Fitting extreme gains and losses of the Prague Stock Exchange Index

Ján Gogola, Ondřej Slavíček

***Abstract:* -** In this paper we focused on the daily log returns of investment in the Prague stock exchange index, PX-Index. Considering an investment trust that takes a „passive" investment strategy and invests its assets in a specified stock-market index - the PX Index. We analysed data from January 1st, 1995 to February 20th, 2014. A popular model for stock market returns is that the log investment returns are independent and identically distributed (i.i.d.) normal random variables. We focused on the daily log returns and analysed the distribution of these returns. By means of the well-known Jarque-Bera test we reject the i.i.d. normal hypothesis of daily log returns. We emphasize this by looking at the data using graphical techniques, such as histogram and Q-Q plot. We can see that the data has fatter left and right-hand tails than the normal distribution. Conclusions of our basic analysis are that the daily log returns are leptokurtic and heavy tailed. They are not i.i.d. and volatility varies over time. Also we can say that extreme daily log returns appear in clusters.

Further we investigated a simple model which incorporates stochastic volatility. We analysed volatility-standardised residuals using a GARCH approach. We can see that standardised residuals do not show any clusters of high and low volatility.

Plotted standardised residuals also show that there are more exceedances of the lower threshold than the upper and that they are larger.

International banking regulations require banks to pay specific attention to the probability of large losses over short periods of time.

We were focusing on the tails of the standardised residual. We fitted tail data separately using a Pareto distribution. Estimated parameters of the Pareto distributions show us that the Pareto distribution gives a generally better fit over the tails than *t* and non-central *t* distribution.

Keywords: - stock exchange index, log returns, normal distribution, GARCH, standardised residuals, ACF

JEL Classification: C13, C18, C63

This work was supported by the research grant: SGSFES_2015001 Ekonomický a sociální rozvoj v soukromém a veřejném sektoru.

Ján Gogola is with the Institute of Mathematics and Quantitative Methods, Faculty of Economics and Administration, University of Pardubice, Studentská 84, 532 10 Pardubice, (e-mail: jan.gogola@upce.cz).

Ondřej Slavíček is with the Institute of Mathematics and Quantitative Methods, Faculty of Economics and Administration, University of Pardubice, Studentská 84, 532 10 Pardubice, (e-mail: ondrej.slavicek@upce.cz).

I. INTRODUCTION

Consider an investment trust that takes a "passive" investment strategy and invests its assets in a specified stock-market index e.g. PX Index.

The PX Index is the official index of major stocks that trade on the Prague Stock Exchange. The index was calculated for the first time on March 20, 2006 when it replaced the PX 50 and PX-D indices. The index took over the historical values of the PX50. The starting day of PX 50 was April 5, 1994 and its opening value was fixed at 1 000 points. [5]

At this time the index included 50 companies on the Prague Stock Exchange.

Figure 1. shows the development of the PX Index from its starting day in 1994 to February 20, 2014. From the middle of 1994 to about 2004 we can see something that looks like business cycles. Business cycles of this type might exist but the cycles are all of different lengths, the timing of the peaks and the lows are difficult to predict. The PX Index reaches its top on October 29, 2007 with 1936 points. As result of financial crisis reached 700 points on October 27, 2008 losing almost 50% of its value in two months. [6]

Since data in the year 1994 are irregular, we decided to analyse data from January 1, 1995 to February 20, 2014.

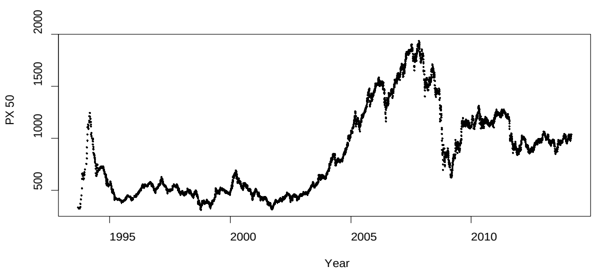

Fig. 1: PX Index (in CZK) from 5. 4. 1994 to 20. 2. 2014. Source: www.pse.cz [6]

What is the distribution of the percentage return (no dividends) over specified period of time?

Suppose that we use one day as our unit of time. The typical approach is to model the log investment return from time $t-1$ to time *t*: which we will denote by $d(t)$.

That is, 1 CZK invested at time $t-1$ will be worth $e^{d(t)}$ CZK at time *t*.

A popular model for stock market returns is that $d(1)$, $d(2)$, … are independent and identically distributed (i.i.d.) normal, $N(\mu,\sigma^2)$, random variables.
Some questions arise.

- Is the assumption of normality appropriate?
- Is the i.i.d. assumption appropriate?

In this paper we are going to focus on the daily log returns (Figure 2.) and analyse the distribution of these returns.

