


Source: According to data on www.zse.hr

On figure 2. different distributions are formed using same data base from sample. However, Kernel density estimate is concerned, as nonparametric density smoothing, rather then classical histogram, to objectively investigate the shape of observed distribution.

Therefore, data are approximated using Normal distribution, as the assumption of most financial analysts. Even so, an empirical finding shows that assumption of Normal distribution is not appropriate and nonrealistic. From the same reason on figure 2. data are approximated using Student's t-distribution with 5 degrees of freedom, and in this case it can be seen that Student's distribution has heavier tails in comparison to Normal distribution. It means that higher probabilities are assigned to extreme values of the distribution, and empirical evidence for heavy tails can be found in high kurtosis.

Figure 2. Density estimation with Normal and Student's t-distribution



Source: According to data on www.zse.hr

In table 1. basic statistics are presented including normality test. Each of shape measures, i.e. skewness and kurtosis are tested separately, indicating that skewness isn't statistically significant whereas excess kurtosis of 4.7 is significantly greater than 3. In general, joint test shows that null hypothesis of normality distribution assumption can't be accepted. This joint test is presented as Jarque-Bera test in table 1.

Table 1. Normality test

Mean	.0005299	Skewness	-.0222934
Std. Dev.	.0178412	Kurtosis	7.7096480

Skewness/Kurtosis tests for Normality

Variable	Pr(Skewness)	Pr(Kurtosis)	joint chi2(2)	joint Prob>chi2
plivaret	0.781	0.000	103.70	0.0000

Source: Tested according to data on www.zse.hr

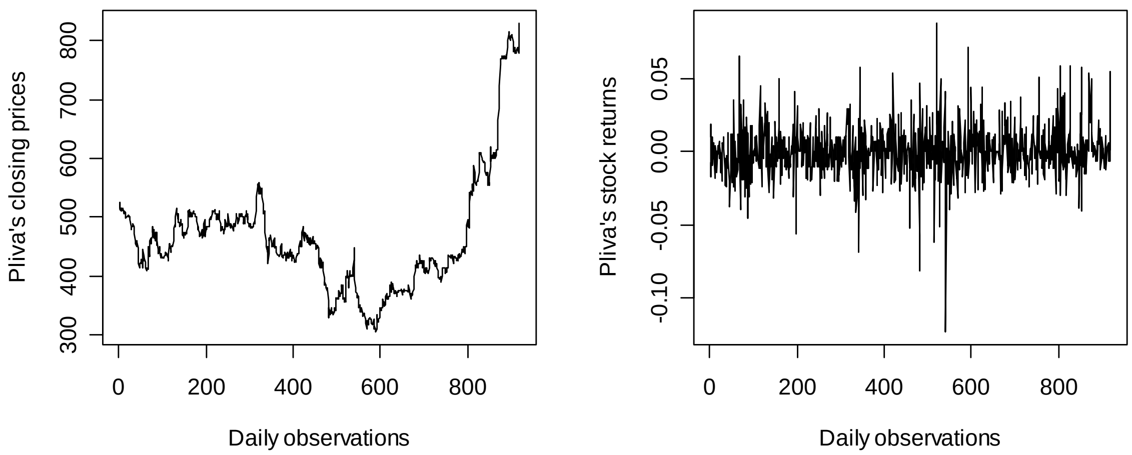
5. Asymmetric GARCH model of Pliva stock returns on Zagreb Stock Exchange

Before we continue to define the model to capture volatility of Pliva returns, presented in figure 3., it is necessary to investigate if there is asymmetry in volatility clustering, i.e. if there is leverage effect. The tendency for volatility to decline when returns rise and to rise when returns fall is called the leverage effect, i.e. "bad" news seems to have a more effect on volatility than does "good" news.

A simple test to investigate the leverage effect is to calculate first-order autocorrelation coefficient between lagged returns and contemporary squared returns:

$$\frac{\sum_{i=2}^{n} r_t^2 r_{t-1}}{\sqrt{\sum_{i=2}^{n} r_t^4 \sum_{i=1}^{n} r_{t-1}^2}} \tag{6}$$

Figure 3. Pliva's closing prices and Pliva's stock returns from 2 January 2003 to 4 September 2006, daily observed on Zagreb Stock Exchange.



