Inzinerine Ekonomika-Engineering Economics, 2019, 30(1), 32–40

One Approach for Backtesting VaR Specifications in the Russian Stock Market

Tamara V. Teplova, Dmitry P. Ruzanov

National Research University Higher School of Economics
Shabolovka str. 26, bldg 4, room 4320, Moscow, Russia
E-mail. tteplova@hse.ru, dmitriy.ruzanov@gmail.com

crossref http://dx.doi.org/10.5755/j01.ee.30.1.17795

Traditionally standard deviation has been considered as the main risk measure of an asset portfolio. The relevance of VaR analysis is widely recognized as an instrument for market risk quantification for investment decisions, asset allocation. Based on common practice VaR is estimated on 10-day basis and using 99 % confidence interval. More accurate VaR estimation requires identifying the optimal VaR parameters. Our paper documents that historical VaR has some limitation for high volatile stock markets. We conduct empirical analysis of statistical tests of VaR estimation with frequency tests, magnitude tests, independence and autocorrelation tests for the Russian stock market. We propose an original algorithm for optimal VaR specification in terms of accuracy of VaR estimates. We used historical and semi parametric VaR (EWMA VaR and volatility adjusted VaR). For each method we consider 16 VaR specifications (which are different combinations of time horizons – 120, 250, 500 and 1000 trading days, and confidence intervals – 90 %, 95 %, 99 %, 99,5 %). We consider the unstable Russian stock market with two main Russian indexes – MICEX and RTS. Backtesting different VaR specifications show that annual 99 % VaR prevails over other VaR specifications for the Russian stock indices. The significance level of confidence 1–5 % are optimal on various time horizons. VaR with our method of algorithmically defined parameters is more effective than commonly used estimation procedure.

Keywords: *Value-at-Risk (VaR); Backtesting; VaR Specification; Semi-Parametric VaR; EWMA Var; Volatility Adjusted VaR; Russian Stock Market, Risk Modeling.*

Introduction

The assessment of potential losses caused by investment decisions and setting deviations limits from expected results is one of the core risk management functions. Traditionally standard deviation of a random variable has been considered as the main risk measure of an asset portfolio. Starting from mid-1990s large banks have actively used a *VaR* (Value-at-Risk) indicator offered by experts of the investment bank J. P. Morgan (1994, 1996) in the technical document "RiskMetrics". RiskMetrics provided a detailed guide to the estimation and application of *VaR* measures in banking. *VaR* is defined as a measure of the maximum potential change in value of a portfolio of financial instruments with a given probability over a pre-set horizon. *VaR* can be applied to any asset class (stocks, bonds, derivatives - Marshall and Siegel (1997)). Soccorsi (2016) concludes that despite the criticism of *VaR* for highly volatile markets, risk assessments are accurate enough. To define *VaR* we need two components: a fixed time horizon and a given confidence level.

Currently *VaR* is widely used not only in the area of financial risk management but also in other scientific fields (reviews are given in Nieto & Ruiz, 2016, Wang & Huang, 2016). Christoffersen et al (2001) report that more than 80 commercial vendors offer risk management systems with *VaR*-like measures. The use of *VaR* for market risk quantification is also mentioned in Basel II regulatory requirements.

The relevance of *VaR* analysis is also supported by the fact that market agents use this measure not only as an instrument for market risk quantification (for example, for the assessment of premiums for the insurance sector by Wang & Huang, 2016) but also for investment decisions, as an optimization criterion for asset allocation (Diamandis *et al.*, 2011; Zhu *et al.*, 2013; Lwin *et al.*, 2014). *VaR* is applied for the analysis of stock markets and indices (Nieto & Ruiz (2016) for the S&P500), and bond markets (Gunay (2016) apply VaR for the analysis of 10-year government bonds of South Korea, Japan, Malaysia and Singapore). The motivation of our study is the fact that *VaR* limits may be set for each trader or trading desk, for insurance companies, which traditionally hold some of their assets in shares. Hence, the accuracy and adequacy of *VaR* is extremely important in financial risk management.

Verifying the accuracy of *VaR* has been discussed in the academic literature since the 2000s (Christoffersen & Diebold, 2000; Berkowitz & O'Brien, 2002; Christoffersen *et al.*, 2001; Berkowitz *et al.*, 2008). After the global financial crisis discussions continue (Wong, 2010; Bernard *et al.*, 2017).

The Basel Committee (2011, 2012) identified the Expected Shortfall (ES) as a possible alternative to VaR, but this decision had opponents. Gneiting (2011) argues that *VaR*, along with any other risk criterion based on quantiles, is an acceptable indicator of risk, but requires a correct specification. Danielsson et al. (2013) point out that *VaR* has limitations as a risk measure and the statistical model can generate inconsistent and biased forecasts of risk, so the use of *VaR* in controlling risks at the level of individual traders or trading desks leads to large biases. Danielsson & Zhou (2015) considered the differences between VAR and ES in relation to sample size. Results for ES (97.5 %) and VAR (99 %), indicated in Basel 3, do not differ with normal distribution. In cases with heavy tails, ES worse than VAR on small samples.

Kellner & Rosch (2016) realized an empirical comparative analysis of the model risk of VAR and ES. The authors compared how VAR and ES react to various sources of model risk in order to better understand the impact of the transition proposed by the Basel Committee (2011, 2012) with VAR (99 %) and ES (97.5 %). It is shown that in times of stressful economic environment ES estimates (a = 0.975) are less reliable than VAR estimates (a = 0.99). ES is more sensitive to regulatory arbitrage and parameter specification errors, ES volatility is also higher.