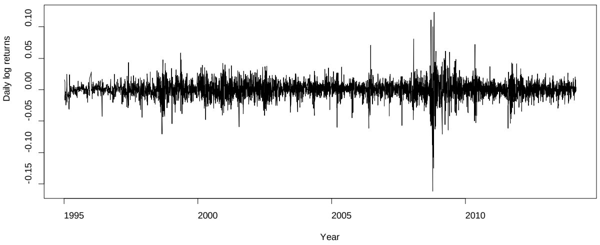

Fig. 2: Percentage daily return on the PX Index (4787 observations). Source: Own calculation

II. Problem Formulation

In this section we will establish certain stylised facts about this returns $d(t)$ series.
Firstly we will assume that returns are i.i.d.. This implies the assumption that volatility is constant.
Our first aspect is to analyse if the $d(t)$ are i.i.d. $N(\mu,\sigma^2)$. We will fit the normal distribution to the data and then conduct a variety of test to see if this model is appropriate. Suppose that we have *T* observations, $d(1)$, $d(2)$, …, $d(T)$. We can go straight ahead and estimate μ and σ using the standard maximum likelihood estimates:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} d(t)\,, \qquad (1)$$

$$\hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T} (d(t)-\hat{\mu})^2\,. \qquad (2)$$

We have 4787 observations in our dataset and we find that $\hat{\mu} = 0.0001247978$ per trading day (or $(1+\hat{\mu})^{252}-1 = 0.03194677$ per annum)[1] and $\hat{\sigma} =$ 0.01400127 per trading day (or $\hat{\sigma}\cdot\sqrt{252} = 0.2222633$ per annum).
The normality we can verify by the coefficients of skewness $\sqrt{b}$ and kurtosis *k*. The normal distribution has skewness 0 and kurtosis 3.

[1] There are approximately 252 trading days per year.

The PX Index daily returns data has a skewness of $\sqrt{b} =$ – 0.44 and a kurtosis of k = 14.687. These empirical coefficients look quite different from 0 and 3 respectively, but are they significantly different?
We can answer this question by means of the Jarque-Bera test.
The Jarque-Bera test gives a test for normality that focuses on both the skewness and kurtosis. Specifically, if the data are i.i.d. $N(\mu,\sigma^2)$ then the Jarque-Bera statistic with *n* observations

$$T_n = \frac{n}{6}\cdot\left(b + \frac{1}{4}(k-3)^2\right), \qquad (3)$$

should have, approximately, a Chi-squared distribution with 2 degree of freedom.
More precisely, if the null hypothesis is true, then the T_n is said to be asymptotically χ_2^2.
For PX Index daily returns data (n = 4787) the Jarque-Bera statistic is equal to 27398.43, which is exceedingly large. The *p*-value is effectively 0 and we reject the i.i.d. normal hypothesis.
The **R** language contains a function called jbtest which performs the Jarque-Bera test. So we just type the command:

```
> jbtest(d)   # Jarque-Bera test of the daily returns #
   Skewness = -0.4403515
   Kurtosis = 14.68711
Jarque-Bera = 27398.43
    p-value = 0
```

We can easily carry out a chi-squared test on our data. This can be done by the following command in **R**:

```
> chi2test.normal(d)
Chi-squared statistic is 564.431 with 97 degrees of freedom.
The p-value for this is 0.
```

The chi-squared statistic works out at 564.431 with 97 degrees of freedom. The *p*-value for this is 0: that is, there is very strong evidence to reject the assumption of normality

The data is clearly non-normal from these analyses. We add to this by looking at the data using graphical techniques, such as histogram and Q-Q plot

We have plotted in Figure 3. the histogram of the daily log returns on the PX Index. We have also drawn in the density function for the $N(\hat{\mu},\hat{\sigma}^2)$ distribution. We can easily see from this that the data exhibit a narrower peak than the *best-fitting* normal distribution. Less obviously, but certainly a feature of the data is, that it has a fatter left and right-hand tails than the *best-fitting* normal distribution. In other words we are looking at a *leptokurtic* distribution.

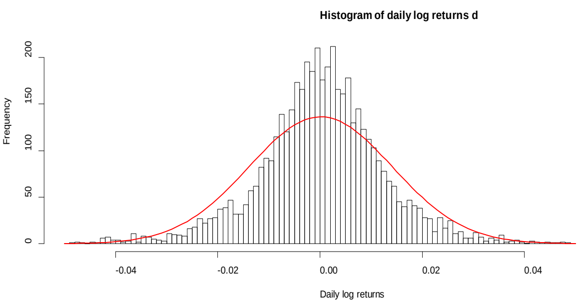

Fig. 3: Histogram of the percentage daily returns on the PX Index. Source: Own calculation
The solid line shows the density function of the $N(\hat{\mu},\hat{\sigma}^2)$ distribution.

A Q-Q plot is a dot plot that plots the ordered sample against the corresponding quantiles of the distribution that we are considering to model the data. Suppose that we have observations $X_1, X_2, \ldots, X_n$. Let $\tilde{X}_1 < \tilde{X}_2 < \ldots < \tilde{X}_n$ be the ordered value of $X_1, X_2, \ldots, X_n$. Now let $q_i = \dfrac{i-0.5}{n}$ for $i = 1,2,\ldots,n$ be theoretical probabilities that are uniformly spread over the range 0 to 1.

