GARCH models in value at risk estimation: empirical evidence from the Montenegrin stock exchange

Julija Cerović Smolović, Milena Lipovina-Božović & Saša Vujošević

Published online: 08 Apr 2017.

Submit your article to this journal ⬚

Article views: 57

View related articles ⬚

View Crossmark data ⬚

Routledge
Taylor & Francis Group

GARCH models in value at risk estimation: empirical evidence from the Montenegrin stock exchange

Julija Cerović Smolović[a], Milena Lipovina-Božović[a] and Saša Vujošević[b]

[a]Quantitative Economics, Faculty of Economics, University of Montenegro, Podgorica, Montenegro; [b]Mathematics and Informatics, Faculty of Economics, University of Montenegro, Podgorica, Montenegro

ABSTRACT

This article considers the adequacy of generalised autoregressive conditional heteroskedasticity (GARCH) model use in measuring risk in the Montenegrin emerging market before and during the global financial crisis. In particular, the purpose of the article is to investigate whether GARCH models are accurate in the evaluation of value at risk (VaR) in emerging stock markets such as the Montenegrin market. The daily return of the Montenegrin stock market index MONEX is analysed for the period January 2004–February 2014. The motivation for this research is the desire to approach quantifying and managing risk in Montenegro more thoroughly, using methodology that has not been used for emerging markets so far. Our backtesting results showed that none of the eight models passed the Kupiec test with 95% of confidence level, while only the ARMA (autoregressive moving-average model) (1,2)–N GARCH model did not pass the Kupiec test with a confidence level of 99%. The results of the Christoffersen test revealed three models (ARMA(1,2)–TS GARCH(1,1) with a Student-t distribution of residuals, the ARMA(1,2)–T GARCH(1,1) model with a Student-t distribution of residuals, and ARMA(1,2)–EGARCH(1,1) with a reparameterised unbounded Johnson distribution [JSU] distribution of residuals) which passed the joint Christoffersen test with a 95% confidence level. It seems that these three models are appropriate for capturing volatility clustering, since all of them failed for a number of exceptions. Finally, none of the analysed models passed the Pearson's Q test, whether with 90%, 95% or 99%.

I. Introduction

The search for appropriate risk measuring methodologies has been followed by increased financial uncertainty worldwide. Financial turmoil and the increased volatility of financial markets have induced the design and development of more sophisticated tools for measuring and forecasting risk. The most well known risk measure is value at risk (VaR), which is defined as the maximum loss over a targeted horizon for a given level of confidence. In

---

CONTACT   Julija Cerović Smolović   ✉ julija@ac.me

other words, it is an estimation of the tails of the empirical distribution of financial losses. It can be used in all types of financial risk measurement. Its application to the analysis of market risk is presented in this article.

VaR was introduced by J.P. Morgan in 1994 as the method of risk management behind its RiskMetrics system. The theoretical background for the VaR method is given by Jorion (1996), Duffie and Pan (1997) and Dowd (1998). Although there are some theoretical studies about the shortcoming of VaR due to its lack of sub-additivity and convexity (Cheng, Liu, and Wang (2004)), there is still no better measure to quantify risk (Orhan & Köksal, 2012).

In this study, we aim to present and test the estimation of VaR based on the econometric approach, with empirical results of measuring VaR in the still developing Montenegrin financial market. This approach consists of the econometric modelling of conditional volatility derived from generalised autoregressive conditional heteroskedasticity (GARCH) models. We decided to use both asymmetric and symmetric GARCH type models with four types of residual distribution: normal, Student-t (both of which are used more frequently), and skewed Student-t as well as a reparameterised Johnson distribution. Then we evaluated the accuracy of the estimation of VaR calculated by the different GARCH models through backtesting. We computed likelihood ratio coverage tests according to the methodology of Christoffersen, both conditional and unconditional, known as the Kupiec and Christoffersen tests, as well as the Pearson's Q test for goodness of fit.

The purpose of this article is to test the relative performance of the GARCH family models in estimating and forecasting VaR in the Montenegrin stock exchange over a long sample period which includes the years of the financial crisis (the period covered is 5 January 2004 to 21 February 2014). In particular, the aim of the article is to investigate whether GARCH models are accurate in the calculation of VaR in emerging stock markets.

The Montenegrin stock market has not been discussed in the empirical literature until recently. To be specific, a very recent paper by Cerovic and Karadzic (2015) considers the adequacy of the methods that are the basis of extreme value theory in the Montenegrin emerging market before and during the global financial crisis. The results of the Kupiec test showed that the peaks-over-threshold method was significantly better than the block maxima method, but both methods failed to pass the Christoffersen independence test and the joint test due to the lack of accuracy in exception clustering when measuring VaR. Although better, the peaks-over-threshold method still cannot be treated as an accurate VaR model for the emerging Montenegrin stock market. Bearing that in mind, the main contribution of this article is to extend the limited empirical research on VaR estimation and forecasting in emerging financial markets and to overcome empirical studies by testing the relative performance of GARCH models in the estimation of Montenegrin stock exchange volatility. Additionally, we intend to make a comparison between the major findings of our study and the study conducted by Cerovic and Karadzic (2015).

The article is organised as follows: the subsequent section presents a brief literature review. The third section reviews the methodology used in an econometric approach to VaR calculation. In the fourth section, the backtesting procedure for examining the appropriate estimation of VaR is given. Data and descriptive statistics are given in the fifth section. The sixth section presents the empirical results and our conclusions are presented in the seventh section.

II. Literature review

The appropriate measuring and forecasting of market losses seems to play an important role in both developed and emerging financial markets. The growing interest of foreign financial investors in investing in emerging financial markets highlights the importance of accurate market risk quantification and prediction. The fundamental characteristics of emerging markets are reflected in lower liquidity, frequent internal and external shocks as well as a higher degree of insider trading which causes the market to be more volatile (Miletic & Miletic, 2013). Emerging markets are characterised by the greater influence of internal trade and high volatility, illiquidity and the shallowness of the market when compared to developed countries, so the evaluation of VaR with standard methods that assume a normal distribution is much more difficult (Zikovic & Aktan, 2009).

There are numerous papers dealing with the appropriate estimation and forecasting of VaR: Alexander and Leigh (1997), Manganelli and Engle (2001), Christoffersen, Hahn, and Inoue (2001), Wong, Cheng, and Wong (2002), Angelidis, Benos, and Degiannakis (2004), McNeil, Frey, and Embrechts (2005), Guermat and Harris (2002) and Tsay (2010). Still, there is no such thing as a universal model adequate for risk measuring.

Unlike the literature dealing with VaR calculation in developed financial markets, both the literature on and empirical research into VaR models in emerging financial market are not very extensive (these include Da Silva, Beatriz, and de Melo Mendes (2003), Gencay and Selcuk (2004), Bao, Lee, and Saltoglu (2006), Zikovic (2007), Zikovic and Aktan (2009), Andjelic, Djakovic, and Radisic (2010), Nikolic-Djoric and Djoric (2011), Mladenovic, Miletic, and Miletic (2012) and Bucevska (2012)). Bucevska (2012) claims that the short historical time-series data did not allow for the performing of reliable econometric analysis (most of the stock markets in these countries were established in the early 1990s).

Zikovic and Aktan (2009) compared the relative performance of VaR models of the returns of the Turkish and Croatian stock exchange indices with the onset of the global financial crisis. In this analysis, they showed that models such as extreme value theory and hybrid historical simulation are best, while other models underestimate the level of risk. Andjelic et al. (2010) analysed four emerging stock markets (Slovenia, Croatia, Serbia and Hungary) and their main findings showed that under stable market conditions, the analysed models (parametric and historical simulation models) gave good forecasts of VaR estimations with a 95% confidence level, while under the conditions of market volatility analysed models give good estimations of VaR parameters with a 99% confidence level.

Nikolic-Djoric and Djoric (2011) investigated the performance of the RiskMetrics method, as well as the GARCH and integrated GARCH (IGARCH) models in VaR forecasting of a stock exchange index in the Serbian financial market. They concluded that GARCH models combined with extreme value theory (the peaks-over-threshold method) decrease the mean value of VaR, and that these models are better than the RiskMetrics method and the IGARCH model. Similarly, Mladenovic et al. (2012) came to the conclusion that the methodology of extreme value theory is slightly better than the GARCH model regarding the calculation of VaR, based on their analysis of stock exchange indices in Central and Eastern European countries (Bulgaria, Czech Republic, Hungary, Croatia, Romania and Serbia), but they generally suggest the use of both approaches for better market risk measurement. Bucevska (2012) showed that the most adequate GARCH family models for estimating volatility in the Macedonian stock market are the asymmetric exponential GARCH (EGARCH) model with Student's t-distribution, the EGARCH model with normal

distribution and the GARCH- Glosten, Jagannathan and Runkle (GARCH-GJR) model. The obtained findings have important implications regarding VaR estimation in turbulent conditions that have to be addressed by investors in emerging capital markets. Miletic and Miletic (2015) implemented GARCH models that involve time varying volatility and heavy tails to the empirical distribution of returns, in selected Central and Eastern European emerging capital markets (Croatia, Czech, Hungary, Romania and Serbia). They showed that GARCH models with a t-distribution of residuals in most analysed cases give a better VaR estimation than GARCH models with normal errors in the case of a 99% confidence level, while the opposite is true in the case of a 95% confidence level. The backtesting results for the crisis period showed that GARCH models with a t-distribution of residuals provide better VaR estimates when compared with GARCH models with a normal distribution, historical simulations or RiskMetrics methods.