Research on new approaches and tests indicates the importance of developing the most appropriate *VaR* in the emerging markets.

These criticisms and different points of view on *VaR* adequacy highlight the significance of our research. Our objective is: to propose an original approach for optimal *VaR* specification for a national emerging financial market. In order to test our algorithm and to identify the optimal *VaR* specification (in terms of *VaR* accuracy) we consider the Russian stock market which is a subject of potential interest for investors. Market volatility is high and there is significant uncertainty of the future stock price dynamics. In this paper we show how an algorithm of optimal *VaR* specification for the two Russian stock indices can be applied.

Based on common practice (and according to Basel II recommendations) *VaR* is estimated on 10-day basis and using 99 % confidence interval. Resent research documents that these expertly defined *VaR* parameters may be not optimal in terms of *VaR* accuracy. There are three main approaches for estimating *VaR*: historical simulation, parametric methods and Monte Carlo simulation. However, almost all risk managers use *VaR* approaches based on past data. More accurate *VaR* estimation requires identifying the optimal *VaR* parameters. Common practice benchmarks for *VaR* parameters are rather questionable. This motivates the search for an algorithm of optimal *VaR* specifications. The paper shows the specification for the Russian stock market.

The motivation of our work is related to the criticism of traditional VaR for risk estimation in volatile markets (Du & Escanciano, 2016; Djakovic & Andjelic, 2017). Testing on developed capital markets proves the acceptability of VaR (Nieto &Ruiz (2016) for S&P500). The adequacy of the *VaR* estimates for emerging markets is less studied (Del Brio *et al.*, 2014; Djakovic & Andjelic, 2017). It is important to back-test decision-making in national markets, for example, as large as the Russian. This work is implemented in our research.

Our paper proposes an empirical analysis of various statistical tests of *VaR* backtesting. The comparative analysis covers various methods such as frequency tests, magnitude tests, independence and autocorrelation tests. We adopt approaches proposed by Kupiec, 1995, Christoffersen & Diebold, 2000; Christoffersen *et al.*, 2001; Engle, 2004, Christoffersen & Pelletier, 2008; Perignon & Smith, 2008; Wong, 2010; Hurlin & Perignon, 2012. We compare historical and semi parametric *VaR* (EWMA *VaR* and volatility adjusted *VaR*). For each method we consider 16 *VaR* specifications (which are different combinations of time horizon – 120, 250, 500, 1000 trading days and confidence intervals – 90 %, 95 %, 99 %, 99.5 %). Based on various models we estimate *VaR* measures for the two main Russian indices – MICEX and RTS and for blue chips.

Backtesting different *VaR* specifications show that annual 99 % *VaR* prevails over other *VaR* specifications for the Russian stock indices. We also propose common algorithm for optimal *VaR* parameter identification for a certain financial market, investment portfolio or asset.

Our paper is organized as follows: in Section 2 we provide a short review of *VaR*. Section 3 analyzes data used for *VaR* estimation and provides comparative analysis of backtesting of various *VaR* measures. In Section 4 we provide an algorithm for the optimal (in terms of the backtests) *VaR* specification for the Russian market, our results and their practical application. In Section 5 we summarize our results.

VaR Estimation and our Hypotheses

Mathematically *VaR* is a measure of the maximum potential loss in value of a financial portfolio with a given probability over a pre-set horizon. *VaR* answers the following question: what is the maximum loss with a certain probability over a specified time horizon?

Using this equation we can define *VaR* explicitly:

$$VaR_{\alpha}(R) = -\inf\{r : P(R_{t+1} < r) \geq \alpha\} = -F_R^{-1}(\alpha), \quad (1)$$

R_t is a random variable which characterizes the result of investment (i.e. returns, logarithmic returns or value of portfolio). Investment outcome may be written as F_R^{-1} (inverse function).

For the correct estimation of *VaR* one should have information about the distribution of the results of investment (i.e. the distribution of the returns of the investment portfolio). A large number of financial models are based on the assumption of the normality of returns. However, the empirical results do not agree with this crucial assumption – real losses occur more frequently and appear to be higher than the values forecast by models with the normality assumption. In order to eliminate this shortcoming *VaR* measures are often based on the assumption that the variables have Student's t-distribution or follow autoregressive conditional heteroscedasticity (ARCH/GARCH) processes.

Following Christoffersen (1998) we set two conditions of *VaR* measure: 1) Unconditional coverage (UC-hypothesis): the probability of ex post return to exceed *VaR* estimate should be equal to *VaR* confidence level α (coverage rate). In case of a violation of this principle we can assert that VaR systematically underestimates (overestimates) risk. 2) Independence hypothesis (IND hypotheses). Rejection of independence hypothesis may lead to clustering effects when a lot of exceptions occur over a short period of time following periods of an absence of such exceptions. For testing the UC hypothesis taking into account the magnitude of exceptions we can use a likelihood ratio test with a slightly different statistic. This gives a resulting statistic for testing two conditions simultaneously (CC — Conditional Coverage Hypothesis).

We test the following hypotheses: *VaR* estimated as a result of the proposed algorithm (algorithmically defined parameters with the UC and the IND hypotheses testing) is more accurate than traditional *VaR* (expertly defined *VaR* parameters or historical *VaR*). The results of our testing show that *VaR* with algorithmically defined parameters is more effective than the commonly used estimation procedure.

