University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 28/2021 (376)

The Effectiveness of Value-at-Risk Models in Various Volatility Regimes

Aleksander Schiffers
Marcin Chlebus

Warsaw 2021

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

The effectiveness of Value-at-Risk models in various volatility regimes

Aleksander Schiffers, Marcin Chlebus

University of Warsaw, Faculty of Economic Sciences, Department of Quantitative Finance
** Corresponding authors: mchlebus@wne.uw.edu.pl, a.schiffers@student.uw.edu.pl*

Abstract: There is an ongoing discussion, what is the most efficient approach to Value-at-Risk estimation. Subsequent studies and meta-analyzes show that there is no scientific consensus in this field and the necessity of further research is frequently underlined. In this study, authors try to assess the comparative performance of models used for Value-at-Risk estimations in changing market volatility regimes. The models considered are: the Historical Simulation, the Risk Metrics®, the GARCH(1,1)-n, the GARCH(1,1)-t, the GARCH(1,1)-st, the GARCH(1,1)-QML. GARCH models are additionally enriched with additional, exogenous regressors in the form of lagged commodity futures contracts returns. The analysis is conducted on a set of utility sector stock indices from six developed countries across the globe: WIG Energia (Poland), Dow Jones Utilities Average (USA), CAC Utilities (France), Tokyo SE Topix-17 Power & Gas (Japan), S&P ASX 200 Utilities (Australia), and DAX All Utilities (Germany). Three samples of different characteristics are distinguished from the last 10 years of data and one of them covers the upsurge in market volatility caused by the Covid-19 pandemic. In order to evaluate the VaR forecasts performance of each model, conditional/unconditional coverage tests of Kupiec and Christoffersen, Dynamic Quantile test, and Diebold-Marino test were used. Empirical results of the study indicate that in the volatile market periods, overall quality of forecasts deteriorates for all models to a varying degree. However, the GARCH(1,1)-st with external regressors is considered the most efficient and robust model due to its ability to capture stylized facts of data distribution. Exogenous variables are worth considering but their contribution to performance improvement may be model and market dependent.

Keywords: risk management, market risk, Value-at-risk, GARCH, Historical Simulation, Risk Metrics®, risk modelling, benchmark, model quality assessment

JEL codes: C51, C52, C53, C58, G15, G32

1 INTRODUCTION

Market risk, especially in periods of crisis, remains at the center of attention for most of the investors and financial institutions. There are multiple measures allowing for estimation of this type of risk, among which the Value-at-Risk (VaR) is one of the most commonly used (Manganelli, Engle (2001)). Value-at-Risk accounts for the maximum loss over specified time horizon with a given confidence interval. The idea behind this approach is to represent complex market risk using one number which is understandable not only for financial professionals (Linsmeier, Pearson (2000)).

The aim of the following study is to test the comparative performance of models used for Value-at-Risk estimations in changing market volatility regimes. Each market crunch generates big price movements in both directions. The newest crisis originated from the spread of virus causing Covid-19 disease, and was probably the most severe market setback ever observed (Sansa (2020)). However, this environment creates a perfect opportunity to test the effectiveness of various approaches to Value-at-Risk modelling as it is crucial for the models to be resilient, robust and reliable regardless of the occurring market volatility.

In order to distinguish three volatility regimes (low, medium, and high) authors decide to set one objective measure which can be applied into all international markets. The exemplary candidate is the Chicago Board Options Exchange's Volatility Index (VIX). Its value reflects the volatility of call and put options on S&P 500 stock index with average time to expiration equal to 30 days. The attractiveness of this gauge as a market volatility proxy stems from its predictive character and the dominant role of the US stock market in the world (Whaley (1995)). *Fear gauge* shows the level of price volatility implied by the option market, not the actual or historical volatility of the index itself. It may be understood as a near term forecast. VIX is directly linked to the S&P 500 index, which groups the biggest companies in the US stock market, but can also be a near-term volatility illustration for other developed countries. Market events in the US may generate spillover effects, determining both the direction and dynamic of changes worldwide (Kim, Kim, Lee (2015)).

Regulatory frameworks specify how the Value-at-Risk models should be constructed, and Basel III provides guidance on how to set parameters such as: minimum historical period for the VaR calculations; data frequency; market risk charges or backtesting methods (Sharma (2012)). Although the Basel committee suggests using the Expected Shortfall (ES) in their newest documents, they still provide demanded Value-at-Risk specifications, as the ES is an expected value of returns below the VaR level at a given confidence interval α ($ES_\alpha =$

$E(r \mid r \leq VaR_{\alpha})$. Value-at-Risk still plays an important role in risk assessment and it should be calculated at 99% confidence interval (97.5% for the Expected Shortfall calculations), in sample period should have at least 250 observations (approx. one calendar year) and forecasts should be made for a minimum 10 trading days. Measures of the Value-at-Risk models performance can be divided into three categories: Inadequate Coverage tests (check the mispredictions), Unconditional Coverage tests (check the precision of forecast), and Independence tests (check whether models adapt to volatility clustering phenomenon). Buczyński and Chlebus (2020) find that for the purpose of Value-at-Risk estimation in-sample period should have a length of about 1000 days to receive the best quality of estimations. This assumption is implemented in this study and it complies with statutory requirements for VaR models.

Validation of comparative performance of Value-at-Risk models in various market environments is regularly conducted by researchers. Abad and Benito (2013) show that semi-parametric and parametric models outperform other, non-parametric techniques. Buczyński and Chlebus (2019) prove that conservative GARCH models with standardized student's t distribution perform finest in terms of a balance between predictive power and a total cost function. Opposing conclusions may be drawn from Totić, Bulajić and Vlastelica (2011) who suggest that techniques based on the Extreme Value Theory outstrip GARCH family and Historical Simulation approaches. Recent trends in VaR modelling concentrate on non-parametric and semi-parametric concepts. Ünal (2011) infers that non-parametric approaches may work better with smaller window sizes. Wang, Jiang and He (2019) show the benefits of using Artificial Neural Networks for VaR estimation, underlining their capability of taking into account many additional explanatory variables. From all above one can draw a conclusion that there is no simple answer which model is superior to others. Dynamics of the volatility, distribution of returns, sample length, model parameters and other details are decisive when it comes to VaR forecasting accuracy. From the comprehensive models review made by Abad, Benito and López (2011) one can notice that optimizing model parameters is equally important as selection of the model itself. Nevertheless, rarely occurring extreme events generating big market moves always create an opportunity to compare techniques of Value-at-Risk estimation. The Covid-19 pandemic caused a major disturbance that has affected the entire world's population and all economies across the globe. Due to increasing globalization over recent decades, negative effects of the pandemic are transmitted even faster than before (Bryson, Van (2020)). In this work authors decide to include this fact and check the performance of Value-at-Risk models in six developed markets far and wide the world.

Models considered represent parametric, semi-parametric and non-parametric techniques of VaR estimation. Historical Simulation falls into the second category and serves as a benchmark for other techniques. It's attractiveness stems from relatively easy implementation and the lack of assumptions regarding the distribution of stock returns. In addition to this, GARCH(1,1) models with varied distributions (normal, student's t and skewed student's t) are assessed. Next, all three aforementioned variants are enriched with exogenous regressors in the form of lagged time series of energy commodities volatility as procedure this may improve VaR forecast accuracy (Kambouroudis, McMillan (2016)). The last model is GARCH (1,1) with a Quasi Maximum Likelihood (QML) estimator based on a non-Gaussian density. In total, there are nine models, six sectoral indices and three time periods to compare the effectiveness of Value-at-Risk forecasts.

This study aims to provide answers to the three main research questions identified by the authors. The first two assume checking whether, in the context of extreme market phenomena, the forecasts of VaR models deteriorate, and if so, to what extent for each of the considered models. Authors state hypotheses that the overall forecasting quality of all models will be lower in the volatile sample and that the GARCH(1,1)-st model, due to its high flexibility (ability to capture stylized facts about the data distribution) will be the most robust to changes in volatility. Lastly, researchers examine the hypothesis, whether additional, exogenous regressors in the parametric GARCH models improve the overall quality of the forecasts.

