

Dynamic Value at Risk Estimation for BELEX15

Emilija Nikolić-Đorić¹ and Dragan Đorić²

Abstract

This paper uses RiskMetrics, GARCH and IGARCH models to calculate daily VaR for Belgrade Stock Exchange index BELEX15 returns based on the normal and Student t innovation distribution. In the case of GARCH and IGARCH models VaR values are obtained applying Extreme Value Theory on the standardized residuals. The Kupiec's LR statistics was used to test the accuracy of risk measurement models. The main conclusions are: (1) when modelling value-at-risk it is very important to have a good model for volatility of stock returns; (2) both stationary and integrated GARCH models outperform RiskMetrics in estimating VaR; (3) although long memory volatility is present in the BELEX15 index, IGARCH models cannot outperform GARCH type models in VaR evaluations for this index.

1 Introduction

Value at Risk (VaR) is a commonly used statistic for measuring potential risk of economic losses in financial markets (Jorion, 1997; Duffie and Pan, 1997; Dowd, 2002; Giot and Laurent, 2003). Using VaR financial institutions can calculate the minimum amount that is expected to lose with a small probability α over a given time horizon k , usually 1-day or 10 days. Typical values for α are 0.1, 0.05 and 0.01. Empirical VaR calculations involve the estimation of lower-order quantiles, for example 10%, 5% or 1% of the return distribution. From a statistical viewpoint, the VaR is nothing else but a given percentile of the profit and loss (P&L) distribution over a fixed horizon. While VaR concept is easy to understand, its measurement is a very challenging statistical problem. Risk analysis can be done in two stages. First, we can express profit-and-loss in terms of returns and then statistically model the returns and estimate the VaR of returns by computing appropriate quantile.

The main problem is related to the estimation of distribution that adequately describes the returns (Đorić and Nikolić-Đorić, 2010). It is well known that the probability distribution of stock returns is fat tailed, which means that extreme price movements occur much

¹ Department of Statistics, Faculty of Agriculture, University of Novi Sad, 21000 Novi Sad, Serbia; emily@polj.uns.ac.rs

² Department of Mathematics, Faculty of Organizational Sciences, University of Belgrade, 11000 Beograd, Serbia; djoric@fon.bg.ac.rs

more often than predicted by a Gaussian model (Fama, 1965). The empirical finding that series of returns often exhibit volatility clustering has led to the development of a variety of univariate time series methods for volatility forecasting.

The simplest approach to volatility forecasting is to estimate the variance as a simple moving average of past squared shocks. The popular GARCH class of models (Bollerslev, 1986) rely on the assumption that the shape of the conditional distribution of returns is fixed over time. The most recent approach to volatility modeling and forecasting is a concept named realized volatility (McAleer and Medeiros, 2008). Models of the volatility allow considering conditional distributions of returns in calculating VaR (Angelidis et al., 2004; So and Yu, 2006; Đorić and Nikolić-Đorić, 2006). Besides the heavy-tailed issue, asymmetry distribution is also often observed in financial time series. This property is very important in risk analysis where the long and short position investments over a given time period relied heavily on the lower and upper tails behaviours. Barndorff-Nielsen (1997) implemented skewed distributions that allowed upper and lower tails to have dissimilar behaviours. Also, recent empirical studies found that many financial return series may exhibit long memory (Christensen and Nielsen, 2007).

In a VaR context, precise prediction of the probability of an extreme movement in a value is essential for risk management. As extreme movements are related to the tails of the distribution of the underlying data generating process, instead of forcing a single distribution for the entire sample, it is possible to investigate only the tails of the return distributions. One such approach to estimating value at risk is the extreme value theory (EVT) which provides some assistance in improving the traditional VaR models (Embrechts et al., 1997; Gençay and Selçuk, 2004).

In recent years there has been a lot of research on VaR estimation of different return series (Giot and Laurent, 2003; Sarma et al., 2003; Huang and Lin, 2004; Kuester et al., 2006), but research papers dealing with VaR calculation in the financial markets EU new member states are very rare. Žiković (2007a) applied VaR methodology and historical simulation on the Croatian stock market indexes in an effort to measure Value-at-Risk. He also (Žiković, 2007) analysed VaR models for ten national indexes: Slovenia - SBI20, Poland - WIG20, Czech Republic - PX50, Slovakia - SKSM, Hungary - BUX, Estonia - TALSE, Lithuania - VILSE, Latvia - RIGSE, Cyprus - CYSMGENL, Malta - MALTEX and concluded that use of common VaR models to forecast VaR is not suitable for transition economies. Also, there is a study on forecasting volatility of the Macedonian Stock Exchange indexes (Kovačić, 2008).

As Serbia is expected to join the European Union, the stylized facts of its stock market could be of interest to potential domestic and foreign investors. The goal of the analysis in this paper was to obtain stylized facts of Belgrade stock exchange series and compare its stochastic behavior with known behavior of stock exchange series of developed economies. The stylized facts are base for choosing the time series models that can be used for estimating volatility and computing VaR as a measure of risk. The relative performance of several alternative VaR models on Belgrade Stock Exchange index BELEX15 was investigated by means of Kupiec likelihood ratio test.

The paper is organized as follows. Section 2 describes the basic concept of VaR and presents various models for VaR. In 2.1 dynamic EWMA models are presented, while in 2.2 GARCH type models and POT model for extreme values are combined. Evaluating VaR model adequacy is given in Section 3. Section 4 presents empirical results obtained

by applying described models to stock index BELEX15. Some concluding remarks are given in Section 5.

2 VaR models

Let P_t be the price of a financial asset on day t . A k -day VaR of a long position on day t is defined by

$$P(P_t - P_{t-k} \leq \text{VaR}(t, k, \alpha)) = \alpha, \quad (2.1)$$

where $\alpha \in (0, 0.5)$. Similarly, a k day VaR of a short position is defined by

$$P(P_t - P_{t-k} \geq \text{VaR}(t, k, \alpha)) = \alpha. \quad (2.2)$$

Holder of the long position suffers a loss when $\Delta_k P_t = P_t - P_{t-k} < 0$, while a holder of a short position loses when $\Delta_k P_t > 0$. Given a distribution of the continuously compounded return $\log(P_t) - \log(P_{t-k})$, VaR can be determined and expressed in terms of a percentile of the return distribution (Dowd, 2002). If q_α is the α th percentile of the return, then VaR of a long position can be written as

$$\text{VaR}(t, k, \alpha) = (e^{q_\alpha} - 1)P_{t-k}. \quad (2.3)$$

From (2.3) it can be seen that good VaR estimates can be produced with accurate forecasts of the percentiles q_α . So, in further we consider only VaR for return series.

