1. A NEW SEMIPARAMETRIC MIRRORED HISTORICAL SIMULATION VALUE-AT-RISK MODEL

Nikola RADIVOJEVIĆ[1]
Luka FILIPOVIĆ[2]
Tomislav D. BRZAKOVIĆ[3]

Abstract

In this paper, the authors have developed and presented a new semiparametric value-at-risk (VaR) model for the assessment of market risk. The model is based on the theoretical foundation of the Historical Simulation (HS) method. The basic intention was to develop a new model that would be easy to implement and able to envelop the empirical features of returns, such as leptokurtosis, asymmetry, autocorrelation, and heteroscedasticity, and also to improve risk estimation in the tail distribution for the sample size and the confidence level prescribed by the Basel III standard. To obtain the answers to the question of whether the new model is an improvement against the popular improvements of the HS method, its performances were tested in terms of adherence to the backtesting rules of the Basel Accord and also compared with the backtesting results of the popular improvements of the HS method. The backtesting results justify the expectations of the new model.

Keywords: risk estimation, emerging markets, conditional value-at-risk, Basel III standard, Berkowitz test, bootstrap method

JEL Classification: G24, C22, C52, C53

1. Introduction

The evaluation of the performance of risk estimation models and the selection of the models that should be considered as efficient are one of the most important tasks of financial risk management. In practice, the most commonly used risk estimation method is the Historical Simulation (HS) method, with the notion that the HS model is used in the paper as the generic name for all the models (the nonparametric and semiparametric models constructed on the

[1] *Technical College at applied studies, Kragujevac, Kosovska 8, 34000 Kragujevac, Serbia. E-mail: radivojevic034@gmail.com.*

[2] *EuroAudit, Enterprise for Audit, Accounting and Financial Consulting, Despota Stefana 12, 11000 Belgrade, Serbia. E-mail: filipovic.luka010@gmail.com.*

[3] *University Business Academy in Novi Sad, Faculty of Applied Management, Economics and Finance in Belgrade, Serbia. E-mail: tomislav.brzakovic@mef.edu.rs.*

theoretical grounds of the HS method). Standard HS models are at the basis of all these models. Hence, it is understandable that many authors have worked on its improvement. The standard HS model is based on the assumption that the trend of the risk factors of previous periods contains all the information necessary to estimate risk, which is equivalent to the assumption of a constant distribution. It only requires the estimation of the appropriate quantile by using quantiles for a set period of the past sample returns. In situations when the assumption is not fulfilled, the application of the HS may at best produce the estimates that will satisfy the 'unconditional coverage' criterion, but not the property of an independent distribution of breaks. More precisely, risk estimations will not meet the conditional coverage criterion. In addition to the above-mentioned issues, the use of extremely long series of data increases the probability of violating the assumption of constant distribution.

Closely related to the foregoing is the fact that, when building empirical density, the standard HS assigns an equal probability weight to each observation, which, according to Radivojevic *et al*., (2017b) is equivalent to the assumption that historically simulated returns are IID through time. However, this manner of weighting creates the problem of adequately capturing the dynamics of returns in time-varying volatility. The implications of this are very clear: risk estimates will be overestimated if the exceptional volatility period is captured, on the one hand, or underestimated if the volatility quiet period is captured, on the other. Another significant drawback of the HS is the fact that, by applying it, the tail distribution for the sample size and the confidence level prescribed by the Basel II and III standards are difficult to estimate. This problem becomes more and more evident as the holding period for which an assessment is being made increases, as is evidenced by numerous studies, such as Şener *et al*. (2012), Rossignolo *et al*. (2012, 2013), Radivojevic *et al*. (2016b, 2017a).

Numerous authors have worked on improving the standard HS. The improvements range from very simple, *i.e.* those focused on solving the tail estimation problem or capturing time-varying volatility, to those very complex solutions focused on reducing both drawbacks of the HS. Simple solutions are only focused on one drawback of the HS, simultaneously ignoring other ones, and they cannot primarily be used in the emerging markets. With the complexity of a solution, the computer-related demanding levels of the implementation of the application of the HS model also grows. Hence, it is imperative that a model able to equally well capture fat tails (*leptokurtosis*), as well as time-varying volatility, which is simultaneously easy to implement, should be developed.

Given these requirements, a new market risk estimation model is introduced in this paper. The model is designed as a new semiparametric mirrored historical simulation value-at-risk model – ARMA(p,q)-GARCH(p,q)MHS. The idea is to take advantage of the MHS model in reducing the tail estimation problem by the HS and the advantages of using the ARMA and GARCH models to transform historical returns into IID returns. This approach should lead to a reduction in both drawbacks of the HS and lead to its improvement without increasing the implementation costs. In order to answer the question whether such a new model is better than the most popular and the most widely used improvements of the HS model, ranging from the simplest, such as the Mirrored Historical Simulation (MHS) and Filtered Historical Simulation (FHS) models, to complex ones, such as the Dynamic Historical Simulation (DHS) model, the backtesting results of the new model will be compared with the backtesting results of the mentioned models in the context of the Basel III standard.

