Violation duration as a better way of VaR model evaluation : evidence from Turkish market portfolio

Kilic, Ekrem
Finecus Financial Software and Consultancy

01. May 2006

Violation Duration As A Better Way of VaR Model Evaluation : Evidence From Turkish Market Portfolio

Ekrem Kilic*

May, 2006

Abstract

Financial crisis those we have been experienced during last two decades encouraged the efforts of both academicians and the market participants to develop clear representations of the risk exposure of a financial institute. As a useful tool for measuring market risk of a portfolio, Value-at-Risk has emerged as the standard. However, there are several alternative Value-at-Risk implementations which may produce significantly different Value-at-Risk forecasts. Thus, evaluation of Value-at-Risk forecasts is as crucial as VaR itself. In this paper I will use the methodology which has described by Christoffersen and Pelletier[6] and I extended the methodology to create duration based analogous of unconditional coverage, conditional coverage and independence tests. I evaluated 14 Value-at-Risk implementation by using a Turkish Market portfolio which contain foreing currency, stock and bonds.

JEL:C52

Keywords: Value-at-Risk, model evaluation, conditional coverage, duration based coverage testing

1 Introduction

Representation of the risk exposure of a financial institute has been a demanding issue for risk managers. Especially, financial crisis those we have

*Finecus Financial Software & Consultancy, phone: +90-216-4678955, fax:+90-216-4678773, e-mail: ekremk@finecus.com

been experienced during last two decades encouraged the efforts of both academicians and the market participants. Financial institutions looked for meaningful information about their risk exposure without the need for further technical explanations. In this situation, J.P. Morgan[16] developed Value-at-Risk (VaR) concept that has emerged as the standard.

VaR is just a single amount that reflects the worst possible loss of an asset portfolio for a given confidence level. In other words, VaR is a percentile of the conditional profit-loss distribution. Although VaR concept intuitively and simply addresses the risk exposure, there is no unique procedure to forecast VaR. There are various VaR implementations those can be classified into two main categories. One approach uses parametric methods and makes a distribution assumption. The other approach simulates the profit-loss distribution and calculates required percentile for this simulated distribution.

Financial risk managers have to select a proper model among the space of possible implementations, because all VaR models do not work well for every markets. The risk of financial risk model is called "model risk"; and is very important phenomenon in risk management. Therefore, evaluation of VaR model among a wide variety of alternative methods is the key element of VaR calculation.

One way to evaluate a VaR model is to employ statistical hypothesis testing methods under the null hypothesis that the model satisfies necessary theoretical conditions. In an early study about model evaluation, Kupiec[19] proposes several tests those are available and finds these tests have very limited power for commonly used sample sizes. Christoffersen[7] underlines the importance of violation clustering and improves testing framework to include conditional coverage. Recently, Christoffersen and Pelletier [6] suggests a new statistical testing framework which is based on duration of violation. They find that these new tests show better power properties with simulated data.

Another way of forecast evaluation is to incorporate a subjective loss functions that reflects the utility maximizing behavior of the financial institution. Lopez[20] formalizes this kind of methods and defines different loss functions for different financial institutions those have different utility functions.

In this paper, two new duration based test are introduced . These new statistical tests are compared with common tests by using 14 VaR implementations.

The rest of the paper is outlined as follows: in the following section, model evaluation methods are described. In the next chapter, performances of the VaR models are compared. Finally, I conclude.

2 Model Evaluation

Let $\tilde{\Phi}$ be the conditional distribution of daily logarithmic returns of a portfolio, R_t, then definition of VaR forecast, v_t^α is;

$$v_t^\alpha = -\tilde{\Phi}^{-1}\left(\alpha|\Im_{t-1}\right) \tag{1}$$

We define violation sequence of VaR forecast as;

$$I_t = \begin{cases} 1, & \text{if } (R_t < -v_t^\alpha) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

A quick theoretical result from these definitions is that for a proper VaR model probability of having a violation should be α. Most statistical evaluation methods exploit this feature of the VaR forecasts.

One important problem for considering violations as an indicator series is that it ignores the magnitude of violation. However the magnitude of violation is very meaningful for the regulatory authorities. At the same time, these techniques do not consider overestimation too. For instance, consider two VaR forecasts v_t^α and $(v_t^\alpha)^+$ where the second one is defined as v_t^α plus a small constant ε which satisfies $\varepsilon < \min\left(I_t \cdot (R_t + v_t^\alpha)\right)$. Then violation sequence of two VaR forecasts will be identical and any testing procedure based only on $\{I_t\}_{t=1}^T$ will produce same results for both VaR forecasts, however first model is more desirable for the firms and the second is more desirable for the regulators. Thus, beyond statistical tests, incorporating loss functions might be useful. Mandira et al.[21] suggests a two step model selection method which contains a first step of statistical evaluation and a second step of loss functions.

In the empirical results, VaR forecasts will be analyzed by employing 4 statistical tests and 2 loss functions. For each portfolio, paper reports selected model or models by following a two step procedure similar to Mandira et al.