III. Methodology

Let a log return series at a moment in time t, be marked as r_t, and it can be expressed in percent. The random variable of loss over the period $[t,t+h]$ is marked as $L_{t+h} = -\left(r_{t+h} - r_t\right) = \Delta r(h)$. Then, F_L is the cummulative function of loss distribution and it holds that $F_L(x) = P(L \leq x)$. VaR at significance level α (most often 1% and 5%) is actually an α-quantile of the distribution function F_L, or in other words, VaR presents the smallest real number satisfying the inequation $F_L(x) \geq \alpha$, i.e.:

$$VaR_\alpha = \inf(x|F_L(x) \geq \alpha). \tag{1}$$

The econometric approach to VaR calculation considers the use of time series econometric models. An autoregressive moving-average model (ARMA) of orders p and q, ARMA (p,q), is used for estimating a log return series, as follows:

$$r_t = \varphi_0 + \sum_{i=1}^{p} \varphi_i r_{t-i} + a_t - \sum_{j=1}^{q} \theta_j a_{t-j}$$

$$a_t = \sigma_t \varepsilon_t \tag{2}$$

The parameters of equation (2) representing the ARMA (p,q) model, are marked as $\varphi_0, \varphi_1, ..., \varphi_p, \theta_1, ..., \theta_q$. The error term of the model, a_t, is the function of ε_t - a series of independent and identically distributed random variables with the properties of white noise.

For modelling the conditional variance of returns r_t we will use following volatility models: The Asymmetric Power ARCH (APARCH), GARCH, Taylor-Schwert GARCH (TS GARCH), Treshold GARCH (T GARCH), Non-linear GARCH (NARCH), Glosten, Jagannathan and Runkle GARCH (GJR-GARCH), Exponential GARCH (EGARCH) and Integrated GARCH (IGARCH).

The APARCH model was introduced by Ding, Granger, and Engle (1993). It is an extension of the standard GARCH model that provides an asymmetric approach to volatility modelling. It considers positive and negative shocks not to have the same effect on volatility. The general structure is as follows:

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{u} \alpha_i(|a_{t-i}| - \gamma_i a_{t-i})^\delta + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^\delta, \tag{3}$$

where the leverage effect is denoted by parameter γ_i and δ is a Box-Cox transformation of the conditional standard deviation. This leverage effect distinguishes positive and negative shocks. Thus, positive innovations have an impact of α, while negative innovations have an impact of $\alpha + \gamma$

The parameters of the model are supposed to satisfy the following conditions:

$$\begin{aligned} &\alpha_0 > 0, \delta \geq 0, \\ &\alpha_i \geq 0, i = 1, 2, ..., u, \\ &\beta_j \geq 0, j = 1, 2, ..., v, \\ &-1 \leq \gamma_i \leq 1, i = 1, 2, ..., u. \end{aligned} \tag{4}$$

This model includes several special versions:

- When $\delta = 2$, $\gamma_i = 0$, $(i = 1, 2, ..., u)$, $\beta_j = 0$, $(i = 1, 2, ..., v)$ the APARCH model is an ARCH model;
- When $\delta = 2$, $\gamma_i = 0$, $(i = 1, 2, ..., u)$, APARCH is a GARCH model;
- When $\delta = 2$, the APARCH model is a GJR GARCH model;
- When $\delta = 1$, $\gamma_i = 0$, $(i = 1, 2, ..., u)$, APARCH is a TS-GARCH model;
- When $\delta = 1$, $\beta_j = 0$, $(i = 1, 2, ..., v)$ APARCH is a T GARCH model;
- When $\gamma_i = 0$, $(i = 1, 2, ..., u)$, $\beta_j = 0$, $(i = 1, 2, ..., v)$, APARCH is a NARCH model; and
- When $\delta \to 0$, the APARCH model is the limiting case in a log-ARCH model (Geweke, 1986; and Pantula, 1986).

The GARCH model that Bollerslev (1986) proposed as a generalisation of the autoregressive conditional heteroscedasticity model – ARCH (Engle, 1982), can be presented as follows (Tsay, 2010):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{u} \alpha_i a_{t-i}^2 + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^2. \tag{5}$$

Parameters $\alpha_0, \alpha_1, ..., \alpha_u, \beta_1, ..., \beta_v$ of the conditional variance equation satisfy the conditions $\alpha_0 > 0, \ \alpha_1, ..., \alpha_u \geq 0, \ \beta_1, ..., \beta_v \geq 0, \ \sum_{i=1}^{\max(u,v)} (\alpha_i + \beta_i) < 1.$

Instead of conditional variance, Taylor (1986) modelled the conditional standard deviation function. Further, Schwert (1989) modelled the conditional standard deviation as a linear function of lagged absolute residuals. So, the TS GARCH(p, q) model is defined as follows:

$$\sigma_t = \alpha_0 + \sum_{i=1}^{u} \alpha_i |a_{t-i}| + \sum_{j=1}^{v} \beta_j \sigma_{t-j}. \tag{6}$$

The T GARCH was introduced by Zakoian (1994), where the conditional standard deviation is modelled by linear combinations of past error terms and past standard deviations. It is very similar to the TS GARCH, and it is given by:

$$\sigma_t = \alpha_0 + \sum_{i=1}^{u} \alpha_i^+ a_{t-i}^+ - \alpha_i^- a_{t-i}^- + \sum_{j=1}^{v} \beta_j \sigma_{t-j}. \tag{7}$$

The positive and negative terms of a_t are $a_{t-i}^+ = \max(a_t, 0)$ and $a_{t-i}^- = \min(a_t, 0)$.

The NARCH model (Higgins & Bera, 1992) can be seen as a Box-Cox power transformation of the terms of the linear ARCH model. Specifically, the model can easily be generalised to include the GARCH model, by adding Box-Cox transformations of lagged values of variance to the right-hand side of its general form, written as:

$$\sigma_t = \left[\alpha_0(\sigma^2)^\delta + \sum_{i=1}^{u} \alpha_i(a_{t-1}^2)^\delta - k_i)^2 \right]^{1/\delta}, \tag{8}$$

or

$$\sigma_t^2 = \alpha_0 + a_{t-1}^2 + \sum_{i=2}^{u} \alpha_i(a_{t-i} - k_i)^2 + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^2. \tag{9}$$

The GJR-GARCH model was introduced in 1993 by Glosten, Jagannathan, and Runkle. In general form it is given by:

$$\sigma_t^2 = \omega + \sum_{i=1}^{u} \alpha_i a_{t-i}^2 + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^2 + \gamma_i I_{t-i} a_{t-i}^2, \tag{10}$$

where α, β and γ are constant parameters, and I is the indicator function that takes the value zero when a_{t-i} is positive, and one when a_{t-i} is negative. So, this dummy variable distinguishes positive and negative shocks, and the asymmetric effects are captured by γ.

The EGARCH (Nelson, 1991) model is defined as follows:

$$\ln(\sigma_t^2) = (\alpha_0 + \sum_{i=1}^{m} \zeta_i v_{it}) + \sum_{i=1}^{u} (\alpha_i a_{t-i} + \gamma_i(|a_{t-i}| - E_i|a_{t-i}|)) + \sum_{j=1}^{v} \beta_j \ln(\sigma_{t-j}^2), \tag{11}$$

where the coefficient α_i captures the sign effect and γ_i captures the size effect. More clearly, EGARCH can be written in the following form:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{\infty} \pi_i g\left(\frac{a_{t-i}}{\sigma_{t-i}}\right), \tag{12}$$

where g is a function of the asymmetric relation between return and volatility, such as:

$$g\left(\frac{a_{t-i}}{\sigma_{t-i}}\right) = \alpha\left(\left|\frac{a_{t-i}}{\sigma_{t-i}}\right| - E\left|\frac{a_{t-i}}{\sigma_{t-i}}\right|\right) + \gamma\left|\frac{a_{t-i}}{\sigma_{t-i}}\right|. \tag{13}$$

This equation represents a zero-mean process, with constants α and γ. For positive error terms ε_t, g is a linear function with the slope coefficient α+γ, while for negative ε_t, g is a linear function with the slope coefficient γ-α.