We define the 1-day logarithmic return (in further text just return) on day t as

$$r_t = \log(P_t) - \log(P_{t-1}) \quad (2.4)$$

and denote the information up to time t by \mathcal{F}_t . That is, for a time series of returns r_t , VaR is such that

$$P(r_t < \text{VaR}_t | \mathcal{F}_{t-1}) = \alpha. \quad (2.5)$$

From this, it is clear that finding a VaR essentially is the same as finding a $100\alpha\%$ conditional quantile. For convention, the sign is changed to avoid negative number in the VaR. Formally, it is possible to model process for the stock returns r_t as follows

$$r_t = \mu_t + \epsilon_t, \quad \epsilon_t = \sigma_t \eta_t, \quad \mu_t = \mu(\mathcal{F}_{t-1} | \theta), \quad \sigma_t^2 = \sigma^2(\mathcal{F}_{t-1} | \theta), \quad (2.6)$$

where \mathcal{F}_{t-1} is the information set available at time $t - 1$, and where μ and σ are functions known up to a finite-dimensional vector of parameter values θ . In this model ϵ_t is the innovation, σ_t is unobservable volatility and η_t is a martingale difference sequence satisfying

$$E(\eta_t | \mathcal{F}_{t-1}, \theta) = 0, \quad V(\eta_t | \mathcal{F}_{t-1}, \theta) = 1. \quad (2.7)$$

As a consequence, we have

$$E(r_t | \mathcal{F}_{t-1}, \theta) = \mu_t, \quad V(r_t | \mathcal{F}_{t-1}, \theta) = \sigma_t^2, \quad \epsilon_t | \mathcal{F}_{t-1} \sim D(0, \sigma_t^2), \quad (2.8)$$

where $D(0, \sigma_t^2)$ represents a conditional distribution with zero mean and variance σ_t^2 .

If the return can be modeled by a parametric distribution, VaR can be derived from the distributional parameters. Unconditional parametric models set $\mu_t = \mu$ and $\sigma_t =$

σ , thereby assuming that the returns are i.i.d. (independent identically distributed) with density given by

$$f_r(x) = \frac{1}{\sigma} f_{r^*} \left(\frac{x - \mu}{\sigma} \right), \quad (2.9)$$

with f_r being density function of the distribution of r_t and f_{r^*} being density function of the standardized distribution of r_t .

RiskMetrics (Risk Metrics, 1996) is the most simple, but still one of the most used models to compute VaR and is considered in 2.1. Most of the VaR methodologies are GARCH type models. In 2.2 GARCH and IGARCH models are applied together with POT method of extreme value theory (EVT). Both conditional normal and conditional t-error distributions are considered.

2.1 Dynamic VaR using EWMA

Since the risk management group at J.P. Morgan developed the RiskMetrics model for measuring VaR in 1994, RiskMetrics has become a benchmark for measuring market risk. The RiskMetrics model (Risk Metrics, 1996) assumes a dynamic EWMA model for the variance,

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda)(r_t - \mu)^2. \quad (2.10)$$

To initialize the recursive variance equations, the sample variance is used

$$\hat{\sigma}_1^2 = \frac{1}{n-1} \sum_{i=1}^n (r_i - \hat{\mu})^2, \quad (2.11)$$

where, following RiskMetrics, we set the smoothing parameter λ to 0.94. The model (2.10) can be written as an exponentially weighted moving average (EWMA) of the past squared innovations or returns,

$$\sigma_t^2 = (1 - \lambda)(r_{t-1}^2 + \lambda r_{t-2}^2 + \lambda^2 r_{t-3}^2 + \dots). \quad (2.12)$$

The smaller the smoothing parameter, the greater the weight is given to recent return data.

Model with normal conditional distribution

If we assume that the conditional distribution of returns is normal with mean zero,

$$r_t = \epsilon_t, \quad \epsilon_t | \mathcal{F}_{t-1} \sim N(0, \sigma_t^2), \quad (2.13)$$

then the 1-day VaR on day t is reduced to $(1 - e^{q\alpha})P_{t-1}$ or approximately $-\sigma_t z_\alpha P_{t-1}$, where z_α is the 100α th percentile of $N(0, 1)$. The dynamic VaR for return r_t at time t for long trading positions is given by

$$\text{VaR}_t^{\text{long}} = \hat{\mu} + z_\alpha \hat{\sigma}_t = \hat{\mu} + \Phi^{-1}(\alpha) \hat{\sigma}_t \quad (2.14)$$

and for short trading positions it is equal to

$$\text{VaR}_t^{\text{short}} = \hat{\mu} + z_{1-\alpha} \hat{\sigma}_t = \hat{\mu} - \Phi^{-1}(\alpha) \hat{\sigma}_t. \quad (2.15)$$

Model with t conditional distribution

If we assume that the conditional distribution of returns is Student t with mean zero, then the 1-day VaR for long trading positions is calculated as

$$\text{VaR}_t^{\text{long}} = \hat{\mu} + st_{\alpha, \nu} \hat{\sigma}_t = \hat{\mu} + T_\nu^{-1}(\alpha) \hat{\sigma}_t \sqrt{\frac{\nu - 2}{\nu}} \quad (2.16)$$

and for short trading positions it is equal to

$$\text{VaR}_t^{\text{short}} = \hat{\mu} + st_{1-\alpha, \nu} \hat{\sigma}_t = \hat{\mu} - T_\nu^{-1}(\alpha) \hat{\sigma}_t \sqrt{\frac{\nu - 2}{\nu}}, \quad (2.17)$$

with $st_{\alpha, \nu}$ being the left quantile at $\alpha\%$ and T being the distribution function for the standardized Student distribution with estimated number of degrees of freedom ν . For the same distribution $st_{1-\alpha, \nu}$ is the right quantile at $\alpha\%$ and the equality $st_{\alpha, \nu} = -st_{1-\alpha, \nu}$ holds.

2.2 Dynamic VaR using GARCH type models and EVT

To obtain value-at-risk estimates two-step estimation procedure called conditional EVT is followed (McNeil and Frey, 2000). First a GARCH-type model is fitted to the return data by quasi-maximum likelihood. Then, we consider the standardized residuals to be realizations of a strict white noise process and use extreme value theory (EVT) to model the tail of innovations using EVT and estimate the quantiles of innovations. Using ARMA and GARCH type models for r_t it is possible to get the conditional mean μ_t , conditional variance σ_t^2 and innovation series ϵ_t . Then, Extreme Value Theory (EVT) calculations are applied on the standardized innovations.

ARMA model for μ_t

Time-varying conditional mean can be captured by ARMA model

$$\mu_t = a_0 + \sum_{i=1}^p a_i r_{t-i} + \sum_{j=1}^q b_j \epsilon_{t-j} \quad (2.18)$$

with $t = 1, \dots, T$. Family of ARMA models are suitable for modelling covariance-stationary processes. Zero-mean ARMA(p, q) model is given by

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \epsilon_t + \sum_{j=1}^q b_j \epsilon_{t-j}, \quad (2.19)$$

where (ϵ_t) is white noise, also known as the process of innovations. Proces (X_t) is an ARMA process with mean μ if the centered series $(X_t - \mu)$ is a zero-mean ARMA process. If the innovations are i.i.d., or themselves form a strictly stationary process, then the ARMA process will also be strictly stationary. ARMA process is causal if X_t has representation of the form

$$X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad (2.20)$$

where coefficients ψ_i are absolutely summable. Every causal ARMA process is covariance stationary (see Proposition 4.9 in McNeil et al., 2005).