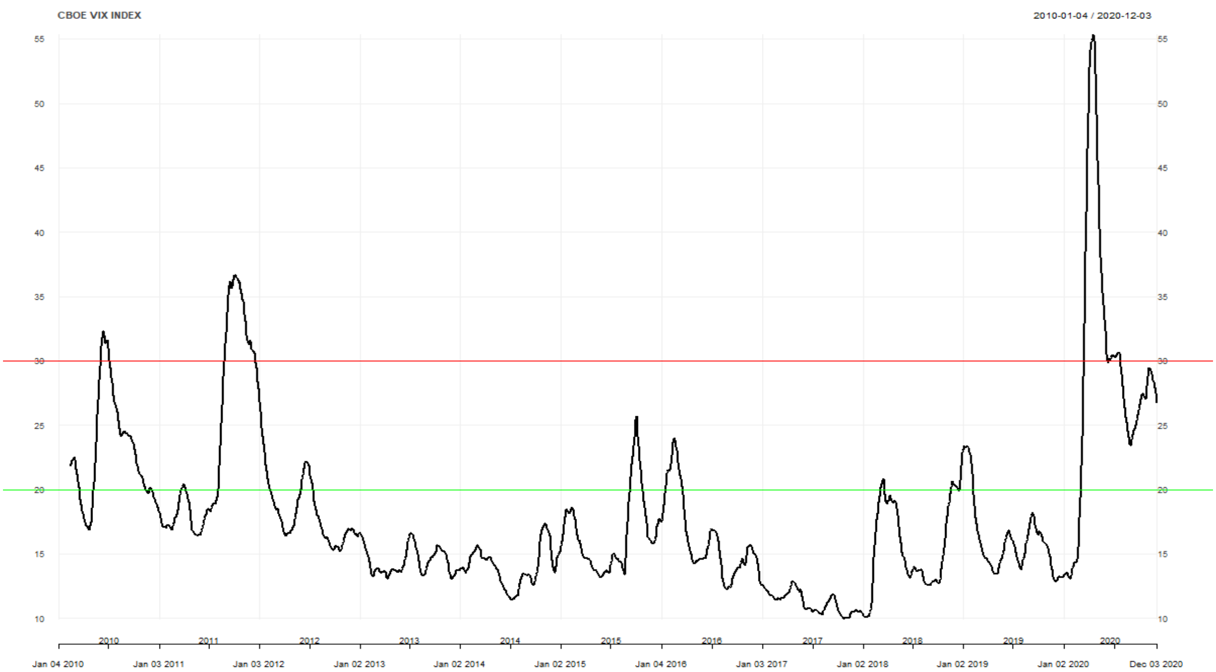
The article is structured as follows: in the next section there is a dataset description. Chapter three contains a detailed methodology review. Subsequent chapter discusses the backtesting framework. Finally, in chapter four, all empirical results are presented and examined. The final part is a short summary and conclusions.

2 DATASET DESCRIPTION

The analysis is conducted on a set of utility sector stock indices from six developed countries across the globe: WIG Energia (Poland), Dow Jones Utilities Average (USA), CAC Utilities (France), Tokyo SE Topix-17 Power & Gas (Japan), S&P ASX 200 Utilities (Australia), and DAX All Utilities (Germany). There are some technical differences between them in terms of number of components, rebalancing frequency or liquidity. However, all belong to the utility sector which groups companies that provide basic amenities, such as water, electricity, heating oil, and natural gas. Selection of this sector was dictated by the relatively high sensitivity to the prices of energy resources which may facilitate the choice of additional exogenous variables.

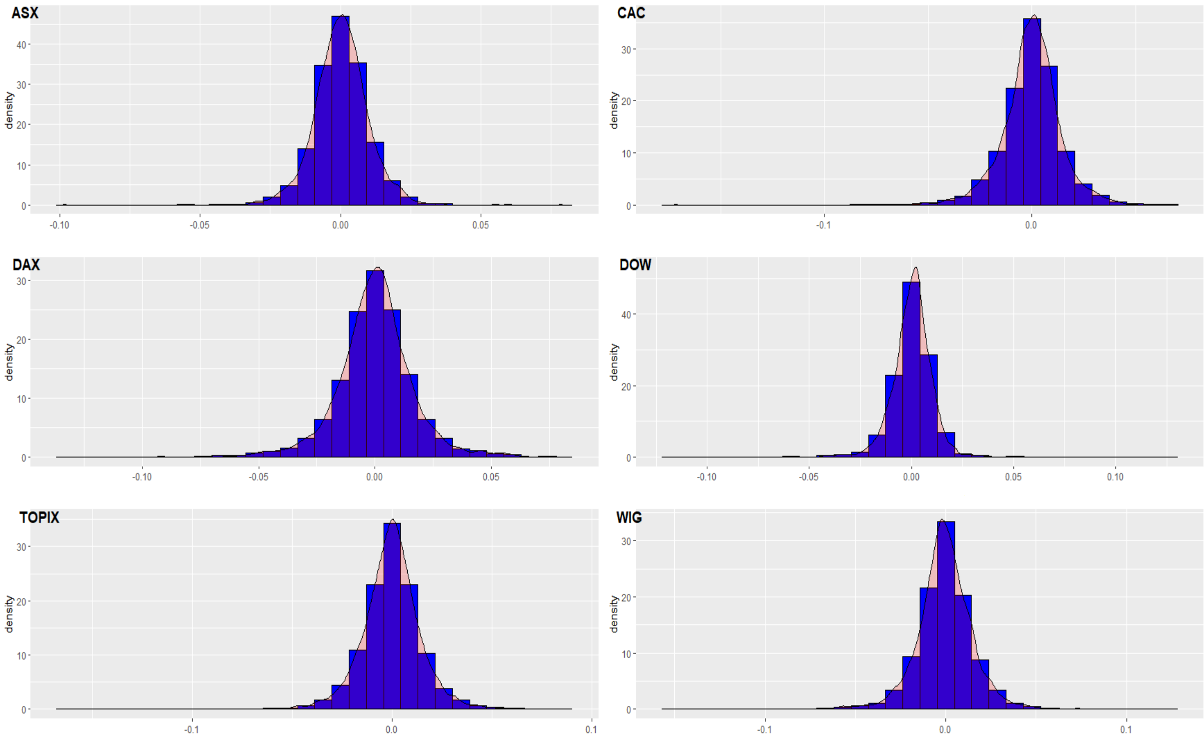
From the last 10 years of data, three samples were distinguished. Each of them contains 1252 daily observations and in-sample and out-of-sample periods are in the proportion of 4:1. In other words, models are built on about four years of data and tested on around one year. As mentioned before, the choice of specific periods was dependent on the level of the VIX index. Values that are greater than 30 indicate high volatility resulting from increased risk and investor's panic behavior. VIX values below 20 generally correspond to stable, stress-free periods in the markets (CBOE VIX White Paper (2019)). Values in between those two thresholds may mark medium volatility. The chart below illustrates the realization of the VIX index over the last ten years. Sample I includes data from Jun. 2012 to Nov. 2016 and all models are fitted on tranquil market surroundings and tested in medium volatility times. Period II is from Jan. 2010 to Aug. 2014 and models are built on relatively high volatility while the testing environment is calm. The last one (Sample no. 3) is between Feb. 2016 and Nov. 2020 and the out-of-sample period covers a large increase of market volatility caused by the Covid-19 pandemic. This specification of the experiment can transfer to a comprehensive assessment of the behavior of the models as multiple possible market scenarios are included.

Figure 1: CBOE VIX index realizations between Jan. 2010 and Dec. 2020. Values above the red horizontal line indicate an increased global volatility and values below the green line indicate a calm period.



For all stock indices, the unconditional mean of daily log-returns is close to zero and the maximum and minimum values are between −16.1% and 12.8%. Record losses for all indices exceed 10% and occur during the Covid-19 pandemic advancement. The skewness statistics are negative for the S&P ASX 200 Utilities (−0.24), CAC Utilities (−0.89), DAX All Utilities (−0.39), Dow Jones Utilities Average (−0.34), Tokyo SE Topix-17 Power & Gas (−0.38) and positive only in case of WIG Energia (0.31). Median is greater than mean for every time series except WIG Energia which, in pair with negative skewness, may indicate possible fat left side tails. All distributions are leptokurtic (kurtosis > 3) so extreme values may occur more often than in the normal distribution and thus present higher market risk. Formal tests for normality of distributions including Shapiro-Wilk, Kolmogorov-Smirnov and Jarque-Bera reject the null hypothesis of normality of all time series at the 5% significance level.