2.1 Unconditional and Conditional Coverage Tests

For a sequence of VaR forecasts that calculated by using a proper model, $\{v_t^\alpha\}_{t=1}^T$, each element of violation sequence, $\{I_t\}_{t=1}^T$, can be modelled as independent draws from a Bernoulli distribution with probability of having a violation is α. Christoffersen[7] suggest a likelihood ratio test for

$$H_0 : \hat{\alpha} = \alpha \tag{3}$$

where $\hat{\alpha}$ ML estimate of α. Likelihood of an i.i.d. Bernoulli distributed sequence can be written as

$$L(\alpha) = \prod_{t=1}^{T} (1-\alpha)^{1-I_t} \alpha^{I_t} = (1-\alpha)^{T_0} \alpha^{T_1} \tag{4}$$

where T_0 is the number of covered days and T_1 is the number of violations. ML estimate of α is

$$\hat{\alpha} = \frac{T_1}{(T_0 + T_1)} \tag{5}$$

Now, we can easily find the likelihood of the sample by plugging the ML estimate into equation 4;

$$L(\hat{\alpha}) = \left(1 - \frac{T_1}{T}\right)^{T_0} \left(\frac{T_1}{T}\right)^{T_1} \tag{6}$$

Then, likelihood ratio test for unconditional coverage is

$$LR_{uc} = 2\left(l(\hat{\alpha}) - l(\alpha)\right) \overset{asy}{\sim} \chi_1^2 \tag{7}$$

where $l(.)$ is the log-likelihood function which defined as $\ln(L)$. Christoffersen[7] showed that LR_{uc} is asymptotically χ^2 distributed with degrees of freedom 1, however most likely we do not have large samples for VaR evaluation. Therefore, it is better to simulate LR_{uc} for finite samples. In this study, I used Monte Carlo simulation technique for p-values[1].

Unconditional coverage test implicitly assumes that the violations are independent over time. This assumption ignores clustering of violation which means that violations can occur closely together. If violations are clustered, probability of having a violation after a violation will be higher than α. In order to test existence of such an effect, we can define a first order Markov sequence with transition matrix

$$A = \begin{bmatrix} \alpha_{00} & \alpha_{01} \\ \alpha_{10} & \alpha_{11} \end{bmatrix} \tag{8}$$

where α_{00} is the probability of having a covered day after a covered day,α_{01} is the probability of having a violation after a covered day and so on... With this setup, independence can be defined as the null hypothesis that

$$H_0 : \alpha_{01} = \alpha_{11} \tag{9}$$

[1] All p-values are calculated by simulating test statistics 10.000 times.

The likelihood function of first order Markov process can be written as

$$L(A) = \alpha_{00}^{T_{00}} \alpha_{01}^{T_{01}} \alpha_{10}^{T_{10}} \alpha_{11}^{T_{11}} \tag{10}$$

The ML estimates of elements in the transition matrix are

$$\hat{\alpha}_{01} = \frac{T_{01}}{T_{00} + T_{01}} \tag{11}$$

$$\hat{\alpha}_{11} = \frac{T_{11}}{T_{10} + T_{11}} \tag{12}$$

$$\hat{\alpha}_{00} = 1 - \hat{\alpha}_{01} \tag{13}$$

$$\hat{\alpha}_{01} = 1 - \hat{\alpha}_{11} \tag{14}$$

Using ML estimates of parameters, we can calculate likelihood of the sample. Now we can test null hypothesis of independence by using another likelihood ratio test as offered by Christoffersen[7].

$$LR_{ind} = 2\left(l\left(\hat{A}\right) - l\left(\hat{\alpha}\right)\right) \overset{asy}{\sim} \chi_1^2 \tag{15}$$

Again test statistic asymptotically a χ^2 with degrees of freedom 1. As final step, test of correct conditional coverage is

$$LR_{cc} = 2\left(l\left(\hat{A}\right) - l\left(\alpha\right)\right) = LR_{uc} + LR_{ind} \overset{asy}{\sim} \chi_2^2 \tag{16}$$

which tests $\alpha_{01} = \alpha_{11} = \alpha$. As it is mentioned before, for finite samples, p-values can be calculated from Monte Carlo simulation.

2.2 Duration Based Tests

Although Christoffersen's conditional coverage test provides a parsimonious procedure for model evaluation, it is limited in the sense that it only considers first order dependence. If the violation sequences exhibit a dependence structure other than first order Markov dependence, test would fail to detect.

In their paper, Christoffersen and Pelletier[6] suggest a new testing framework which based on duration of violations rather than sequence of violations itself. The motivation behind this approach is that if violations are clustered, there would be relatively short durations and relatively long durations as many as it is unlikely to occur under a proper duration distribution.

No-hit duration between two violations can be simply defined as

$$D_i = t_i - t_{i-1}, \qquad t_0 = 0 \tag{17}$$

where t is the day of violation and i is the number of violation. Therefore first duration is equal to day of first violation.

As it is remarked above, each element of violation sequence comes from a Bernoulli distribution. Thus, if we consider duration d, as $(d-1)$ times consecutive non-violations and one violation at d^{th} trial, probability distribution of no-hit duration is

$$\Pr(D=d) = (1-\alpha)^{d-1}\,\alpha \tag{18}$$

which is called geometric distribution. Expected duration for geometrically distributed random variable is $\frac{1}{\alpha}$ and the variance is $\frac{1-\alpha}{\alpha^2}$. Dealing with duration distributions, hazard functions are also important, they identifies the characteristics of the distribution. Hazard function of duration distribution is defined as

$$\lambda\left(d\right) = \frac{\Pr(D=d)}{1-\Pr(D<d)} \tag{19}$$

As a special case geometric distribution has a constant hazard function as follows,

$$\lambda\left(d\right) = \frac{(1-\alpha)^{d-1}\,\alpha}{1-\sum\limits_{i=0}^{d-2}(1-\alpha)^{i}\,\alpha} \tag{20}$$