If the GARCH model (1,1) satisfies the parameters sum $\alpha_1 + \beta_1 = 1$, then the model describes a process of unlimited growth of the conditional variability. Such a model is known as IGARCH (1,1). It is the basis of VaR estimation, representing the standard approach to risk measuring – RiskMetrics. In general form, it can be written as (Orhan & Köksal, 2012):

$$\sigma_t^2 = \alpha_0 + a_{t-1}^2 + \sum_{i=2}^{u} \alpha_i(a_{t-i}^2 - a_{t-1}^2) + \sum_{j=1}^{v} \beta_j(\sigma_{t-j}^2 - a_{t-1}^2). \tag{14}$$

This methodology was developed by the company J.P. Morgan (Longerstaey & More, 1995), and it implies that the conditional distribution of the series of log daily returns is $r_t|\Omega_{t-1}:N(\mu_t, \sigma_t^2)$, where μ_t is the conditional mean, and σ_t^2 is the conditional variance of a series r_t. The following relations are valid for them:

$$\mu_t = 0, \quad \sigma_t^2 = \alpha\sigma_{t-1}^2 + (1-\alpha)\ r_{t-1}^2, \quad 0 < \alpha < 1. \tag{15}$$

In the case of long trading positions, the risk of a loss arises when the traded asset price declines and inversely, in the case of short trading positions, the risk occurs when the asset price rises. So, the left and right tails of the distribution are modelled for the long and short positions, respectively.

If the series ε_t is a random variable with the standardised normal distribution, then the conditional distribution of a random variable r_{h+1} for the available data with the moment h inclusive, also has a normal distribution with the mean $\hat{r}_h(1)$ and variance $\hat{\sigma}_h^2(1)$. Then, the 5% quantile of the conditional distribution, representing the estimation of VaR at a 95% confidence level and for a forecast horizon 1 step ahead, is computed as:

$$VaR_{long} = \hat{r}_h(1) - 1,65\hat{\sigma}_h(1)$$
$$VaR_{short} = \hat{r}_h(1) + 1,65\hat{\sigma}_h(1). \tag{16}$$

If the random variable ε_t has a Student-t distribution, with v degrees of freedom, then the 5% quantile of the conditional distribution is:

$$VaR_{long} = \hat{r}_h(1) + \frac{t_v\alpha}{\sqrt{\frac{v}{v-2}}}\hat{\sigma}_h(1)$$
$$VaR_{short} = \hat{r}_h(1) + \frac{t_v(1-\alpha)}{\sqrt{\frac{v}{v-2}}}\hat{\sigma}_h(1), \tag{17}$$

where $t_v(1-\alpha)$ is the corresponding critical value of $(1-\alpha)$ quantile from the t distribution with v degrees of freedom.

Lambert and Laurent (2001) proposed a skewed Student-t distribution to capture both skewness and fat tails. The log-likelihood function of the skewed Student-t distribution has the following form[1]:

$$L_{skst} = T(\ln\Gamma(\frac{v+1}{2}) - \ln\Gamma(\frac{v}{2}) - \frac{1}{2}\ln(\pi(v-2)) + \ln\left(\frac{2}{\xi + 1/\xi}\right) + \ln s$$
$$-\frac{1}{2}\sum_{t=1}^{T}\left[\ln(\sigma_t^2) + (1+v)\ln\left[1 + \left(\frac{(sz_t + m)^2}{v-2}\right)\xi^{-2I_t}\right]\right] \tag{18}$$

where z_t is a random variable with a skewed Student-t distribution with v degrees of freedom, ξ is the asymmetry (skewness) parameter, m is the mean and s is the standard deviation, and $I_t = 1$ when $z_t \geq -m/s$, while $I_t = -1$ when $z_t < -m/s$.

If a random variable ε_t has a skewed Student-t distribution, with υ degrees of freedom, then the 5% quantile of the conditional distribution is:

$$
\begin{aligned}
VaR_{long} &= \hat{r}_h(1) + skst_{\alpha,\upsilon,\varepsilon}\hat{\sigma}_h(1) \\
VaR_{short} &= \hat{r}_h(1) + skst_{1-\alpha,\upsilon,\varepsilon}\hat{\sigma}_h(1),
\end{aligned}
\tag{19}
$$

where $skst_{\alpha,\upsilon,\varepsilon}$ and $skst_{1-\alpha,\upsilon,\varepsilon}$ are the left and right quantiles of α% from the skewed t distribution with υ degrees of freedom.

The reparameterised Johnson SU distribution, discussed in Rigby and Stasinopoulos (2005), is a four parameter distribution denoted by $JSU(\mu, \sigma, \upsilon, \tau)$ with the mean μ and standard deviation σ for all values of the skew and shape parameters υ and τ respectively. Using the general form of Johnson densities, the log-likelihood function is:

$$
L_{jsu} = n\ln\tau - n\ln\sigma - \frac{n}{2}\ln(2\pi) + \sum_{i=1}^{n}g\left(\frac{x-\mu}{\sigma}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(\upsilon + \tau g\left(\frac{x-\mu}{\sigma}\right)\right)^2. \tag{20}
$$

Accordingly, if a random variable ε_t has a reparameterised JSU distribution, then the 5% quantile of the conditional distribution is:

$$
\begin{aligned}
VaR_{long} &= \hat{r}_h(1) + jsu_{\alpha,\upsilon,\varepsilon}\hat{\sigma}_h(1) \\
VaR_{short} &= \hat{r}_h(1) + jsu_{1-\alpha,\upsilon,\varepsilon}\hat{\sigma}_h(1),
\end{aligned}
\tag{21}
$$

where $jsu_{\alpha,\upsilon,\varepsilon}$ and $jsu_{1-\alpha,\upsilon,\varepsilon}$ are the left and right quantiles of α% from the reparameterised JSU distribution.

III. Backtesting

When a VaR model is estimated it is important to check its reliability and accuracy. The statistical procedure for examining the appropriate estimation of VaR is called backtesting. The aim of backtesting is to estimate whether the amount of losses predicted by VaR is correct. That process implements unconditional or conditional coverage tests for the correct number of exceedances. The unconditional tests check whether the frequency exceptions, during the selected time interval, are in accordance with the chosen confidence level. The most commonly used test in this group is the Kupiec test. On the other hand, conditional coverage tests examine conditionality and changes in data over time, and the most famous test in this group is the Christoffersen independence test. Campbell (2005) showed that tests that examine several quantiles are the most successful in identifying inaccurate Var models, so we will include the Pearson's Q test.

(a) The Kupiec test

Let N be the observed number of exceedances in the sample, or, in other words, $N = \sum_{t=1}^{T}I_t$ is a number of days over a T period of time when the portfolio loss over a fixed interval $r_{t,t+1}$ was larger than the VaR estimate, where (Campbell, 2005):

$$
I_{t+1} = \begin{cases} 1, & if \quad r_{t,t+1} \geq -VaR_t \\ 0, & if \quad r_{t,t+1} < VaR_t. \end{cases}
\tag{22}
$$

The failure number follows a binomial distribution where the expected exception frequency is $p = N/T$. The ratio of failures, N, to trials, T, under the Null hypothesis should be p. The appropriate likelihood ratio statistic is:

$$LR_c = 2\ln[(1 - \frac{N}{T})^{T-N}(\frac{N}{T})N] - 2\ln[(1-p)^{T-N}p^N]. \tag{23}$$

The Kupiec test has a chi-square distribution, asymptotically, with one degree of freedom. This test can reject a model for both high and low failures but, as stated by Kupiec (Kupiec, 1995), its power is generally poor, so conditional coverage tests, such as the Christoffersen test, can be used for the further examination of VaR model reliability.

(b) The Christoffersen test

The conditional coverage test, under the Christoffersen approach, is important insofar as it detects whether the exceptions occur in clusters or not. In other words, it can account for volatility clustering. If the existence of clustering can be proved, the model is mis-specified and needs to be recalibrated. It is important to test for volatility clustering from a practical point of view. For example, if a bank allocates capital for 50 exceedances over a period of four years it may not be able to stay liquid if a majority of the exceedances appear during the course of two months. A new indicator building on the exception indicator above is calculated which defines n_{ij} to be the amount of days that j (exception) occurred when it was i (no exception) the day before. The probability of state j being observed given that state i was observed the previous day is noted by π_{ij} (Jorion, 2007). The test statistic testing independence is (Dowd, 2005):

$$LR_{cc} = -2\ln[(1-p)^{T-N}p^N] + 2\ln[(1 - \pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1 - \pi_{11})^{n_{10}}\pi_{11}^{n_{11}}], \tag{24}$$

where the corresponding probabilities are $\pi_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$, so π_{01} is the probability of a non-exception being followed by an exception, and π_{11} is the probability of an exception being followed by an exception. The absolute probability of a non-exception or exception being followed by an exception is denoted by p. The test statistic is distributed as χ^2 with two degrees of freedom.