GARCH model for σ_t^2

By GARCH type models we model σ_t^2 . The GARCH(p, q) model for integers $p > 0$ and $q \geq 0$ is defined as

$$\sigma_t^2 = \omega + \beta(B)\sigma_t^2 + \alpha(B)\epsilon_t^2, \quad (2.21)$$

where $\omega > 0$, α and β are lag polynomials,

$$\alpha(B) = \alpha_1 B + \dots + \alpha_q B^q, \quad \beta(B) = \beta_1 B + \dots + \beta_p B^p,$$

with $\alpha_i \geq 0$ for $i = 1, \dots, q$ and $\beta_j \geq 0$ for $j = 1, \dots, p$. The variance equation (2.21) of the GARCH(p, q) process can be expressed as

$$\sigma_t^2 = \frac{\omega}{1 - \beta(B)} + \frac{\alpha(B)}{1 - \beta(B)} \epsilon_t^2. \quad (2.22)$$

Note that the RiskMetrics model can be viewed as a special case of GARCH(1,1) model with $\mu = \omega = 0$ and $\lambda = \beta_1 = 1 - \alpha_1$. Also, the variance equation of the GARCH model can be written in ARMA form for ϵ_t^2 ,

$$(1 - \alpha(B) - \beta(B))\epsilon_t^2 = \omega + (1 - \beta(B))v_t, \quad v_t = \epsilon_t - \sigma_t^2, \quad (2.23)$$

where v_t are martingale differences and usually are interpreted as the innovations to the conditional variance.

The GARCH model was proposed by Bollerslev (1986) and it successfully captures several characteristics of financial time series, such as heavy tailed returns and volatility clustering. In empirical investigation we estimate VaR and determine the density of the one-day ahead VaR under the GARCH(1,1) dynamics with both Normal and Student-t disturbances. It is known (Bollerslev, 1986) that the GARCH process is covariance stationary if and only if $\alpha_1 + \beta_1 < 1$.

IGARCH model for σ_t^2

In some applications of GARCH models, the estimated lag polynomial $1 - \alpha(B) - \beta(B)$ is found to have a significant unit root. Factoring this polynomial as $(1 - B)\phi(B)$, where $\phi(B)$ has all the roots outside the unit circle, Engle and Bollerslev (1986) proposed the following integrated GARCH, or IGARCH(p, q) model

$$\phi(B)(1 - B)\epsilon_t^2 = \omega + (1 - \beta(B))v_t, \quad v_t = \epsilon_t^2 - \sigma_t^2, \quad (2.24)$$

where

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_q B^q.$$

The variance equation can be rewritten as

$$\sigma_t^2 = \frac{\omega}{1 - \beta(B)} + \frac{(1 - B)(1 - \phi(B))}{1 - \beta(B)} \epsilon_t^2. \quad (2.25)$$

When $\mu = 0$, the IGARCH(1,1) model reduces to RiskMetrics model with $\lambda = \beta_1$. Therefore, IGARCH(1,1) can be at least as good model as RiskMetrics is. In addition, parameter β_1 can be estimated from the data. Also, assuming $\mu = 0$ in RiskMetrics may generate biased VaR estimates. GARCH type models allow estimation of parameter μ .

Estimation of GARCH-type models

For the error term ϵ_t two conditional distributions $D(0, \sigma_t^2)$ were considered: (1) normal distribution $N(0, \sigma_t^2)$; (2) a standardized t distribution with ν degrees of freedom and variance σ_t^2 . The parameters under GARCH and IGARCH models with normal and t -distributed errors can be estimated by the maximum likelihood method. Given a time series of returns r_t over a period of T days, ($t = 1, \dots, T$) the maximum likelihood estimate of μ under any GARCH-type model is the sample mean

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T r_t.$$

With $\mu = \hat{\mu}$, in the case of normal errors, the log-likelihood function is

$$l_n = -\frac{1}{2}T \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \left[\ln \sigma_t^2 + \frac{\hat{\epsilon}_t^2}{\sigma_t^2} \right], \quad (2.26)$$

where $\hat{\epsilon}_t = r_t - \hat{\mu}$ and where σ_t is given in (2.21).

Similarly, the log-likelihood function under t errors is given by

$$l_t = T \ln \frac{\Gamma(\nu/2 + 1/2)}{\sqrt{\pi(\nu - 2)}\Gamma(\nu/2)} - \frac{1}{2} \sum_{t=1}^T \left[\ln \sigma_t^2 + (\nu + 1) \ln \left(1 + \frac{\hat{\epsilon}_t^2}{(\nu - 2)\sigma_t^2} \right) \right], \quad (2.27)$$

where σ_t^2 is given in (2.25). Once the maximum likelihood estimates of the model parameters in GARCH and IGARCH models are obtained, we can forecast the 1-day ahead variance σ_t^2 .

The POT method for residuals

We fit GARCH(1,1) model to r_t to get the estimates for conditional mean μ_t , conditional variance σ_t^2 and innovation ϵ_t series. Then, Extreme Value Theory (EVT) calculations are applied on the estimates of standardised innovations η_t (Embrechts et al., 1997).

Suppose we have a sequence of i.i.d. observations X_1, \dots, X_n from an unknown distribution function F . We are interested in excess losses over a high threshold u . We define the distribution function of the excesses over the threshold u by

$$F_u(x) = P(X - u \leq x \mid X > u) = \frac{F(x + u) - F(u)}{1 - F(u)}, \quad 0 \leq x < x_0 - u, \quad (2.28)$$

where x_0 is right endpoint of the distribution F . For modelling distribution of F_u one can use the generalized Pareto distribution (GPD) which is usually expressed as a two parameter distribution with distribution function

$$G_{\xi,a}(x) = \begin{cases} 1 - (1 + \xi x/a)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-x/a) & \text{if } \xi = 0. \end{cases} \quad (2.29)$$

This distributional choice is motivated by a theorem which states that, for a certain class of distributions, the GPD is the limiting distribution for the distribution of the excesses. Using trial-and-error, a negative threshold is chosen, below which the empirical mean excess function of the tail distribution of standardised residuals starts increasing (in absolute value) linearly. This is called threshold u . Let the number of points in the tail distribution corresponding to u be n_u . We fit a Generalised Pareto Distribution (GPD) to the tail. For points in the tail of the distribution it can be noted that

$$F(x) = (1 - P(X \leq u))F_u(x - u) + P(X \leq u). \quad (2.30)$$

It is now known that $F_u(x - u)$ can be estimated by $G_{\xi,a}(x - u)$ for u large. We can also estimate $P(X \leq u)$ from the data by the empirical distribution function $F_n(u)$. As $F_n(u) = 1 - n_u/n$, for $x \geq u$ we have tail estimate

$$\widehat{F}(x) = (1 - F_n(u))G_{\xi,a}(x - u) + F_n(u) = 1 - \frac{n_u}{n} \left(1 + \xi \frac{x - u}{\widehat{a}}\right)^{-1/\widehat{\xi}}. \quad (2.31)$$