The Russian Market Data

For assessment of historical *VaR* we used daily returns of two Russian stock indices – MICEX and RTS and blue chips from 30th November 2000 to 15th May 2015 (1,862 observations). The choice of these stocks and indices can be explained by the high level of liquidity and the availability of data. We also note that time horizon used covers a period of significant financial stress in 2009.

The return of each financial instrument was estimated as follows:

$$R_t = \ln(\frac{P_t}{P_{t-1}}), t = 1,...2869$$

where P_t is the index close value at day t. For two series of stock returns we have estimated descriptive statistics provided in Table 1. The investigated time series have negative skewness and positive excess kurtosis which implies the leptokurtic nature of returns distribution ("fat tails").

Table 1

Descriptive Statistics of Russian Stock Indexes Returns

Parameter	MICEX	RTSI
Number of observations	1862	1862
Minimum value	–0.207	–0.2120
Maximum value	0.252	0.2020
Range	0.459	0.4140
Skewness	–0.231	–0.5305
Kurtosis	16.809	11.5562
Shapiro-Wilk statistics	0.863	0.8925
p-value of Shapiro-Wilk statistics	0	0

Figure 1 presents QQ-plots for stock indexes MICEX and RTS in comparison with a normal distribution (straight diagonal line). The graphs demonstrate that empirical distributions of returns are far from normal. Significant deviation of QQ-plots from the diagonal line is a visual proof for fat tails.

We have also performed Dicky-Fuller stationarity test which revealed that the investigated time series are not stationary, hence the application of parametric *VaR* is problematic as parameters of the theoretical distribution are constantly changing.

Estimation of h day historical VaR for two Russian indices, testing VaR accuracy

We performed assessment of historical *VaR* for different confidence levels: α = 0,1, 0.05, 0.01, 0.005, 0,001, and time horizons *h* = 125, 250, 500, 750 and 1000 for two Russian stock indexes MICEX and RTS, and also for a set of Russian stock market blue chips. Taking into account the fact that on average in one year there are 250 trading days *VaR* assessment has been performed on semiannual, annual, 2-year, 3-year and 4-year data.

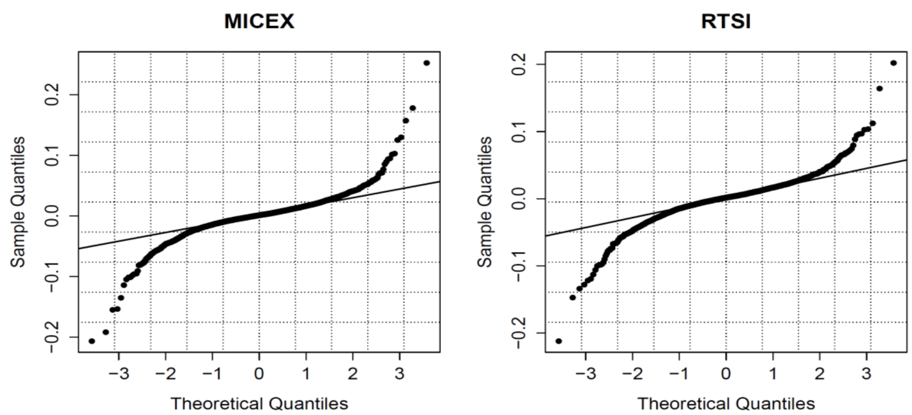

Figure 1. QQ-Plots for Two Russian Stock Indexes: MICEX and RTS (in Comparison With Normal Distribution)

As a result of testing the UC hypothesis with the help of likelihood ratio test (LR) we do not reject the null $H_{0,UC} : E\,(I_t(\alpha)) = \alpha$ in most cases[1]. For MICEX the UC hypothesis is rejected at significance level 0.5 % and 0.1 % in semiannual VaR and at 10 % and 5 % in 4-year *VaR*. For RTS we reject the UC-hypothesis only once – for 0.1 % semiannual *VaR*. In other *VaR* modifications the frequency of exceptions on average matched the coverage rate α. Thus, we can conclude that for large samples even historical *VaR* being the most straightforward method to apply, gives accurate estimates (including low confidence intervals 1 %, 0.5 %, 0.1 %), hence using simple LR test for large historical samples is not effective – in most cases we will observe a non-rejection of the UC hypothesis.

[1] *Testing has been conducted on a 95 % confidence interval.*

Proposed Algorithm for Optimal VaR Specification of Russian Stock market

The results of our test demonstrate a slightly different pattern: the best time horizon for *VaR* estimation is semiannual. Nevertheless the share of instruments which passed the CC test has maximum on this time horizon of 95 % *VaR*. *VaR* measures estimated at other confidence levels do not demonstrate an obvious dependency between *VaR* accuracy and time horizon (Figure 2). We can also make some conclusions regarding preferable time horizon (h).

Figure 2 characterizes a deviation of observed frequency of *VaR* exceptions $\frac{H}{T}$ from confidence level α. We note that the minimum magnitude of deviations for both indices can be seen for 250-day (annual) *VaR*. Thus, maximum accuracy is achieved for *VaR* models estimated over average time horizons whereas for long time horizons overfitting effects lead to lower model accuracy. It is also worth noting that samples of small size (i.e. semiannual *VaR* with h = 125) systemically overestimate *VaR* and total risk. Samples of higher size (h = 500, 750, 1000) underestimate *VaR* (especially for h = 750 and h = 1000). Therefore, the historical *VaR* estimated on 1,862 observations (2007–2015) is an adequate market risk measure given that time horizon h is chosen correctly.