$$\lambda\left(d\right) = \alpha \tag{21}$$

A constant hazard function means that duration distribution has no memory. As it will be mentioned later, Christoffersen and Pelletier[6] tests this feature of duration sequence by incorporating Weibull distribution. However, before proceeding through this way, I will propose another duration based test which tests the null hypothesis that duration sequence is from a geometric distribution that has a violation probability equal to α. For this purpose, first it is necessary to define the likelihood function of the geometrically distributed durations. Log-likelihood function which considers censored and uncensored durations can be written as follows;

$$\begin{aligned} l_{cn} = {} & C_1 \ln\left(1-\Pr(D<d)\right) + (1-C_1)\ln\left(\Pr(D=d)\right) + \sum_{i=2}^{N-1}\ln\left(\Pr(D=d)\right) \\ & + C_N \ln\left(1-\Pr(D<d)\right) + (1-C_N)\ln\left(\Pr(D=d)\right) \end{aligned} \tag{22}$$

where $\{C_i\}_{i=t}^{T}$ is the sequence of indicators, it shows a duration is censored $(C_i=1)$ or not $(C_i=0)$. Thus, for all durations this indicator will be 0, except first and last durations. If the first element of violation sequence is 1, then $C_1=0$, otherwise $C_1=1$, meaning first duration is left censored. Similarly, if the last element of violation sequence is 1, then $C_N=0$, otherwise

$C_N = 1$, which means that last duration is right censored. $(1 - \Pr(D < d))$ is also called as survival function and for geometric distribution it is defined as

$$S(d) = (1-\alpha)^{d-1} \tag{23}$$

Inserting equation 18 and 23 into 22 and rearranging, we will have

$$l_{cn}(\alpha) = -C_1 \ln(\alpha) - C_N \ln(\alpha) + N \ln(\alpha) + \ln(1-\alpha) \sum_{i=1}^{N} (D_i - 1) \tag{24}$$

and ML estimate can be found as

$$\hat{\alpha} = \frac{N - C_1 - C_N}{\left(\sum_{i=1}^{N} D_i\right) - C_1 - C_N} \tag{25}$$

Now we can test the null hypothesis that claim

$$H_0 : \hat{\alpha} = \alpha \tag{26}$$

by using the following likelihood ratio,

$$LR_{geo} = 2\left(l_{cn}(\hat{\alpha}) - l_{cn}(\alpha)\right) \tag{27}$$

For finite sample inference, again we can benefit from the advantages of Monte Carlo techniques. A useful description of the Monte Carlo procedure can be found in Christoffersen and Pelletier[6].

After defining this simple duration based test, let us return back to memory-free nature of the geometric distribution and the test suggested by Christoffersen and Pelletier. First of all, geometric distribution will be substituted with its continues-time limit, exponential distribution. Thus, distribution of no-hit duration under the null now becomes[2]

$$f_{\exp}(D) = \alpha \exp(-\alpha D) \tag{28}$$

[2]At this point it is possible to replicate the geometric distribution test by having exponential distribution as the null and alternative instead of geometric distribution. Using equation 22, ML estimate might be found as

$$\hat{\alpha} = \frac{N - C_1 - C_N}{\sum_{i=1}^{N} D_i}$$

The results of this alternative test are unsurprisingly quite similar, since both distributions are the same at the limit.

To be able to test, memory of hazard function, Christoffersen and Pelletier incorporates Weibull distribution as an alternative, because Weibull distribution allows for duration dependence and independence due to parameter choice. Probability density function of Weibull distribution is

$$f_W(D) = a^b b D^{b-1} \exp\left(-(aD)^b\right) \tag{29}$$

And its hazard function can be formalized as

$$\lambda_w(D) = a^b b D^{b-1} \tag{30}$$

An important property of Weibull function is when $b = 1$, the hazard function becomes a constant function and moreover Weibull distribution reduces to exponential distribution. Therefore, the independence of no-hit duration can be tested using following null hypothesis

$$H_0 : b = 1 \tag{31}$$

Log-likelihood function of the durations again follows the general form given in equation 22. However, this time ML estimates of Weibull parameters a and b are needed to be optimized by using a numerical optimization procedure. Fortunately, it is possible to find following relation between a and b by derivating log-likelihood function

$$\hat{a} = \left(\frac{N - C_1 - C_N}{\sum_{i=1}^{N} D_i^b}\right)^{\frac{1}{b}} \tag{32}$$

Then, optimization problem becomes a univariate unconstraint maximization. When $b = 1$, equation 32 turns to ML estimate of exponential distribution, $\hat{\alpha}$ (see footnote 2). Hence the null hypothesis 31 implicitly says that

$$D_i \sim Exponential\left(\hat{\alpha}\right) \tag{33}$$

In this paper, going one step further, the null hypothesis is substituted with

$$D_i \sim Exponential\left(\alpha\right) \tag{34}$$

where α is the original coverage of VaR forecast. This approach turns hypothesis 31 to simultaneous hypothesis

$$H_0 : b = 1,\ \alpha = \hat{\alpha} \tag{35}$$

The original Weibull test of Christoffersen and Pelletier is the analogous of independence test. Similarly, the test with exponential distribution is the duration based analogous of unconditional coverage test. Therefore, extending the hypothesis 31 to 35, I prepared the the analogous of conditional coverage test.

Once again, test statistic consists of a likelihood ratio test and p-values of this statistics are generated using Monte Carlo methods.