2. Literature Review

As has already been mentioned in the Introduction, the HS is one of the most widely used market risk estimation methods. Therefore, the fact that this method was many authors' matter of interest is not surprising. The first improvements of the HS were made by Holton (1998), Boudoukh *et al.* (1998) and Hull and White (1998). Holton's solution indicates that collected data are doubled according to the 'mirror effect' principle, in which manner the number of scenarios doubles with a reduction in the standard error. Boudoukh *et al.* (1998) developed the so-called hybrid market risk estimation approach. Their solution implies a combination of the standard HS model and an exponential smooth approach to VaR estimation. In theory, the hybrid approach eliminates the listed drawbacks of the HS. However, many empirical studies have not shown significant improvements. Hull and White (1998) introduced the so-called Filtered Historical Simulation model (FHS), which is based on a combination of the HS approach and the EWMA/GARCH approach to forecasting volatility. By combining the hybrid model with Hull-White's model, Zikovic tried to develop a model of the HS that would be adequate for the emerging markets. His research studies showed that the model performed better on the Croatian capital market than the two above-mentioned VaR models. Zikovic and Prohaska (2010) were the first who tried to improve the applicability of the hybrid approach by developing a procedure for determining the optimal decline factor, which was tested on a sample of nine Mediterranean stock markets. The results were very weak. Zikovic (2013) suggested the Hybrid Historical Simulation model (HHS), which is based on a combination of a modified recursive bootstrap procedure and the parametric GARCH approach to volatility forecasting. Barone-Adesi and Giannopoulos (2001) were the first attempting to improve the HS by using the bootstrap method. They implemented the bootstrap method known as the Filtered Historical Simulation, random drawing with the replacement from the original sample of standardized residuals, the parameters being constant in all bootstrap replicates. They claim that this approach does not take the existence of volatility clusters into consideration as a consequence of the IID assumption. A similar methodology was proposed by Brandolini and Colucci (2012). Starting from the ideas of Babu and Singh, Radivojevic *et al.* (2017a) proposed a new Historical Bootstrap VaR model. Alemany *et al.* (2013) proposed an interesting nonparametric model for VaR estimates. Their model is based on the double transformation of the kernel estimation of the cumulative distribution function. However, the model is more useful for measuring operating risk rather than market risk. In order to capture heavy tails and heteroscedasticity in financial data, Bee (2012) presented the Dynamic Historical Simulation model (DHS), which is quite similar to the FHS model proposed by Fernandez. The model performs very well at an extremely high confidence level, but the research study covers developed markets.

3. The Theoretical Background of the New Semiparametric Mirrored HS VaR Model

The new semiparametric model is named ARMA-GARCH-Bootstrap HS model. The development of the model itself started with the idea of developing a new model that will be easy to implement and which will reduce the main drawbacks of the HS. Namely, the idea is to take advantage of the MHS model in reducing the tail estimation problem by the HS and the advantages of the use of the ARMA and GARCH models to transform historical returns into IID returns. This approach should lead to a reduction in both drawbacks of the HS and to its improvement, without increasing the implementation costs. Given the fact that, to a certain extent, the MHS model solves the problem of the insufficient number of the observations that fall into tail distribution, the idea is to improve the standard HS model by incorporating in the dataset the returns that will satisfy the assumption of an IID. The basic intention implies the transformation of original data into IID returns by using the model that can capture autocorrelation both in returns and in squared returns. In other words, the intention is to successfully capture both dependencies by the simple ARMA(p,q)-GARCH(p,q) model and then multiply the returns obtained in such a manner by applying the mirror principle. In this way, a sufficient number of observations needed for risk estimation (tail distribution) will be obtained according to the rules and requirements of the Basel II and III standards.

The implementation of the model indicates a few steps. The first step implies fitting the ARMA(p,q) model into a series of historical returns in order to ensure that residuals are IID, as follows:

$$r_i = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i} + \sum_{i=1}^{q} \theta \varepsilon_{t-i} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = \eta_t \sqrt{\sigma_t^2} \tag{2}$$

where: $\eta_t \sim$ IID N(0,1) .

The next step implies fitting the GARCH(p,q) model into the obtained residuals:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{3}$$

The third step implies that the residuals (ε_t) are divided by the corresponding conditional GARCH(p,q) volatility forecast (σ_t), with the aim of obtaining standardized residuals (z_t), as follows:

$$z_t = \frac{\varepsilon_t}{\sigma_t} \tag{4}$$

In the next, fourth step, a series of historical returns are generated by such standardized residuals:

$$\hat{r}_{t+1} = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i+1} + \sum_{i=1}^{q} \theta z_{t-i+1} + z_{t+1} \tag{5}$$

In this way, the returns that will satisfy the assumption of IID are obtained. Furthermore, these returns will be multiplied according to the mirror principle. VaR estimation is obtained as follows:

$$VaR_{\hat{N}=1|\hat{N}}^{cl} \equiv r_w((\hat{T}+1)cl) \tag{6}$$

while/where $r_w((\hat{T}+1)cl)$ is taken from the ordered series of returns {$r_w(1)$, $r_w(2)$..$r_w(\hat{T})$}, noting that $\hat{T}$ represents a set of real and mapped values.

By applying this solution, not only are the returns adequate for the use of the HS obtained, but those eliminating another lack of the HS with a rapidly declining number of observations with an increase in the holding period are also obtained. In other words, the presented model exploits the advantages of the nonparametric models because it is freed from the assumption of distribution, which enables it to capture fat tails and other deviations of the empirical distribution. On the other hand, the application of the GARCH model allows for taking advantage of the application of the conditional volatility forecast, which allows us to comprehend time-varying volatility.

4. The Data and Methodology of Analysis

As early as in the 1960s, Mandelbrot (1963) and Fama (1965) demonstrated that the return series of the daily returns of securities had a distribution deviating from normal distribution and the assumptions of the identical and independent distribution, *i.e.* that returns are not independent incidental variables following the marginal process, but they are significantly intercorrelated variables showing a strong tendency to group into clusters. Today, these characteristics are typical of return series independently of the financial market development degree, yet noting that the deviation degree is still determined by the market development degree. This is testified to by numerous studies of both developed and undeveloped markets, such as: Bekaert *et al*. (2002, 2003), Patel (2003), Barry and Rodriguez (2004), Tokat and Wikas (2004), Dunis and Shanon (2005), Nuti (2009), Zikovic and Atkan (2009), Zikovic and Filer (2013), Radivojevic *et al*. (2016b, 2017b), Kostadinovic, Radojcic (2017). Hence, the examination of the validity of a VaR model implies testing the performances of the model on the financial markets of a different development degree. Bearing in mind this requirement, the model was tested on a sample of 15 capital markets, namely on the capital markets of Serbia, Croatia, Romania, Bulgaria, Slovenia, Turkey, Slovakia, the Czech Republic, Hungary, Montenegro, Bosnia and Herzegovina, Estonia, Latvia, and Lithuania. The first six markets represent the frontier markets, the second four markets are the emerging markets, the third two markets are the undeveloped markets, whereas the last three markets are the developed markets. The markets were selected bearing in mind the common historical heritage with respect to the social, political and economic order from the Second World War to the beginning of the 1990s and the commencement of transition process in these states, as well as the degree of the successfulness of the implementation of the transition process(es) and the development of the financial market. These markets are characterized by the fact that there were frequent structural changes during the last few decades. It is generally known that the VaR parametric models are not recommendable for the markets where there are frequent changes in the correlation matrices and structural changes. Therefore, it is important that a model founded on the theoretical grounds of the HS method as much as possible should be developed for markets like these.