The POT estimator of x_p is obtained from (2.31) by solving equation $\widehat{F}(\widehat{x}_p) = p$ and substituting unknown parameters of the GPD by estimates $\widehat{\xi}$ and \widehat{a} . As $\text{VaR}_p = x_p$, the VaR for the tail of the standardised residuals is found using the expression

$$\text{VaR}_\eta = \widehat{x}_p = u + \frac{\widehat{a}}{\widehat{\xi}} \left(\left(\frac{n}{n_u} (1 - p) \right)^{-\widehat{\xi}} - 1 \right). \quad (2.32)$$

VaR by conditional EVT

Once the one-step-ahead forecasts of the conditional mean $\widehat{\mu}_t$ and conditional variance $\widehat{\sigma}_t$ are computed conditionally on past information \mathcal{F}_t and $\text{VaR}_\eta^{\text{long}}$ and $\text{VaR}_\eta^{\text{short}}$ estimated, VaR for returns can be calculated. The dynamic VaR for return r_t at time t computed in $t - 1$ for long trading positions is given by

$$\text{VaR}_t^{\text{long}} = \widehat{\mu}_t - \widehat{\sigma}_t \text{VaR}_\eta^{\text{long}} \quad (2.33)$$

and for short trading positions it is equal to

$$\text{VaR}_t^{\text{short}} = \widehat{\mu}_t + \widehat{\sigma}_t \text{VaR}_\eta^{\text{short}}. \quad (2.34)$$

3 Evaluating VaR model adequacy

Various methods and tests have been suggested for evaluating VaR model accuracy. Performance of the VaRs for different pre-specified level α can be evaluated by computing their failure rate for the price series. Statistical adequacy could be tested based on Kupiec likelihood ratio test which examines whether the average number of violations is statistically equal to the expected rate.

3.1 Failure rate

The failure rate is widely applied in studying the effectiveness of VaR models. The definition of failure rate is the proportion of the number of times the observations exceed the forecasted VaR to the number of all observations. If the failure rate is very close to the pre-specified VaR level, it could be concluded that the VaR model is specified very well.

3.2 Kupiec likelihood ratio test

For the purpose of testing VaR models more precisely, the Kupiec LR test is adopted to test the effectiveness of our VaR models. A likelihood ratio test developed by Kupiec (1995) will be used to find out whether a VaR model is to be rejected or not. The number n of VaR violations in a sample of size T has a binomial distribution, $n \sim B(T, p)$. The failure rate is n/T and, ideally, it should be equal to the left tail probability, p . The null H_0 and alternative H_1 hypotheses are:

$$H_0 : \frac{n}{T} = p, \quad H_1 : \frac{n}{T} \neq p, \quad (3.1)$$

where

$$p = P(r_t < \text{VaR}_p \mid \mathcal{F}_{t-1}) \quad (3.2)$$

for all t . Then, the appropriate likelihood ratio statistic is

$$LR = 2 [\log (q^n (1 - q)^{T-n}) - \log (p^n (1 - p)^{T-n})], \quad (3.3)$$

where $q = n/T$. This likelihood ratio is asymptotically χ_1^2 distributed under the null that p is the true probability the VaR is exceeded. With a certain confidence level we can construct nonrejection regions that indicate whether a model is to be rejected or not.

Therefore, the risk model is rejected if it generates too many or too few violations. However, Kupiec test can accept a model which incurs violation clustering but in which the overall number of violations is close to the desired coverage level. For other ways of testing VaR models see Christoffersen (1998).

4 Empirical Results

4.1 Data

The data used in the paper are the market index BELEX15 of the Belgrade Stock Exchange and are obtained from the BELEX website. BELEX15, leading index of the Belgrade Stock Exchange, describes the movement of prices of the most liquid Serbian shares (includes shares of 15 companies) and is calculated in real time.

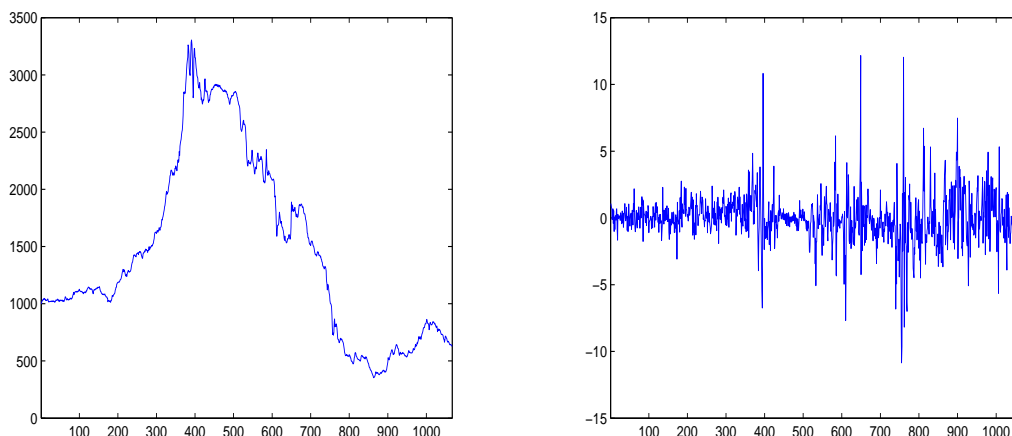


Figure 1: Evolution of BELEX15 daily index (on the left) and daily return (on the right) for period from 4 October 2005 to 25 December 2009.

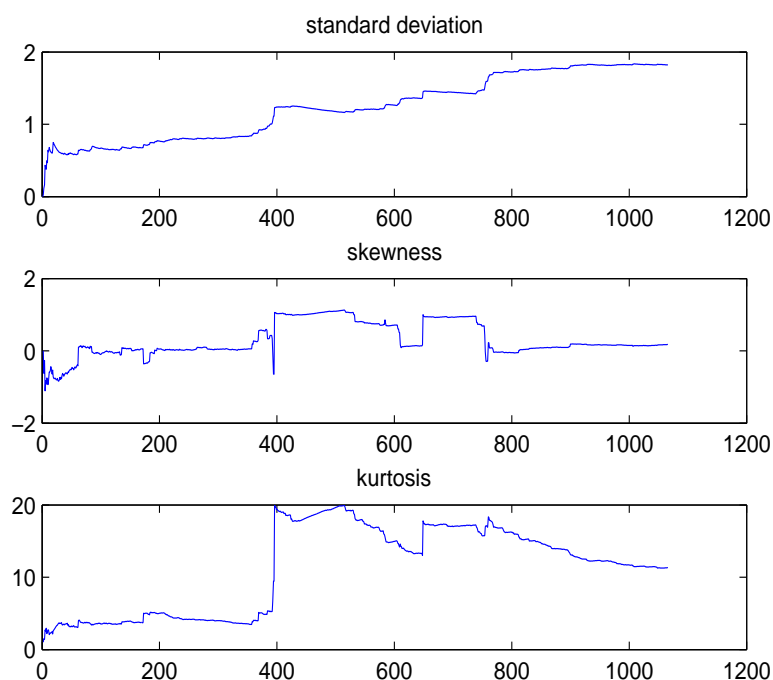


Figure 2: Standard deviation, skewness and kurtosis for BELEX15 up to point

The sample period covers 1067 trading days from 4 October 2005 to 25 December 2009. The plots of the BELEX15 index and returns are given in Figure 1. In this section, the return r_t is expressed in percentages, i.e. $r_t = 100(\log P_t - \log P_{t-1})$.