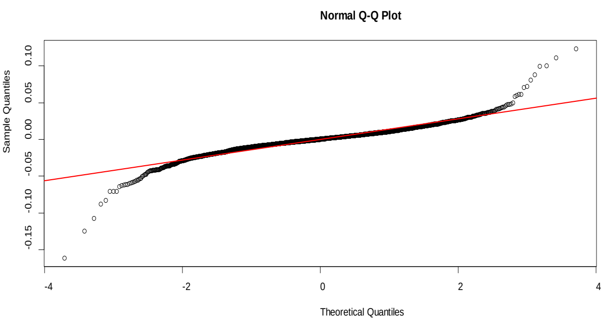

Fig 4: Q-Q plot of daily returns on the PX Index.
Source: Own calculations

Let $Y_i = \Phi^{-1}(q_i)$ be the corresponding theoretical quantile of the standard Normal distribution.
If the data were genuinely normally distributed then we would expect to see the 4787 points much more in a straight line. The fact that Figure 4. actually exhibits an inverted "S" shape means that the data has fat left and right-hand tails. The downturn in the plot at the left-hand end means that the left-hand tail is fatter than the normal distribution: in other words we should expect rather more large losses over time than we would predict using the Normal distribution. This inverted "S" shape therefore points to the data being leptokurtic. We can use the shape of the Q-Q plot to guide our next choice of distribution.
The formal hypothesis tests and the less-formal graphical/diagnostic tests clearly indicate that the assumption that returns are normally distributed is not valid.
Additionally Figure 2. also suggests that the daily log returns are not i.i.d.. Instead, it looks like there are clear clusters of high and low volatility. The PX Index log returns have clusters of high volatility (e.g. in 2008) and low volatility (e.g. 2013).
Conclusions of our basis analysis are:

1) Log returns are leptokurtic and heavy tailed.
2) Log returns are not i.i.d.
3) Volatility appears to vary over time.
4) Extreme log returns appear in clusters.

III. Problem Solution

In Figure 5. we investigate the evidence for non-i.i.d. log returns in a more systematic way. Here we look at the autocorrelation function.

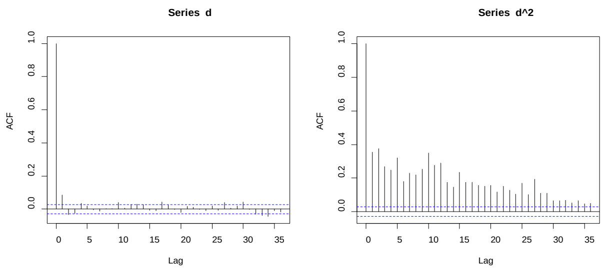

Fig. 5: Left: Sample autocorrelation function for PX Index daily log returns.
Right: Sample autocorrelation function for squared PX Index daily log returns.
Horizontal dashed lines give the 95% confidence interval.
Source: Own calculation

In the left-hand plot we have plotted the sample autocorrelation function for the daily log returns $d(t)$ for $t = 1,2,\ldots,n$. The values of the $\rho(k)$ ($\rho(k) = cor(d(t), d(t+k))$, where *k* is referred to as *the lag* are all fairly close to zero (except, a small but significant positive correlation at lag 1). This initial observation is consistent with log returns being i.i.d. but, of course, it does not imply that log returns are i.i.d.
The fact that the ACF is close to zero implies that a high log return one day does not give us any information about the expected log return the next day.
In right-hand plot we have plotted the sample autocorrelation function for the squares of the daily log returns $d(t)$ for $t = 1,2,\ldots,n$. This plot is very different. There is a moderate, but nevertheless highly significant, correlation between the $d(t)^2$ on different days. It tells us that if $d(t)^2$ was high on day *t* (that is, a large positive or negative log return) then it is likely that $d(t+1)^2$ will also be above average. The left-hand plot tells us, though, that we cannot be precise in any way about the sign of $d(t+1)$ or its conditional expected value. The significant autocorrelations in $d(t)^2$ imply that the market goes through phases of high and low volatility. The fact that the autocorrelation function decline very slowly means, that these phases can last for some time.
We now propose a simple model which incorporates stochastic volatility of the form:

$$d(t) = \mu + \sigma(t) \cdot Z(t), \quad (4)$$

where μ is constant, $\sigma(t)$ is some stochastic volatility process and $Z(1), Z(2), \ldots$ are i.i.d. *volatility-standardised residuals*.

An important observation about the equation (4) is that the value of $\sigma(t)$ must be known at time $t-1$ based on information available up to and including time $t-1$. It is usually assumed that $Var[Z(t)] = 1$ which means that $\sigma(t)^2 = Var[d(t) \mid \Im_{t-1}]$ where $\Im_{t-1}$ represents the market information available up to and including time $t-1$. In other words $\sigma(t)$ is the conditional standard deviation of $d(t)$ given the market information up to $t-1$. Usually we also assume that $E[Z(t)] = 0$.

Now define the variance process to be
$v(t) = \sigma(t)^2$
and propose the simple model

$$v(t+1) = \theta \cdot v(t) + (1-\theta) \cdot (d(t) - \mu)^2 = \theta \cdot v(t) + (1-\theta) \cdot v(t) \cdot Z(t)^2 . \quad (5)$$

It is straightforward to show that this implies that $v(t) = (1-\theta) \cdot \sum_{k=0}^{t-1} (d(t-k) - \mu)^2 + \theta^t \cdot v(0)$. In fact, this process we have defined for $v(t)$ (equation (5)) is a special case of what is called a GARCH(1,1) time series process.
Now that we have estimated the volatility process $v(t)$ we can calculate the volatility-standardised residuals

$$\hat{Z}(t) = \frac{d(t) - \hat{\mu}}{\sqrt{\hat{v}(t)}}, \quad (6)$$

which we can now analyse.

In Figure 6. we have plotted the volatility-standardised residuals. As a reminder, in the upper plot (Figure 2.) there are clear clusters of large gains and losses and other clusters of small gains and losses. The Figure 6. shows the standardised residuals, and we see nothing of the clustering. We can conclude that the standardisation has passed our first visual diagnostic test.