The daily logarithmic returns of the stock indices from these markets were used for the analysis of the performance of the models. The tested stock indices are: the BELEXline index (Serbia), the BET index (Romania), the SOFIX index (Bulgaria), the CROBEX index (Croatia), the SBITOP index (Slovenia), the XU100 index (Turkey), the SAX index (Slovakia), the PX index (the Czech Republic), the BUX index (Hungary), the MONEX index (Montenegro), the SAXS10 index (Bosnia and Herzegovina), the OMXT index (Estonia), the OMXR index (Latvia), and the OMXV index (Lithuania). The research period was from February 1st, 2014, to February 1st, 2017. The calculated VaR and CVaR figures are for the one-day ahead horizon for the period between February 1st, 2016, and February 1st, 2017, according to the Basel III standard (see Kellner and Rösch, 2016).

As the representative of the HS, the HS500 model was applied. As a representative of the FHS, the FHS500 model was applied, as prescribed by Hull and White (1998), with the

GARCH model volatility. As representative of the DHS, the model proposed by Bee (2012) was applied in this paper. By applying the new model, VaR estimates were obtained, as described in the previous part of the paper. CVaR estimates were obtained as prescribed by Acerbi and Tasche (2002).

In order to determine the degree of the compatibility between the characteristics of the real market conditions and the assumptions underlying the models, the basic characteristics of the distribution of the daily logarithmic returns of the selected indices were analyzed at the beginning of the paper. The descriptive statistics and the normality tests for the entire analyzed sample for the returns of the selected indices are presented in Table 1.

The descriptive statistics of the daily logarithmic returns of the selected indices partially confirm the results of the previous empirical research study, according to which the characteristics of the real market conditions deviate from the assumptions of the HS model. The standard deviations are in line with the average of the EU developing countries (see Radivojevic *et al.*, 2016b). This implies a relatively high level of fluctuations in the value of daily returns, which is confirmed by the difference between the minimum and maximum values. An analysis of the mean returns shows that there is a presence of a relatively low level of average returns, which is surprising since the prevailing attitude is that frontier and emerging markets have higher returns and offer a possibility of achieving high-risk premiums. One of the explanations may be found in a reduction in investment activities, primarily of foreign investors, during the observed period. The analysis further reveals that the indices have a leptocentric distribution, which is confirmed by the analysis of the coefficient asymmetry and the coefficients of kurtosis. The coefficients of kurtosis range from 1.875 to 27.90, which means that there is a higher likelihood of occurrence of extreme returns than those predicted by the normal distribution. The asymmetry coefficients range from -1.189, in the case of the OMXV index, to 0.069, in the case of the SOFIX index. The values of these coefficients indicate a higher likelihood of achieving extreme positive returns on the capital markets of Serbia and Bulgaria and extreme negative returns on the rest of the capital markets. In order to formally examine whether the returns follow a normal distribution, three tests were used, namely the Jarque-Bera, Doornik-Hansen, and Shapiro-Wilk tests. These tests were used to obtain robust estimations. The values of all three tests indicate that the null hypothesis of normality should be rejected. Engel's test was used to assess the presence of ARCH effects, which is based on a Lagrange Multiplier for the ARCH(1) model. The results are shown in Table 2.

The results of Engel's test are surprising – they reveal that the presence of the ARCH effect is not connected with the degree of development of these markets, given the fact that no presence of the ARCH effect was recorded on the undeveloped market, such as the Montenegrin market, whereas the presence of that effect was recorded on developed markets, such as Lithuania's capital market.

The results obtained on the previously described tests are indicative of the fact that the returns are not IID. It is necessary to model the returns as the ARMA-GARCH process. The estimates of the ARMA(p,q)-GARCH(p,q) model parameters are given in Table 4 (noting that the presented rates of the GARCH model are those with normal distribution, *i.e.* Student's t-distribution, depending on the value of the Log-Likelihood criterion), while the estimated parameters of the volatility models used to estimate risk by applying the FHS and DHS models are given in Table 3.

Table 1

The Descriptive Statistics of the Daily Logarithmic Returns of the Selected Indices`

	BELEX line	SOFIX	BET	MONEX	CROBEX	ATHEX	SBITOP	SAXS10	XU100	SAX	PX	BUX	OMXT	OMXR	OMXV
Mean	0.000	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.001	0.000
Stand.dev.	0.006	0.008	0.008	0.007	0.007	0.025	0.008	0.008	0.013	0.013	0.010	0.011	0.006	0.010	0.005
Kurtosis	2.102	8.065	8.392	2.897	2.897	7.842	3.417	6.556	1.857	13.409	2.792	3.304	9.259	27.90	9.228
Skewness	0.033	0.069	-1.059	-0.061	-0.061	-0.667	-0.547	-0.211	-0.250	-0.031	-0.356	-0.37	-0.979	2.539	-1.189
Range	0.057	0.104	0.098	0.063	0.063	0.289	0.073	0.087	0.125	0.184	0.092	0.112	0.074	0.175	0.057
Min. values	-0.032	-0.047	-0.065	-0.030	-0.030	-0.177	-0.047	-0.042	-0.071	-0.093	-0.047	-0.063	-0.052	-0.059	-0.038
Max. values	0.025	0.056	0.033	0.032	0.032	0.112	0.026	0.045	0.054	0.091	0.045	0.050	0.022	0.116	0.018
No. obs.	755	740	755	755	755	754	754	754	754	754	754	754	754	753	751
Jarque-Bera test	136.24	1975.22	2322.78	259.65	2296.35	1958.29	397.69	1334.85	67.16	5568.05	256.21	353.84	2773.05	24893.5	2801.17
	(1,8e-03)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Doornik-Hansen test	88.67	592.58	335.23	145.14	653.53	452.37	136.99	449.47	113.87	1124.95	118.99	153.29	436.68	571.38	334.30
	(4,6e-02)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Shapiro-Wilk test	0.98	0.91	0.07	0.96	0.70	0.91	0.96	0.89	0.98	0.87	0.97	0.97	0.93	0.80	0.91
	(1,3e-008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Source: Authors' calculations.
Note: The p-values are shown in parentheses.