Figure 2: The histograms and empirical density (red line) of returns of stocks indices.



In order to provide selected models with additional, exogenous variables, four lagged time series representing prices of future contracts for commodities, that are widely used in a process of energy delivery, are included. Futures are tradable commitment contracts that are written on an underlying asset and are offered by an organized futures exchange. They are commonly used for risk hedging or speculation on asset price movements (Parameswaran, (2011)). In this work, one-day lagged prices of future contracts for heating oil, crude oil, natural gas, and coal available on CME Group exchanges are used for estimation of Value-at-Risk models.

3 METHODOLOGY

The focus of this paper is the measurement of the effectiveness of Value-at-Risk calculation methods. VaR is a probability-based market risk measure that attempts to quantify the losses generated by moves of financial market variables, which has been receiving public endorsement since 1990s. (Jorion, (2001)) It is accessible and may be readily interpreted by non-financial specialists as it illustrates complex market risk with one simple number without decomposing it. VaR is defined as a maximum possible loss in a given time frame for a particular stock or portfolio, at a given level of confidence. Formally, the concept may be expressed by the mathematical formula below:

$$P(r_t < VaR_\alpha(t)|\Omega_{t-1}) = \alpha; \qquad [1]$$

where r_t denotes financial returns at time t, α is the confidence interval, t is time interval and Ω_{t-1} stands for information set available in period t-1.

3.1 Historical Simulation

The first model considered in this study is the Historical Simulation which is one of the most commonly applicable techniques for Value-at-Risk calculations and may act as a reference point for the rest of the methods (Butler, Schachter (1996)). It belongs to the non-parametric family of models which tries to measure the VaR without making strong assumptions about the distribution of financial returns. All approaches of this kind are based on a conviction that the near-future market risk may be successfully estimated using past realizations of market losses. As stated by Down (2002), the relative easiness of implementation and the ability to include specific features of a particular time series as wide tails, skewness and other unique characteristics, are the main advantages of this approach. On the other hand the weakness of the model stems from the relative slowness of adaptation to the new market events. As quintiles demand extreme observations to change their values, the scenario of constant overestimation of risk, when the model is trained in a volatile period and tested in a calm environment, may be realized. Historical Simulation is also relatively sensitive to in-sample window sizes and may provide more consistent estimations with a longer time horizon included for built. This method is often referred to as *Histogram approach* because the actual Value-at-Risk is a α quantile of the empirical distribution of historical returns:

$$VaR_{\alpha}(t) = q_{\alpha}; \quad [2]$$

where q_{α} denotes α quantile of the in-sample distribution of returns.

3.2 Risk Metrics®

The Risk Metrics® model is a concept put forward by J.P. Morgan investment bank in the late 80s and intended to be available to all market participants. It belongs to the parametric family of models and is based on the assumption that the daily changes in market variables are normally distributed (Risk Metrics Technical Document (1996)). However, this simplistic assumption is not firmly grounded in reality. Stock returns and market-model residuals can convey significant evidence of nonnormality in both the marginal and joint distributions of these variables (Richardson and Smith (1993)). Aforementioned formal tests applied to data used in this study show that the hypothesis of normality of returns distributions must be rejected for all indices. The existence of fat tails in data may lead to a significant underestimation of risk. Thus Risk Metrics® may not provide accurate and efficient Value-at-Risk forecasts. It does not change the fact that the model is widely used by financial institutions and academics due to its relative simplicity of implementation and established position in the history of risk management. Moreover, when there is no sufficiently long set of returns to use a non-parametric model such as Historical Simulation or there is a need to closely model an ideal distribution, Risk Metrics® finds its application. Formally, VaR from the model can be depicted using the formula below:

$$VaR_{\alpha}(t) = \bar{r}_t + \sqrt{\delta_t} \cdot z_{\alpha}; \quad [3]$$

where $\bar{r}_t$ is the mean return at time t, δ_t represents the variance at time t, and z_{α} denotes the α-quantile of the standard normal distribution.

3.3 GARCH models

The Garch family of models was introduced by Engle (1982) and Bollerslev (1986) and was intended for market volatility modelling. It was important to create a concept being able to capture a volatility clustering phenomenon which is thought to be an immanent feature of financial data (Orhan and Köksal (2011)). Over the years, many variations of the GARCH models have emerged, allowing for adjusting the set of parameters to find a better fitness for empirically observed market conditions, and became the vital approach to VaR modelling. In this study, authors settle on checking the performance of the GARCH(1,1) model with various assumptions about the distribution of underlying instruments (normal, t-student's, and skewed t-student's). In addition to this, standard GARCH (1,1) is fitted using an unbiased QML estimator. First three models are also present with external regressors to include in the conditional variance equation. Formally, the standard GARCH process can be described by the following equations:

$$r_t|\Omega_{t-1} \sim IID(\mu_t, h_t), \quad [4]$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}; \quad [5]$$

where IID (μ_t, h_t) is an identical and independent distribution with conditional mean (μ_t) and conditional variance (h_t), r_t is a financial return at time t, α_0 is a constant, $\alpha_1$is a parameter corresponding to the stochastic process realization at time t -1, ε_{t-1}^2 represents squared random error in the period t -1, and $\beta_1$stands for parameter corresponding to the conditional variance realization at time t -1.

Thus, having two previous equations, Value at Risk can be represented by the formula below:

$$VaR_\alpha(t) = \widehat{\mu_t} + \sqrt{\widehat{h_t}} \cdot q_\alpha; \quad [6]$$

where $\widehat{\mu_t}$ is a forecast of the conditional mean at time t, $\widehat{h_t}$ is a forecast of the conditional variance at the time t, and q_α is a α quantile of the assumed underlying distribution of returns.

The necessity of testing multiple assumptions about the data distribution is indicated by many researchers including Angelidis, Benos and Degiannakis (2004). They find that the normal distribution assumption provides the least robust outcomes as it misspecifies some stylized facts about returns on indices. It is also shown that this model variant is more sensitive to the sample size. In order to capture the leptokurticism of rates of return, VaR estimates using GARCH are performed with both t-students' and skewed t-students' distributions of residuals too.

GARCH methodology allows for including exogenous regressors in the form of other financial data time series which are interconnected with the base instrument being modelled. The idea behind this process is to provide additional information which may help to explain the behavior of financial data, and in this study, provide more accurate Value-at-Risk forecasts (Nana, Korn and Erlwein-Sayer (2013)). Exogenous variables may be included in both conditional mean and conditional variation equations of the GARCH model. The selection of a place of input should stem from economic intuition as additional regressors in the mean component generate different fitted and predicted values of μ_t but the same variance forecasts around the different points and vice-versa. In this paper authors believe that events in the energy resources market affect the volatility of utility stocks returns and hence should be included in the conditional variation equation which is formally expressed by the formula below (Duro (2020)):

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \sum_{i=1}^{m} \vartheta_i v_{it}; \qquad [7]$$

where the additional component $\sum_{i=1}^{m} \vartheta_i v_{it}$ represents the m possible number of external regressors v_i at time t.

QML-GARCH (1,1) model is presented in this study to answer the problem of financial data not being normally distributed. The model states that the correct dynamic form is fitted but the innovations are incorrectly assumed to be Gaussian. Here, the likelihood is treated as an objective function to be maximized rather than a proper likelihood. The correction of empirical error procedure was proposed by Engle and Manganelli (1999), and it is based on the estimation of the conditional variance using GARCH (1,1) model, and subsequently finding Value-at-Risk figures as the empirical distribution quantile of the standardized residuals of this model. The robustness of the QML estimator in the context of the moment's properties and stability of the variables is a abundantly attractive feature as there is no need to perform a preliminary analysis of a given covariate before estimating the GARCH model (Han and Kristensen (2015)).