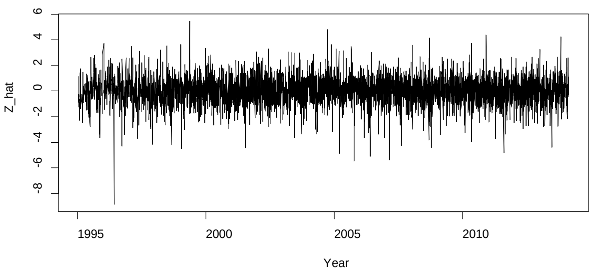

Fig. 6: Volatility-standardised residuals.
Source: Own calculation

Our next plot, Figure 7, shows the autocorrelation functions for $\hat{Z}(t)$ and $\hat{Z}(t)^2$. The right-hand plot for $\hat{Z}(t)^2$ shows a dramatic improvement over Figure 5.

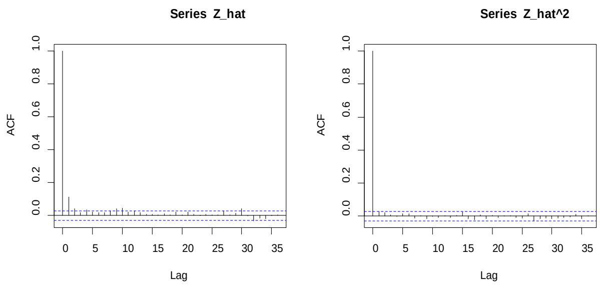

Fig. 7: Left: Sample autocorrelation function for the volatility-standardised residuals.
Right: Sample autocorrelation function for squared volatility-standardised residuals.
Source: Own calculation

The model (equation (5)) for stochastic volatility using exponential weighting is a special case of a GARCH process. A GARCH process (Generalized Auto-Regressive Conditionally Heteroscedastic) is defined as follows:

- Let $X(t) = d(t) - \mu$.
- Let $Z(1), Z(2), \ldots$ be a sequences of i.i.d. random variables with mean 0 and variance 1.
- For integers $p, q \geq 1$, the GARCH(p,q) model is governed by the equations

$$X(t) = \sigma(t) \cdot Z(t), \quad (7)$$

$$\sigma(t)^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \cdot X(t-i)^2 + \sum_{j=1}^{q} \beta_j \cdot \sigma(t-j)^2 , \quad (8)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1, 2, \ldots, p$ and $\beta_j \geq 0$ for $j = 1, 2, \ldots, q$.
If $p = q = 1$, $\alpha_0 = 0$, $\alpha_1 = 1-\theta$ and $\beta_1 = \theta$ then we have our original model (equation (5)).

The GARCH(1,1) is perhaps the most widely used of all the GARCH models, being relatively simple as well as providing a statistically good model for stochastic volatility.
We can note the following when $p = q = 1$:

- From equation (8), it follows that

$$\sigma(t)^2 = \alpha_0 + (\alpha_1 \cdot Z(t-1)^2 + \beta_1) \cdot \sigma(t-1)^2 . \quad (9)$$

- If $E[\log(\alpha_1 \cdot Z(t-1)^2 + \beta_1)] < 0$ then the model for $X(t)$ is strictly stationary.
- If $\alpha_1 + \beta_1 < 1$ then the model for $X(t)$ is covariance stationary, with unconditional variance $Var[X(t)] = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$.

Estimation of parameters is dealt with in McNeil, Frey and Embrechts [2].

Suppose that we wish to fit the GARCH(1,1) model. One approach to parameter estimation is to use maximum likelihood. We have a set of observations $\{d(t): t = 1, 2, \ldots, n\}$, with parameter vector $\theta = (\alpha_0, \alpha_1, \beta_1, \sigma(1), \mu)$.

Construction of the likelihood function proceeds as follows:

- Let $h(z,\phi)$ be the density function of the i.i.d. $Z(t)$. Recall that $E[Z(t)] = 0$ and $Var[Z(t)] = 1$, so the parameter vector ϕ defined the remaining parameters of the distribution. For example, for *t* distribution, ϕ is simply the number of degrees of freedom. For the non-central *t* distribution ϕ has two elements: the number of degrees of freedom and the non-centrality parameter.
- $\dfrac{(d(1)-\mu)}{\sigma_\theta(1)}$ has density $h(z,\phi)$.
- For *t* = 2, 3, …, *n*:
  - Let $Z(t-1) = \dfrac{(d(t-1)-\mu)}{\sigma(t-1)}$.
  - Define $\sigma_\theta(t)^2 = \alpha_0 + (\alpha_1 \cdot Z(t-1)^2 + \beta_1) \cdot \sigma(t-1)^2$.
  - $d(t)-\mu$ conditional on $\sigma_\theta(1)$ and $d(1), d(2), \ldots, d(t-1)$ has the same density as $d(t)-\mu$ conditional on $\sigma_\theta(t)$: that is, knowledge of $\sigma_\theta(t)$ is sufficient. Thus, $\dfrac{d(t)-\mu}{\sigma_\theta(t)}$ has density $h(z,\phi)$.

The observations are obviously not independent, so we need to build up the likelihood sequentially.