Table 2

ARCH effects

	BELEX line	SOFIX	BET	MONEX	CRO BEX	ATH EX	SBI TOP	SAXS 10	XU100	SAX	PX	BUX	OMXT	OMXR	OMXV
Engel's test	0.499	9.000	29.565	0.081	45.7963	2.982	6.883	6.467	0.507	0.415	78.864	3.789	35.335	0.033	11.585
p-value	0.479	0.002	0.000	0.775	0.000	0.084	0.008	0.010	0.476	0.519	0.000	0.052	0.000	0.854	0.000

Source: Authors' calculations.
Note: The p-values are shown in parentheses.

Table 3

The Estimates of the Parameters of the GARCH(1,1) Models

		BELEX line	SOFIX	BET	MONEX	CRO-BEX	ATHEX	SBI-TOP	SAXS 10	XU 100	SAX	PX	BUX	OMXT	OMXR	OMXV
Parameters of GARCH(1.1)																
□	value	0.081	0.379	0.185	0.037	0.300	0.145	0.173	0.030	-	0.328	0.134	0.055	0.310	0.284	0.121
	p-value	0.025	0.136	0.077	0.03	0.000	0.000	0.003	0.002	-	0.000	0.000	0.010	0.000	0.000	0.000
□	value	0.867	0.510	0.727	0.928	0.700	0.852	0.513	0.954	-	0.361	0.759	0.834	0.162	0.493	0.631
	p-value	0.043	0.099	0.084	0.060	0.000	0.000	0.001	0.001	-	0.000	0.000	0.000	0.049	0.000	0.000
□	value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000	0.000	0.000	0.000	0.000	0.000
	p-value	0.000	0.000	0.000	0.000	0.000	0.022	0.000	0.024	-	0.000	0.000	0.000	0.000	0.000	0.000
Parameters of GARCH(1.1)-t(d)																
α	value	0.097	0.165	0.145	0.078	0.496	0.15	0.206	-	-	-	0.122	0.059	0.206	0.062	0.135
	p-value	0.033	0.078	0.042	0.044	0.000	0.000	0.003	-	-	-	0.000	0.010	0.000	0.024	0.000
β	value	0.833	0.660	0.739	0.851	0.601	0.8189	0.56	-	-	-	0.767	0.851	0.341	0.909	0.732
	p-value	0.065	0.107	0.075	0.077	0.000	0.000	0.000	-	-	-	0.000	0.000	0.018	0.000	0.000
ω	value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-	-	-	0.000	0.000	0.000	0.000	0.000
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-	-	-	0.000	0.021	0.000	0.000	0.000
υ	value	6.394	4.599	4.635	3.195	3.750	4.730	4.760	-	-	-	7.510	6.810	4.570	4.430	3.710
	p-value	1.526	-0.720	-0.829	-0.374	0.000	0.000	0.000	-	-	-	0.000	0.000	0.000	0.000	0.000

Source: Authors' calculations.

Note: The parameters of the GARCH models were rated by the maximum likelihood estimation and the quasi-maximum likelihood estimation.

Table 4

The Estimates of the ARMA(p,q)-GARCH(p,q) Model Parameters

	ATHEX	SBITOP	SAXS10	XU100	SAX
AR(1)	1.294	-0.469	0.863	0.986	0.556
AR(2)	-0.744	-0.590			
MA(1)	-1.213	0.612	-0.948	-1.000	-0.687
MA(2)	0.610	0.688			
α	0.147	0.184	0.033	-	0.289
β	0.821	0.598	0.952	-	0.451
ω	0.000	0.000	0.000	-	0.000
η	4.560	4.860	-	-	-
	PX	BUX	OMXT	OMXR	OMXV
AR(1)	0.980	-0.891	-1.061	-0.821	-0.078
AR(2)			-0.948		
MA(1)	-1.000	0.876	1.082	0.746	
MA(2)			0.993		
α	0.130	0.056	0.218	0.058**	0.126
β	0.739	0.856	0.323**	0.916	0.742
ω	0.000	0.000	0.000	0.000	0.000
η	7.380	8.560	4.610	3.350	3.640
	BELEXline	SOFIX	BET	MONEX	CROBEX
AR(1)	0.139	0.745	0.968	0.075	0.659
AR(2)		-1.035	-1.442		-0.986
AR(3)		0.480	0.374		-0.691
AR(4)		-0.651	-1.199		1.000
AR(5)			0.718		
AR(6)			-0.860		
MA(1)		-0.698	-0.967		
MA(2)		0.933	1.492		
MA(3)		-0.415	-0.449		
MA(4)		0.700	1.347		
MA(5)			-0.820		
MA(6)			0.932		
α	0.085	0.121	0.142	0.081	0.171
β	0.852	0.722	0.728	0.837	0.829
ω	0.000	0.000	0.000	0.000	0.000
η	6.680	4.610	4.990	3.180	-

Source: Authors' calculations.

All the rated parameters are statistically significant at the 1% confidence level, except for those marked with **. ** denotes significance at the 10% level. The parameters of the GARCH models were rated by the maximum likelihood estimation and the quasi-maximum likelihood estimation, whereas the estimates of the parameters of the ARMA(p,q) model were obtained by applying the Levenberg-Marquardt algorithm.

5. The Results of the Backtesting Procedure

Unlike VaR backtesting, CVaR backtesting is significantly more complex. This is the reason why the Basel III standard is not the prescribed way to backtest the validity of CVaR assessments. Many authors, such as Emmer *et al*. (2013), Terzic and Milojevic (2016) recommend different methods for CVaR backtesting and most of them agree that the first step in testing the validity of risk models implies VaR backtesting. For this purpose, Kupiec's unconditional coverage test (LR_{uc} test) and Christofferson's conditional coverage test (LR_{cc} test) were used in this paper. Both tests were done at a 5% significance level. The results of both tests are presented in Table 5.