The Christoffersen test needs several hundred observations in order to be accurate, and the main advantage of this procedure is that it can reject a VaR model that generates either too many or too few clustered violations. These two likelihood ratio (LR) tests can be combined, thereby creating a complete test for coverage and independence, which is also distributed as a $\chi^2(2)$:

$$LR_{cc} = LR_{uc} + LR_{ind}. \tag{25}$$

This is the Christoffersen approach to check the predictive ability and accuracy of a VaR model. The advantage of this test is in the combination of two tests which can be tested separately to backtrack if the model fails due to the wrong coverage or due to exception clustering. Altogether, these tests provide the necessary tools to evaluate and compare the VaR models mentioned above. A full specification of null and alternative hypotheses for these tests is given in Table 1.

Table 1. Null and alternative hypotheses for the three-step Christoffersen test.

	LR_{uc}	LR_{ind}	LR_{cc}
H_0	Correct unconditional coverage	Exceedances are independent	Correct conditional coverage
H_1	Incorrect unconditional coverage	Exceedances are not independent	Incorrect conditional coverage

Source: Authors' calculations.

(c) The Pearson's Q test

The Pearson's Q test can be applied for more rigorous backtesting, as Campbell (2005) has noted. It is based upon the number of observed violations across different quantiles. The procedure requires the unit interval to be parted into k different sub-intervals (bins). The particular partition depends on the particular set of quantiles of interest. For example, the partition [0.00, 0.01], [0.01, 0.05], [0.05, 0.1] and [0.1, 1] is used in cases when quantiles more extreme than the 10th are of interest.[2] The number of VaR violations that occur within each bin is used in a Chi square test statistic, which is calculated as follows:

$$Q = \sum_{i=1}^{k} \frac{(N_{(l_i, u_i)} - N(u_i - l_i))^2}{N(u_i - l_i)}. \tag{26}$$

The number of VaR violations in the ith bin is marked by $N_{(l_i, u_i)}$, and N is the total number of days used to conduct the test, while l_i and u_i are the lower and upper bound of the ith bin. The Pearson's Q test is distributed according to the chi squared distribution with k-1 degrees of freedom.

IV. Data and descriptive statistics

The capital market in Montenegro is characterised by a relatively simple structure. At the end of 2012, there were 13 intermediaries – brokers and/or dealer companies, authorised for trading on the Montenegro Stock Exchange. Capital is traded on the Montenegro Stock Exchange, although one other stock market functioned for 10 years, which merged with the Montenegro Stock Exchange in 2011. The Central Depository Agency has the role of depositing shares, which is also the operator of the system for clearing and the settlement of securities. The Commission for Securities operates in a regulatory and supervisory role. Moreover, companies that are often mentioned when talking about the capital market are: (1) closed and open-ended investment funds (developed through previous mutual funds). There were eight of them at the end of 2012 (four closed and four open – resulting in four of the six former collective investment funds (CIF). The Securities Commission is in a dispute with two other former CIF); and (2) voluntary pension funds (two voluntary pension funds at the end of 2012).

If there is a need to extract a key event or trend in the Montenegrin economy in the last decade, it is certainly the strong inflow of foreign direct and other mid-term investment, which has acted as a key positive shock to economic growth, but also as a shock to many other macroeconomic variables. A positive shock inflow of foreign direct investment hit both the capital market and the real estate market, as the following figure shows (Figure 1). Note that the positive trend from 2009 is mostly due to part of the privatisation of Elektropivreda Crne Gore, a. d. Niksic (EPCG). If there was not the set of these transactions

(which affected the sales and prices of other securities), the state in the capital market of Montenegro would hardly have deviated significantly from the rest of the period 2008–2013. The year 2009 was actually a recession year and this year showed a high rate of decline in economic activity.

Strong growth in the stock market began in 2005, continued in 2006, and finally in 2007 (especially the first nine months) the state of the capital market in Montenegro has grown into a collective 'madness'. After that, the bubble began to burst and what followed was a drastic fall in prices (80–85%) at the end of 2007 and in 2008. The general 'lethargy' of the market continued until the end of 2013.

The best indicator of the state of the Montenegrin stock market is the Montenegrin stock index MONEX, so its daily log returns are observed. The stock market index MONEX is the general (benchmark) index of the Montenegro Stock Exchange aimed at describing the price movement of the most representative shares on the Official and Free market segment of the Montenegro Stock Exchange. The time series of observed log returns of the MONEX stock index on a daily basis consists of 2508 points of data in total (from 5 January 2004 to 21 February 2014), and they are presented in Figure 2. The log daily returns (or continuously compounded returns) represent the difference between the logarithmic levels of prices on two successive days. It can also be expressed in percent, when these differences are multiplied by 100.[3]

The expressed volatility of the Montenegrin stock index MONEX can be seen in Figure 2. It is also evident that this series is stationary. Its empirical distribution deviates from normal distribution, as the Q-Q plot (Figure 3) shows. Specifically, the quantiles of an empirical distribution are plotted against the quantiles of a normal distribution. From Figure 3 it is clear that the Q-Q plot is not linear and that the empirical distribution differs from the
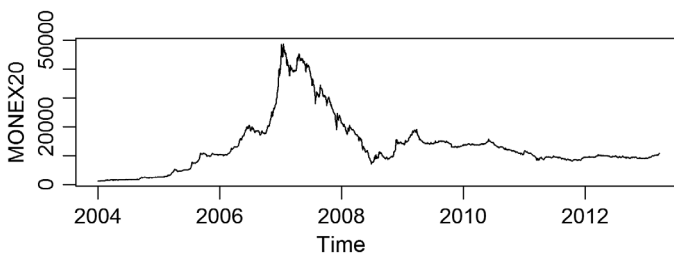

Figure 1. Time series of the stock index MONEX from January 2004 to February 2014. Source: Montenegro Stock Exchange and authors' calculations.
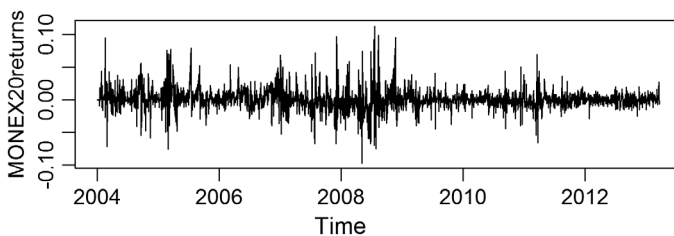

Figure 2. Daily log return of MONEX from January 2004 to February 2014. Source: Montenegro Stock Exchange and authors' calculations. Note: *P*-values of corresponding test statistics are given in parentheses.[5]
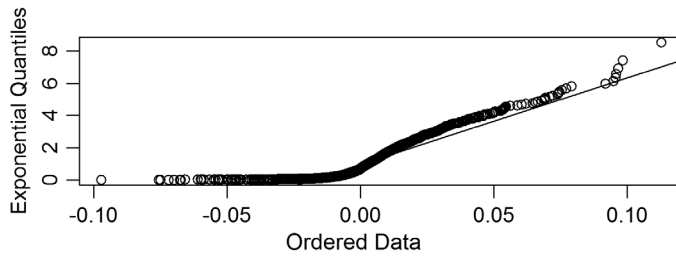
Figure 3. Q-Q plot of daily return of MONEX relative to normal distribution. Source: Authors' calculations.

Table 2. Basic descriptive statistics of daily logarithmic return for MONEX.

Variance	Skewness	Kurtosis	JB	Box-Ljung (m=10)	Box-Ljung (a_t^2)
2.869291	0.686277	6.536753	4672.7013 (<2.2e-16)	219.6358 (<2.2e-16)	1003.793 (<2.2e-16)

Source: Authors' calculations.

hypothesised normal distribution. So, the fat-tail nature of the observed logarithmic return series is expressed, as well as the skewness coefficient and Jarque-Bera (JB) test-statistics has shown. These descriptive statistics are given in Table 2 with the corresponding *p*-values in parentheses.

The skewness shows that the series is not sharply asymmetric, but there is a particular positive asymmetry. The normality deviation is mostly due to high kurtosis, which means the existence of 'fat tails' – tails are heavier than normal distribution tails. The Box-Ljung statistics indicate evidence of autocorrelation in both the standardised residuals and the squared standardised residuals.