Results of Augmented Dickey-Fuller test with exogenous constant, linear trend and autocorrelated terms of order selected by Schwarz information criterion, applied on series of daily index, confirm the presence of a unit root ($ADF = -1,1461$, $p = 0.9194$). Null

hypothesis of presence of unit roots for returns is rejected ($ADF = -22.0975$, $p = 0.0000$). Visual inspection of returns shows that the variances change over time around some level, with large (small) changes tending to be followed by large (small) changes of either sign (volatility tends to cluster). Periods of high volatility can be distinguished from low volatility periods. In order to check if moments of order two to four are finite, up to date samples are used to calculate standard deviation, skewness and kurtosis (Figure 2). It is evident that after approximately 800 observations these sample moments became stable, which supports conclusion about finiteness of corresponding population values.

Descriptive statistics

Summary statistics of returns are given in Table 1. The return series exhibit a positive skewness (0.1752) and high excess kurtosis (11.3420), indicating that the returns are not normally distributed. These findings are consistent with plots of the normal QQ plot, box-plot, histogram and empirical density function (Figure 3). Also, from the Q-Q plot and box plot it is obvious that outliers and extreme values cause fat tails.

Table 1: Descriptive Statistics of BELEX 15 index returns.

	mean	median	min	max	variance	skewness	kurtosis
BELEX 15	-0.0433	-0.0326	-10.8614	12.1576	3.3144	0.1752	11.3420

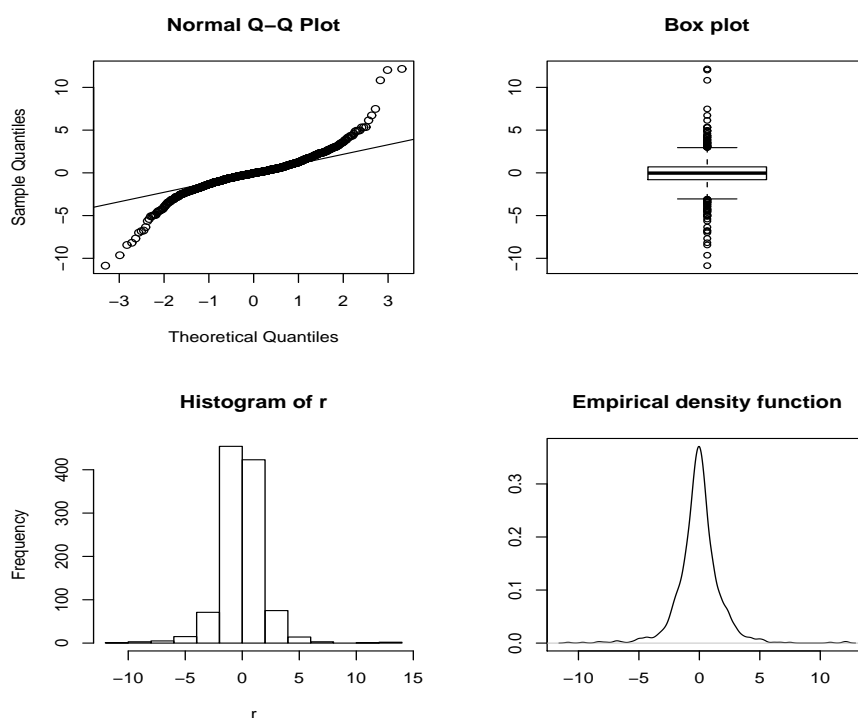


Figure 3: Empirical distribution for BELEX15 returns.

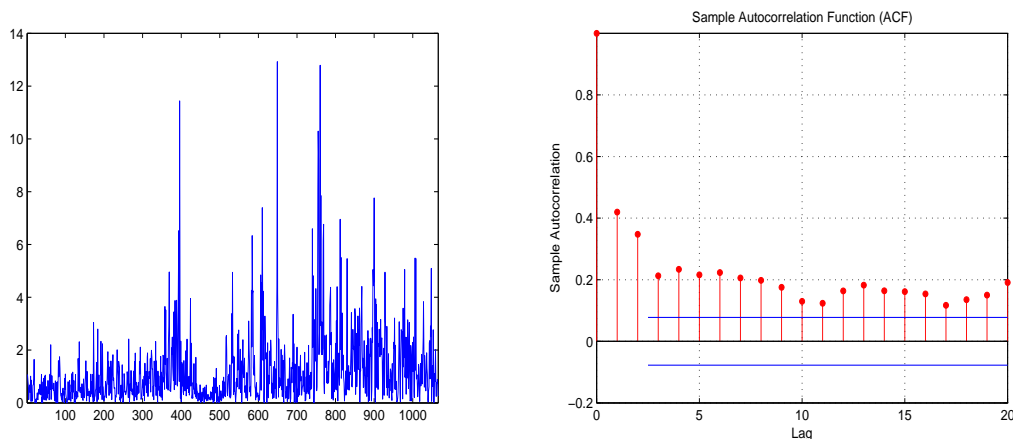
Table 2: Statistical tests for distribution of returns.

Tests	Test statistics	p-values
Jarque-Bera test	3099.259	0.0000
Doornik i Hansen test for independent observations	282.9798	0.0000
Doornik i Hansen for weakly dependent observations	105.3553	0.0000
D'Agostino test of skewness	1.5389	0.1238
D'Agostino omnibus test	183.2526	0.0000
Anderson-Darling test	21.7437	0.0000
Cramer-von Mises test	3.8743	0.0000
Lilliefors test	0.0934	0.0000
Shapiro-Wilk test	0.8975	0.0000
Cabilio-Masaro test of symmetry	0.3135	0.7539
Mira test	0.3131	0.7542
MGG test	0.3818	0.7026

The significant deviation from normality is confirmed by means of statistical tests that are based on the fact that skewness and excess kurtosis are both equal to zero for normal distribution (Jarque-Bera test, D'Agostino omnibus test, Doornik i Hansen test). The same conclusion is for the tests based on density functions (Anderson-Darling test, Lilliefors test) or properties of ranked series (Shapiro-Wilk test). Several applied tests of symmetry (D'Agostino test of skewness, Cabilio-Masaro test of symmetry, Mira test, MGG test) are consistent in conclusion that asymmetry of returns is not statistically significant (Table 2).