4 BACKTESTING FRAMEWORK

Assessing the comparative performance of each model must be always set on solid foundations. Nobody should draw any conclusions in terms of model accuracy based on its level of sophistication, theoretical features or rationality of assumptions. The only way to check the comprehensive quality of Value-at-Risk forecasts is to conduct a complete and versatile backtesting of models. Quote from Aaron Brown, prominent risk manager and financial knowledge propagator, may illustrate this thought as he used to say: *"VaR is only as good as its backtest. When someone shows me a VaR number, I don't ask how it is computed, I ask to see the backtest."* The VaR models are eminently attractive in terms of comparisons, because the forecasts are expressed with a single digit, which enables a quantitative assessment.

There are multiple approaches to VaR backtesting and those well-established frameworks can be broken down into three major categories. First class consists of statistics which compare Value-at-Risk performance with realized P&Ls' of the underlying instruments. Coverage tests are another group of backtesting methods intended to check whether the number of exceedances is consistent with the theoretical confidence interval of a specified VaR model. Widely present in the literature of subject Kupiec, Christoffersen and Dynamic Quantile tests fall into this category. The last group includes indicators constructed to check which of Value-at-Risk forecasts are statistically better than others (e.g. Diebold-Marino test).

The excess ratio is a useful figure illustrating what is the percentage of VaR violations. In this study, all daily breaches of returns over Value-at-Risk forecasts are summed up and divided by a total number of days in the testing sample. The excess ratio can then be described as the percentage of model failure, and for the correctly foreseeing model, should be below or equal the significance level at which VaR was calculated. Formally, this statistic may be expressed by the formula below:

$$\mathrm{ER} = \frac{\sum_{t=1}^{n} 1_{r_t} < VaR_\alpha(t)}{n}; \qquad [8]$$

where n is the number of periods in the test sample and $1_{r_t} < VaR_\alpha(t)$ represents the number of daily PnLs' that breach the VaR forecasts.

Another testing approach presented in this study is the Kupiec proportion of failure (PF) test which measures, similar to excess ratio, whether the number of exceptions is consistent with the confidence level of VaR forecast. Kupiec (1995) showed that if the probability of a VaR violation is constant, the overall number of breaches follows the binomial distribution. The test statistics is much better in terms of model quality assessment than the aforementioned excess ratio as it penalizes the upward and downward deviations from VaR forecasts. One can imagine a model which is constantly overestimating the risk so the excess ratio is zero but the overall quality of forecast is also at a very low level. The null hypothesis of the test is that α theoretical excess ratio (confidence level α of VaR forecast) is equal to observed one ($\hat{\alpha}$): $H_o: \alpha = \hat{\alpha} = ER$. The test statistic follows an asymptotic chi-squared distribution with one degree of freedom and is represented by the formula below:

$$LR_{UC} = 2 \ln \left(\frac{1-\hat{\alpha}}{1-\alpha}\right)^{n-n_e} \left(\frac{\hat{\alpha}}{\alpha}\right)^{n_e} \sim \chi^2; \qquad [9]$$

where α stands for theoretical excess ratio, $\hat{\alpha}$ is the empirical excess ratio, n accounts for the number of periods forecasted, and n_e is the number of VaR breaches.

An alternative VaR testing method developed a few years after the Kupiec test which, in a way, is its extension, is Chrisoffersen (1998) test. This Markov chain involving approach enriches Kupiec's test statistic with a separate statistic accounting for an independence of VaR transgressions. Additionally to the correct rate of coverage, the test scrutinizes whether the probability of an exception on one day depends on the outcome of the previous day (Nieppola (2009)). If the measure accurately predicts the market risk of the underlying instrument then the chance of breaching VaR forecast on day t should be independent of whether the VaR on day t-1 was exceeded or not. Similarly to the previous technique, test statistic comes from an asymptotic chi-squared distribution but this time, with two degrees of freedom and is described by the formula below:

$$LR_{CC} = LR_{UC} + LR_{IND} \sim \chi^2; \qquad [10]$$

where LR_{UC} is the Kupiec test statistic of the unconditional coverage, and LR_{IND} stands for the Christoffersen statistic of VaR forecast independence.

The second component (LR_{IND}) is responsible for testing the null hypothesis of the independence of VaR transgressions against an alternative hypothesis of them being characterized by the first order Markov chain. In general, Markov chain is a stochastic process which illustrates the sequence of possible events in which the probability of each event depends only on the state attained in the previous event (Oxford Dictionary (2017)). This sequence in relation to the test can be described by the following formula:

$$LR_{IND} = 2\left(ln((1-\pi_{01})^{N_{00}}\pi_{01}{}^{N_{01}}(1-\pi_{11})^{N_{10}}\pi_{11}{}^{N_{11}}) - ln((1-\hat{\alpha})^{N_{00}+N_{01}}\hat{\alpha}^{N_{00}+N_{01}})\right);[11]$$

where N_{ij} is the number of observations, where the j state (two possible variants: 0 – not an exceedance, 1 – exceedance) occurred after observing the i state in the previous period; $\pi_{01} = \frac{N_{01}}{N_{01}+N_{10}}$ represents the probability of exceedance when the VaR forecast was not beached in the previous period; $\pi_{11} = \frac{N_{11}}{N_{11}+N_{10}}$ stands for the probability of exceedance when there was an exceedance in the previous period.

Christoffersen test has two major flaws. Due to its form it cannot detect dependencies of order higher than one and it considers only past realizations of the time series which does not allow for inclusion of exogenous variables. A concept which may be the solution to those issues was presented by Engle and Manganelli (2004) who have developed the Dynamic Quantile test. This backtesting method uses a linear regression model which links VaR violations at time t with the past breaches in order to test the conditional efficiency hypothesis. The main goal of this test is to examine whether there is an autocorrelation among the Value-at-Risk exceedances and whether the total number of transgressions is equal to the confidence level of VaR forecasts. The null hypothesis of the DQ test assumes that all parameters in the regression equation are equal to zero. The joint nullity of coefficients means that the VaR exceedance at time t is uncorrelated with the past exceedances. An alternative hypothesis is that at least one of the parameters of the regression equation is significantly different from zero. The regression equation is:

$$I_t = \beta_0 + \sum_{i=1}^{p} \beta_i I_{t-1} + \sum_{j=1}^{q} \mu_j X_j + \varepsilon_t; \qquad [12]$$

where X_j represents all explanatory variables in the information set, p is the number of lags of the dependent variable, q is the number of lags of the independent variable, and $I_t = \begin{cases} 1-\alpha, & if\ r_t < VaR_t^{\alpha} \\ -\alpha, & else \end{cases}$.

The null hypothesis is therefore:

$$H_0: \beta_1 = \beta = \cdots = \beta_p = \mu_1 = \cdots = \mu_q = 0; \qquad [13]$$

A form of the linear regression permits an inclusion of additional, explanatory variables to the information set in order to check the their influence on the distribution of exceedances. Dynamic Quantile test statistic follows the chi-squared distribution with $p + q$ degrees of freedom and is described by the formula below:

$$\mathrm{DQ} = \frac{I_t' \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}' I_t}{\alpha(1-\alpha)} \sim \chi^2; \qquad [14]$$

The last measure presented in this study is the Diebold-Mariano test which compares the predictive accuracy of two Value-at-Risk forecasts. It differs from previous methods because it involves two time series of predicted VaRs' to produce an output. In empirical applications it is often the case that two or more time series models are available for forecasting a particular variable of interest. Using the DM test, a researcher can compare techniques directly and categorize them from best to worst. The null hypothesis is that the two predictions have the same accuracy, and the alternative one states that two models accuracy is at the different level. The test statistics of the DM test comes from the standard normal distribution and is as follows:

$$DM = \frac{\mu_d}{\sqrt{\frac{1}{n}\sigma_d^2}} \sim N(0,1); \qquad [15]$$

where d is a difference of squared differences between the realized rate of return, and the predicted level of VaR of two models $((r_1 - VaR_1)^2 - (r_2 - VaR_2)^2)$, μ_d stands for an average of a process d, σ_d^2 is a variation of process d, and n accounts for a total number of forecasts. $H_0: E(d) = 0$.