Thus,

$$L(\theta,\phi;d) =$$

$$f(d(1) \mid \sigma_\theta(1)) \cdot \prod_{t=2}^{n} f\big(d(t) \mid \sigma_\theta(1), d(1), \ldots, d(t-1)\big) =$$

$$\prod_{t=1}^{n} \frac{1}{\sigma_\theta(t)} \cdot h\left(\frac{d(t)-\mu}{\sigma_\theta(t)};\phi\right). \qquad (10)$$

Full maximum likelihood (MLE) is implemented by simultaneously maximizing $L(\theta,\phi;d)$ over all elements of θ and ϕ.

As an alternative to full MLE we will take an approximate 2-stage procedure which generally delivers good results and indeed makes the process of finding a good distribution for the *Z*(*t*) easier. This procedure is called *quasi maximum likelihood* (QML).

- Stage 1: Assume that the *Z*(*t*) are i.i.d. *N*(0,1), so that ϕ is empty. Maximize the likelihood $L(\theta,\phi;d)$, over all elements of θ. This stage outputs not just the quasi maximum likelihood estimate for θ but it also outputs a set of standardised residuals $\hat{Z}(1), \hat{Z}(2), \ldots, \hat{Z}(n)$.
- Stage 2: Analyse the $\hat{Z}(t)$ and determine what is the best distribution for them.

We know that the distribution of the daily log returns *d*(*t*), when treated as being a sequence of i.i.d. random variables, exhibited fatter tails than the normal distribution. These apparent fat tails can be caused by a combination of two features:

- stochastic volatility;
- a fat-tailed distribution for *Z*(*t*).

Note that the inclusion of stochastic volatility means that the observed distribution of the *d*(*t*) will have fatter tails than the underlying i.i.d. *Z*(*t*).

Under QML the estimated GARCH(1,1) parameters are (using **R** commands):

```
> res1<-fit.garch11.normal(d)
14561.43  # maximum likelihood #
> res1$par    # The vector of parameters #
sigma(1)          alpha0             alpha1
beta1           mu
1.253475e-02   3.464359e-06   1.303942e-01
8.552684e-01 5.963464e-04
```

This results in an unconditional standard deviation ($\alpha_1 + \beta_1 = 0.9856626 < 1$) for *d*(*t*) equals to $\sqrt{\dfrac{\alpha_0}{1-(\alpha_1+\beta_1)}}$ = 0.0155445 per trading day or 24.67% per annum (252 trading days).

Standardised residuals, *Z*(*t*) are plotted in Figure 8., where we cannot see any obvious clusters of high and low volatility. The horizontal dashed lines give the 1% and 99% quantiles of standard normal. There seems to be more exceedances of the lower threshold than the upper, and these seem to be larger.

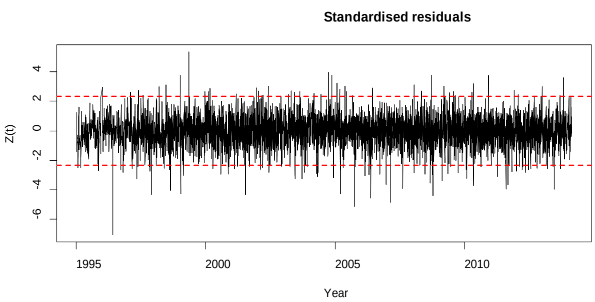

Fig. 8: Daily standardised residuals, *Z*(*t*), for the stochastic volatility model. Source: Own calculation

We now proceed onto an analysis of the *Z*(*t*) on the assumption that they are i.i.d..

First we are looking at the possibility that *Z*(*t*) are normally distributed. Two graphical diagnostics are provided in Figure 9.

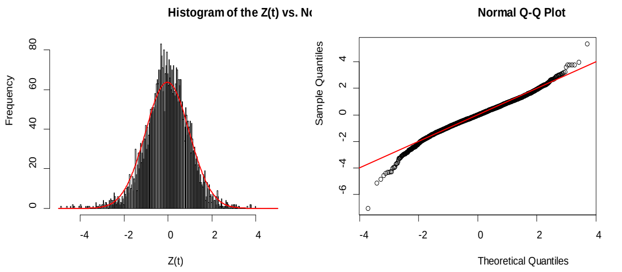

Fig. 9: Left: Empirical histogram of the $Z(t)$ (bars) with a fitted Normal overlaid (solid line). Right: Q-Q plot of standard normal quantiles vs. empirical quantiles. Source: Own calculation

We can compare the histogram of the $Z(t)$ (left-hand plot) with Figure 3., and conclude that the $Z(t)$ appear to be closer to a normal distribution than the original $d(t)$. The kurtosis of the $d(t)$ is 14.687 reducing to 4.862 for the $Z(t)$. However, the histogram still provides evidence that the data have a narrower peak and (by inference) fatter tails than the standard normal.

The Q-Q plot (right-hand plot) leads us to a similar conclusion, that the Normal distribution is better than before, but that it still does a bad job in modelling the tails of the data. Specifically both tails in the data are fatter than the Normal, especially the left-hand tail. The Q-Q plot also shows some skewness in the data, and, indeed, the coefficient of skewness is – 0.27, suggesting a long left-hand tail.