Stylized facts

Due to a large body of empirical evidence, stock returns time series display stylized facts such as volatility clustering, high kurtosis, low starting and slow decaying autocorrelation function of squared returns, asymmetric reaction on 'good' and 'bad news' and many others.

**Figure 4:** Absolute returns and their autocorrelation function.

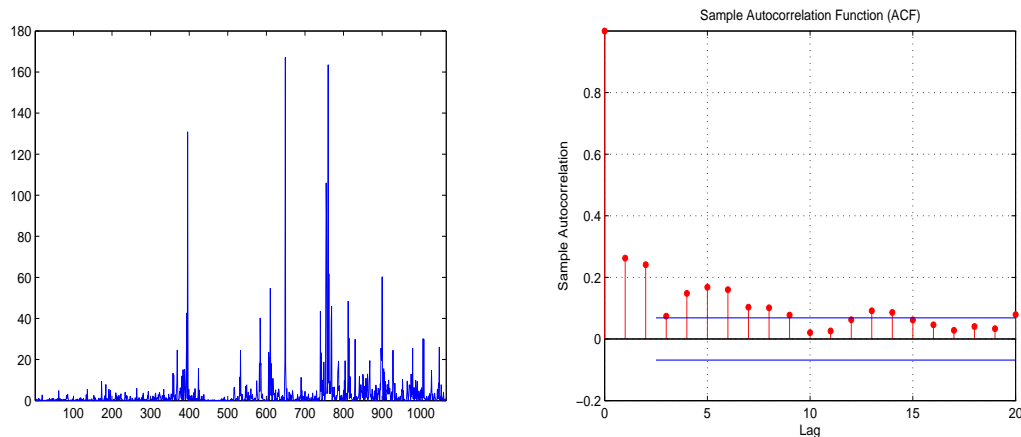


Figure 5: Square returns and their autocorrelation function.

To observe conditional volatility in BELEX indexes absolute or squared returns are used. In Figure 4 and Figure 5 absolute and squared returns are shown with their autocorrelation functions. These plots show variation in conditional volatility. Thus, the use of GARCH-type models for the conditional variance is justified. Also, slow decay of autocorrelation in absolute and squared returns is evident from the autocorrelation plots. This may be interpreted as a sign of long-range dependence of variance.

4.2 Dynamic VaR using EWMA

The dynamic VaR values using EWMA and Normal distribution assumption for long position and two values of α are given in Figure 6. Figure 7 gives a comparison of a Normal and t conditional VaR for $\alpha = 0.05$. The parameter $\nu = 3.01$ of t distribution is estimated by fitting t distribution to returns using maximum likelihood method.

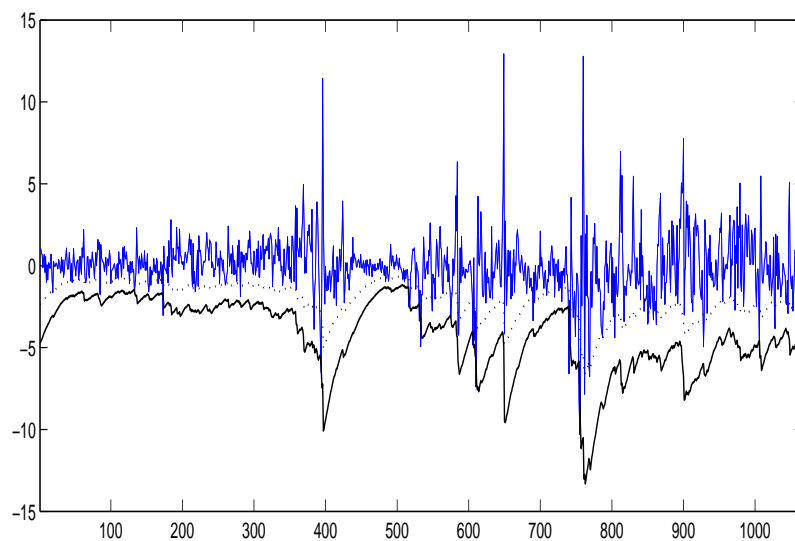


Figure 6: EWMA Normal VaR for $\alpha = 0.1$ (dotted line) and $\alpha = 0.05$.

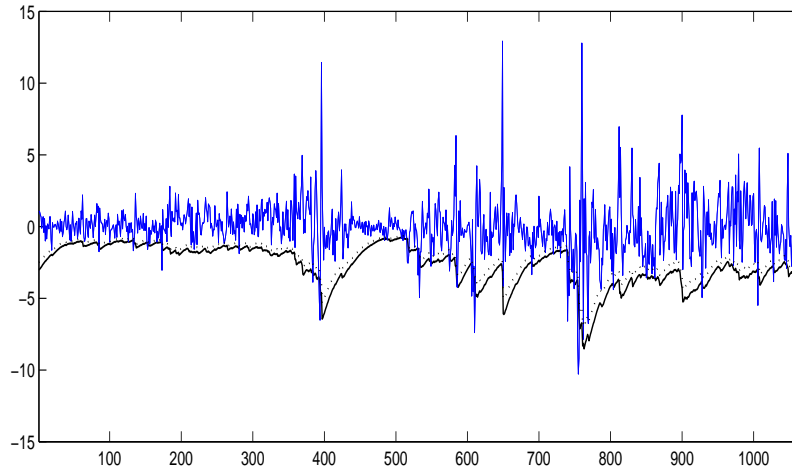


Figure 7: EWMA VAR Normal (dotted line) and t for $\alpha = 0.05$.

Table 3: Mean dynamic VaRs using EWMA, violations and failure rates.

Dynamic VaR						
α	10%	5%	2%	1%	0.5%	0.1%
	Mean Value at risk					
EWMA-N	-2.099	-2.686	-3.347	-3.788	-4.192	-5.024
EWMA- t	-1.555	-2.224	-3.277	-4.266	-5.480	-9.560
	Violations					
EWMA-N	83	44	25	18	13	9
EWMA- t	145	72	28	13	8	1
	Failure rates					
EWMA-N	0.0778	0.0412	0.0234	0.0168	0.0121	0.0084
EWMA- t	0.1360	0.0675	0.0262	0.0121	0.0075	0.0009

Table 4: Kupiec test for EWMA models.