The application of more complicated approaches to *VaR* estimation allows for the rejection of the UC hypothesis for some more pairs (k, h). We additionally tested the UC hypothesis using LR magnitude test. Technically, this means that we tested two UC hypotheses for each time horizon h – directly for confidence level α (which accounts for the exceptions of *VaR*) and also for confidence level 0,2α (which accounts for the super-exceptions of *VaR*). Value 0,2α was defined expertly according to Colletaz *et al.* (2012). The application of this method did not give us any benefits compared to the standard LR test (Table 2). According to Table 2 we can conclude that the UC hypothesis can be rejected for the following model specifications:

MICEX: h = 1000, α = 10 %;
MICEX: h = 1000, α = 5 %;
MICEX, RTS: h = 125, α = 1 %;
MICEX: h = 125, α = 0.5 %;
MICEX: h = 500, α = 0.5 %.

Testing the UC hypothesis has shown that optimal results are achieved over annual time horizons even for rather small significance level of confidence. However failure to reject the UC hypothesis does not guarantee accuracy of the *VaR* model. For complete verification we should be sure that the model doesn't have cluster effects which imply that probability of *VaR* violation is conditional on history of recent violations. In other words, we should test our model for independence of *VaR* exceptions. In addition to testing the independence hypothesis we have also tested joint conditional coverage (CC) hypothesis:

$$H_{0,CC} : \begin{cases} UC: E\,(I_t(\alpha)) = \alpha \\ IND: I_t, I_s - i.i.d \end{cases}.$$

The CC hypothesis has been tested with the use of statistics which is a simple sum of the two previous ones (Christoffersen et al., 2011):

$$LR_{CC} = LR_{UC} + LR_{IND} = -2\ln\left[(1-\alpha)^{T-H}(\alpha)^{H}\right] + \\ + 2\ln\left[(1-\hat{q}_{01})^{n_{00}}\hat{q}_{01}^{\,n_{01}}(1-\hat{q}_{11})^{n_{10}}\hat{q}_{11}^{\,n_{11}}\right] \xrightarrow{T\to\infty} \chi^2(2)$$

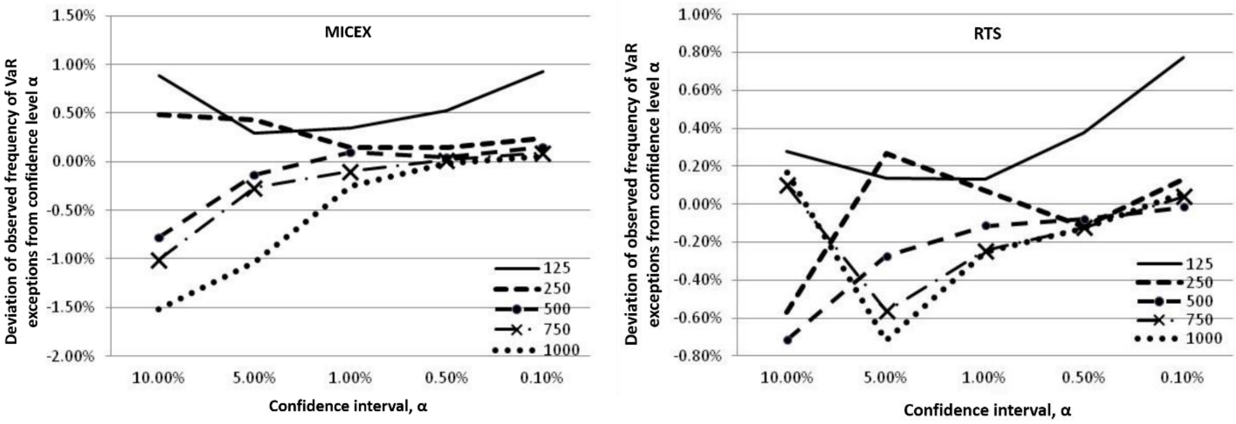

Figure 2. Deviations of Observed Frequency of VaR Exceptions from Significance Level *α* for Various Time Horizons (125, 250, 500, 750, 1000 Trading Days)

Table 2 presents the results of testing UC, CC and IND hypotheses («+» —not rejected, «–» —rejected) for various model parameters (time horizon h and significance level α, Equation 2). Table 2 shows that taking into account both the UC and independence criteria, the CC hypothesis is rejected quite frequently. This fact gives rise to the problem of the choice of *VaR* optimal parameters. We observe the following relationship between the sample size and confidence interval: the CC hypothesis is not rejected for time horizons of average length (h = 250) and average confidence intervals (1 %) or for large sample (h = 500, 1000) and low confidence *VaR* (0.5 %, 0.1 %).

Taking into account the fact that for financial instruments other than MICEX and RTS large samples may be not available, it is logical to use average time horizon and confidence intervals as a set of optimal parameters for *VaR*. We also note that the choice of optimal parameters for *VaR* estimation corresponds with the results of testing the UC hypothesis with the use of magnitude-based tests which is an additional argument in favor of our statement. Further, we have estimated *VaR* based on an exponential weighting of observations (EWMA VaR). Estimation has been conducted for three values of weighting constant λ, which accounts for weight distribution between more recent and older observations (the higher the value of lambda, the lower weight assigned to older observations): λ = 0.99, 0.995, 0.999.