These graphical diagnostics can be backed up by formal hypothesis tests.

```
> jbtest(ZQ)  # ZQ is the vector of
              standardized residuals Z(t)
   Skewness = -0.2725932
   Kurtosis = 4.862287
Jarque-Bera = 751.0284
    p-value = 0
> chi2test.normal(ZQ)
Chi-squared statistic is 181.686 with 97
degrees of freedom.
The p-value for this is 4.221595e-07.
```

The Jarque-Bera test results in rejection of the i.i.d. normal hypothesis. The Chi-squared test compares the number of observations in 100 bands, each with probability 0.01 under H_0. The test statistic of 181.686 is not disastrously high, but it is big enough to result in a very low p-value, so again H_0 is rejected.

The well-known alternative distribution on the real line to the normal is the t distribution.

Suppose that

- Z and Y are independent random variables,
- $Z \sim N(0,1)$ has a standard normal distribution,
- $Y \sim \chi^2_\nu$ has a standard chi-squared distribution with ν degrees of freedom,
- The random variable X is defined as $X = \dfrac{Z}{\sqrt{\dfrac{Y}{\nu}}}$,

then the random variable X has a standard t distribution with ν degrees of freedom, and its probability density function equals

$$f(x) = \frac{\Gamma\left(\frac{1}{2}(\nu+1)\right)}{\sqrt{\pi\nu}\cdot\Gamma\left(\frac{1}{2}\nu\right)} \cdot \left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

By the method of moments we can estimate parameters of the t distribution by matching specified moments of the distribution to the sample moments. Our model is that the volatility standardised residuals, $Z(t)$, are given by $Z(t) = m + s\cdot X(t)$, where the $X(t)$ are i.i.d. standard t random variables with ν degrees of freedom. From the moments of the standard t distribution we can infer that

$$E[Z(t)] = m, \tag{11}$$

$$Var[Z(t)] = s^2 \cdot \frac{\nu}{\nu-2}, \text{ (for } \nu > 2) \tag{12}$$

$$Kurtosis[Z(t)] = \frac{3\cdot(\nu-2)}{\nu-4}, \text{ (for } \nu > 4). \tag{13}$$

Thus we use the sample mean, variance and kurtosis to estimate m, s and ν.

We have the following statistics:

$E[Z(t)] = -0.0300086$, $Var[Z(t)] = 0.999772$, $k = 4.862287$.

Matching these three moments we get:

$\hat{\nu} = 7.221846$, $\hat{s} = 0.850145$, $\hat{m} = -0.0300086$.

Under maximum likelihood the estimated t distribution parameters are (using **R** command mle.t):

```
> mle.t(ZQ)
mu    = -0.01648076
sigma = 0.8535201
nu    = 7.442079
log-likelihood = -6696.788
[1] -0.01648076  0.85352007  7.44207945
```

The summary Table 1. contains the log-likelihood for three methods we have investigated.

It is clear that the t distribution fits much better than the normal. Additionally, we can see that the method of moments produces a worse fit using log-likelihood as a measure.

Previously, in the Jarque-Bera test, we saw that the empirical coefficient of skewness is – 0. 2725932, so it makes sense to investigate some skewed, fat tailed distributions. There are many such distributions to choose. The one we will investigate is called the *non-central t distribution* (NCT).

Suppose that

- $D \in$ R is some constant
- Z and Y are independent random variables,
- $Z \sim N(0,1)$ has a standard normal distribution,
- $Y \sim \chi^2_\nu$ has a standard chi-squared distribution with ν degrees of freedom,
- The random variable X is defined as $X = \dfrac{Z+D}{\sqrt{\dfrac{Y}{\nu}}}$,

then the random variable X has a non-central t distribution with ν degrees of freedom and non-centrality parameter, D.

The NCT distribution is now fitted to the volatility standardised residuals (using **R** command mle.nct):

```
> pv3<-mle.nct(ZQ)
mu    = 0.2815168
sigma = 0.8533425
nu    = 7.637477
ncp   = -0.3267568
log-likelihood = -6692.568
```

Maximum likelihood estimates for the parameter estimates of the normal and t distributions are repeated in the Table 2. along with those for the non-central t.

The normal distribution is a special case of the t which in turn is a special case of the NCT, so we can see an increase each time in the log-likelihood.

IV. Fitting Extreme Values

International banking regulations require banks to pay specific attention to the probability of large losses over short periods of time (typically 1 or 10 trading days). More generally we may wish to pay specific attention to the possibility of large gains or losses for general risk management purposes.[2]

We focus here on analysis of the tails of the $Z(t)$ resulting from the QML estimation used in fitting a GARCH(1,1) model to the daily log returns data.

In Figures 11. to 13. we take the $Z(t)$ output by the QML approach, plot the empirical CDF for the data and compare this with the fitted CDF's for normal, t and NCT distributions.

In Figures 12. and 13. we have zoomed in on the 5% and 1% left and right hand tails of the cumulative distribution. The left-hand plots give us information about the probability of large losses and the right-hand plots about the probability of large gains on the index.

In general, the t and NCT look rather better than the normal. The NCT generally seems better than the t, but by a smaller margin. Skewness in the empirical data shows up in fatter left-hand tails, and this means that the t and NCT are better.