In this paper this new test will be called as modified Weibull test and this test can be shown also as

$$LR_{weibull^*} = LR_{weibull} + LR_{\exp} \tag{36}$$

where $LR_{weibull^*}$ is the modified Weibull test statistic, $LR_{weibull}$ is the Weibull test statistic of Christoffersen and Pelletier and $LR_{\exp}$ is the exponential distribution test (see footnote 2).

3 Empirical Results

In this section, I present application results of the VaR evaluation methods to the simulation based VaR models. For this purpose, I employed the portfolio which contains Turkish market instruments[3].

Firstly, let us investigate details of the portfolio. Turkish portfolio includes 5 instruments; two zero bonds of Turkish Treasury with 117-day and 453-day maturities[4], two fx positions (USD/TRY and EUR/TRY), and one stock exchange index (ISE100 Index of Turkey). Portfolio has homogeneous present value distribution, in other words each position has 20% weight in the portfolio.

In this study, I worked roughly 500 VaR results for the portfolio from November 2003 and November 2005. Since each VaR estimation requires past data, the observations start from November 2002. Another point is parameter estimation of volatility and correlation models.Each models are re-estimated with the observations of the related VaR. Thus, GARCH(1,1) parameters or DCC(1,1) model parameters are estimated by using a 252-day length moving window of observations.

3.0.1 Results of 99% VaR

Table 1 shows the results of the unconditional coverage, independence, and conditional coverage tests for 99% VaR forecasts of Turkish portfolio . First

[3]All calculations are made in terms of TRY (New Turkish Lira) and all instruments TRY denominated.

[4]Their ISIN codes are TRB220206T14 and TRT240107T12, respectively.

column gives the names of the VaR models. Following three columns provide LR statistics of the tests. Although distribution of these tests are known, as it mentioned before, I preferred applying Monte Carlo method for finite sample inference. Probabilities of LR statistics are given below the LR statistics. Next column gives estimated unconditional coverage probability. And the last two columns shows estimated conditional coverage probabilities.

Table 1: Results of the unconditional coverage, independence and conditional coverage tests for 99% VaR estimation of Turkish portfolio.

METHODS	LR_{uc}	LR_{ind}	LR_{cc}	$\widehat{\alpha}$	$\widehat{\alpha}_{01}$	$\widehat{\alpha}_{11}$
HS	0.001	0.100	0.101	0.010	0.010	-
	(0.818)		(0.827)			
HS-EVT	0.237	0.060	0.297	0.008	0.008	-
	(0.507)		(0.659)			
HS-Kernel	0.984	0.028	1.011	0.006	0.006	-
	(0.257)		(0.286)			
FHS-EWMA	0.001	0.100	0.101	0.010	0.010	-
	(0.818)		(0.827)			
FHS-GARCH	0.237	0.060	0.297	0.008	0.008	-
	(0.507)		(0.659)			
FHS-EWMA-EVT	0.237	0.060	0.297	0.008	0.008	-
	(0.507)		(0.659)			
FHS-GARCH-EVT	0.170	0.148	0.319	0.012	0.012	-
	(0.677)		(0.528)			
FHS-EWMA-Kernel	0.001	0.100	0.101	0.010	0.010	-
	(0.818)		(0.827)			
FHS-GARCH-Kernel	0.984	0.028	1.011	0.006	0.006	-
	(0.257)		(0.286)			
WHS	0.237	0.060	0.297	0.008	0.008	-
	(0.507)		(0.659)			
MC-EWMA	0.001	0.100	0.101	0.010	0.010	-
	(0.818)		(0.827)			
MC-CCC	2.414	0.004	2.417	0.004	0.004	-
	(0.106)		(0.143)			
MC-DCC	2.414	0.004	2.417	0.004	0.004	-
	(0.106)		(0.143)			

For this portfolio, any method is rejected. For this case unconditional coverage reject the model but conditional coverage is slightly failed to reject. It is also interesting to notice that all HS variants of US portfolio are rejected.

Now let us examine higher order dependence by incorporating duration based tests. First, I start with geometric distribution test. Table 2 gives the results of the geometric distribution test. In the first column of the table, method names are given. Next column shows the test statistics and the last gives the coverage probability of the geometric distribution.

Table 2: Results of the geometric distribution test for 99% VaR estimation of Turkish portfolio.

METHODS	LR_{geo}	$\widehat{\alpha}$
HS	0.229	0.008
	(0.618)	
HS-EVT	0.967	0.006
	(0.470)	
HS-Kernel	2.389	0.004
	(0.182)	
FHS-EWMA	0.229	0.008
	(0.579)	
FHS-GARCH	0.967	0.006
	(0.337)	
FHS-EWMA-EVT	0.967	0.006
	(0.337)	
FHS-GARCH-EVT	0.000	0.010
	(0.936)	
FHS-EWMA-Kernel	0.229	0.008
	(0.618)	
FHS-GARCH-Kernel	2.389	0.004
	(0.248)	
WHS	0.967	0.006
	(0.470)	
MC-EWMA	0.229	0.008
	(0.618)	
MC-CCC	4.862**	0.002
	(0.049)	
MC-DCC	4.862**	0.002
	(0.049)	

Geometric distribution test rejects MC-CCC and MC-DCC forecasts of Turkish portfolio. In the estimation of CCC and DCC models, GARCH(1,1) model is used as univariate volatility specification and normality is assumed. Then we can say, conditional correlation models that employs GARCH(1,1) with normality is not capable of reflecting correlation structure of Turkish markets for the analysis period, because MC-EWMA is survived with a high probability although it has the same features with MC-CCC and MC-DCC except covariance modelling. For US portfolio, 3 models are failed to reject (FHS-GARCH, FHS-EWMA-EVT, and MC-EWMA), other models are rejected.