Source: According to data on www.zse.hr

The results of asymmetric volatility testing are given in table 2.

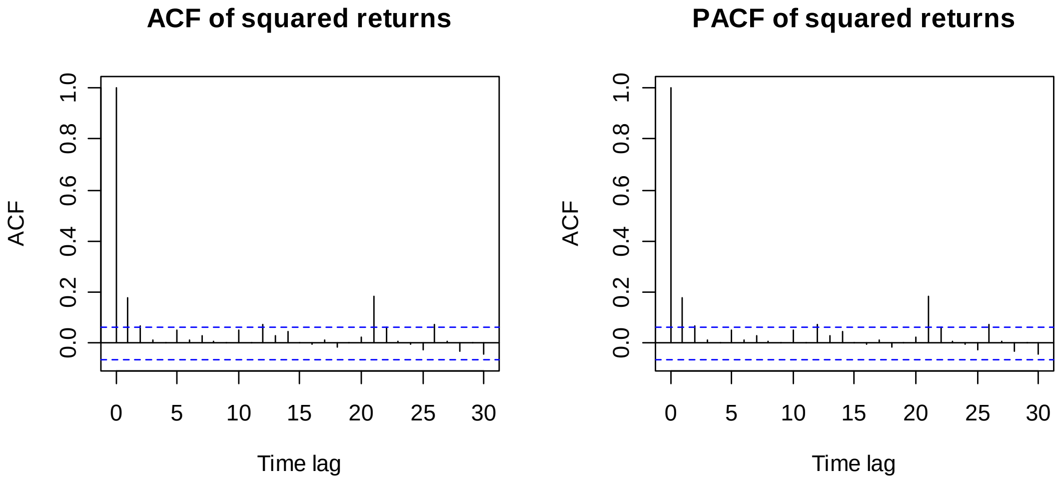
Table 2. Testing for leverage effects

		Lagged returns
Squared returns	Pearson Correlation	-,163
	Sig. (2-tailed)	,046
	N	918

Source: Tested according to data on ZSE

It can be concluded that there is asymmetric volatility clustering of Pliva returns at p-value less than 1%, because the above autocorrelation coefficient is negative and significantly different from zero.

Figure 4. Autocorrelation and partial autocorrelation functions of squared Pliva returns



Source: According to data on www.zse.hr

From figure 4. it is obvious that there is significant autocorrelation in squared return series of Pliva stocks for almost each time lag. It means that return series contain ARCH effects. These ARCH effects are also tested using Lagrange multiplier test, which results are given in table 3.

Table 3. Lagrange multiplier (LM) test

ARCH Test:

F-statistic	4.400118	Probability	0.004399
Obs*R-squared	13.06909	Probability	0.004490

Dependent Variable: RESID^2

Included observations: 916 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000188	2.17E-05	8.673513	0.0000
RESID^2(-1)	0.097833	0.033114	2.954462	0.0032
RESID^2(-2)	0.054730	0.033227	1.647132	0.0999
RESID^2(-3)	0.014804	0.033271	0.444954	0.6565

Source: Tested according to data on www.zse.hr

From table 3. it can be seen that variance is heteroscedastic because the square unexpected returns follow AR(1) process, i.e. the lagged squared returns parameter is statistically significant at empirical p-value 0.32%, or even more LM test value, for large samples, is significant at 0.449%. It means that variance is time-varying.

Table 4. White heteroskedasticity test

F-statistic	62.26114	Probability	0.000000
Obs*R-squared	109.9794	Probability	0.000000

Dependent Variable: RESID^2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000178	1.79E-05	9.967315	0.0000
CROBEXRET	-0.004771	0.001636	-2.916349	0.0036
CROBEXRET^2	0.472505	0.044129	10.70728	0.0000

Source: Tested according to data on www.zse.hr

Results from table 4. confirms that variance of estimated residuals is heteroskedastic, at empirical significance less then 1%, according to White's regression.