The obtained descriptive statistics allow the modelling of the tails of the empirical distribution of the MONEX rates of return and the estimation of the parameters of VaR using an econometric approach.

V. Empirical results

Based on the information criteria – the Akaike Information Criterion (AIC) in Table 3 – it is estimated that the best model for modelling the logarithmic return series of MONEX is ARMA(1,2).[4]

Estimates of the parameters from joint ARMA with different specifications of conditional heteroscedasticity models (APARCH, GARCH, TS-GARCH, T GARCH, NARCH, GJR-GARCH, EGARCH and IGARCH) together with the associated tests of residual autocorrelation and conditional heteroscedasticity are given in Tables A1–A8 in the Appendix. We have four versions of these eight models regarding the distribution of the standardised residuals: the normal distribution of residuals, the Student-*t* distribution of residuals, the skewed Student-*t* distribution of residuals and the reparameterised JSU distribution of residuals.

The accuracy of the estimated models considered in the study can be assessed by counting the number of actual returns that are larger than the estimated VaR, and comparing this figure with the theoretically expected number of excesses for a determined probability. Of course, the closer the empirically observed number of excesses is to the theoretically expected amount, the more preferable the method is for estimating risk measures.

Table 3. Criteria for ARMA(p,q) selection.

ARMA model	AIC
ARMA(0,0)	-13335.54
ARMA(1,0)	-13476.8
ARMA(2,0)	-13475.04
ARMA(3,0)	-13474.03
ARMA(0,1)	-13467.64
ARMA(0,2)	-13473.83
ARMA(0,3)	-13472.12
ARMA(1,1)	-13475.1
ARMA(1,2)	**-13499.51**
ARMA(1,3)	-13498.15
ARMA(2,1)	-13472.96
ARMA(2,2)	-13498
ARMA(2,3)	-13496.4

Source: Authors' calculations.

We decided to backtest the best models of those eight analysed, since every model has four different distributions of residuals (Table A1–A8). Once again, we used AIC criteria (modified with Bayesian information criterion [BIC] criteria) to choose one model: for example, we have the ARMA(1,2)A–PARCH(1,1) model with a normal distribution of residuals, the ARMA(1,2)–APARCH(1,1) model with a Student-t distribution of residuals, the ARMA(1,2)–APARCH(1,1) model with a skewed Student-t distribution of residuals and the ARMA(1,2)–APARCH(1,1) model with a reparameterised JSU distribution of residuals. The chosen models are: the ARMA(1,2)–APARCH(1,1) model with a JSU distribution of residuals, the ARMA(1,2)–GARCH(1,1) model with a JSU distribution of residuals, the ARMA(1,2)–TS GARCH(1,1) model with a Student-t distribution of residuals, the ARMA(1,2)–T GARCH(1,1) model with a Student-t distribution of residuals, the ARMA(1,2)–NARCH(1,1) model with a Student-t distribution of residuals, the ARMA(1,2)–GJR-GARCH(1,1) model with a JSU distribution of residuals, the ARMA(1,2)–EGARCH(1,1) model with a JSU distribution of residuals, and the ARMA(1,2)–IGARCH(1,1) model with a JSU distribution of residuals (parameter estimates of these chosen models are marked in bold in Tables A1–A8).

In our backtesting and forecasting methodology we analysed the following approach of a sliding window of 100 daily returns data as the basis for model estimation. This methodology allows us to perform a rolling estimation and forecasting of a proposed model, returning the VaR at specified levels. More importantly, it returns the distributional forecast parameters necessary to calculate any required measure on the forecast density. In the estimation of the parameters of the model, as the daily returns of the following day were added, the oldest daily returns were dropped from the observation window. For example, with a window size of 100, the window is placed between the 1st and the 100th data points, the model is estimated, and the return forecast is obtained for the 101st day at different quantiles. Next, the window is moved one day ahead to the 2nd and 101st data points to obtain a forecast of the 102nd day return with updated parameters from this new sample.

In the case of no-convergence in some or all the windows, there is NA (not applicable, due to insufficient number of convergences). Non-convergence here implies either a failure of the solver to converge to a solution (a global failure), or a failure to invert the resulting Hessian (a local failure).

Table 4. Backtesting results for MONEX stock index daily returns.

Model	The Kupiec test 95%		The Kupiec test 99%		The Christoffersen test 95%	The Christoffersen test 99%
	Actual number of exceedances (expected is **50**)	Test statistic	Actual number of exceedances (expected is **10**)	Test statistic	Test statistic	Test statistic
ARMA(1,2)–APARCH(1,1)	34	6.01 (0.014)	8	0.43 (0.512)*	6.034 (0.049)	NA
ARMA(1,2)–GARCH(1,1)	32	7.739 (0.005)	8	0.43 (0.512)*	7.74 (0.021)	NA
ARMA(1,2)–TS-GARCH(1,1)	35	5.237 (0.022)	9	0.103 (0.749)*	5.285 (0.071)*	NA
ARMA(1,2)–T GARCH(1,1)	35	5.237 (0.022)	9	0.103 (0.749)*	5.285 (0.071) *	NA
ARMA(1,2)–NARCH(1,1)	25	15.995 (0)	4	4.706 (0.03)	NA	NA
ARMA(1,2)–GJR-GARCH(1,1)	32	7.739 (0.005)	8	0.43 (0.512)*	7.74 (0.021)	NA
ARMA(1,2)–EGARCH(1,1)	36	4.524 (0.033)	8	0.43 (0.512)*	4.604 (0.1)*	NA
ARMA(1,2)–IGARCH(1,1)	32	7.739 (0.005)	8	0.43 (0.512)*	7.74 (0.021)	NA

Source: Authors' calculations (p-values are given in parentheses).

In the first part of Table 4, the results of the Kupiec test are presented for eight GARCH models. Besides the theoretically expected number of excesses for the 5% and 1% level of significance, the number of actual excesses is presented for selected quantiles associated with the distribution. The VaR forecast for the MONEX stock index daily returns obtained by the eight GARCH models gives unsatisfactory results and consequently fails this test for the 95% level of confidence. In other words, the null hypothesis of correct unconditional coverage can safely be rejected for the 95% level of confidence. We can see that the actual number of excesses is far too small compared to theoretically 50 expected excesses. On the other hand, VaR violations for the 99% level of confidence are far better than in the case of the 95% level of confidence. To be specific, seven of the analysed models the (joint ARMA[1,2] with APARCH, GARCH, TS-GARCH, T GARCH, GJR-GARCH, EGARCH and IGARCH) passed the Kupiec test with a 99% confidence level. We can conclude that the ARMA(1,2)–NARCH(1,1) model overestimates the VaR forecast for MONEX stock index daily returns, while the other seven models give correct VaR estimates regarding unconditional coverage (the actual number of exceedances is close to the expected number).

In the second part of Table 4. we determine whether the estimated GARCH models suffer from volatility clustering or not by use of the Christoffersen test. Three models, ARMA(1,2)–TS GARCH(1,1), ARMA(1,2)–T GARCH(1,1) and ARMA(1,2)–EGARCH(1,1) passed the joint Christoffersen test with a 95% confidence level. It seems that these three models are appropriate for capturing volatility clustering, since all of them failed for a number of exceptions. The other models reject the null hypothesis of correct exceedances and the independence of failures and are therefore disqualified, with a confidence level of 95%. It follows from the rejection of *LRuc* at the same level of 95% that a combined hypothesis of correct unconditional coverage and independence can be safely rejected. The joint Christoffersen test with a 99% confidence level could not be carried out due to no-convergence in some of all the possible windows.

Finally, the Pearson Q test was conducted, at three levels of confidence: 90%, 95% and 99%. We isolated violations of VaR for the selected eight models in two bins (with 90%), three bins (95%) and four bins (99%), and calculated Q statistics according to Formula (26). The results of the Pearson's Q test are shown in Table 5. As already stated, this statistic Q is distributed according to the chi squared distribution with k-1 degrees of freedom, so the critical values for this test are 6.635, 5.991 and 6.251 for 1% (k=2), 5% (k=3) and 10% (k=4) level of significance, respectively. Consequently, none of the analysed model passes the Pearson's Q test, whether at 90%, 95% or 99%.

Table 5. Pearson's Q test backtesting results of selected models.