The results of the estimation are shown in Figure 3 and in Table 3. Exponential weighting does not lead to more accurate estimates of *VaR* in comparison with the simple historical simulation. *VaR* accuracy rises with an increase of weighting parameter λ.

Values are significantly higher for EWMA *VaR* rather than for simple historically simulated *VaR*. The same applies to UC and IND statistics. A decrease of model accuracy can also be seen due to a decrease of value H/T which characterized observed frequency of *VaR* violations. For EWMA *VaR* this value deviates to the right from the significance level of confidence α = 1 %, which means systematic risk underestimation.

Hereby, use of EWMA *VaR* doesn't lead to an increase of model accuracy (for its optimal specification *h*=250, α=1 %).

Table 2

Comparative Characteristics of EWMA *VaR* and Simple Historical *VaR*

Statistics	λ = 0.99	λ = 0.995	λ = 0.999	λ = 0.9995	Unfiltered VaR (h = 250, α = 1%)
			MICEX		
H/T	0.03367777	0.0199	0.01378	0.01339457	0,01148106
UC statistic	91.46016216	20.08818	3.369016	2.748927495	0,552590601
IND statistic	6.434735	5.24157	2.69957	2.879208	0,674
CC statistic	97.89489716	25.32975	6.068588	5.628135495	1,226852401
$T_{critical}$			5.991464547		
			RTS		
H/T	0.03665521	0.02138	0.01145	0.01107293	0,0106911
UC statistic	112.0065612	25.97991	0.552431	0.307312694	–0,003102288
IND statistic	5.201076	3.17351	0.41125	0.4717792	0,5372567
CC statistic	117.2076372	29.15342	0.963685	0.779091894	0,534154412
$T_{critical}$			5.991464547		

Table 3

Comparative Characteristics of EWMA *VaR* and Simple Historical *VaR*

Statistics	λ = 0.99	λ = 0.995	λ = 0.999	λ = 0.9995	Unfiltered VaR (h = 250, α = 1%)
			MICEX		
H/T	0.03367777	0.0199	0.01378	0.01339457	0,01148106
UC statistic	91.46016216	20.08818	3.369016	2.748927495	0,552590601
IND statistic	6.434735	5.24157	2.69957	2.879208	0,674
CC statistic	97.89489716	25.32975	6.068588	5.628135495	1,226852401
$T_{critical}$			5.991464547		
			RTS		
H/T	0.03665521	0.02138	0.01145	0.01107293	0,0106911
UC statistic	112.0065612	25.97991	0.552431	0.307312694	–0,003102288
IND statistic	5.201076	3.17351	0.41125	0.4717792	0,5372567
CC statistic	117.2076372	29.15342	0.963685	0.779091894	0,534154412
$T_{critical}$			5.991464547		

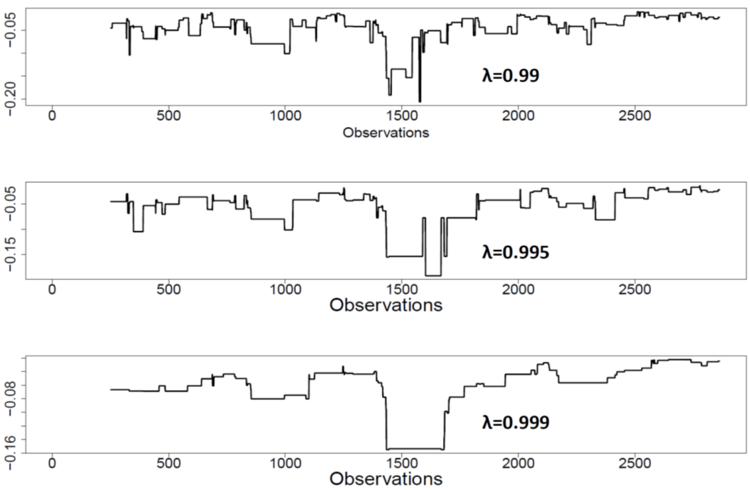

Figure 3. Exponentially Weighted Annual 1 % *VaR* Estimated for MICEX and Different Weighting Constant Values

Volatility adjustment is the third and the most complicated modification of historical *VaR*. This approach is based on the assumption that price returns follow a known stochastic process with autoregressive conditional heteroskedasticity. First, parameters of the GARCH model or its specification need to be estimated. Then price returns are adjusted to volatility in the following way:

$$r_{adj}^{t} = \frac{r_t}{\hat{\sigma}_t}\hat{\sigma}_T, t = 1, 2..., T.$$

The logic of such a procedure is quite straightforward: as time elapses characteristics of stochastic prices underlying price movement can change dramatically. That is why before estimating *VaR* one needs to adjust the time series so that adjusted price returns are stationary (Figure 4).

For the estimation of conditional volatility we implemented an exponential specification of the GARCH model proposed by Nelson (1990) – EGARCH(1,1). The main motivation of using such a specification was the ability of EGARCH model to take into account asymmetric effects.

Estimated parameters of the model are shown in Table 4 (the estimation has been conducted in R with the use of ugarch package). Then the results were compared with the previously tested models.

Estimations reveal that historical volatility adjusted *VaR* is more accurate in terms of the CC hypothesis than traditional historical *VaR*. More accurate frequency of *VaR* violations and lower levels (in comparison to other models) of UC and IND statistics give evidence to this statement (Table 5).