The expected value of return series is calculated from the simple linear regression model, usually taking constant as regressor. If there is significant autocorrelation in returns, best fitted ARMA models are usually used, following Box-Jenkins procedure. In this research it has been found most appropriate to enter market return series of Crobex index as regressor. It has been shown that ARCH(p) process with infinite number of parameters is equivalent to generalized ARCH process, GARCH(p,q) process, which is very well approximated by simple GARCH(1,1). As the time lag increases in an ARCH(p) model it becomes more difficult to estimate parameters. Besides it is recommended to use parsimonious model as GARCH(1,1) that is much easier to identify and estimate.

But if there is asymmetric volatility clustering Ding, Granger and Engle (1993) proposed Asymmetric Power ARCH (APARCH) model:

$$
\begin{aligned}
& r_t = \beta \cdot m_t + \varepsilon_t \\
& \sigma_t^{\delta} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \cdot \left(\left| \varepsilon_{t-i} \right| - \gamma_i \cdot \varepsilon_{t-i} \right)^{\delta} + \sum_{j=1}^{q} \beta_j \cdot \sigma_{t-j}^{\delta},
\end{aligned} \tag{7}
$$

The APARCH model is the most promising one, because this model nests at least seven ARCH type models, according to the estimated parameters. In model (7) parameter δ plays the role of a Box-Cox transformation of the time-varying conditional standard deviation σ_t, while γ_i reflects leverage effect, i.e. asymmetric information influence.

In financial theory it is confirmed that parameter β, in the mean conditional equation, determines volatility of stock, i.e. stock risk. If $\beta > 1$, the stock is highly risked. If $0 < \beta < 1$, the stock is lowly risked, while free-risk stock assumes that $\beta = 0$.

Among the whole family of APARCH models, the most appropriate in this case is GJR - GARCH(1,1) model, proposed by Glosten, Jagannathan and Runkle (1993). This model is special case of APARCH(1,1) model when parameter $\delta = 2$.

Estimated model, assuming Normal distribution, using Maximum Likelihood (ML) method is given in table 5.

Table 5. Conditional mean and conditional variance equations

Method: ML - ARCH (Marquardt) - Normal distribution

Sample: 1/02/2003 9/04/2006

Convergence achieved after 22 iterations

GARCH = C(2) + C(3)*(ABS(RESID(-1)) - C(4)*RESID(-1)) + C(5)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
CROBEXRET	0.948885	0.022167	42.80645	0.0000
Variance Equation				
C(2)	3.92E-06	8.08E-07	4.853776	0.0000
C(3)	0.025012	0.009820	2.547087	0.0109
C(4)	0.606337	0.264914	2.288807	0.0221
C(5)	0.949479	0.008156	116.4150	0.0000
Adjusted R-squared	0.288067	S.D. dependent var		0.017841
S.E. of regression	0.015054	Akaike info criterion		-5.628536
Sum squared resid	0.207126	Schwarz criterion		-5.602294
Log likelihood	2591.312	Durbin-Watson stat		2.019460

Source: Estimated according to data on www.zse.hr

The same model is estimated in table 6. by maximizing likelihood function, assuming Student's distribution of Pliva's stock returns with 5.3 degrees of freedom, whereas degrees of freedom are estimated according to equation (2).

Table 6. Conditional mean and conditional variance equations

Method: ML - ARCH (Marquardt) - Student's t distribution

Convergence achieved after 23 iterations

t-distribution degree of freedom parameter fixed at 5.3

GARCH = C(2) + C(3)*(ABS(RESID(-1)) - C(4)*RESID(-1)) + C(5)*GARCH(-1)