Table 3 and 4 present the results of the modified Weibull, Weibull, and exponential distribution tests for 99% VaR forecasts of Turkish portfolio and US portfolio, respectively. Again, first column gives the names of the VaR models and next three columns provide LR statistics of the tests. Following column gives estimated coverage probability for exponential distribution. And the last two columns shows estimated a and b parameters of Weibull distribution.

Test statistics of the exponential distribution is quite similar. However, I observed that distribution of test statistics differs. Distribution of geometric test statistic has a longer right tail while distribution of exponential test statistic is flatter at the center of distribution. For Turkish portfolio, exponential distribution test rejects MC-CCC and MC-DCC models as they rejected by the geometric distribution test, however exponential distribution test rejects at 90% significance while the geometric distribution test rejects at 95% significance. For US portfolio, 3 models are failed to reject (FHS-GARCH, FHS-EWMA-EVT, and MC-EWMA), other models are rejected by at least one of the tests. Another example for the difference between geometric distribution test and exponential distribution test is WHS forecast of US portfolio; for this case geometric distribution test rejects the null, however exponential is failed to reject. Weibull test and modified Weibull test produces totally different results as they supposed to; Weibull test rejects only WHS forecast, however modified Weibull test rejected 7 models. The reason for difference is, Weibull test deals with the dependence between violations, on the other hand modified Weibull test consider coverage too. A final remark is that when a estimate is zero there is no optimal b, thus solution of b is set of real numbers.

Table 3: Results of the exponential distribution, Weibull and modified Weibull tests for 99% VaR estimation of Turkish portfolio.

METHODS	LR_{exp}	$LR_{weibull}$	$LR_{weibull^*}$	$\widehat{\alpha}$	$\widehat{a}$	$\widehat{b}$
HS	0.235	0.097	0.333	0.008	0.008	1.141
	(0.643)	(0.759)	(0.876)			
HS-EVT	0.975	1.885	2.860	0.006	0.006	2.048
	(0.356)	(0.251)	(0.359)			
HS-Kernel	2.395	2.429	4.824	0.004	0.005	2.773
	(0.200)	(0.198)	(0.216)			
FHS-EWMA	0.235	0.702	0.937	0.008	0.008	1.387
	(0.643)	(0.467)	(0.730)			
FHS-GARCH	0.975	1.647	2.623	0.006	0.006	1.867
	(0.442)	(0.284)	(0.397)			
FHS-EWMA-EVT	0.975	1.647	2.623	0.006	0.006	1.867
	(0.442)	(0.284)	(0.397)			
FHS-GARCH-EVT	0.000	0.533	0.534	0.010	0.010	1.276
	(0.942)	(0.521)	(0.822)			
FHS-EWMA-Kernel	0.235	0.226	0.461	0.008	0.008	1.206
	(0.643)	(0.656)	(0.838)			
FHS-GARCH-Kernel	2.395	0.006	2.401	0.004	0.004	0.959
	(0.157)	(0.911)	(0.457)			
WHS	0.975	1.885	2.860	0.006	0.006	2.048
	(0.356)	(0.251)	(0.359)			
MC-EWMA	0.235	1.747	1.982	0.008	0.008	1.671
	(0.584)	(0.270)	(0.506)			
MC-CCC	4.861*	0.149	5.010	0.002	0.001	0.729
	(0.083)	(0.713)	(0.194)			
MC-DCC	4.861*	0.149	5.010	0.002	0.001	0.729
	(0.083)	(0.713)	(0.194)			

Table 4: Results of the exponential distribution, Weibull and modified Weibull tests for 99% VaR estimation of US portfolio.

METHODS	LR_{exp}	$LR_{weibull}$	$LR_{weibull^*}$	$\widehat{\alpha}$	$\widehat{a}$	$\widehat{b}$
HS	20.080**	-	20.080**	-	-	$\{\mathbb{R}\}$
	(0.000)	(0.962)	(0.001)			
HS-EVT	20.080**	-	20.080**	-	-	$\{\mathbb{R}\}$
	(0.000)	(0.962)	(0.001)			
HS-Kernel	20.080**	-	20.080**	-	-	$\{\mathbb{R}\}$
	(0.000)	(0.962)	(0.001)			
FHS-EWMA	6.955**	0.567	7.522	0.024	0.023	1.197
	(0.039)	(0.514)	(0.112)			
FHS-GARCH	0.695	3.627	4.322	0.014	0.014	1.965
	(0.536)	(0.133)	(0.249)			
FHS-EWMA-EVT	2.548	2.243	4.791	0.018	0.017	1.553
	(0.143)	(0.225)	(0.228)			
FHS-GARCH-EVT	10.040**	-	10.040**	-	-	$\{\mathbb{R}\}$
	(0.008)	(0.962)	(0.056)			
FHS-EWMA-Kernel	4.813**	0.678	5.491	0.002	0.001	0.544
	(0.044)	(0.479)	(0.169)			
FHS-GARCH-Kernel	10.040**	-	10.040**	-	-	$\{\mathbb{R}\}$
	(0.008)	(0.962)	(0.056)			
WHS	4.813	9.496**	14.309**	0.002	0.003	81.801
	(0.122)	(0.037)	(0.036)			
MC-EWMA	1.496	0.211	1.707	0.016	0.016	1.138
	(0.306)	(0.676)	(0.546)			
MC-CCC	4.813**	0.678	5.491	0.002	0.001	0.544
	(0.044)	(0.479)	(0.169)			
MC-DCC	10.040**	-	10.040**	-	-	$\{\mathbb{R}\}$
	(0.008)	(0.962)	(0.056)			