Table 5

The Backtesting Results for the LR_{uc} Test and the LR_{cc} Test

	The LR_{uc} test for 99%VaR								
	HS500			FHS500			MHS500		
Stock index	No. of breaks	Critical value of LR_{uc}	p-value	No. of breaks	Critical value of LR_{uc}	p-value	No. of breaks	Critical value of LR_{uc}	p-value
BELEXline	2	0.125	0.724	2	0.125	0.724	0	-	-
SOFIX	3	0.076	0.783	2	0.125	0.724	1	1.213	0.271
BET	8	7.297	0.007	3	0.076	0.783	2	0.125	0.724
MONEX	3	0.076	0.783	2	0.125	0.724	2	0.125	0.724
CROBEX	0	-	-	1	0.062	0.803	0	-	-
ATHEX	1	0.062	0.803	2	0.359	0.549	1	0.062	0.803
SBITOP	3	1.709	0.191	1	0.062	0.803	3	1.709	0.191
SAXS10	2	0.359	0.549	2	0.359	0.549	2	0.359	0.549
XU100	4	3.748	0.053	4	3.748	0.053	5	6.300	0.012
SAX	3	1.709	0.191	5	6.300	0.012	1	0.062	0.803
PX	5	6.300	0.012	-	-	-	5	6.300	0.012
BUX	3	1.709	0.191	3	1.709	0.191	3	1.709	0.191
OMXT	2	0.359	0.549	3	1.709	0.191	2	0.062	0.803
OMXR	5	6.300	0.012	5	6.300	0.012	2	0.062	0.803
OMXV	2	0.359	0.549	1	0.062	0.803	2	0.062	0.803
	DHS500			ARMA(p,q)-GARCH(p,q)-MHS500			ARMA(p,q)-GARCH(p,q) with t-distribution-MHS500		
Stock index	No. of breaks	Critical value of LR_{uc}	p-value	No. of breaks	Critical value of LR_{uc}	p-value			
BELEXline	3	0.076	0.783	3	0.076	0.783	1	1.213	0.271
SOFIX	4	0.698	0.403	3	0.076	0.783	2	0.125	0.724
BET	2	0.125	0.724	1	1.213	0.271	2	0.125	0.724
MONEX	2	0.125	0.724	2	0.125	0.724	1	1.213	0.271
CROBEX	1	0.062	0.803	1	1.213	0.271	1	1.213	0.271
ATHEX	1	0.062	0.803	2	0.125	0.724	2	0.125	0.724
SBITOP	2	0.125	0.724				2	0.125	0.724
SAXS10	2	0.125	0.724				2	0.125	0.724
XU100	4	0.698	0.403	4	3.748	0.053	4	3.748	0.053
SAX	5	6.300	0.012	5	6.300	0.012	4	3.748	0.053

PX	-	-	-	-	-	-	-	-	-
BUX	3	0.076	0.783	3	0.076	0.783	3	0.076	0.783
OMXT	2	0.125	0.724	3	0.076	0.783	3	0.076	0.783
OMXR	4	3.748	0.053	4	3.748	0.053	4	3.748	0.053
OMXV	1	0.062	0.803	1	0.062	0.803	1	0.062	0.803
The LR_{cc} test for 99%VaR									
	HS500			FHS500			MHS500		
Stock index	No. of breaks	Critical value of LR_{cc}	p-value	No. of breaks	Critical value of LR_{cc}	p-value	No. of breaks	Critical value of LR_{cc} test	p-value
BELEXline	2	0.353	0.838	2	0.353	0.838	0	-	-
SOFIX	3	1.696	0.428	2	0.353	0.838	1	0.066	0.968
BET	8	16.114	0.000	3	1.696	0.428	2	0.353	0.838
MONEX	3	1.696	0.428	2	0.353	0.838	2	0.353	0.838
CROBEX	0	-	-	1	0.066	0.968	0	-	-
ATHEX	1	0.064	0.969	2	0.359	0.836	1	0.064	0.969
SBITOP	3	1.709	0.425	1	0.064	0.969	3	1.709	0.425
SAXS10	2	0.359	0.836	2	0.359	0.836	2	0.359	0.836
XU100	4	3.748	0.154	4	3.748	0.154	5	6.300	0.043
SAX	3	1.709	0.425	5	6.300	0.043	1	0.064	0.969
PX	5	6.300	0.043	0	-	-	5	6.300	0.043
BUX	3	1.709	0.425	3	1.709	0.425	3	1.709	0.425
OMXT	2	0.359	0.836	3	7.174	0.028	2	0.359	0.836
OMXR	5	6.329	0.042	5	6.329	0.042	2	0.364	0.833
OMXV	2	0.378	0.828	1	0.057	0.972	2	0.378	0.828
	DHS500			ARMA(p,q)-GARCH(p,q)-MHS500			ARMA(p,q)-GARCH(p,q) with t-distribution-MHS500		
Stock index	No. of breaks	Critical value of LR_{cc}	p-value	No. of breaks	Critical value of LR_{cc}	p-value			
BELEXline	3	1.696	0.428	3	1.696	0.428	1	0.066	0.968
SOFIX	4	3.726	0.155	3	1.696	0.428	2	0.353	0.838
BET	2	0.353	0.838	1	0.066	0.968	2	0.353	0.838
MONEX	2	0.353	0.838	2	0.353	0.838	1	0.066	0.968
CROBEX	1	0.066	0.968	1	0.066	0.968	1	0.066	0.968
ATHEX	1	0.064	0.969	2	0.359	0.836	2	0.359	0.836
SBITOP	2	0.359	0.836	2	0.359	0.836	2	0.359	0.836
SAXS10	2	0.359	0.836	2	0.359	0.836	2	0.359	0.836
XU100	4	3.748	0.154	4	3.748	0.154	4	3.748	0.154
SAX	5	6.300	0.043	5	6.300	0.043	4	3.748	0.154
PX	0	-	-	0	0	-	-	-	-
BUX	3	1.709	0.425	3	1.709	0.425	3	1.709	0.425
OMXT	2	0.359	0.836	3	1.709	0.425	3	1.709	0.425
OMXR	4	3.770	0.152	4	3.770	0.152	4	3.770	0.152
OMXV	1	0.057	0.972	1	0.057	0.972	1	0.057	0.972

Source: Authors' calculations.