Thus, historical volatility adjusted *VaR* can be considered as the most effective way for market risk assessment comparing with other methods of *VaR*. Volatility adjustments make the return series stationary which results in *VaR* model accuracy. Nevertheless we note that in our research we have considered various methods of *VaR* estimation with specified parameters (250-day 1 % *VaR*). The choice of this specification was motivated by results of testing *VaR* accuracy at different significance level and time horizons.

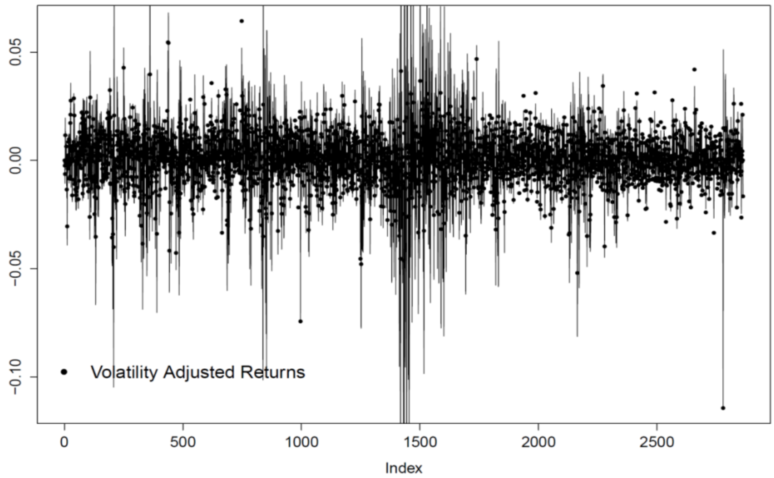

Figure 4. Diagram of Initial and Volatility Adjusted Returns of MICEX

Table 4

Estimates of the EGARCH Model for Russian Stock Indices MICEX and RTS (2007–2015)

Estimated Parameter	Parameter's value	Standard error	t-statistic	p-value
		MICEX		
μ	0.0007	0.0003	2.5264	0.0115
ω	–0.2778	0.0355	–7.8342	0.0000
α	–0.6420	0.0113	–5.6998	0.0000
β	0.9630	0.0044	219.3860	0.0000
γ	0.2440	0.0189	12.9346	0.0000
		RTS		
μ	0.0010	0.0000	20.3897	0.0000
ω	–0.2945	0.0070	–41.8103	0.0000
α	–0.0653	0.0092	–7.1142	0.0000
β	0.9621	0.0010	927.8616	0.0000
γ	0.1926	0.0161	11.9616	0.0000

Table 5

Results of Testing UC, CC and IND Hypotheses of Volatility Adjusted Historical VaR (Compared with Simple Unfiltered Historical VaR) for Russian Stock Indices MICEX and RTS

h	Tests	MICEX		RTS	
		α=1%			
		EGARCH	Unfiltered VaR	EGARCH	Unfiltered VaR
h = 125	UC statistic	1,35344	3,076	0,171269	0,447762
	IND statistic	0,394	0,400	1,352279	2,67778

h	Tests	MICEX		RTS	
		α=1%			
		EGARCH	Unfiltered VaR	EGARCH	Unfiltered VaR
	CC statistic	1,74744	3,476	1,523548	3,126
	CC hypothesis testing result	+	+	+	+
h = 250	UC statistic	0,240555	0,553	0,111211	0,123568
	IND statistic	0,264545	0,674	0,264601	0,53726
	CC statistic	0,5051	1,227	0,375812	0,661
	CC hypothesis testing result	+	+	+	+
h = 500	UC statistic	0,213778	0,233	0,230976	0,3208
	IND statistic	3,14141	4,586	11,01446	14,2122
	CC statistic	3,355188	4,818	11,24543	14,533
	CC hypothesis testing result	+	+	+	–
h = 750	UC statistic	0,08344	0,224	0,368216	1,402729
	IND statistic	3,570295	6,551	14,74563	16,7564
	CC statistic	3,653735	6,776	15,11385	15,615
	CC hypothesis testing result	+	–	–	–
h = 1000	UC statistic	0,96354	1,272	0,390541	1,301802
	IND statistic	7,55259	8,534	7,007968	17,4111
	CC statistic	8,51613	9,806	7,398508	18,713
	CC hypothesis testing result	–	–	–	–

Thus, we have conducted backtesting of various *VaR* risk metrics (historical *VaR*, EWMA *VaR*, volatility adjusted *VaR*) for different time horizons and confidence levels. By comparing UC and CC statistics we have identified optimal parameters of *VaR* specification for the Russian market. However such approach is not accurate as we compare only several *VaR* metrics based on certain *VaR* parameters (time horizon of 125, 250, 500, etc. days and confidence levels of 10 %, 5 %, 1 %, 0.1 %). For solving this problem we propose to test UC and CC hypothesis for a large number of combinations of confidence level and time horizons for a simulated stock portfolio. Such procedure allows to identify optimal *VaR* parameters for a random Russian stock market portfolio.

Our algorithm consists of the following steps:

1. To generate randomly coefficients of simulated portfolio which consists of liquid stocks included into the MICEX index. Estimate historical *VaR* based on the simulated portfolio for various significance level of confidence and time horizons. In the algorithm the following parameter values are considered:
   a. Time horizons: 50–1000 days (with step of 10 days);
   b. Confidence interval: 0.5 %–10 % (with step of 0.5 %).
2. To backtest the simulated portfolio with the use of UC and CC tests;
3. To repeat the iteration multiple times. In this work we have simulated portfolio 5000 times.

This allows us to compare various specification of *VaR* risk metrics and identify the optimal parameters. Table 6 shows the results of our final simulations.