Models	Under Reporting Level		
	1%	5%	10%
ARMA(1,2)–APARCH(1,1)	56.667 (<10-4)	61.788 (<10-4)	63.021 (<10-4)
ARMA(1,2)–GARCH(1,1)	91.476 (<10-4)	128.118 (<10-4)	128.757 (<10-4)
ARMA(1,2)–TS-GARCH(1,1)	79.769 (<10-4)	112.356 (<10-4)	112.601 (<10-4)
ARMA(1,2)–T GARCH(1,1)	44.717 (<10-4)	48.620 (<10-4)	49.688 (<10-4)
ARMA(1,2)–NARCH(1,1)	69.289 (<10-4)	71.099 (<10-4)	71.185 (<10-4)
ARMA(1,2)–GJR-GARCH(1,1)	54.908 (<10-4)	60.434 (<10-4)	61.767 (<10-4)
ARMA(1,2)–EGARCH(1,1)	54.908 (<10-4)	60.434 (<10-4)	61.767 (<10-4)
ARMA(1,2)–IGARCH(1,1)	80.056 (<10-4)	81.807 (<10-4)	82.223 (<10-4)

Source: Authors' calculations (p-values are given in parentheses).

VI. Conclusion

This article evaluates the performance of symmetric and asymmetric GARCH models based on the normal distribution, Student-t distribution, skewed Student-t distribution and reparameterised JSU distribution of residuals in estimating and forecasting market risk in the Montenegrin stock market. The relative performance of GARCH type models was tested using the daily returns of the Montenegrin stock index MONEX. The period of examination was from 5 January 2004 to 21 February 2014, which is a sufficiently long period, including tranquil periods as well as crisis years. The growing interest of foreign investors in investing in emerging financial markets and the increased turbulence in these markets in times of crisis have both highlighted the importance of adequate market risk quantification and forecasting.

Descriptive statistics for the MONEX stock index showed the presence of skewness and kurtosis. Using AIC criteria, the ARMA(1,2) model was chosen as adequate for modelling the series of logarithmic returns. The significant presence of ARCH and clustering effects were detected in the residuals. The residuals obtained from different estimated GARCH models showed no presence of the ARCH effect using the ARCH-LM test. The models have appropriate statistical characteristics in terms both of autocorrelation and eliminating ARCH effects.

The results of backtesting showed that none of the eight models passed the Kupiec test with a 95% confidence level. On the other hand, seven analysed models, out of eight (specifically, the joint ARMA(1,2) with APARCH, GARCH, TS-GARCH, T GARCH, GJR-GARCH, EGARCH and IGARCH) passed the Kupiec test with a 99% confidence level. The results of the Christoffersen test revealed three models, ARMA(1,2)–TS GARCH(1,1) with a Student-t distribution of residuals, the ARMA(1,2)–T GARCH(1,1) model with a Student-t distribution of residuals, and ARMA(1,2)–EGARCH(1,1) with a reparameterised JSU distribution of residuals that passed the joint Christoffersen test with a 95% confidence level. It seems that these three models are appropriate for capturing volatility clustering, since all of them failed for a number of exceptions. The other models rejected the null hypothesis of correct exceedances and the independence of failures and are therefore disqualified. The conclusion is that the ARMA(1,2)–NARCH(1,1) model is not adequate for capturing volatility clustering (the Christoffersen test was failed at a 95% confidence level). It also failed the correct number of exceptions under the Kupiec test at a 95% and 99% confidence. That is the reason to believe that the ARMA(1,2)–NARCH(1,1) model is not accurate in measuring VaR for use in relation to the Montenegrin stock market. Finally, none of the analysed model passed the Pearson's Q test, whether at 90%, 95% or 99%.

Cerovic and Karadzic (2015) showed that the peaks-over-threshold model passed the Kupiec test with 95% and 99% confidence levels. However, this model fails to pass the Christoffersen joint test for both 95% and 99%. From these results, we conclude that the peaks-over-threshold model for measuring VaR is not appropriate for capturing volatility clustering on the Montenegrin stock market and should be improved, though it is accurate as far as correct exceedances are concerned. In this article, we found that ARMA(1,2)–TS GARCH(1,1) with a Student-t distribution of residuals, the ARMA(1,2)–T GARCH(1,1) model with a Student-t distribution of residuals, and ARMA(1,2)–EGARCH(1,1) with a reparameterised JSU distribution of residuals passed the joint Christoffersen test with a 95% confidence level. We can conclude that these three models are appropriate for capturing

volatility clustering, while all of them failed Kupiec test for a number of exceptions at a 95% confidence level.

Once again, we will stress what Campbell said: any evidence of an accurate VaR model can be described only by hit sequences that satisfy both unconditional coverage and independence properties (Campbell, 2005). Perhaps the key to an accurate VaR model for the Montenegrin stock market lies in a combination of these three GARCH models and a Pareto distribution of residuals, similar to that used in the peaks-over-threshold method (extreme value theory). That is an idea for further analysis, while it would also be interesting to include more stock exchange indices from similar markets in South-Eastern and Central Europe in order to see whether this conclusion can be generalised to other frontier and emerging markets.

Notes

1. For comparison, the log-likelihood function of the Student-t distribution is as follows:

$$L_{st} = T(\ln \Gamma(\tfrac{v+1}{2}) - \ln \Gamma(\tfrac{v}{2}) - \tfrac{1}{2}\ln(\pi(v-2)) - \tfrac{1}{2}\sum_{t=1}^{T}\left[\ln(\sigma_t^2) + (1+v)\ln\left[1 + \left(\tfrac{z_t^2}{\sigma_t^2(v-2)}\right)\right]\right].$$

2. For quantiles more extreme than the 5th percentile, the partition should be [0.00, 0.01], [0.01, 0.05] and [0.05, 1].
3. For the purpose of this analysis, we used the official data available on the website of the Montenegro Stock Exchange (http://www.montenegroberza.com).
4. All empirical results expressed in this article are calculated using the R program package.
5. JB represents the Jarque-Bera statistics for normality testing and Box-Ljung represents the Box-Ljung statistics for testing autocorrelation in standardised residuals and squared standardised residuals, respectively.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alexander, C. O., & Leigh, C. T. (1997). On the covariance models used in Value at Risk models. *The Journal of Derivatives, 4*, 50–62. doi:10.3905/jod.1997.407974.

Andjelic, G., Djakovic, V., & Radisic, S. (2010). Application of VaR in emerging markets: A case of selected Central and Eastern European Countries. *African Journal of Business Management, 4*, 3666–3680. Retrieved from http://www.academicjournals.org/ajbm/

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical Methodology, 1*, 105–128. doi:10.1016/j.stamet.2004.08.004.

Bao, Y., Lee, T., & Saltoglu, B. (2006). Evaluating predictive performance of value-at-risk models in emerging markets: A reality check. *Journal of Forecasting, 25*, 101–128. doi:10.1002/for.977.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327. doi:10.1016/0304-4076(86)90063-1.

Bucevska, V. (2012). An Empirical evaluation of GARCH models in value-at-risk estimation: Evidence from the Macedonian stock exchange. *Business Systems Research, 4*, 49–64. doi:10.2478/bsrj-2013-0005.

Campbell, S. (2005). *A review of backtesting and backtesting procedures* (Discussion Paper No. 2005-21), Washington, D.C.: Finance and Economics Discussion Series, Federal Reserve Board.

Cerovic, J., & Karadzic, V. (2015). Extreme value theory in emerging markets: Evidence from montenegrin stock exchange. *Economic Annals, 60*, 87–116. doi:10.2298/EKA1506087C.

Cheng, S., Liu, Y., & Wang, S. (2004). Progress in risk measurement. *Advanced Modelling and Optimization, 6*(1), 1–20. Retrieved from: http://camo.ici.ro/journal/jamo.htm

Christoffersen, P., Hahn, J., & Inoue, A. (2001). *Testing and comparing value-at-risk measures* (Report No. 2001s-03), Montreal: CIRANO Scientific Series.

Da Silva, A., Beatriz, V., & de Melo Mendes, B. (2003). Value-at-Risk and extreme returns in Asian stock markets. *International Journal of Business, 8*, 17–40. Retrieved from http://www.craig.csufresno.edu/ijb/index.htm

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance, 1*, 83–106. doi:10.1016/0927-5398(93)90006-D.

Dowd, K. (1998). *Beyond value at risk: The new science of risk management.* New York, NY: John Wiley & Sons.

Dowd, K. (2005). *Measuring market risk* (2nd ed.). Chichester: John Wiley & Sons.

Duffie, D., & Pan, J. (1997). An overview of value at risk. *The Journal of Derivatives, 4*, 7–49. doi:10.3905/jod.1997.407971.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007. doi:10.2307/1912773.

Gencay, R., & Selcuk, F. (2004). Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of Forecasting, 20*, 287–303. doi:10.1016/j.ijforecast.2003.09.005.

Geweke, J. (1986). A Comment. *Econometric Reviews, 5*, 57–61. doi:10.1080/07474938608800097.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance, 48*, 1779–1801. doi:10.1111/j.1540-6261.tb05128.x.

Guermat, C., & Harris, D. F. (2002). Forecasting value at risk allowing for time and kurtosis of portfolio returns. *International Journal of Forecasting, 18*, 409–419. doi:10.1016/S0169-2070(01)00122-4.