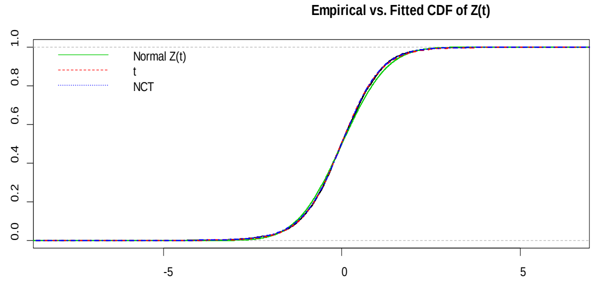

Fig. 11: Normal, t and NCT distributions fitted to QML $Z(t)$. Source: Own Processing

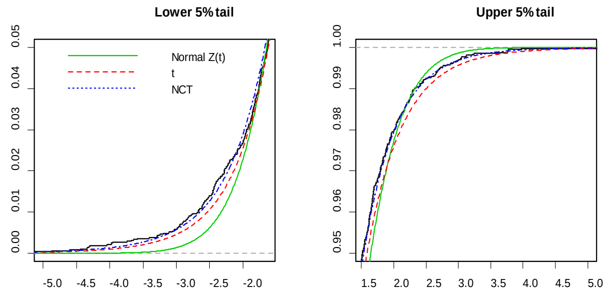

Fig. 12:. Detail of Figure 11. Source: Own Processing

To avoid problems with the tails a good compromise is to use the above mentioned distributions within the main body of the data (say between 5% and 95% quantiles) and to fit each tail data separately using a standard distribution. We will refer to this as the *hybrid approach*. [4]

Above the 95% quantile and below the 5% quantile we will fit a Pareto distribution to the excess returns over the 95% quantile and below the 5% quantile, respectively.

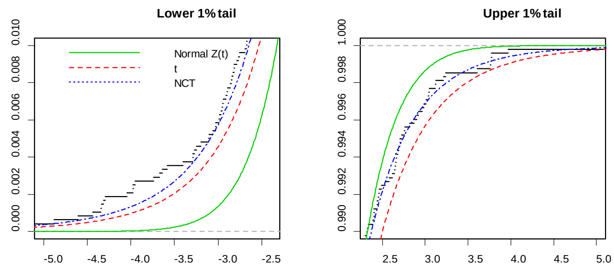

Fig. 13: Detail of Figure 11. Source: Own Processing

Random variable X has a Pareto distribution [3] with parameters $\lambda > 0$ and $\alpha > 0$ if it has the pdf

$$f(x) = \frac{\alpha\lambda^\alpha}{(\lambda + x)^{\alpha+1}}, \text{ for } x > 0. \tag{14}$$

Its cumulative function is

[2] Large gains in a stock-market index can cause losses, e.g. for banks that have sold call options on that index.

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x} \right)^{\alpha}, \text{ for } x > 0. \qquad (15)$$

Now suppose that we have n observations in total. The 95% quantile will be denoted by q. Further suppose that there are m observations out of the n that exceed q, and that these take the values $x_1, x_2, \ldots, x_m$.

Now fit a Pareto distribution using the method of moments or maximum likelihood to the excess returns $y_i = x_i - q$, returning the parameter estimates $\hat{\alpha}$ and $\hat{\lambda}$. [1]

For $x > q$ the cumulative distribution function is then:

$$F(x) = 1 - 0.05 \cdot \left(\frac{\hat{\lambda}}{\hat{\lambda} + (x - q)} \right)^{\hat{\alpha}}. \qquad (16)$$

For the data below the 5% quantile we will follow similarly, but for $x < q$ the cumulative distribution function is $F(x) = 0.05 \cdot \left(\frac{\hat{\lambda}}{\hat{\lambda} + (q - x)} \right)^{\hat{\alpha}}$.

Figure 14. shows results for fitting a Pareto distribution to the lower 5% and 1% of the data, and to the upper 5% and 1% of the data.

Estimated values for λ and α are given in Table 3. For the lower 1% we saw that the estimated value of α is 19.84, while the upper 1% tail has $\hat{\alpha}$ = 155.55. This means that the left-hand tail is fatter, contrasting with a relatively thin right-hand tail. This means that the right-hand tail of the data we can smoothly fit by NCT distribution.

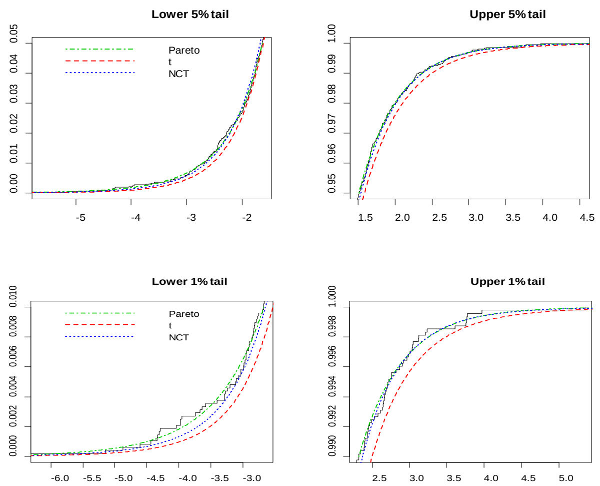

Fig. 14: Upper and lower tail characteristics of the empirical and fitted CDF's. Source: Own Processing

V. Conclusion

For our analysis we proposed a model which incorporates stochastic volatility. One of the most used model for daily return series is the GARCH model. Under QML we estimated GARCH(1,1) parameters and obtained standardised residuals. Standardised residuals do not show any clusters of high and low volatility. We took them as i.i.d.

The GARCH(1,1) model is the first and foremost a model for short-term risk assessment. Longer-term predictions will be less reliable.