As one may see in Table 5, the worst performers according to these tests are the HS500 and FHS500 models. These models did not satisfy the tests in three capital markets.

Surprisingly, the MHS500 model achieved better performances as compared to the FHS500 model, but this can be explained by the fact that, due to the data doubling, some extreme value that had remained in the sample due to the rolling window exerted an influence on the VaR value. The good results were also achieved in the DHS500 model and the ARMA(p,q)-GARCH(p,q) model with normal distribution-MHS500, whereas the ARMA(p,q)-GARCH(p,q) model with T-distribution-MHS500 showed the best results.

Generally, it is known that the main drawback of both tests is their questionable statistical power when applied to finite samples (see Wied *et al.*, 2013, etc.). For that reason, the validity of the obtained backtesting results for the conditional coverage test was verified by following the next Monte Carlo procedure: first, 9,999 samples of random IID Bernoulli (*p*) variables were generated, where the sample size equaled the actual sample; after that, based on these artificial samples, 9,999 simulated LR_{cc} tests were calculated and named $\{L\tilde{R}_{cc}(i)\}_{i=1}^{9.999}$. The last step implied the calculation of simulated p-values as a share of the simulated LR_{cc} values larger than those actually obtained by the LR_{uc} test:

$$p - value = \frac{1}{10.000}\left\{1 + \sum_{i=1}^{9.999} I\left(L\tilde{R}_{cc}(i) > LR_{cc}\right)\right\} \quad (7)$$

where: I(.) assumes the value one, if the argument is true, and the value zero otherwise. The results of these simulations are presented in Table 6.

Table 6

The Backtesting Results Based on the Monte Carlo Procedure for the LR_{cc}

	HS500	FHS500	MHS500	DHS	ARMA(p,q)-GARCH(p,q)-MHS500	ARMA(p,q)-GARCH(p,q) with t-distribution-MHS500
BELEXline	0.009	0.003	0.001	0.135	0.086	0.531
SOFIX	0.096	0.514	0.089	0.081	0.167	0.346
BET	0.032	0.030	0.000	0.136	0.007	0.214
MONEX	0.000	0.051	0.165	0.035	0.140	0.119
CROBEX	0.215	0.345	0.689	0.647	0.457	0.222
ATHEX	0.366	0.258	0.437	0.373	0.286	0.793
SBITOP	0.147	0.46	0.037	0.211	0.103	0.143
SAXS10	0.501	0.338	0.237	0.121	0.318	0.191
XU100	0.024	0.047	0.491	0.03	0.011	0.025
SAX	0.378	0.028	0.204	0.029	0.044	0.331
PX	0.044	0.467	0.311	0.197	0.357	0.468
BUX	0.357	0.631	0.019	0.246	0.247	0.139
OMXT	0.112	0.018	0.347	0.179	0.166	0.177
OMXR	0.037	0.441	0.147	0.011	0.036	0.049
OMXV	0.411	0.276	0.554	0.505	0.202	0.183

Source: Authors' calculations.
Note: The average feasible rates of the tests range from 0.711 to 0.648.

The results of the simulation confirmed the findings of the previous analysis, noting that the tested models did not meet the conditional coverage criterion for a larger number of the markets. The best performances were again those demonstrated by the ARMA(p,q)-GARCH(p,q) wit T-distribution-MHS500.

Having in mind the issue raised by Gneiting (2012), Berkowitz CVaR backtesting was used in this paper. Berkowitz (2001) proposed a test based on the Levy-Rosenblatt transformation. The test implies a double transformation of the observed losses: the first transformation involves the replacement of the loss in *t* with the predicted probability of observing this or a smaller loss. This probability is obtained by inserting the loss into the *CFDF*(L_t):$p_t = F(L_t)$. This transformation produces numbers between 0 and 1. If the predicted distribution is correct, the numbers should be uniformly distributed between 0 and 1; the second transformation involves the transformation of p_t by applying the inverse cumulative standard normal distribution function: $z_t = \Phi^{-1}(L_t)$. If the model is valid, the series (z_t) should follow *IID N(0,1)*. Berkowitz (2001) suggested that the test should be restricted to the hypothesis asserting that (z_t) has a zero mean and a unit variance. He suggested that the joint hypothesis should be tested by using the following likelihood ratio test:

$$LR_B = 2[ln\, L\,(\mu = \hat{\mu}_{ML}, \sigma^2 = \hat{\sigma}_{ML}) - ln\, L\,(\mu = 0, \sigma^2 = 1)] \quad (8)$$

The LR_B test is asymptotically distributed as χ2 with two degrees of freedom. The Berkowitz test applied in the paper compares the shape of the forecasted density tail with the observed tail. Any observations that did not fall into the tail were truncated, noting that the threshold was defined as follows:

$$TH_{i,t} = max\{CVaR_1, CVaR_2 .... CVaR_t\} \quad (9)$$

The results of the Berkowitz test are presented in Table 7.

Table 7

The Backtesting Results for the LB_B for 97.5%CVaR

	HS500	FHS500	MHS500	DHS	ARMA(p,q)-GARCH(p,q)-MHS500	ARMA(p,q)-GARCH(p,q) with t-distribution-MHS500
BELEXline	0.000	0.069	0.000	0.068	0.493	0.468
SOFIX	0.000	0.252	0.000	0.939	0.026	0.248
BET	N/A	0.000	0.229	0.127	0.000	0.333
MONEX	0.000	0.102	0.001	0.363	0.096	0.108
CROBEX	N/A	N/A	N/A	0.579	0.093	0.099
ATHEX	0.118	0.087	0.103	0.068	0.203	0.036
SBITOP	0.057	0.061	0.545	0.192	0.12	0.102
SAXS10	0.066	0.011	0.104	0.056	0.055	0.054
XU100	0.025	0.026	0.059	0.039	0.011	0.311
SAX	0.047	0.066	N/A	0.101	0.103	0.021
PX	N/A	N/A	0.013	0.024	N/A	N/A
BUX	0.173	0.029	0.009	0.000	0.062	0.048
OMXT	0.034	0.000	0.041	0.049	0.021	0.096
OMXR	0.063	0.173	0.069	0.088	0.471	0.062
OMXV	0.027	0.77	0.221	0.041	0.029	0.033

Source: Authors' calculations.