5 RESULTS

Obtained results are presented in tables 1-4. All computations were conducted in R programming environment. Tables one and two summarize the results of Kupiec, Christoffersen, and Dynamic-Quantile tests described in the previous section as well as the total number of exceedances and excess ratio for all models and indices in three tested periods. In the two subsequent records, the outcomes of the Diebold-Mariano test are broken down into individual indices for each testing sample. All statistical tests are performed with a 95% confidence interval.

In the testing period I, the best performing Value-at-Risk model (calculated at the 99% confidence interval) in terms of forecasting accuracy (Table 1.) was the GARCH(1,1) with a skewed t-student's distribution and four exogenous explanatory variables in the form of commodity futures returns. The model, as the only one, gives no reason to reject the null hypotheses of the conditional coverage and unconditional coverage tests in case of all six tested indices (p-value > 0.05). Moreover, the excess ratio stays in the fairly narrow range between 1.58% and 2.37% which indicates that the model is accurately forecasting market volatility for all markets. Its relative strength may be attributed to the external regressors which provide additional information to a model, increasing the overall precision. On the other hand, the worst behavior belongs to the Risk Metrics® approach. The model is underestimating the risk of downward market moves and has the highest average excess ratio of 2.96% - almost three times higher than the theoretical threshold of 1%. As a result, it was able to pass Kupiec and Christoffersen tests only for Topix-17 Power & Gas and ASX 200 Utilities. The presence of fat tails in the Sample I may lead to mispredictions of models which assume normality of data distributions. In general, most of the models included in this comparison were providing accurate forecasts. Historical Simulation, being one of the simplest techniques, also scores really well as its ability to reflect empirically observed, non-standard features of instrument returns may result in a better fitness of the models form. Furthermore, HS approach does not overshoot the risk due to its relative slowness of adaptation (low volatility in IS period and increasing in OOS period). Diebold-Mariano test results shown in Table 3. indicate that at the significance level of 5%, GARCH(1,1) models with and without exogenous regressors produce the VaR forecasts of the same accuracy for ASX and Dow indices (p-value > 0.05). For ASX and TOPIX records, Historical Simulation was equally good as GARCH-T, GARCH-ST, GARCH-T-CMD, and GARCH ST-CMD.

For the less conservative VaR predictions at 97.5% confidence level, the best model is GARCH(1,1) with skewed student's t distribution. This time, exogenous regressors do not bring the improvement of results but Diebold-Mariano test suggests that both GARCH-ST and GARCH-ST-CMD generate the same level of forecast accuracy for 5 indices. The model failed only once in terms of coverage tests. In contrary to the 99% VaR forecasts, Historical Simulation method exhibit a comparatively poor performance as it was unable to generate a robust forecast in 4 out of 6 times based on conditional/unconditional coverage tests. It was underestimating the market risk of an underlying instruments leading to the large number of VaR transgressions. GARCH(1,1) with the QML estimator shows reasonably good outcomes as it fails mainly in terms of Dynamic-Quantile test which indicates that VaR breaches may be intercorrelated. The worst performing model is again the Risk Metrics®, recording the highest number of exceedances of all models.

Period II is characterized by the decreasing volatility – models are trained during a moderately turbulent window and tested in a calm environment. This sort of specification should requite models that adapt fast to new market events to prevent overestimation of risk. For both Value-at-Risk confidence levels, this assumption is visible as parametric models from GARCH family generate highly accurate forecasts while non-parametric HS approach exhibits a relatively weak performance. Even though all types of GARCH forecasts do not give a reason to reject the null hypothesis of Kupiec, Christoffersen and Dynamic-Quantile tests at 95% confidence level (Table 1.), they do not evince the same level of accuracy which is marked by the DM test outcomes (Table 3.). For TOPIX and ASX indices there is no difference between models with and without exogenous variables but for other crosses the null of same accuracy of models must be rejected. Thus, to determine the best model, a standard deviation of excess ratio from the confidence level of VaR forecast is calculated. Finally, the best performing 99% model is GARCH(1,1) with normal standard distribution, and for 97.5% level is the GARCH(1,1) with skewed student's t distribution using external regressors. The worst models in Period II for both confidence intervals of VaR forecasts are the Riskmetrics® (1) and Historical Simulation (2) due to respectively: fallacious distribution model (1), and slow adaptation of modelled quantiles (2).

Period III is the most challenging one in the context of VaR assessment because the IS period is relatively calm but in the OOS time interval there is an upsurge of volatility caused by the Covid-19 pandemic. The results of this study correspond to this initial thought as overall results are deficient. Among the Value-at-Risk forecasts obtained at the confidence level of 99%, the GARCH-QML, GARCH-T, GARCH-ST, and GARCH-ST-CMD models should be distinguished. The first of listed approaches which involves QML estimator was performing superior to others in terms of coverage tests – it was able to pass Kupiec and Christoffersen trials for 5 out of 6 indices at the significance interval of 5%. Nevertheless, the null hypothesis of DQ test must be rejected for all tested instruments in case of the GARCH-QML. This indicates that there is an autocorrelation among VaR transgressions. GARCH-QML exhibit a strong underestimation of the market risk of ASX index, noting an excess ratio of 9.52% (global maximum). Other aforementioned techniques were able to forecast accurately (coverage tests p-values > 0.05) for 4 of 6 markets and were providing slightly less deviated excess ratio than the GARCH-QML. Diebold Mariano test shows that the GARCH models and their counterparts with additional variables generated equally accurate forecasts at 95% confidence level for half of the instruments. The worst performing model, not being able to meet any theoretical requirements of tests was the Riskmetrics®. Fat-tailed distributions of the volatile Covid-19 era are deceiving the model which is based on normal standard distribution assumption. In case of VaR forecasts at 97.5% confidence level, outcomes are similar with the only difference that the GARCH-QML fared worse compared to the 99% level. GARCH-ST-CMD and GARCH-ST took *ex-aqueo* first place, and Riskmetrics® and HS methods seemed to be inaccurate.

Table 1: Statistical benchmarking results for the considered Value-at-Risk models at 99% confidence level in each considered time frame. Green fields indicate p-values greater than 5% for Kupiec, Christoffersen and Dynamic-Quantile tests.