3.0.2 Results of 95% VaR

In this section, evaluation test results of 95% VaR forecasts are analyzed. Table 5 and 6 show the results of the unconditional coverage, independence, and conditional coverage tests for 95% VaR forecasts of Turkish portfolio and US portfolio, respectively. First column gives the names of the VaR models. Following three columns provide LR statistics of the tests. Probabilities of LR statistics are given below the LR statistics. Next column gives estimated unconditional coverage probability. And the last two columns shows estimated conditional coverage probabilities.

Table 5: Results of the unconditional coverage, independence and conditional coverage tests for 95% VaR estimation of Turkish portfolio.

METHODS	LR_{uc}	LR_{ind}	LR_{cc}	$\widehat{\alpha}$	$\widehat{\alpha}_{01}$	$\widehat{\alpha}_{11}$
HS	0.459	2.005	2.464	0.044	0.046	-
	(0.476)		(0.387)			
HS-EVT	0.797	1.823	2.620	0.042	0.043	-
	(0.354)		(0.311)			
HS-Kernel	5.094**	0.919	6.013**	0.030	0.031	-
	(0.019)		(0.041)			
FHS-EWMA	0.023	0.320	0.343	0.051	0.050	0.077
	(0.843)		(0.815)			
FHS-GARCH	0.023	0.320	0.343	0.051	0.050	0.077
	(0.843)		(0.815)			
FHS-EWMA-EVT	0.217	0.002	0.220	0.046	0.046	0.043
	(0.610)		(0.878)			
FHS-GARCH-EVT	0.561	0.344	0.904	0.057	0.059	0.034
	(0.416)		(0.659)			
FHS-EWMA-Kernel	0.561	0.344	0.904	0.057	0.059	0.034
	(0.416)		(0.659)			
FHS-GARCH-Kernel	0.125	0.169	0.294	0.053	0.055	0.037
	(0.688)		(0.850)			
WHS	0.561	3.536	4.096	0.057	0.061	-
	(0.416)		(0.120)			
MC-EWMA	0.459	0.002	0.461	0.044	0.044	0.045
	(0.476)		(0.754)			
MC-CCC	1.774	1.486	3.260	0.038	0.039	-
	(0.152)		(0.203)			
MC-DCC	1.774	1.486	3.260	0.038	0.039	-
	(0.152)		(0.203)			

Table 6: Results of the unconditional coverage, independence and conditional coverage tests for 95% VaR estimation of US portfolio.

METHODS	LR_{uc}	LR_{ind}	LR_{cc}	$\widehat{\alpha}$	$\widehat{\alpha}_{01}$	$\widehat{\alpha}_{11}$
HS	0.420	0.001	0.421	0.044	0.044	0.045
	(0.472)		(0.789)			
HS-EVT	0.745	1.834	2.579	0.042	0.044	-
	(0.353)		(0.331)			
HS-Kernel	3.089*	1.192	4.281	0.034	0.035	-
	(0.060)		(0.101)			
FHS-EWMA	3.68**	0.992	4.672*	0.070	0.067	0.114
	(0.049)		(0.084)			
FHS-GARCH	2.392*	0.015	2.407	0.066	0.066	0.061
	(0.099)		(0.431)			
FHS-EWMA-EVT	0.609	0.353	0.962	0.058	0.059	0.034
	(0.417)		(0.640)			
FHS-GARCH-EVT	0.000	0.057	0.057	0.050	0.050	0.040
	(0.922)		(0.974)			
FHS-EWMA-Kernel	0.190	0.003	0.193	0.046	0.046	0.043
	(0.608)		(0.883)			
FHS-GARCH-Kernel	1.169	0.052	1.221	0.040	0.040	0.050
	(0.262)		(0.609)			
WHS	0.051	0.022	0.073	0.048	0.048	0.042
	(0.762)		(0.942)			
MC-EWMA	3.005	0.225	3.229	0.068	0.066	0.088
	(0.083)		(0.207)			
MC-CCC	0.000	0.057	0.057	0.050	0.050	0.040
	(0.922)		(0.974)			
MC-DCC	0.190	0.003	0.193	0.046	0.046	0.043
	(0.608)		(0.883)			

For Turkish portfolio, only HS-Kernel model is rejected. The results of the US portfolio are much more different than the results of 99% VaR forecasts. It is also interesting to notice that FHS-GARCH model which is one of the 4 surviving models among 99% VaR forecasts, is rejected at 95%. HS-Kernel and FHS-EWMA models are rejected at both coverage levels.

Table 7 and 8 gives the results of the geometric distribution test. In the first column of the table, method names are given. Next column shows the test statistics and the last gives the coverage probability of the geometric distribution.

Table 7: Results of the geometric distribution test for 95% VaR estimation of US portfolio.