Note: The test was done at a 5% significance level.

As can be observed in Table 7, the worst performer is the HS500 model, only to be followed by the MHS500 and FHS, ARMA(p, q)-GARCH(p,q)-MHS and DHS500 models. The best performer is the ARMA(p,q)-GARCH(p,q) with T-distribution-MHS500. The reason for this can be found in the fact that the model is based on Student's t-distribution, and as such it captures fat tails well.

As shown in Table 7, it was impossible to conduct the Berkowitz test in a few cases, first of all, due to a lack of a sufficient number of exceedances. The test suffers from two drawbacks. It requires parametric assumptions and large samples. The need for a parametric assumption does not have to be an issue if VaR is calculated by using parametric distribution. However, it is important to be able to distinguish a bad model from a bad parametric assumption. The major drawback of the test is the need for large samples, an unrealistic assumption in the backtesting of the ES since the number of losses at hand is always just a few.

Righi and Ceretta (2013) agreed that the LR_B test might be inaccurate in samples, such as those used in practice and for regulation purposes, as these tend to be small. Bearing this in mind, the backtesting results of the LR_B test are subject to verification. For this purpose, the bootstrap method was used. By using the parametric bootstrap method, 10,000 samples were generated, the size of which was equal to the actual sample for which the risk estimation was made, for each model and each market separately. Then, the LR_B value was calculated for each simulated sample and the average values of LR_B obtained by the arithmetic mean are shown in Table 8.

Table 8

The LB_B Backtesting Results Based on the Parametric Bootstrap Method

	HS500	**FHS500**	**MHS500**	**DHS**	**ARMA(p,q)-GARCH(p,q)-MHS500**	**ARMA(p,q)-GARCH(p,q) with t-distribution-MHS500**
BELEXline	0.000	0.006	0.000	0.001	0.493	0.111
SOFIX	0.000	0.096	0.000	0.146	0.096	0.023
BET	0.000	0.000	0.000	0.051	0.000	0.084
MONEX	0.000	0.002	0.000	0.238	0.135	0.063
CROBEX	0.067	0.179	0.059	0.093	0.312	0.209
ATHEX	0.083	0.019	0.066	0.179	0.088	0.077
SBITOP	0.026	0.066	0.032	0.006	0.039	0.181
SAXS10	0.049	0.18	0.111	0.014	0.021	0.203
XU100	0.000	0.095	0.099	0.051	0.075	0.046
SAX	0.028	0.000	0.058	0.2	0.033	0.222
PX	0.012	0.108	0.000	0.156	0.001	0.001
BUX	0.091	0.039	0.021	0.005	0.066	0.000
OMXT	0.001	0.065	0.055	0.008	0.001	0.058
OMXR	0.036	0.057	0.79	0.081	0.055	0.09
OMXV	0.009	0.001	0.006	0.117	0.000	0.001

Source: Authors' calculations.
Note: The Average feasible rates of the tests range from 0.631 to 0.708.

The LB_B backtesting results based on the parametric bootstrap method are confirmed by Righi and Ceretta's claims. The results show that very popular nonparametric and semiparametric models of the HS cannot be reliable, not even in a situation of absence of an ARCH effect on markets. According to these results, the best performers are again the ARMA(p,q)-GARCH(p,q) with t-distribution-MHS500 and DHS models. This does not surprise when we have in mind the fact that these models are capable of coopting ARCH effects and leptokurtosis.

6. Conclusion

As is shown above, a new semiparametric market risk estimation VaR model is presented and tested. The model is so designed to successfully reduce the drawbacks of the HS method in terms of capturing the empirical features of returns, such as leptokurtosis, asymmetry, autocorrelation, and heteroscedasticity, i.e. solve the problem of reliable VaR estimates for an extremely high confidence level for the sample size defined by the Basel III standard. Namely, the idea is to take advantage of the MHS model in reducing the tail estimation problem by the HS and the advantages of the use of the ARMA and GARCH models to transform historical returns into IID returns. This model leads to a reduction in both drawbacks of the HS model and its improvement, without increasing implementation costs. The proposed model is quite easy to understand and implement.

In order to obtain an answer to the question whether the proposed model performs better than the most popular and the most widely used improvements of the HS or not, the backtesting results of the model were compared with the backtesting results of the most popular and the most widely used improvements of the HS model. Since VaR does not fulfill all the characteristics of coherent risk measures, the Basel Committee has recently proposed fundamental changes in the regulatory treatment of financial institutions′ trading book positions, among which the replacement of 99% VaR with 97.5% CVaR for the quantification of market risk is recommended. For this reason, the performances of the model were assessed in relation to its ability to generate reliable CVaR estimates. For this purpose, the Berkowitz test was used for backtesting. Since many authors agree that the first step in testing the validity of risk models implies VaR backtesting, the VaR backtesting procedure by doing Kupiec and Christofferson tests was carried out first, after which the Berkowitz test for CVaR was performed. In order to verify the backtesting results, the Dufour Monte Carlo testing technique and the parametric bootstrap method were applied. The obtained backtesting results justify the expectations of the new model.

Bearing in mind the research results and the complexity of the proposed model, on the one hand, and the generally known fact that no use of parametric models is recommended for the markets susceptible to frequent structural changes and in which the matrix of correlation between securities is instable, on the other, clearly indicate the justification for using the proposed model. The fact that the model works better in relation to the nonparametric models of the HS clearly shows that it has taken advantage of their strengths when speaking about the application of these models to such markets. The same conclusion may also be drawn regarding the comparison of performances of the new model in relation to the semiparametric models constructed on the basis of the HS method.