Higgins, M. L., & Bera, A. K. (1992). A class of nonlinear arch models. *International Economic Review, 33*, 137–158. doi:10.2307/2526988.

Jorion, P. (1996). Risk^2: Measuring the risk in value at risk. *Financial Analysts Journal, 52*, 47–56. doi:10.2469/faj.v52.n6.2039.

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). New York, NY: McGraw-Hill.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives, 3*, 73–84. doi:10.3905/jod.1995.407942.

Lambert, P., & Laurent, S. (2001). *Modelling financial time series using GARCH-type models with a skewed student distribution for the innovations* (Discussion Paper No. 0125). Institute de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Longerstaey, J. & More, L. (1995). *Introduction to riskmetrics.* New York, NY: Morgan Guaranty Trust Company.

Manganelli, S., & Engle, R. F. (2001). *Value at risk models in finance* (Working Paper No. 75). European Central Bank Working Paper Series, Frankfurt am Main.

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management.* Princeton: Princeton University Press.

Miletic, M., & Miletic, S. (2013). Measuring value at risk on emerging markets: Empirical evidence from Serbian stock exchange. *Facta Universitatis, Series: Economics and Organization, 10*, 25–37. Retrieved from http://facta.junis.ni.ac.rs/eao/eao.html

Miletic, M. & Miletic, S. (2015). Performance of value at risk models in the midst of the global financial crisis in selected CEE emerging capital markets. *Economic Research-Ekonomska Istraživanja, 28*, 132–166. doi:10.1080/1331677X.2015.1028243.

Mladenovic Z., Miletic M., & Miletic S. (September 2012). *Value at risk in European emerging economies: An empirical assessment of financial crisis period.* Paper presented at the conference From Global Crisis to Economic Growth. Which Way to Take?, Belgrade, University of Belgrade, Faculty of Economics.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370. doi:10.2307/2938260.

Nikolic-Djoric, E., & Djoric, D. (2011). Dynamic value at risk estimation for BELEX15. *Metodološki zvezki, 8*, 79–98. Retrieved from http://www.stat-d.si/mz/

Orhan, M., & Köksal, B. (2012). A comparison of GARCH models for VaR estimation. *Expert Systems with Applications, 39*, 3582–3592. doi:10.1016/j.eswa.2011.09.048.

Pantula, S. G. (1986). Comment. *Econometric Reviews, 5*, 71–74. doi:10.1080/07474938608800099.

Rigby, R. A. & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society Series C (Applied Statistics), 54*, 507–554. doi:10.1111/j.1467-9876.2005.00510.x.

Schwert, W. G. (1989). Why does stock market volatility change over time? *The Journal of Finance, 44*, 1115–1153. doi:10.1111/j.1540-6261.1989.tb02647.x.

Taylor, S. (1986). *Modelling financial time series*. New York, NY: Wiley.

Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). Hoboken: John Wiley & Sons.

Wong, C. S., Cheng, Y. W., & Wong, Y. P. (2002). Market risk management of banks: Implications from the accuracy of VaR Forecast. *Journal of Forecasting, 23*, 22–33. doi:10.1002/for.842.

Zakoian, J.-M. (1994). Threshold Heteroskedastic models. *Journal of Economic Dynamics and Control, 18*, 931–955. doi:10.1016/0165-1889(94)90039-6.

Zikovic, S. (2007, June). *Measuring market risk in EU new member states*. Paper presented at the 13th Dubrovnik Economic Conference of Croatian National Bank, Dubrovnik, Dubrovnik, Croatia. Croatian National Bank.

Zikovic, S., & Aktan, B. (2009). Global financial crisis and VaR performance in emerging markets: A case of EU candidate state - Turkey and Croatia. *Proceedings of Rijeka Faculty of Economics, 27*, 145–170. Retrieved from http://www.efri.uniri.hr/en/proceedings

Appendix

Table A1. Parameter estimates of the ARMA(1,2)–APARCH(1,1) model with different distributions of the standardised residuals for MONEX.

Parameters	Normal	Student	Skew Student	Jsu
			Distribution	
Mu	0.000226 (0.654307)	-0.000157 (0.670126)	0.000189 (0.68701)	**0.000213 (0.66202)**
Ar1	0.963061 (0)	0.963209 (0)	0.963267 (0)	**0.964305 (0)**
ma1	-0.794060 (0)	-0.832269 (0)	-0.833189 (0)	**-0.836361 (0)**
Ma2	-0.125885 (0)	-0.083940 (0.000012)	-0.082885 (0)	**-0.081405 (0)**
Omega	0.000001 (0.94946)	0.000288 (0.269128)	0.000279 (0.27073)	**0.000243 (0.27801)**
Alpha1	0.138138 (0.34158)	0.302179 (0)	0.296572 (0)	**0.273527 (0)**
Gamma1	0.000196 (0.99685)	-0.082326 (0.111380)	-0.084803 (0.10182)	**-0.078912 (0.12665)**
Beta1	0.799418 (0)	0.738675 (0)	0.743112 (0)	**0.747698 (0)**
delta	2.591597 (0.46071)	1.305979 (0)	1.308111 (0)	**1.321624 (0)**
LogLikelihood	7256.651	7450.212	7450.994	**7453.546**
AIC	-5.7796	-5.9332	-5.9330	**-5.9350**
BIC	-5.7587	-5.9099	-5.9075	**-5.9095**
Ljung-Box test lag[1]	1.661 (0.1974)	0.01589 (0.8997)	0.02716 (0.8691)	**0.04479 (0.8324)**
ARCH LM test lag[7]	1.34267 (0.8523)	1.0267 (0.9092)	1.0308 (0.9085)	**1.0458 (0.9060)**

Source: Authors' calculations.

Table A2. Parameter estimates of the ARMA(1,2)–GARCH(1,1) model with different distributions of the standardised residuals for MONEX.

Parameters	Distribution			
	Normal	Student	Skew Student	**Jsu**
Mu	0.000219 (0.66922)	-0.000258 (0.512480)	0.000073 (0.880466)	**0.000107 (0.828856)**
Ar1	0.973912 (0)	0.962897 (0)	0.963024 (0)	**0.963660 (0)**
ma1	-0.809540 (0)	-0.826028 (0)	-0.827816 (0)	**-0.831292 (0)**
Ma2	-0.131377 (0)	-0.089206 (0)	-0.088141 (0)	**-0.085389 (0)**
Omega	0.000008 (0)	0.000014 (0.000019)	0.000014 (0.000021)	**0.000013 (0.000019)**
Alpha1	0.157764 (0)	0.288995 (0)	0.284567 (0)	**0.279777 (0)**
Beta1	0.817578 (0)	0.710003 (0)	0.714432 (0)	**0.719127 (0)**
LogLikelihood	7254.978	7445.05	7445.822	**7449.058**
AIC	-5.7799	-5.9307	-5.9305	**-5.9331**
BIC	-5.7636	-5.9121	-5.9096	**-5.9121**
Ljung-Box test lag[1]	2.240 (0.1345)	0.02202 (0.8820)	0.01446 (0.9043)	**0.01731 (0.8953)**
ARCH LM test lag[7]	1.16722 (0.8849)	1.7104 (0.7783)	1.7244 (0.7754)	**1.7625 (0.7674)**

Source: Authors' calculations.

Table A3. Parameter estimates of the ARMA(1,2)–TS GARCH(1,1) model with different distributions of the standardised residuals for MONEX.

Parameters	Distribution			
	Normal	**Student**	Skew Student	Jsu
Mu	0.018959 (0)	**-0.000316 (0.614214)**	0.000001 (0.997415)	0.000009 (0.975570)
Ar1	1 (0)	**0.964580 (0)**	0.964130 (0)	0.965226 (0)
ma1	-0.849621 (0)	**-0.836363 (0)**	-0.837860 (0)	-0.841420 (0)
Ma2	-0.129780 (0)	**-0.084316 (0.025644)**	-0.081768 (0.000050)	-0.079997 (0.000720)
Omega	0.000739 (0)	**0.001070 (0.000007)**	0.001048 (0.000007)	0.000977 (0.000007)
Alpha1	0.201094 (0)	**0.284288 (0)**	0.279542 (0)	0.259339 (0)
Beta1	0.804166 (0)	**0.749560 (0)**	0.753314 (0)	0.758752 (0)
LogLikelihood	7249.842	**7447.336**	7448.034	7450.695
AIC	-5.7758	**-5.9325**	-5.9322	-5.9344
BIC	-5.7595	**-5.9139**	-5.9113	-5.9135
Ljung-Box test lag[1]	10.09 (0.001488)	**0.1652 (0.6844)**	0.1960 (0.6580)	0.2627 (0.6083)
ARCH LM test lag[7]	1.44919 (0.8315)	**0.3488 (0.5548)**	0.9371 (0.9236)	0.9247 (0.9255)

Source: Authors' calculations.