Analysis of the standardised residuals showed that the non-central t distribution (NCT) fits them much better than the t distribution or the normal distribution.

After all we analysed the tails of the standardised residuals. Above 95% quantile and below the 5% quantile we used Pareto distribution for fitting.

We can see that Pareto distribution provides a generally better fit over the tails than t and non-central t distribution.

We conclude this paper by looking at Q-Q plots of the tails of the data versus the theoretical Pareto distribution. These are plotted in Figure 15. for both the 5% and 1% tails. In all cases the Q-Q plot looks reasonably linear suggesting that the Pareto is an appropriate choice for modelling excess gains and losses.

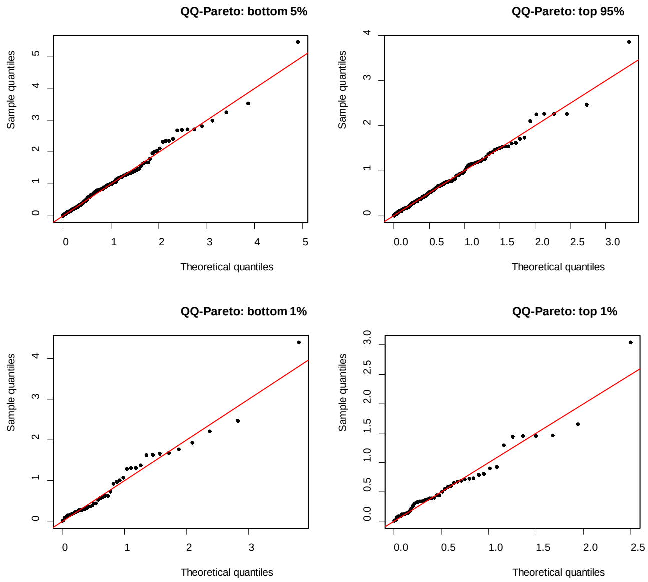

Fig. 15: Q-Q plot for the excess losses and gains for the 1%, 5%, 95% and 99% quantiles versus theoretical Pareto quantiles. Source: Own Processing

Acknowledgments

This research could be performed due to research grant: SGSFES_2015001 Ekonomický a sociální rozvoj v soukromém a veřejném sektoru

References

[1] Jindrová P., Jakubinský R., *Significance and Possibilities of Major Accident Insurance*, E+M Ekonomie a management, XVII, no. 4., p. 121-131, ISSN 1212-3609, 2015.

[2] McNeil A. J., Frey R., Embrechts P., *Quantitative Risk Management*, Princeton University Press, Princeton and Oxford,, ISBN 13: 978-0-691-12255-7, 2005.
[3] Pacáková V., Gogola J., *Pareto distribution in Insurance and Reinsurance*, 9th International Scientific Conference Financial Management of Firms and Financial Institution Ostrava, p. 648-657, ISBN 978-80-248-317, 2013.
[4] Pacáková V., Linda B., *Simulation of Extreme Losses in Non-life insurance*, E+M Ekonomie a management, XII, no. 4., p. 97-103, ISSN 1212-3609, 2009.
[5] www.bloomberg.com/quote/PX:IND
[6] www.pse.cz/dokument.aspx?k=Burzovni-Indexy

RNDr. Ján Gogola, Ph.D. graduated in Mathematics (2000) at Comenius University in Bratislava (Slovakia), Faculty of Mathematics, Physics and Informatics.
2002 - RNDr. in Mathematical Analysis at the same university, degree Ph.D (2008) at the same university.
He was working at Department of Mathematics, Faculty of Economic Informatics, University of Economics in Bratislava since 2004 to 2013. At the present he has been working at Faculty of Economics and Administration in Pardubice (Czech Republic) since 2013. RNDr. Gogola has been concentrated on actuarial science and management of financial risks.

Mgr. Ondřej Slavíček graduated in Applied Mathematics (2008) at Palackého University Olomouc.
At the present he has been working at Faculty of Economics and Administration in Pardubice since 2008.
Mgr. Slavíček has been concentrated on actuarial science and management of financial risks.

Estimate	*m*	*s*	*ν*	log-likelihood
Normal: MLE	– 0.0300086	0.999886	∞	– 6791.914
t: Method of moments	– 0.0300086	0.850145	7.221846	– 6706.331
t: MLE	– 0.0164807	0.853520	7.442079	– 6696.788

Table 1: Log-likelihood of normal and *t* distribution. Source Own calculation

Estimate	*m*	*s*	*ν*	NCP	log-likelihood
Normal: MLE	– 0.0300086	0.999886	∞		– 6791.914
t: MLE	– 0.0164807	0.853520	7.442079		– 6696.788
NCT: MLE	0.2815168	0.8533425	7.637477	– 0.3267	– 6692.568

Table 2: Log-likelihood of normal, *t* distribution and non-central *t* distribution. Source: Own calculation

Tail	Volatility-standardised residuals, *Z*(*t*)		
	Quantile cut-off	$\hat{\lambda}$	$\hat{\alpha}$
5% worst losses	– 1.631937	8.140316	12.95973
1% worst losses	– 2.673547	14.88173	19.84448
5% top gains	1.511473	81.37704	153.8542
1% top gains	2.319732	90.08061	155.5489

Table 3: Estimated values for λ and α for the Pareto tail distributions. Source: Own calculation