References

Acerbi, C. and Tasche, D., 2002. Expected Shortfall: A natural coherent alternative to Value at Risk. *Economic Notes*, 31(2), pp.379–388.

Alemany, R., Bolancé, M. and Guillén, M., 2013. A nonparametric approach to calculating Value-at-Risk. *Insurance: Mathematics and Economics*, 52(2), pp.255-262.

Barone-Adesi, G. and Giannopoulos, K., 2001. Non-parametric VaR techniques. Myths and realities, *Economic Notes,* 30(2), pp.167–181.

Barry, C.B., Rodriguez, M., 2004. Risk and Return Characteristics of Property Indices in Emerging Markets. *Emerging Markets Review*, 5(2), pp.131-159.

Bee, M., 2012. Dynamic value-at-risk models and the peaks-over-threshold method for market risk measurement: an empirical investigation during a financial crisis. *Journal of Risk Model Validation*, 2(2), pp.3–45.

Bekaert, G., Harvey, C.R., 2003. Emerging market finance. *Journal of Empirical Finance,* 10(1-2) pp.3-55.

Bekaert, G., Harvey, C.R., Lumsdaine, R.L., 2002. The dynamics of emerging market equity flows. *Journal of International Money and Finance*, 21 (3), pp.295-350.

Berkowitz, J., 2001. Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics*, 19(4), pp.465–474.

Boudoukh, J., Richardson, M., and Whitelaw, F.R., 1998. The Best of Both Worlds: A hybrid Approach to Calculating Value at Risk. *Risk,* 11(5), pp.64-67.

Brandolini, D. and Colucci, S., 2012. Backtesting value-at-risk: a comparison between filtered bootstrap and historical simulation. *Journal of Risk Model Validation,* 6(4), pp.3–16.

Dunis, C.L., Shanon, G., 2005. Emerging markets of South-East and Central Asia: Do they still offer a diversification benefit? *Journal of Asset Management*, (6)3, pp.168-90.

Emmer, S., Kratz, M., and Tasche, D., 2013. What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk* ,18(2), pp.31–60.

Gneiting, T., 2012. Making and evaluating point forecasts. *Journal of the American Statistical Association,* 106 (494), pp.746-762.

Holton, A.G., 1998. Simulating value at risk. *Risk* 11(5), pp.60–63.

Hull, J. and White, A., 1998. Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1(1), pp.1-19.

Kellner, R. and Rösch, D., 2016. Quantifying market risk with value-at-risk or expected shortfall? – Consequences for capital requirements and model risk. *Journal of Economic Dynamics and Control*, 68, pp.45–63.

Kostadinovic, I. and Radojicic, J., 2017. Banking Sector in the Context of Sustainable Development. *Economics of Sustainable Development,* 1(2), pp. 109-120.

Kresta, A. and Tichy, T., 2016. Selection of efficient market risk models: backtesting results evaluation with DEA approach. *Computers & Industrial Engineering*, 102, pp.331–339.

Patel, J.B. 2003. Inter-Temporal Relationship between the U.S. Stock Market and the Emerging Markets of Asia. *Journal of Business and Economics Research*, 10(1), pp.1-6.

Nuti, M. D. 2009. The impact of the global crisis on transition economies. *Economic Annals*, (54)181, pp.7-20.

Radivojevic, N., Cvjetkovic, M. and Stepanov, S., 2016a. The new hybrid VaR approach based on EVT. *Estudios de Economia,* 43(1), pp.29-52.

Radivojevic, N., Curcic. V.N., Milojkovic, D. and Miletic, S.V., 2016b. Testing value at risk models in emerging markets during crises: a case study on South Eastern European countries. *Journal of Risk Model Validation,* 10(2), pp. 57-81.

Radivojevic, N., Sabot-Matic, Z. and Mirjanic, B., 2017a. New historical bootstrap value at risk model. *Journal of Risk Model Validation*, 11(4), pp. 57-75.

Radivojevic, N., Curcic, V. N., Vukajlovic, Dj. 2017b. Hull-White's value at risk model: case study of Baltic equities market. *Journal of business economics and management*, 18(5), pp. 1023–1041.

Righi, M., Ceretta, P., 2013. Individual and flexible expected shortfall backtesting. *Journal of Risk Model Validation*, 7(3), pp.3–20.

Rossignolo, F.A., Fethib, M.D. and Shaban, M., 2012. Value-at-Risk models and Basel capital charges Evidence from Emerging and Frontier stock markets. *Journal of Financial Stability,* 8, pp.303-319.

Rossignolo, F.A., Fethib, M.D. and Shaban, M., 2013. Market crises and Basel capital requirements: Could Basel III have been different? Evidence from Portugal, Ireland, Greece and Spain (PIGS). *Journal of Banking and Finance,* 37, pp.1323-1339.

Şener, E., Baronyana, S. and Mengütürk, L., 2012. Ranking the predictive performances of value-at-risk estimation methods. *International Journal of Forecasting,* 28, pp.849–873.

Terzic, I. and Milojevic, M., Risk model backtesting 2016. *Ekonomika,* 62(1), pp.151-162

Tokat, Y., Wicas, N. 2004. Investing in Emerging Stock Markets. *The Journal of Wealth Management*, 7(2), pp.68-80

Zikovic, S. and Filer, R.K., 2013. Ranking of VaR and ES models: performance in developed and emerging markets, *Czech Journal of Economics and Finance,* 63(3), pp.327–359.

Zikovic, S. and Prohaska, Z., 2010. Optimization of decay factor in time weighted (BRW) simulation, implications for VaR performance in Mediterranean countries, *Ekonomska istraživanja,* 23(1), pp.73–85.

Zikovic, S., Aktan, B., 2009. Global financial crisis and VaR performance in emerging markets: A case of EU candidate states – Turkey and Croatia", *Journal of Economics and Business*, 27(1), pp.149-170.

Wied, D., Weiß, G. and Ziggel, D., 2016. Evaluating Value-at-Risk forecasts: A new set of multivariate backtests. Journal of Banking & Finance, 72, pp.121-132.