METHODS	LR_{geo}	$\widehat{\alpha}$
HS	0.762	0.042
	(0.393)	
HS-EVT	1.190	0.040
	(0.315)	
HS-Kernel	6.156**	0.028
	(0.016)	
FHS-EWMA	0.001	0.050
	(0.942)	
FHS-GARCH	0.001	0.050
	(0.942)	
FHS-EWMA-EVT	0.433	0.044
	(0.507)	
FHS-GARCH-EVT	0.328	0.056
	(0.617)	
FHS-EWMA-Kernel	0.328	0.056
	(0.617)	
FHS-GARCH-Kernel	0.030	0.052
	(0.907)	
WHS	0.328	0.056
	(0.607)	
MC-EWMA	0.762	0.042
	(0.393)	
MC-CCC	2.365	0.036
	(0.117)	
MC-DCC	2.365	0.036
	(0.128)	

Table 8: Results of the geometric distribution test for 95% VaR estimation of US portfolio.

Methods	LR_{geo}	$\widehat{\alpha}$
HS	0.711 (0.408)	0.042
HS-EVT	1.127 (0.313)	0.040
HS-Kernel	3.888** (0.044)	0.032
FHS-EWMA	3.081* (0.073)	0.068
FHS-GARCH	1.903 (0.174)	0.064
FHS-EWMA-EVT	0.365 (0.577)	0.056
FHS-GARCH-EVT	0.043 (0.778)	0.048
FHS-EWMA-Kernel	0.394 (0.492)	0.044
FHS-GARCH-Kernel	1.647 (0.199)	0.038
WHS	0.173 (0.655)	0.046
MC-EWMA	2.459 (0.116)	0.066
MC-CCC	0.043 (0.791)	0.048
MC-DCC	0.394 (0.492)	0.044

Geometric distribution test rejects HS-Kernel forecast of Turkish portfolio. For US portfolio, again HS-Kernel and FHS-EWMA models are rejected, geometric distribution test reject these models at 99% coverage too.

Table 9 and 10 present the results of the modified Weibull, Weibull, and exponential distribution tests for 95% VaR forecasts of Turkish portfolio and US portfolio, respectively. Again, first column gives the names of the VaR models and next three columns provide LR statistics of the tests. Following column gives estimated coverage probability for exponential distribution. And the last two columns shows estimated a and b parameters of Weibull distribution.

For Turkish portfolio, all HS variants are rejected by at least one of the three tests. All tests rejects HS-Kernel. For US portfolio, 4 models are rejected (HS-Kernel, FHS-EWMA-EVT, FHS-EWMA-Kernel, and MC-EWMA) by at least one of the three tests. Among 4 EWMA related forecasts, 3 models are rejected. Then, it might be an evidence to claim that EWMA model is not a proper model for volatility modelling of US markets.

4 Conclusion

This paper have suggested new tests for model evaluation. To investigate performances of the test, I applied 13 simulation based VaR models to two portfolios which contain fx, bond, and stock positions.

The new statistical tests use the setup that described by Christoffersen and Pelletier [6]. I modefied their test to get duration based analogous of unconditional coverage, conditional coverage and independence tests. Empirical results showed that modified version of Weibull test get enabled to detect coverage problem too. For all of the p-values, Monte Carlo analysis are used.

Table 9: Results of the exponential distribution, Weibull and modified Weibull tests for 95% VaR estimation of Turkish portfolio.

Methods	LR_{exp}	$LR_{weibull}$	$LR_{weibull^*}$	$\widehat{\alpha}$	$\widehat{a}$	$\widehat{b}$
HS	0.759	3.740*	4.499	0.042	0.050	0.761
	(0.373)	(0.082)	(0.137)			
HS-EVT	1.176	5.161**	6.337*	0.040	0.050	0.724
	(0.273)	(0.040)	(0.062)			
HS-Kernel	5.986**	4.090*	10.077**	0.028	0.034	0.695
	(0.015)	(0.068)	(0.011)			
FHS-EWMA	0.002	0.000	0.003	0.050	0.049	1.002
	(0.949)	(0.993)	(0.999)			
FHS-GARCH	0.002	0.000	0.003	0.050	0.049	1.002
	(0.949)	(0.993)	(0.999)			
FHS-EWMA-EVT	0.438	0.130	0.568	0.044	0.045	0.942
	(0.498)	(0.754)	(0.786)			
FHS-GARCH-EVT	0.289	0.005	0.294	0.055	0.055	1.010
	(0.599)	(0.949)	(0.883)			
FHS-EWMA-Kernel	0.289	0.005	0.294	0.055	0.055	1.010
	(0.599)	(0.949)	(0.883)			
FHS-GARCH-Kernel	0.022	0.019	0.041	0.051	0.051	1.021
	(0.894)	(0.904)	(0.982)			
WHS	0.289	1.394	1.683	0.055	0.061	0.864
	(0.585)	(0.296)	(0.497)			
MC-EWMA	0.759	0.281	1.040	0.042	0.040	1.097
	(0.415)	(0.643)	(0.652)			
MC-CCC	2.316	0.642	2.958	0.036	0.034	1.157
	(0.112)	(0.486)	(0.276)			
MC-DCC	2.316	0.671	2.987	0.036	0.034	1.157
	(0.111)	(0.475)	(0.273)			

Table 10: Results of the exponential distribution, Weibull and modified Weibull tests for 95% VaR estimation of US portfolio.