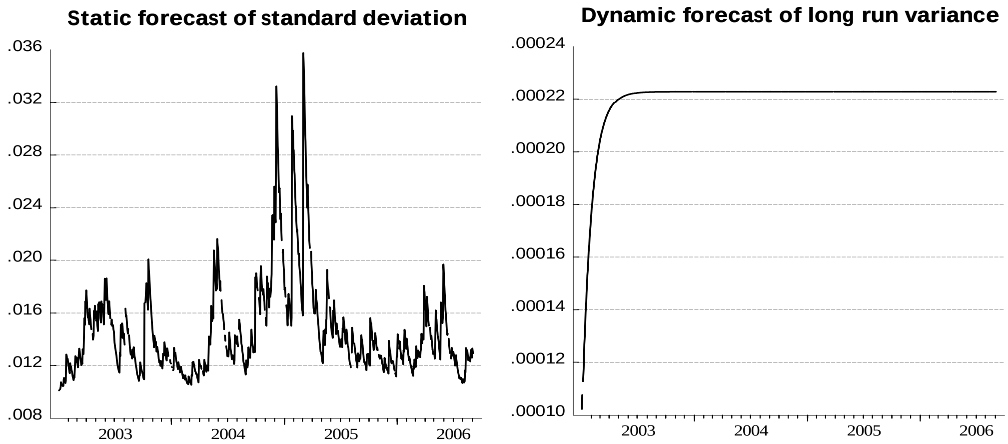
	Coefficient	Std. Error	z-Statistic	Prob.
CROBEXRET	1.052386	0.035928	29.29130	0.0000
Variance Equation				
C(2)	9.93E-06	3.94E-06	2.520250	0.0117
C(3)	0.053097	0.019865	2.672932	0.0075
C(4)	0.463518	0.191284	2.423194	0.0154
C(5)	0.890947	0.030276	29.42759	0.0000
S.E. of regression	0.015117	Akaike info criterion		-5.719580
Sum squared resid	0.208871	Schwarz criterion		-5.693338
Log likelihood	2633.147	Durbin-Watson stat		2.022856

Source: Estimated according to data on www.zse.hr

Comparing these two models it can be seen that in table 5., assuming Normal distribution, parameter beta in conditional mean equation is less than one. From the modern financial market models aspect it can be interpreted that investments in Pliva stocks are low risky. However, as it can be seen it table 6., assuming Student's distribution with estimated non-integer degrees of freedom, parameter beta is higher than one, which means that Pliva's stock investments are highly risky. This change in parameter value can be explained by distribution shape with heavy tails.

Even so, information criteria are more representative in Student's distribution model, as well as maximal value of likelihood function. Also sum of parameters $\alpha_1 + \beta_1$ (in table 6. $c(3)+c(5)$), according to equations (7), indicates that there is persistence volatility, i.e. conditional variance decays slowly, not far from long-memory model. In addition "bad" news have higher impact in increasing volatility then the "good" news, represented by leverage parameter $\gamma_1 = 0.4635$, i.e. coefficient $c(4)$ in table 6.

Figure 5. Static and dynamic forecasting



Source: According to estimated APARCH(1,1) model in table 6.

Figure 5. shows static forecast of conditional standard deviation and dynamic forecast of unconditional long run variance, using Student's distribution model in table 6. Diagnostic checking for estimated model in table 6. including investigation of standardized residuals, with satisfied covariance-stationary condition, are as follows:

Table 7. Diagnostic tests for estimated standardized residuals

			Stat.	p-value
Jarque-Bera Test	R	Chi^2	392.7961	0
Shapiro-Wilk Test	R	W	0.9578688	1.346745e-15
Ljung-Box Test	R	Q(10)	11.5202	0.3184496
Ljung-Box Test	R	Q(15)	17.28357	0.3021972
Ljung-Box Test	R	Q(20)	19.07808	0.5167562
Ljung-Box Test	R^2	Q(10)	11.27558	0.3364544
Ljung-Box Test	R^2	Q(15)	14.22036	0.5088865
Ljung-Box Test	R^2	Q(20)	15.21380	0.7640459
LM Arch Test	R	TR^2	13.60740	0.3264783

Source: Tested according to estimated APARCH(1,1) model

6. Value at Risk forecasting for long trading position

As introduced in first section Value at Risk, based on the Normal distribution, can be calculated as:

$$VaR_t(\alpha) = \hat{\mu}_t + z \cdot \hat{\sigma}_t, \tag{8}$$

where expected mean and expected standard deviation, at time *t*, are predicted from estimated GJR-GARCH(1,1) model. VaR expressed in equation (8) can be interpreted as expected minimal percentage loss within probability of α, when z is left percentile from Normal distribution.

$$VaR_t(\alpha) = p_t\left[\hat{\mu}_t + z \cdot \hat{\sigma}_t\right] \ . \tag{8a}$$

Equation (8a) defines minimal loss in relevant currency according to p_t i.e. current stock price. This is the case when investor holds "long" position, i.e. if he has bought an asset, in which case he incurs the risk of a loss of value of the asset. When investor holds "short" position (he has sold an asset, in which case he incurs a positive opportunity cost if the asset value increases), variable z presents the right percentile from Normal distribution.