Dynamic VaR						
	10%		5%		2%	
	LR	p	LR	p	LR	p
EWMA-N	6.2355	0.0125	1.8114	0.1783	0.6144	0.4331
EWMA- t	13.9765	0.0001	6.2525	0.0124	1.9461	0.1630
	1%		0.5%		0.1%	
	LR	p	LR	p	LR	p
EWMA-N	4.2306	0.0397	7.8971	0.0049	22.5909	2.0e-6
EWMA- t	0.4849	0.4862	1.1641	0.2806	0.0041	0.9484

The mean values of dynamic 1-day VaRs, the number of VaR violations for entire sample and corresponding failure rates for different α are reported in Table 3. For $\alpha \geq 2\%$ failure rates obtained with assumption of normality are closer to α , while for $\alpha < 2\%$ closer to α are failures of the model that assumes t distribution. Results of Kupiec test support this conclusion (Table 4). Risk model EWMA-N is accepted for $\alpha = 5\%$ and $\alpha = 2\%$, while EWMA- t is accepted for $\alpha < 2\%$.

4.3 Dynamic VaR using GARCH type models and EVT

We fit GARCH(1,1) model to r_t to get the conditional mean $\hat{\mu}_t$, conditional variance $\hat{\sigma}_t^2$ and innovation $\hat{\epsilon}_t$ series. Then we apply Extreme Value Theory (EVT) calculations on the standardised innovations $\hat{\eta}_t$.

ARMA - GARCH type models fitting

A common empirical finding is that the sum of GARCH coefficients $\alpha + \beta$ is close to 1 implying persistent volatility. In this case the impact of past information on future volatility forecasts decays very slowly.

Persistence in variance may be overstated because of the existence of deterministic structural shifts in the model. In order to alleviate effect of sudden changes and time shifts on estimates of GARCH parameters, returns are standardized with up to point mean and standard deviation.

Estimates of ARMA-GARCH parameters, its standard deviations and corresponding t values are obtained by an Ox Package for GARCH models (Laurent and Peters, 2002) and presented in the Table 5. In the same table are presented parameter estimates of GPD distribution fitted to tail distribution of standardized model residuals. The other approach was to model returns by means of ARMA+IGARCH model and then fit GPD distribution as in the previous case (Table 6).

Table 5: Parameter estimates of ARMA+GARCH with Normal and Student innovations.

	ARMA + GARCH					
	<i>Normal distribution</i>			<i>t distribution</i>		
	Coefficient	St. error	t-value	Coefficient	St. error	t-value
C(M)	-0.0516	0.02589	-1.9937	-0.0674	0.02340	-2.8833
AR(1)	0.5461	0.0739	7.3870	0.5687	0.0704	8.0708
MA(1)	-0.2232	0.0945	-2.3607	-0.2577	0.0873	-2.9496
C(V)	0.1931	0.0332	5.8150	0.1757	0.0443	3.9582
ARCH(1)	0.3516	0.0374	9.3996	0.3733	0.0642	5.8134
GARCH(1)	0.5631	0.0369	15.2281	0.5713	0.0538	10.6042
ν	–	–	–	5.7290	1.0972	5.2216
GPD shape	0.1278	–	–	0.1777	–	–
GPD scale	0.5254	–	–	0.4923	–	–

Table 6: Parameter estimates of ARMA+IGARCH with Normal and Student innovations.

ARMA + IGARCH						
	<i>Normal distribution</i>			<i>t distribution</i>		
	Coefficient	St. error	t-value	Coefficient	St. error	t-value
C(M)	0.0494	0.0510	0.9687	-0.0086	0.0497	-0.1728
AR(1)	0.5409	0.0746	7.2460	0.59081	0.0821	7.1960
MA(1)	-0.2145	0.0866	-2.4770	-0.2686	0.0969	-2.7700
C(V)	0.1448	0.0327	4.4180	0.1393	0.0434	3.2110
ARCH(1)	0.4075	0.0447	9.1140	0.3934	0.0599	6.5590
GARCH(1)	0.5925	–	–	0.6065	–	–
ν	–	–	–	5.7666	0.9356	6.1630
GPD shape	0.2376	–	–	0.2719	–	–
GPD scale	0.4675	–	–	0.4570	–	–

VaR estimation results

Dynamic VaRs for returns assuming long trading position and two values of α are shown in Figure 8.

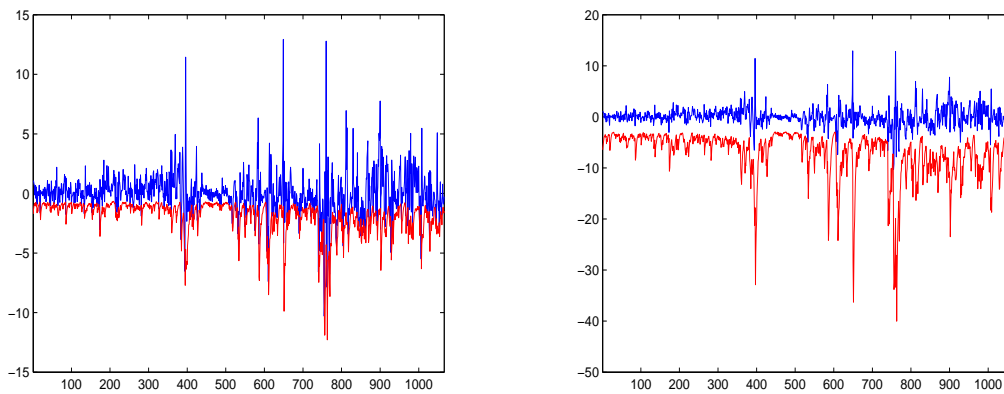


Figure 8: Dynamic VaR for BELEX15 daily returns where $\alpha = 0.1$ (on the left) and $\alpha = 0.001$ (on the right).

Table 7: Mean VaR estimates.

GARCH type models with POT						
α	10%	5%	2%	1%	0.5%	0.1%
GARCH-N	-1.898	-2.342	-3.001	-3.571	-4.200	-5.939
GARCH- <i>t</i>	-1.937	-2.372	-3.044	-3.640	-4.323	-6.317
IGARCH-N	-1.918	-2.361	-3.072	-3.725	-4.497	-6.874
IGARCH- <i>t</i>	-1.917	-2.358	-3.078	-3.750	-4.557	-7.106

Table 8: VaR failure rates.

GARCH type models with POT						
α	10%	5%	2%	1%	0.5%	0.1%
GARCH-N	0.0891	0.0516	0.0215	0.0103	0.0065	0.0018
GARCH- t	0.0873	0.0507	0.0215	0.0103	0.0065	0.0018
IGARCH-N	0.0928	0.0506	0.0206	0.0112	0.0065	0.0009
IGARCH- t	0.0947	0.0497	0.0215	0.0112	0.0065	0.0009

VaR mean values of dynamic VaR for all chosen values of α based on four applied GARCH type models with POT are presented in Table 7. Table 8 gives the sample failure rates for different α . The discrepancy between the failure rates and α 's depends on α , so none of applied models is of superior performance. The same conclusion holds if the number of violations of VaR values is considered. According to Kupiec test (Table 9) all applied models are acceptable.

Table 9: Kupiec test for GARCH type models with POT.