VaR 99% confidence level		Period I					Period II					Period III				
		No. exceedances	Excess ratio (%)	Kupiec	Christoffersen	Dynamic Quantile	No. exceedances	Excess ratio (%)	Kupiec	Christoffersen	Dynamic Quantile	No. exceedances	Excess ratio (%)	Kupiec	Christoffersen	Dynamic Quantile
HISTORICAL	ASX	3	1.19%	77.30%	92.52%	52.81%	0	0.00%	2.41%	7.87%	92.60%	8	3.17%	0.58%	1.11%	0.00%
	CAC	7	2.77%	2.03%	5.54%	10.19%	0	0.00%	2.41%	7.87%	92.60%	11	4.37%	0.01%	0.01%	0.00%
	DAX	4	1.58%	39.18%	64.98%	8.11%	0	0.00%	2.41%	7.87%	92.60%	10	3.97%	0.04%	0.02%	0.00%
	DOW	4	1.58%	39.18%	66.04%	6.36%	1	0.40%	27.08%	54.31%	98.97%	11	4.37%	0.01%	0.00%	0.00%
	TOPIX	1	0.40%	27.08%	54.31%	99.39%	1	0.40%	77.30%	6.29%	40.82%	6	2.38%	6.25%	5.19%	0.00%
	WIG	6	2.37%	6.25%	15.23%	0.66%	2	0.79%	77.30%	92.52%	96.89%	11	4.37%	0.01%	0.01%	0.00%
RISKMETRICS	ASX	3	1.19%	77.30%	92.52%	80.28%	0	0.00%	2.41%	7.87%	92.60%	11	4.37%	0.01%	0.03%	0.00%
	CAC	10	3.95%	0.04%	0.02%	0.00%	0	0.00%	2.41%	7.87%	92.60%	17	6.75%	0.00%	0.00%	0.00%
	DAX	10	3.95%	0.04%	0.12%	0.00%	0	0.00%	2.41%	7.87%	92.60%	11	4.37%	0.01%	0.01%	0.00%
	DOW	9	3.56%	0.15%	0.05%	0.00%	2	0.79%	72.81%	92.64%	97.05%	20	7.94%	0.00%	0.00%	0.00%
	TOPIX	5	1.98%	16.85%	7.91%	0.06%	3	1.19%	77.30%	6.29%	0.00%	7	2.78%	2.03%	0.23%	0.00%
	WIG	8	3.16%	0.58%	1.72%	0.19%	3	1.19%	77.30%	92.52%	93.66%	12	4.76%	0.00%	0.00%	0.00%
GARCH-N	ASX	4	1.58%	39.18%	64.98%	39.09%	1	0.40%	27.08%	54.31%	98.50%	9	3.57%	0.15%	0.39%	0.00%
	CAC	9	3.56%	0.15%	0.47%	0.08%	2	0.79%	72.81%	92.64%	44.33%	9	3.57%	0.15%	0.39%	0.00%
	DAX	5	1.98%	16.85%	35.01%	34.69%	2	0.79%	72.81%	92.64%	96.41%	8	3.17%	0.58%	1.11%	0.00%
	DOW	9	3.56%	0.15%	0.36%	0.00%	3	1.19%	77.30%	92.52%	82.13%	10	3.97%	0.04%	0.11%	0.00%
	TOPIX	5	1.98%	16.85%	35.01%	46.61%	2	0.79%	72.81%	92.64%	100.00%	5	1.98%	16.85%	35.01%	0.00%
	WIG	6	2.37%	6.25%	15.23%	0.44%	3	1.19%	77.30%	92.52%	96.17%	7	2.78%	2.03%	0.23%	0.00%
													0.00%			
GARCH-T	ASX	3	1.19%	77.30%	92.52%	99.05%	1	0.40%	27.08%	54.31%	98.63%	7	2.78%	2.03%	2.66%	0.00%
	CAC	8	3.16%	0.58%	1.72%	0.29%	2	0.79%	72.81%	92.64%	48.03%	6	2.38%	6.25%	5.19%	0.00%
	DAX	5	1.98%	16.85%	35.01%	37.73%	1	0.40%	27.08%	54.31%	97.91%	6	2.38%	6.25%	5.19%	0.00%
	DOW	8	3.16%	0.58%	1.78%	0.05%	3	1.19%	77.30%	92.52%	87.07%	6	2.38%	6.25%	15.23%	0.47%
	TOPIX	5	1.98%	16.85%	35.01%	48.87%	2	0.79%	72.81%	92.64%	99.99%	6	2.38%	6.25%	15.23%	0.00%
	WIG	5	1.98%	16.85%	35.01%	0.29%	2	0.79%	72.81%	92.64%	99.69%	7	2.78%	2.03%	0.23%	0.00%
													0.00%			
GARCH-ST	ASX	3	1.19%	77.30%	92.52%	98.55%	1	0.40%	27.08%	54.31%	98.73%	7	2.78%	2.03%	2.66%	0.00%
	CAC	8	3.16%	0.58%	1.72%	0.36%	2	0.79%	72.81%	92.64%	46.06%	4	1.59%	39.18%	8.79%	0.00%
	DAX	5	1.98%	16.85%	35.01%	36.74%	1	0.40%	27.08%	54.31%	98.04%	4	1.59%	39.18%	64.98%	89.38%
	DOW	5	1.98%	16.85%	35.73%	85.38%	2	0.79%	72.81%	92.64%	81.73%	4	1.59%	39.18%	64.98%	0.02%
	TOPIX	5	1.98%	16.85%	35.01%	44.55%	2	0.79%	72.81%	92.64%	99.99%	6	2.38%	6.25%	15.23%	0.00%
	WIG	5	1.98%	16.85%	35.01%	0.28%	2	0.79%	72.81%	92.64%	99.75%	8	3.17%	0.58%	0.13%	0.00%
GARCH-QML	ASX	4	1.58%	39.18%	64.98%	45.54%	1	0.40%	27.08%	54.31%	98.63%	24	9.52%	0.00%	0.00%	0.00%
	CAC	6	2.37%	6.25%	15.23%	35.23%	2	0.79%	72.81%	92.64%	39.91%	5	1.98%	16.85%	7.91%	0.00%
	DAX	5	1.98%	16.85%	35.01%	31.39%	1	0.40%	27.08%	54.31%	98.50%	6	2.38%	6.25%	5.19%	0.00%
	DOW	8	3.16%	0.58%	1.78%	0.00%	3	1.19%	77.30%	92.52%	77.70%	4	1.59%	39.18%	64.98%	0.02%
	TOPIX	4	1.58%	39.18%	64.98%	61.07%	1	0.40%	27.08%	54.31%	99.51%	5	1.98%	16.85%	35.01%	0.00%
	WIG	6	2.37%	6.25%	15.23%	0.33%	2	0.79%	72.81%	92.64%	99.52%	5	1.98%	16.85%	7.91%	0.00%
GARCH-N-CMD	ASX	4	1.58%	39.18%	64.98%	61.16%	1	0.40%	27.08%	54.31%	97.52%	8	3.17%	0.58%	1.11%	0.00%
	CAC	10	3.95%	0.04%	0.11%	0.00%	2	0.79%	72.81%	92.64%	27.39%	7	2.78%	2.03%	2.66%	0.00%
	DAX	5	1.98%	16.85%	35.01%	27.73%	2	0.79%	72.81%	92.64%	94.02%	8	3.17%	0.58%	1.11%	0.00%
	DOW	8	3.16%	0.58%	0.99%	0.01%	4	1.59%	39.18%	64.98%	33.46%	10	3.97%	0.04%	0.11%	0.00%
	TOPIX	5	1.98%	16.85%	35.01%	47.93%	2	0.79%	72.81%	92.64%	99.99%	6	2.38%	6.25%	15.23%	0.00%
	WIG	6	2.37%	6.25%	15.23%	0.41%	3	1.19%	77.30%	92.52%	95.77%	8	3.17%	0.58%	0.13%	0.00%
GARCH-T-CMD	ASX	4	1.58%	39.18%	64.98%	64.24%	1	0.40%	27.08%	54.31%	97.71%	7	2.78%	2.03%	2.66%	0.00%
	CAC	6	2.37%	6.25%	15.23%	38.13%	2	0.79%	72.81%	92.64%	28.34%	5	1.98%	16.85%	7.91%	0.00%
	DAX	5	1.98%	16.85%	35.01%	32.38%	1	0.40%	27.08%	54.31%	97.46%	4	1.59%	39.18%	64.98%	88.88%
	DOW	8	3.16%	0.58%	1.78%	0.08%	3	1.19%	77.30%	92.52%	86.63%	6	2.38%	6.25%	15.23%	0.00%
	TOPIX	5	1.98%	16.85%	35.01%	49.47%	2	0.79%	72.81%	92.64%	99.97%	5	1.98%	16.85%	35.01%	0.00%
	WIG	6	2.37%	6.25%	15.23%	0.47%	2	0.79%	72.81%	92.64%	99.58%	7	2.78%	2.03%	0.23%	0.00%
GARCH-ST-CMD	ASX	4	1.58%	39.18%	64.98%	61.90%	1	0.40%	27.08%	54.31%	97.84%	7	2.78%	2.03%	2.66%	0.00%
	CAC	6	2.37%	6.25%	15.23%	39.25%	2	0.79%	72.81%	92.64%	27.85%	5	1.98%	16.85%	7.91%	0.00%
	DAX	5	1.98%	16.85%	35.01%	31.75%	1	0.40%	27.08%	54.31%	97.62%	4	1.59%	39.18%	64.98%	91.92%
	DOW	5	1.98%	16.85%	35.73%	84.63%	2	0.79%	72.81%	92.64%	78.42%	4	1.59%	39.18%	64.98%	0.02%
	TOPIX	5	1.98%	16.85%	35.01%	44.92%	2	0.79%	72.81%	92.64%	99.97%	6	2.38%	6.25%	15.23%	0.00%
	WIG	5	1.98%	16.85%	35.01%	0.30%	2	0.79%	72.81%	92.64%	99.64%	8	3.17%	0.58%	0.13%	0.00%

Table 2: Statistical benchmarking results for the considered Value-at-Risk models at 97.5% confidence level in each considered time frame. Green fields indicate p-values greater than 5% for Kupiec, Christoffersen and Dynamic-Quantile tests.