The distribution of backtesting results (UC statistics given certain confidence interval and time horizon) is shown below (Figure 5).

The surface (Figure 5) demonstrates significant nonlinearity. Nevertheless we can interpret some of its patterns to identify the optimal parameters of *VaR*. The key observation is that UC and CC statistics have high values for average results of significance level of confidence such as 4 %–6 %. We also observe low confidence intervals (1 % and less). For such intervals UC and CC statistics have high values and hence the corresponding *VaR* models are not optimal. We also provide a heatmap plot which allows us to identify areas of optimal *VaR* parameters (Figure 6). The best results are observed for long time horizons starting from 750 days (3 years) and also for a 1- year period. Significance level of confidence 1–5 % are optimal over various time horizons.

Table 6

Illustration of Simulation of a Random Portfolio

No. of simulation	Alpha, %	Time horizon, k (days)	uc.LRstat	cc.LRstat	actual.exceed	expected.exceed
1	0.0525	505	1.474858	10.2857	61	70
2	0.0525	510	1.39989	10.23529	61	70
3	0.0525	515	1.048589	9.721678	62	70
4	0.0525	520	0.985039	9.671713	62	69
5	0.0525	525	0.92332	9.621633	62	69
6	0.0525	530	0.863449	9.571437	62	69
7	0.0525	535	0.805439	9.521125	62	69
8	0.0525	540	0.987822	9.93031	61	68
9	0.0525	545	1.190463	10.3528	60	68
10	0.0525	550	0.86518	9.827677	61	68

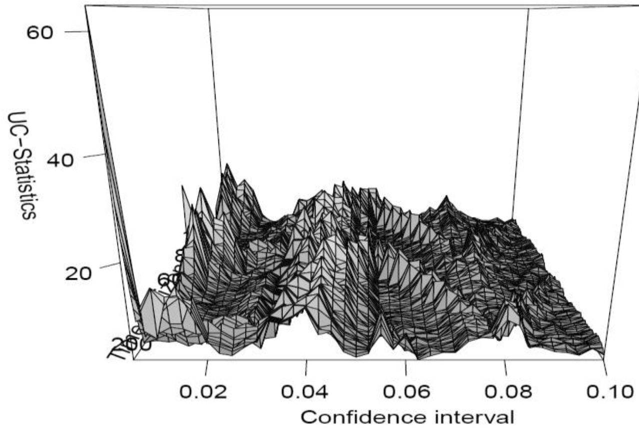

Figure 5. Backtesting Results for the Russian Stock Market

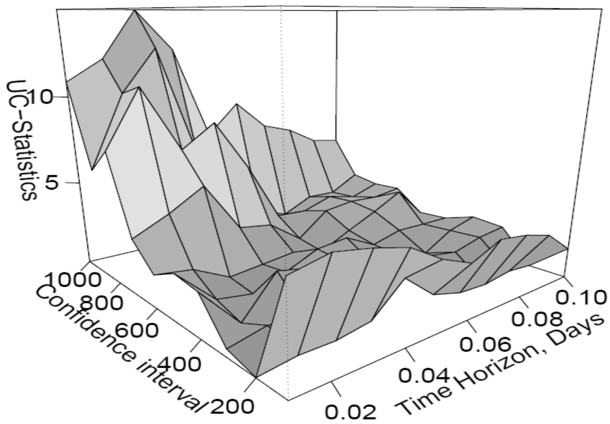

Figure 6. Testing Results for the Russian Stock Market (Rescaled Surface)

Conclusions

Nowadays there is no consensus on the preferred approach to *VaR* calculation. All the existing approaches are based upon certain assumptions and simplifications. Certain methods and models may be appropriate in certain situations and not applicable to others. As national markets are different from one another, the development of risk analysis mechanisms should take into account the specific behavior of market risk. Emerging stock markets often demonstrate excess kurtosis when compared to normal distribution (extreme returns are more likely) and a significant degree of autocorrelation and heteroscedasticity. So some popular and widely used *VaR* backtesting approaches are based on false simplifications. The transition to non-parametric methods does not always solve the problem. Nonparametric approaches may be effective in capturing kurtosis and fat tails, but they will not be successful for heteroscedasticity.

Our empirical study analyzes various statistical tests of *VaR* backtesting (UC and CC tests) on a sample of national stock market dynamics – the Russian market. Although a lot of procedures are available for forecasting and testing *VaR*, no consensus has been reached as to which procedures are best for certain national market. Our comparative analysis covers various methods of back-testing such as frequency tests, magnitude tests, independence and autocorrelation tests. We extend approaches proposed by Christoffersen & Diebold (2000), Christoffersen et al. (2001), Engle (2004), Christoffersen & Pelletier (2008), Hurlin & Perignon (2012). We used historical and semi parametric *VaR* (EWMA *VaR* and volatility adjusted *VaR*). For each method we have considered 16 *VaR* specifications (which are different combinations of time horizon – 120, 250, 500, 1000 trading days, and confidence intervals – 90 %, 95 %, 99 %, 99.5 %).

The proposed original algorithm allows us to select the optimal parameters (UC and CC tests) for *VaR* for Russian indices. As a practical result of our study we find optimal *VaR* parameters for two Russian stock indices – MICEX and RTS. This estimation based on the proposed algorithm which allows us to find optimal parameters of *VaR* estimation for a randomly simulated portfolio in terms of the independence of *VaR* and its accuracy.