GARCH type models with POT						
	10%		5%		2%	
	LR	p	LR	p	LR	p
GARCH-N	1.4504	0.2284	0.0599	0.8066	0.1349	0.7133
GARCH t	1.9777	0.1596	0.0110	0.9162	0.1349	0.7133
IGARCH N	0.6152	0.4328	0.0096	0.9217	0.0219	0.8823
IGARCH t	0.3320	0.5644	0.0017	0.9663	0.1317	0.7166
	1%		0.5%		0.1%	
	LR	p	LR	p	LR	p
GARCH-N	0.0114	0.9146	0.4816	0.4876	0.6515	0.4195
GARCH t	0.0114	0.9146	0.4816	0.4876	0.6515	0.4195
IGARCH N	0.1635	0.6859	0.4784	0.4891	0.0041	0.9484
IGARCH t	0.1635	0.6859	0.4784	0.4891	0.0041	0.9484

5 Concluding remarks

The purpose of this paper has been to make a comparative study of the predictive ability of VaR estimates for BELEX15 index for various estimation techniques. First, the data are analysed in order to get an idea of the stylized facts of stock market returns. Throughout the analysis, a holding period of one day was used. Several Value-at-Risk models were presented and empirically evaluated. ARMA(1,1) model was used to remove the autoregression from the data. Various values for the left tail probability level were considered, ranging from the very conservative level of 0.01 percent to the less cautious 10 percent.

Evaluation of applied methods was done by means of back-testing for the whole sample. It was not possible to perform out of sample analysis because of the lack of data.

Based on the analysis in this paper, the following conclusions could be made.

1. In the case of BELEX15 index asymmetric behaviour was not discovered, although it is typical for many stock return indexes. Autocorrelation function of absolute and squared returns exhibits very slow decay that is an indication of variance long memory.
2. Return distribution is characterized by high frequency of values close to zero which is the result of smaller stocks trading volume compared with developed markets.
3. The dynamic VaR model using EWMA assuming the normal distribution of innovations is accepted for $\alpha = 0,05$ and $\alpha = 0.02$, while with t distribution is accepted for $\alpha < 0.02$.
4. The most important characteristic of BELEX15 index returns for modelling Value-at-Risk is volatility clustering. Modelling the volatility with GARCH type models, combined with POT method for residuals, reduces the average VaR.
5. IGARCH models cannot outperform GARCH type models in this VaR evaluations, so it seems that long memory volatility is not crucial in VaR estimation.
6. VaR models obtained using heavy tailed and asymmetric distributions, especially those based on EV approach, should be more capable of capturing the true level of risk since they focus on the tail regions of the return distribution.
7. Besides Kupiec's LR test which is based on frequency of failures, the accuracy of VaR model should be evaluated by more sensitive test that measures the frequency, size and independence of losses.

Acknowledgment

The first author is supported in part by the Ministry of Education and Science of the Republic of Serbia (grant no. 43007).

References

- [1] Angelidis, T., Benos, A., and Degiannakis, S. (2004): The use of GARCH models in VaR estimation. *Statistical Methodology*, **1**(2), 105-128.
- [2] Barndorff-Nielsen, O.E. (1997): Normal inverse Gaussian distributions and the modelling of stock returns. *Scandinavian Journal of Statistics*, **24**, 1-13.
- [3] Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- [4] Christensen, B. J. and Nielsen, M. O. (2007): The effect of long memory in volatility on stock market fluctuations. *Review of Economics and Statistics*, **89**, 684-700.

-
- [5] Christoffersen, P. (1998): Evaluating interval forecasts. *International Economic Review*, **39**, 841-862.
- [6] Đorić, D. and Nikolić-Đorić, E. (2006): Risk analysis using GARCH models. *Statistical Review*, LV, 36-47 (in Serbian).
- [7] Đorić, D. and Nikolić-Đorić, E. (2010): Return distribution and value at risk estimation for BELEX15. Accepted for publication in *YUJOR*.
- [8] Dowd, K. (2002): *Measuring Market Risk*. John Wiley & Sons Ltd., New York.
- [9] Duffie, D. and Pan, J. (1997): An overview of value at risk. *The Journal of Derivatives*, **4**, 7-49.
- [10] Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997): *Modelling Extremal Events*. Springer, Berlin.
- [11] Engle, F. and Bollerslev, T. (1986): Modelling the persistence of conditional variances. *Econometric Reviews*, **5**, 1-50.
- [12] Fama, E. (1965): The behaviour of stock prices. *Journal of Bussines*, **47**, 244-280.
- [13] Gençay, R. and Selçuk, F. (2004):. Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, **20**, 287-303.
- [14] Giot, P. and Laurent, S. (2003): Value-at-Risk for long and short trading positions. *Journal of Applied Econometrics*, **18**, 641-664.
- [15] Huang, Y.C and Lin, B-J. (2004): Value-at-Risk analysis for Taiwan stock index futures: Fat tails and conditional asymmetries in return innovations. *Review of Quantitative Finance and Accounting*, **22**, 79-95.
- [16] Jorion, P. (1997): *Value at Risk: The New Benchmark for Controlling Market Risk*. McGraw-Hill.
- [17] Kovačić, Z. (2008): Forecasting volatility: evidence from the Macedonian stock exchange. *International Research Journal of Finance and Economics*, **18**, 182-212.
- [18] Kuester, K., Mittnik, S. and Paolella, M.S. (2006): Value-at-Risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, **4**, 53-89.
- [19] Kupiec, P. H. (1995): Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, **2**, 73-84.
- [20] Laurent, S. and Peters, J. P. (2002): G@RCH 2.2: An Ox package for estimating and forecasting various ARCH models. *Journal of Economic Surveys*, **16**, 447-485.
- [21] McAleer, M. and Medeiros, M. C. (2008): Realized volatility: A review. *Econometric Reviews*, **27**, 10-45.

-
- [22] McNeil, A. J., Frey, R. and P. Embrechts (2005): *Quantitative Risk Management - Concepts, Techniques and Tools*, Princeton University Press, Princeton.
- [23] McNail, A. and Frey, R. (2000): Estimation of tail-related risk measures for heteroscedastic financial times series: An extreme value approach. *Journal of Empirical Finance*, **7**, 271–300.
- [24] RiskMetrics Group, (1996): *RiskMetrics - Technical Document*, Morgan J.P.
- [25] Sarma, M., Thomas, S., and Shah, A. (2003): Selection of VaR models. *Journal of Forecasting*, **22**, 337-358.
- [26] So, M.K.P. and Yu, P.L.H. (2006): Empirical analysis of GARCH models in value at risk estimation. *Journal of International Markets, Institutions and Money*, **16**, 180-197.
- [27] Živković, S. (2007): *Measuring market risk in EU new member states*. Paper presented at the 13th Dubrovnik Economic Conference, Dubrovnik, Croatia.
- [28] Živković, S. (2007): Testing popular VaR models in EU new member and candidate states. *The Proceedings of Rijeka Faculty of Economics*, **25**, 325-346.