Methods	LR_{exp}	$LR_{weibull}$	$LR_{weibull^*}$	$\widehat{\alpha}$	$\widehat{a}$	$\widehat{b}$
HS	0.709	0.892	1.601	0.042	0.040	1.187
	(0.394)	(0.408)	(0.506)			
HS-EVT	1.115	2.419	3.534	0.040	0.037	1.338
	(0.287)	(0.165)	(0.212)			
HS-Kernel	3.791**	0.934	4.725	0.032	0.030	1.222
	(0.045)	(0.396)	(0.124)			
FHS-EWMA	2.838	1.774	4.611	0.068	0.064	1.216
	(0.107)	(0.235)	(0.130)			
FHS-GARCH	1.744	4.214*	5.958*	0.064	0.059	1.372
	(0.191)	(0.064)	(0.070)			
FHS-EWMA-EVT	0.323	4.830**	5.153*	0.056	0.052	1.444
	(0.598)	(0.048)	(0.099)			
FHS-GARCH-EVT	0.049	1.001	1.050	0.048	0.046	1.179
	(0.841)	(0.379)	(0.639)			
FHS-EWMA-Kernel	0.400	0.818	1.218	0.044	0.042	1.171
	(0.560)	(0.428)	(0.589)			
FHS-GARCH-Kernel	1.620	0.075	1.694	0.038	0.037	1.051
	(0.223)	(0.812)	(0.480)			
WHS	0.181	0.474	0.655	0.046	0.044	1.122
	(0.636)	(0.550)	(0.759)			
MC-EWMA	2.260	2.917	5.177*	0.066	0.062	1.296
	(0.115)	(0.126)	(0.098)			
MC-CCC	0.049	0.621	0.670	0.048	0.046	1.139
	(0.835)	(0.493)	(0.755)			
MC-DCC	0.400	1.509	1.909	0.044	0.042	1.241
	(0.541)	(0.275)	(0.438)			

References

[1] Bollerslev, T., (1990), "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model", *Review of Economics and Statistics*,No.72, pp.498-505.

[2] Boudoukh, J., M. Richardson , and R. Whitelaw, (1998). "The best of both worlds". *Risk*, May, pp. 64-67.

[3] Burden, R.L., J.D. Faires, (2005). *Numerical Analysis*. Eigth Edition, Chp. 4, pp.196-203, Thomson Higher Education, Belmont, CA,USA

[4] Butler, J.S. and B. Schachter, (1998). "Estimating Value-at-Risk with a Precision Measure by Combining Kernel Estimation with Historical Simulation". *Review of Derivatives Research*, No.1, pp.371-390.

[5] Christoffersen, P., J. Hahn and A.Inoue, (2001). "Testing, Comparing, and Combining Value-at-Risk Measures". *Center for Financial Institutions Working Papers 99-44*, Wharton School Center for Financial Institutions, University of Pennsylvania.

[6] Christoffersen, P., D. Pelletier, (2004). "Backtesting Value-at-Risk: A Duration-Based Approach". *Journal of Financial Econometrics*, Vol.2, No.1, pp.84-108, Oxford University Press

[7] Christoffersen, P., (1998)."Evaluating Interval Forecasts". *International Economic Review* No.39, pp.841-862.

[8] Christoffersen, P., (2003). *Elements of Financial Risk Management*.San Diego, Academic Press.

[9] Danielsson, J., C. G. de Vries, (1998). "Beyond the Sample: Extreme Quantile and Probability Estimation," *FMG Discussion Papers* dp298, Financial Markets Group

[10] Danielsson, J., C. G. de Vries, (1998). "Value-at-Risk and Extreme Returns," *Tinbergen Institute Discussion Papers 98-017/2*, Tinbergen Institute

[11] Engle, R., (2002). "Dynamic Conditional Correlation : A Simple Class of Multivariate GARCH". *Journal of Business and Economic Statistics*, No.20, pp.339-350.

[12] Glasserman, P., P. Heidelberger, P. Shahabuddin, (2000). "Efficient Monte Carlo Methods for Value-at-Risk". *Risk Management Report 2000.*

[13] Glasserman, P., (2004). *Monte Carlo Methods in Financial Engineering.* New York, Springer-Verlag

[14] Hull J. and White, (1998). "Incorporating Volatility Updating Into the Historical Simulation Method for Value-at-Risk".*Journal of Derivatives*, Vol. 6, No. 1, (Fall 1998), pp. 5-19

[15] Jorion, P., (2001). *Value at Risk: The New Benchmark for Controlling Market Risk.* 2 edn, McGraw-Hill, New York.

[16] J. P. Morgan, (1996). *RiskMetrics Technical Manual.* Fourth Edition. J. P. Morgan

[17] J. P. Morgan, (1999). *Risk Management: A Practical Guide.* J. P. Morgan

[18] Kilic, E., (2004). "Forecasting Volatility of Turkish Markets: A Comparison of Thin and Thick Models".*Econometrics 0510007*, Economics Working Paper Archive EconWPA.

[19] Kupiec,P.H., (1995). "Techniques for verifying the accuracy of risk measurement models". *Finance and Economics Discussion Series 95-24*, Board of Governors of the Federal Reserve System

[20] Lopez, J.A., (1998). "Methods For Evaluating Value-at-Risk Estimates". *Research Paper No.9802*, Federal Reserve Bank of New York.

[21] Mandira, S., S. Thomas, and A.Shah, (2003). "Selection of Value-at-Risk Models". *Journal of Forecasting*, No.22, pp.337-358, John Wiley & Sons, Ltd.

[22] Pagan, A., A. Ullah, (1999). *Nonparametric Econometrics.* First Edition, Chp.1, pp.71-77, Cambridge University Press

[23] Rosenblatt, M., (1956). "Remarks on Some Nonparametric Estimates of a Density Function". *Annals of Mathematical Statistics*, Vol.27,N.3,pp.:832-837.