VaR 97.5% confidence level		Period I					Period II					Period III				
		No. exceedances	Excess ratio (%)	Kupiec	Christoffersen	Dynamic Quantile	No. exceedances	Excess ratio (%)	Kupiec	Christoffersen	Dynamic Quantile	No. exceedances	Excess ratio (%)	Kupiec	Christoffersen	Dynamic Quantile
HISTORICAL	ASX	9	3.57%	31.06%	35.85%	33.60%	2	0.79%	4.24%	12.54%	86.47%	15	5.95%	0.29%	1.18%	0.00%
	CAC	14	5.56%	0.75%	0.30%	0.00%	0	0.00%	0.03%	0.17%	49.56%	21	8.33%	0.00%	0.00%	0.00%
	DAX	11	4.37%	1.83%	2.26%	0.00%	2	0.79%	4.24%	12.54%	80.64%	14	5.56%	0.75%	1.28%	0.00%
	DOW	12	4.76%	4.16%	3.04%	0.09%	4	1.59%	31.58%	7.67%	10.55%	24	9.52%	0.00%	0.00%	0.00%
	TOPIX	9	3.57%	31.06%	0.42%	0.00%	4	1.59%	31.58%	0.13%	0.00%	10	3.97%	17.19%	5.76%	0.01%
	WIG	12	4.76%	4.16%	6.89%	1.82%	4	1.59%	31.58%	56.69%	94.24%	21	8.33%	0.00%	0.00%	0.00%
RISKMETRICS	ASX	11	4.37%	8.78%	18.32%	30.32%	2	0.79%	4.24%	12.54%	87.23%	13	5.16%	1.83%	5.72%	0.00%
	CAC	16	6.35%	0.11%	0.10%	0.00%	0	0.00%	0.03%	0.17%	49.56%	21	8.33%	0.00%	0.00%	0.00%
	DAX	16	6.35%	0.11%	0.02%	0.00%	2	0.79%	4.24%	12.54%	80.49%	13	5.16%	1.83%	2.26%	0.00%
	DOW	15	5.95%	0.29%	0.15%	0.00%	4	1.59%	31.58%	7.67%	9.33%	22	8.73%	0.00%	0.00%	0.00%
	TOPIX	9	3.57%	31.06%	0.42%	0.00%	4	1.59%	31.58%	0.13%	0.00%	10	3.97%	17.19%	5.76%	0.01%
	WIG	13	5.16%	1.83%	3.05%	1.10%	4	1.59%	31.58%	56.69%	96.48%	19	7.54%	0.00%	0.01%	0.00%
GARCH-N	ASX	10	3.97%	17.19%	27.42%	9.45%	3	1.19%	13.63%	31.80%	97.05%	11	4.37%	8.78%	18.32%	0.03%
	CAC	11	4.37%	8.78%	14.10%	15.92%	4	1.59%	31.58%	56.69%	63.02%	11	4.37%	8.78%	18.32%	1.38%
	DAX	8	3.17%	51.68%	62.34%	54.40%	5	1.98%	57.99%	77.54%	96.52%	12	4.76%	4.16%	3.53%	0.00%
	DOW	14	5.56%	0.75%	2.67%	0.54%	7	2.78%	78.92%	79.00%	76.90%	12	4.76%	4.16%	10.83%	0.00%
	TOPIX	8	3.17%	51.68%	62.34%	70.57%	3	1.19%	13.63%	31.80%	96.20%	8	3.17%	51.68%	62.34%	37.03%
	WIG	10	3.97%	17.19%	26.01%	0.62%	4	1.59%	31.58%	56.69%	96.48%	11	4.37%	8.78%	4.08%	0.03%
GARCH-T	ASX	10	3.97%	17.19%	27.42%	9.91%	3	1.19%	13.63%	31.80%	97.05%	11	4.37%	8.78%	18.32%	0.03%
	CAC	10	3.97%	17.19%	26.01%	20.06%	4	1.59%	31.58%	56.69%	60.34%	11	4.37%	8.78%	18.32%	1.41%
	DAX	9	3.57%	31.06%	42.85%	22.73%	4	1.59%	31.58%	56.69%	91.80%	12	4.76%	4.16%	3.53%	0.00%
	DOW	13	5.16%	1.83%	3.23%	1.63%	7	2.78%	78.92%	79.00%	77.85%	11	4.37%	8.78%	18.32%	0.00%
	TOPIX	7	2.78%	78.92%	79.00%	85.00%	4	1.59%	31.58%	56.69%	96.95%	7	2.78%	78.92%	79.00%	24.04%
	WIG	10	3.97%	17.19%	26.01%	0.17%	4	1.59%	31.58%	56.69%	97.21%	13	5.16%	1.83%	0.35%	0.00%
GARCH-ST	ASX	10	3.97%	17.19%	27.42%	9.45%	3	1.19%	13.63%	31.80%	97.06%	10	3.97%	17.19%	27.42%	0.01%
	CAC	10	3.97%	17.19%	26.01%	21.89%	4	1.59%	31.58%	56.69%	58.48%	10	3.97%	17.19%	27.42%	0.71%
	DAX	8	3.17%	51.68%	62.34%	56.44%	4	1.59%	31.58%	56.69%	91.88%	12	4.76%	4.16%	3.53%	0.00%
	DOW	12	4.76%	4.16%	7.25%	5.38%	5	1.98%	57.99%	77.54%	97.24%	11	4.37%	8.78%	18.32%	0.00%
	TOPIX	7	2.78%	78.92%	79.00%	83.04%	4	1.59%	31.58%	56.69%	96.97%	8	3.17%	51.68%	62.34%	32.69%
	WIG	10	3.97%	17.19%	26.01%	0.17%	4	1.59%	31.58%	56.69%	96.80%	14	5.56%	0.75%	0.25%	0.00%
			0.00%													
GARCH-QML	ASX	8	3.17%	51.68%	40.25%	14.44%	3	1.19%	13.63%	31.80%	97.06%	28	11.11%	0.00%	0.00%	0.00%
	CAC	10	3.97%	17.19%	26.01%	27.23%	4	1.59%	31.58%	56.69%	68.44%	9	3.57%	31.06%	35.85%	0.20%
	DAX	8	3.17%	51.68%	62.34%	55.95%	5	1.98%	57.99%	77.54%	96.03%	13	5.16%	1.83%	2.26%	0.01%
	DOW	12	4.76%	4.16%	7.25%	0.17%	5	1.98%	57.99%	77.54%	94.81%	10	3.97%	17.19%	26.01%	1.32%
	TOPIX	9	3.57%	31.06%	42.85%	24.62%	3	1.19%	13.63%	31.80%	96.46%	8	3.17%	51.68%	62.34%	29.33%
	WIG	9	3.57%	31.06%	42.85%	0.31%	4	1.59%	31.58%	56.69%	94.59%	12	4.76%	4.16%	3.04%	0.03%
GARCH-N-CMD	ASX	10	3.97%	17.19%	27.42%	18.58%	5	1.98%	57.99%	77.54%	0.00%	12	4.76%	4.16%	10.83%	0.06%
	CAC	10	3.97%	17.19%	26.01%	31.03%	4	1.59%	31.58%	56.69%	56.39%	11	4.37%	8.78%	18.32%	1.41%
	DAX	10	3.97%	17.19%	26.01%	10.61%	6	2.38%	89.50%	85.64%	89.35%	13	5.16%	1.83%	2.26%	0.01%
	DOW	14	5.56%	0.75%	2.67%	0.44%	7	2.78%	78.92%	79.00%	67.48%	13	5.16%	1.83%	5.72%	0.00%
	TOPIX	9	3.57%	31.06%	42.85%	44.92%	3	1.19%	13.63%	31.80%	95.30%	7	2.78%	78.92%	79.00%	24.01%
	WIG	11	4.37%	8.78%	14.10%	0.00%	5	1.98%	57.99%	77.54%	79.49%	13	5.16%	1.83%	1.99%	0.00%
GARCH-T-CMD	ASX	9	3.57%	31.06%	35.85%	17.32%	4	1.59%	31.58%	56.69%	14.30%	12	4.76%	4.16%	10.83%	0.00%
	CAC	10	3.97%	17.19%	26.01%	33.72%	4	1.59%	31.58%	56.69%	50.04%	11	4.37%	8.78%	18.32%	1.41%
	DAX	8	3.17%	51.68%	62.34%	31.77%	5	1.98%	57.99%	77.54%	95.10%	13	5.16%	1.83%	2.26%	0.01%
	DOW	13	5.16%	1.83%	3.23%	1.40%	7	2.78%	78.92%	79.00%	53.21%	11	4.37%	8.78%	18.32%	0.00%
	TOPIX	9	3.57%	31.06%	42.85%	49.17%	5	1.98%	57.99%	77.54%	87.05%	9	3.57%	31.06%	42.85%	0.74%
	WIG	9	3.57%	31.06%	42.85%	0.41%	5	1.98%	57.99%	77.54%	87.23%	14	5.56%	0.75%	0.25%	0.00%
GARCH-ST-CMD	ASX	9	3.57%	31.06%	35.85%	17.56%	4	1.59%	31.58%	56.69%	14.30%	10	3.97%	17.19%	27.42%	0.01%
	CAC	10	3.97%	17.19%	26.01%	33.87%	4	1.59%	31.58%	56.69%	48.10%	10	3.97%	17.19%	27.42%	0.71%
	DAX	9	3.57%	31.06%	42.85%	18.17%	5	1.98%	57.99%	77.54%	95.44%	12	4.76%	4.16%	3.53%	0.00%
	DOW	13	5.16%	1.83%	3.23%	2.34%	5	1.98%	57.99%	77.54%	94.53%	11	4.37%	8.78%	18.32%	0.00%
	TOPIX	10	3.97%	17.19%	26.01%	13.65%	5	1.98%	57.99%	77.54%	87.26%	7	2.78%	78.92%	79.00%	22.99%
	WIG	9	3.57%	31.06%	42.85%	0.45%	5	1.98%	57.99%	77.54%	84.16%	14	5.56%	0.75%	0.25%	0.00%