Assuming Student's t-distribution VaR can be calculated as follows:

$$VaR_t(\alpha) = \hat{\mu}_t + t_\alpha^{df} \cdot \left[\hat{\sigma}_t \cdot \sqrt{\frac{3+\hat{k}^*}{3+2\cdot\hat{k}^*}}\right], \tag{9}$$

where expected standard deviation is corrected to get unbiased estimator of standard Student's scale parameter, according to equations (4) and (5).

So, if today is 4 September 2006 (the last day of observed period), and if investor has bought Pliva stocks at price of 820,00 kunas, it can be predicted, for example for two days ahead, that his loss wouldn't exceed 20,67 kunas per stock with probability of 95%. Moreover, investor can be sure that his loss wouldn't exceed 43,32 kunas per stock with confidence level of 99%:

$$\begin{aligned} VaR_{t+2}(0.05) &= 820 \cdot \left[0 - 1.9901 \cdot 0.020396 \cdot \sqrt{(3+4.7)/(3+2\cdot 4.7)}\right] = 20.67 \\ VaR_{t+2}(0.01) &= 820 \cdot \left[0 - 3.2867 \cdot 0.020396 \cdot \sqrt{(3+4.7)/(3+2\cdot 4.7)}\right] = 43.32 \end{aligned} \tag{10}$$

7. Conclusion remarks

For any investor on stock market is very important to predict possible loss, depending on if he has bought or sold stocks. By forecasting Value at Risk investor can protect himself "a priori" from estimated market risk, using financial derivatives, i.e. options, forwards, futures and other instruments. In that sense we find financial econometrics as the most useful tool for modeling conditional mean and conditional variance of nonstationary financial time series. The assumption of heavy tailed distribution, such as Student's t-distribution with non-integer degrees of freedom is used in asymmetric GARCH(p,q) model. It becomes more adequate with much precisely forecast which is shown on the example of Pliva stocks for the first time on Croatian Capital Market.

References

[1] A. ANDREEV, A. KANTO, *Conditional value-at-risk estimation using non-integer values of degrees of freedom in Student's t-distribution,* Journal of Risk, Vol. 7, No. 2 (2005), 55-61.

[2] L. BAUWENS, S. LAURENT, *A New Class of Multivariate Skew Densities, With Application to Generalized Autoregressive Conditional Heteroscedasticity Models***,** Journal of Business & Economic Statistics, Vol. 23, No. 3 (2005), 346-353.

[3] A. DE SERVIGNY, O. RENAULT, *Measuring and Managing Credit Risk,* McGraw-Hill, Cambridge, 2004.

[4] R. ENGLE, *GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics*, Journal of Economic Perspectives, Vol. 15, No. 4 (2001), 157-168.

[5] R. ENGLE, S. MANGANELLI, *Value at Risk Models in Finance*, Working Paper No. 75, European Central Bank (2001), 1-39.
[6] R. ENGLE, *Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskcedasticity Models*, Journal of Business and Economic Statistics, Vol. 20, No. 3 (2002), 339-350.
[7] P. HALL, Q. YAO, *Inference in ARCH and GARCH Models with Heavy-Tailed Errors*, Econometrica, Vol. 71, No. 1 (2003), 285-317.
[8] S. LAURENT, *Analytical derivates of the APARCH model*, Computational Economics, Vol. 24, No. 1 (2004), 51-57.
[9] S. H. POON, *A Practical Guide to Forecasting Financial Market Volatility*, John Wiley & Sons, Ltd. Chichester, 2005.
[10] S. T. RACHEV, C. MENN, F. J. FABOZZI, *Fat-Tailed and Skewed Asset Return Distributions*, John Wiley & Sons, Inc. New Jersey, 2005.