Table A4. Parameter estimates of the ARMA(1,2)–T GARCH(1,1) model with different distributions of the standardised residuals for MONEX.

Parameters	Distribution			
	Normal	**Student**	Skew Student	Jsu
Mu	0.018906 (0)	**-0.000154 (0.706278)**	0.000163 (0.481959)	0.000155 (0.991346)
Ar1	1 (0)	**0.963351 (0)**	0.963432 (0)	0.964731 (0.233186)
ma1	-0.849586 (0)	**-0.835300 (0)**	-0.835633 (0)	-0.840334 (0.388743)
Ma2	-0.129833 (0)	**-0.081843 (0)**	-0.080679	-0.078417 (0.873226)
Omega	0.000737 (0)	**0.001073 (0.000005)**	0.001050 (0.000005)	0.000980 (0.000016)
Alpha1	0.200809 (0)	**0.280778 (0)**	0.275704 (0)	0.256452 (0)
Beta1	0.804482 (0)	**0.751695 (0)**	0.756025 (0)	0.760522 (0)
Gamma1	0.005402 (0.88445)	**-0.099301 (0.085636)**	-0.102598 (0.060254)	-0.095363 (0.872292)
LogLikelihood	7249.853	**7448.866**	7449.617	7452.07
AIC	-5.7750	**-5.9329**	-5.9327	-5.9347
BIC	-5.7564	**-5.9120**	-5.9095	-5.9114
Ljung-Box test lag[1]	10.20 (0.001401)	**0.1873 (0.6652)**	0.2267 (0.6339)	0.2842 (0.5939)
ARCH LM test lag[7]	1.440802 (0.8332)	**0.7655 (0.9484)**	0.7601 (0.9492)	0.7620 (0.9489)

Source: Authors' calculations.

Table A5. Parameter estimates of the ARMA(1,2)–N GARCH(1,1) model with different distributions of the standardised residuals for MONEX.

	Distribution			
Parameters	Normal	**Student**	Skew Student	Jsu
Mu	-0.003055 (0.002408)	**-0.000704 (0.258908)**	-0.000749 (0.280155)	-0.000606 (0.36875)
Ar1	0.986783 (0)	**0.971276 (0)**	0.971215 (0)	0.972163 (0)
ma1	-0.676030 (0)	**-0.824786 (0)**	-0.824608 (0)	-0.825966 (0)
Ma2	-0.272560 (0)	**-0.102437 (0.000058)**	-0.102444 (0.000077)	-0.102102 (0)
Omega	0 (0)	**0.000682 (0.442602)**	0.000689 (0.443160)	0.000493 (0.46238)
Alpha1	0.224662 (0)	**1 (0.000004)**	1 (0.000004)	0.779278 (0)
delta	3.861294 (0)	**1.677522 (0)**	1.675495 (0)	1.698972 (0)
LogLikelihood	7000.218	**7362.766**	7362.773	7365.475
AIC	-5.5767	**-5.8650**	-5.8643	-5.8664
BIC	-5.5605	**-5.8465**	-5.8433	-5.8455
Ljung-Box test lag[1]	0.09187 (7.618e-01)	**3.907 (4.807e-02)**	3.903 (4.819e-02)	3.846 (4.987e-02)
ARCH LM test lag[7]	87.44 (0)	**55.91 (1.810e-14)**	55.810 (1.921e-14)	57.566 (6.439e-15)

Source: Authors' calculations.

Table A6. Parameter estimates of the ARMA(1,2)–GJR-GARCH(1,1) model with different distributions of the standardised residuals for MONEX.

	Distribution			
Parameters	Normal	Student	Skew Student	**Jsu**
Mu	0.000226 (0.67073)	-0.000145 (0.800840)	0.000187 (0.75267)	**0.000225 (0.627538)**
Ar1	0.973959 (0)	0.961769 (0)	0.961598 (0)	**0.962438 (0)**
ma1	-0.809590 (0)	-0.824731 (0)	-0.825575 (0)	**-0.829398 (0)**
Ma2	-0.131398 (0)	-0.088067 (0.062830)	-0.087199 (0.11836)	**-0.084478 (0.019141)**
Omega	0.000008 (0)	0.000014 (0.000019)	0.000014 (0.00002)	**0.000013 (0.000016)**
Alpha1	0.157758 (0)	0.288868 (0)	0.283024 (0)	**0.278763 (0)**
Beta1	0.817567 (0)	0.709084 (0)	0.713733 (0)	**0.718015 (0)**
Gamma1	-0.001498 (0.96210)	-0.060252 (0.194572)	-0.060375 (0.21902)	**-0.057975 (0.158239)**
LogLikelihood	7254.98	7446.037	7446.811	**7449.946**
AIC	-5.7791	-5.9307	-5.9305	**-5.9330**
BIC	-5.7605	-5.9097	-5.9072	**-5.9097**
Ljung-Box test lag[1]	2.239 (0.1346)	0.02146 (0.8835)	0.01272 (0.9102)	**0.01705 (0.8961)**
ARCH LM test lag[7]	1.16613 (0.8851)	1.548 (0.8118)	1.5584 (0.8096)	**1.6105 (0.7989)**

Source: Authors' calculations.

Table A7. Parameter estimates of the ARMA(1,2)–EGARCH(1,1) model with different distributions of the standardised residuals for MONEX.

	Distribution			
Parameters	Normal	Student	Skew Student	**Jsu**
Mu	0.000757 (0.097021)	-0.000198 (0.323847)	0.000194 (0.65746)	**0.000234 (0.621608)**
Ar1	0.982859 (0)	0.963156 (0)	0.962543 (0)	**0.963860 (0)**
ma1	-0.817860 (0)	-0.830713 (0)	-0.830929 (0)	**-0.834861 (0)**
Ma2	-0.136818 (0)	-0.086628 (0.031434)	-0.085010 (0)	**-0.082865 (0)**
Omega	-0.387195 (0)	-0.552945 (0)	-0.537073 (0)	**-0.540089 (0.000001)**
Alpha1	0.009713 (0.388725)	0.030401 (0.158961)	0.031829 (0.15174)	**0.028345 (0.178914)**
Beta1	0.951753 (0)	0.934343 (0)	0.936240 (0)	**0.936843 (0)**
Gamma1	0.299417 (0)	0.453802 (0)	0.444478 (0)	**0.419602 (0)**
LogLikelihood	7234.642	7444.543	7445.496	**7448.455**
AIC	-5.7629	-5.9295	-5.9294	**-5.9318**
BIC	-5.7443	-5.9085	-5.9062	**-5.9085**
Ljung-Box test lag[1]	3.278 (0.0702)	0.0134 (0.9079)	0.0299 (0.8627)	**0.04689 (0.8286)**
ARCH LM test lag[7]	0.51961 (0.9766)	0.9378 (0.9235)	0.9267 (0.9252)	**0.9128 (0.9273)**

Source: Authors' calculations.

Table A8. Parameter estimates of the ARMA(1,2)–IGARCH(1,1) model with different distributions of the standardised residuals for MONEX.

Parameters	Distribution			
	Normal	Student	Skew Student	**Jsu**
Mu	0.000203 (0.6873)	-0.000258 (0.511791)	0.000073 (0.880137)	**0.000107 (0.828874)**
Ar1	0.974963 (0)	0.962885 (0)	0.963009 (0)	**0.963657 (0)**
ma1	-0.811121 (0)	-0.826066 (0)	-0.827838 (0)	**-0.831353 (0)**
Ma2	-0.132347 (0)	-0.089157 (0)	-0.088087 (0)	**-0.085310 (0)**
Omega	0.000006 (0)	0.000014 (0.000016)	0.000014 (0.000018)	**0.000013 (0.000013)**
Alpha1	0.174981 (0)	0.289902 (0)	0.285588 (0)	**0.280856 (0)**
Beta1	0.825019 (NA)	0.710098 (NA)	0.714412 (NA)	**0.719144 (NA)**
LogLikelihood	7251.571	7445.081	7445.851	**7449.058**
AIC	-5.7780	-5.9315	-5.9313	**-5.9339**
BIC	-5.7640	-5.9152	-5.9127	**-5.9153**
Ljung-Box test lag[1]	1.698 (0.1926)	0.02265 (0.8804)	0.01519 (0.9019)	**0.01827 (0.8925)**
ARCH LM test lag[7]	1.4480 (0.8317)	1.7129 (0.7778)	1.7270 (0.7748)	**1.765 (0.7668)**

Source: Authors' calculations.