For the assessment of historical *VaR* we used daily returns of two the indices in 1862 observations. Annual 99 % *VaR* prevails over other *VaR* specifications. The best results are observed for long time horizons starting from 750 days (3 years) and also for a 1-year period. The significance level of confidence 1–5 % are optimal on various time horizons.

Such *VaR* estimates can be widely used in practice for the assessment of the market risk of various financial instruments, assets portfolios and financial markets as a whole. These estimates are able to quantify market risk more accurately which can potentially reduce bank losses and increase net trading returns.

References

Basel Committee on Banking Supervision. (2011). Messages from the Academic Literature on Risk Measurement for the Trading Book, WP[19], Basel, Switzerland. Available at: http://www.bis.org/publ/bcbs_wp19.pdf

Basel Committee on Banking Supervision. (2012). Consultative Document: Fundamental Review of the Trading Book, Basel, Switzerland. Available from internet: http://www.bis.org/publ/bcbs219.pdf 1

Berkowitz, J., Christoffersen, P. F., & Pelletier, D. (2013). Evaluating Value-at-Risk Models with Desk Level Data. *Management Science, Informs*, 57(12), 2213–2227. https://doi.org/10.1287/mnsc.1080.0964

Bernard, C., Ruschendorf, L., Vanduffel S., & Yao, J. (2017). How robust is the value-at-risk of credit risk portfolios?. *The European Journal of Finance*, 23 (6), 507–534. https://doi.org/10.1080/1351847X.2015.1104370

Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review*, 39, 841–862. https://doi.org/10.2307/2527341

Christoffersen, P. F. (2003). Elements of Financial Risk Management. Boston: Academic Press.

Christoffersen, P. F. (2009). Value-at-Risk models. In: T.G. Andersen, R.A. Davis., J. P. Kreiss, and T. Mikosch. *Handbook of Financial Time Series. Springer Verlag*. https://doi.org/10.1007/978-3-540-71297-8_33

Colletaz, G., Hurlin, C., & Perignon, C. (2012). The Risk Map: A New Tool for Backtesting Risk Models. Working Paper, HEC Paris.

Danielsson, J., & Zhou, C. (2015) Why risk is so hard to measure, SRC Discussion Paper, No 36. Systematic Risk Centre, The London School of Economics and Political Science, London, UK.

Diamandis, F. R., Anastassios A., Kouretas, G. P., & Zarangas, L. (2011). Value at Risk for long and short trading positions: Evidence from developed and emerging equity markets. *International Review of Financial Analysis*, 20, 165–176. https://doi.org/10.1016/j.irfa.2011.02.009

Gneiting, T. (2011). Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, 106(494), 746–762. https://doi.org/10.1198/jasa.2011.r10138

Gunay, S. (2016). Alteration of risk in Asian bond markets during and after mortgage crisis: evidence from VaR analysis. *Asian Academy of Management Journal of Accounting & Finance*, 12 (2), 159–182. https://doi.org/10.21315/aamjaf2016.12.S1.7

Hull, J., & White, A. (1998). Incorporating Volatility Updating into the Historical Simulation Method for Value-at-Risk. *Journal of Risk*, 1(1), 5–19. https://doi.org/10.21314/JOR.1998.001

Morgan, J. P. (1994). Riskmetrics TM Technical Document, Morgan Guaranty Trust Company, Global Research, New York.

Jorion, P. (2007). Value at Risk: The New Benchmark for Managing Financial Risk, 3 ed. McGraw-Hill.

Kellner, R., & Rosch, 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, 45–63. https://doi.org/10.1016/j.jedc.2016.05.002

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives*, 3, 73–84. https://doi.org/10.3905/jod.1995.407942

Nelson, D. (1990). Conditional Heteroscedasticity in Asset Returns: a New Approach. *Econometrica, 59*, 347–370. https://doi.org/10.2307/2938260

Nieto, M. R., & Ruiz, E. (2016). Frontiers in VaR Forecasting and Backtesting. *International Journal of Forecasting*, 32 (2), 475–501. https://doi.org/10.1016/j.ijforecast.2015.08.003

Perignon, C., & Smith, D. R. (2008). A New Approach to Comparing VaR Estimation Methods. *Journal of Derivatives*, 16, 54–66. https://doi.org/10.3905/JOD.2008.16.2.054

Perignon, C., & Smith, D. R. (2010). The Level and Quality of Value-at-Risk Disclosure by Commercial Banks. *Journal of Banking and Finance*, 34, 362–377. https://doi.org/10.1016/j.jbankfin.2009.08.009

Soccorsi S. (2016). Measuring nonfundamentalness for structural VARs. *Journal of Economic Dynamics and Control*, 71, 86–101. https://doi.org/10.1016/j.jedc.2016.08.001

Wang C. P., & Huang H. H. (2016). Optimal insurance contract under VaR and CVaR constraints. *The North American Journal of Economics and Finance*, 37, 110–127. https://doi.org/10.1016/j.najef.2016.03.007

Wong, W. K. (2010). Backtesting value-at-risk based on tail losses .*Journal of Empirical Finance*, 17(3), 526–538. https://doi.org/10.1016/j.jempfin.2009.11.004

Zhu, S., Sun, X., & Li, D. (2013). Nonlinear portfolio selection using approximate parametric Value-at-Risk, *Journal of Banking & Finance*, 37, 2124–2139. https://doi.org/10.1016/j.jbankfin.2013.01.036

The article has been reviewed.

Received in March 2017; accepted in February 2019.