Table 3: Diebold–Mariano test results calculated for 99% VaR for each individual index in considered time intervals. Null hypothesis is that the two forecasts are of equal accuracy and alternative hypothesis of different quality. Green fields indicate p-values greater than 5%.

Period I — VaR 99% confidence level

ASX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	57.56%	18.23%	1.28%	0.00%	52.60%	23.13%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	19.59%	0.00%	0.00%
GARCH-T	57.56%	0.00%	0.00%	-	0.00%	0.00%	0.00%	85.49%	27.85%
GARCH-ST	18.23%	0.00%	0.00%	0.00%	-	0.00%	0.00%	32.86%	95.96%
GARCH-QML	1.28%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	19.59%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	52.60%	0.00%	0.00%	85.49%	32.86%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	23.13%	0.00%	0.00%	27.85%	95.96%	0.00%	0.00%	0.00%	-

CAC	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	74.62%	0.00%	0.00%	0.00%	89.60%	0.01%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	74.62%	0.00%	-	0.00%	0.00%	0.00%	8.56%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.09%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.08%
GARCH-N-CMD	89.60%	0.00%	8.56%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.09%	0.00%	0.08%	0.00%	0.00%	-

DAX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.12%	0.00%	0.00%	0.00%	0.48%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.12%	0.00%	-	0.00%	0.00%	0.00%	0.34%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.25%	74.85%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.01%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.48%	0.00%	0.34%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.25%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	74.85%	0.01%	0.00%	0.00%	0.00%	-

DOW	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.61%	0.00%	0.00%	0.00%	0.91%	0.00%	0.00%
GARCH-N	0.00%	0.61%	-	0.00%	0.00%	0.00%	32.74%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	34.33%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	10.26%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.91%	32.74%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	34.33%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	10.26%	0.00%	0.00%	0.00%	-

TOPIX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	33.83%	18.89%	4.46%	0.00%	13.76%	9.52%
Riskmetrics	0.00%	-	9.47%	0.02%	0.03%	0.00%	18.61%	0.07%	0.09%
GARCH-N	0.00%	9.47%	-	0.00%	0.00%	0.00%	3.17%	0.00%	0.00%
GARCH-T	33.83%	0.02%	0.00%	-	0.00%	0.00%	0.00%	0.71%	0.13%
GARCH-ST	18.89%	0.03%	0.00%	0.00%	-	0.00%	0.00%	30.48%	7.21%
GARCH-QML	4.46%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	18.61%	3.17%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	13.76%	0.07%	0.00%	0.71%	30.48%	0.00%	0.00%	-	0.01%
GARCH-ST-CMD	9.52%	0.09%	0.00%	0.13%	7.21%	0.00%	0.00%	0.01%	-

WIG	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.06%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.06%	0.00%	-	0.00%	0.00%	0.00%	4.50%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.04%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	23.40%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.03%	0.00%	4.50%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.04%	23.40%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-

Period II — VaR 99% confidence level

ASX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	74.03%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	89.65%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	65.73%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	74.03%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	89.65%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	65.73%	0.00%	0.00%	0.00%	-

CAC	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	3.00%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.19%	0.00%	0.11%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.19%	0.00%	-	0.00%	0.01%	0.00%
GARCH-N-CMD	0.00%	0.00%	3.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.11%	0.00%	0.01%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-

DAX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	1.66%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.10%	0.00%	0.00%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.49%
GARCH-QML	0.00%	0.00%	0.00%	0.10%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	1.66%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.49%	0.00%	0.00%	0.00%	-

DOW	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.93%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	16.41%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	4.82%	0.00%
GARCH-ST	0.00%	1.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	89.00%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	16.41%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	4.82%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.93%	0.00%	0.00%	89.00%	0.00%	0.00%	0.00%	-

TOPIX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.25%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	89.11%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	49.01%	4.16%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	63.11%	26.96%
GARCH-QML	0.00%	0.25%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	89.11%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	49.01%	63.11%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	4.16%	26.96%	0.00%	0.00%	0.00%	-

WIG	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.77%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	65.34%	74.84%	0.12%	0.00%	91.08%	15.56%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.30%	0.00%	0.00%
GARCH-T	0.00%	65.34%	0.00%	-	0.00%	0.00%	0.00%	0.78%	0.00%
GARCH-ST	0.00%	74.84%	0.00%	0.00%	-	0.00%	0.00%	27.65%	0.00%
GARCH-QML	0.77%	0.12%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	0.30%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	91.08%	0.00%	0.78%	27.65%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	15.56%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-

Period III VaR 99% confidence level

ASX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.01%	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
GARCH-N	0.01%	0.00%	-	30.47%	96.32%	0.00%	6.53%	31.54%	95.79%
GARCH-T	0.00%	0.00%	30.47%	-	0.00%	0.00%	52.56%	60.13%	0.00%
GARCH-ST	0.00%	0.00%	96.32%	0.00%	-	0.00%	75.65%	0.00%	12.14%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.05%	0.01%	6.53%	52.56%	75.65%	0.00%	-	54.16%	68.68%
GARCH-T-CMD	0.00%	0.00%	31.54%	60.13%	0.00%	0.00%	54.16%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	95.79%	0.00%	12.14%	0.00%	68.68%	0.00%	-

CAC	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	15.70%	0.16%	0.00%	0.03%	18.29%	0.12%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	15.70%	0.00%	-	0.00%	0.00%	0.00%	14.34%	0.00%	0.00%
GARCH-T	0.16%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	86.28%	0.00%	0.00%	0.00%
GARCH-QML	0.03%	0.00%	0.00%	0.00%	86.28%	-	0.00%	0.00%	46.10%
GARCH-N-CMD	18.29%	0.00%	14.34%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.12%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.00%	46.10%	0.00%	0.00%	-

DAX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.27%	0.04%	0.00%	0.00%	0.20%	0.04%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.27%	0.00%	-	0.00%	0.00%	0.00%	5.57%	0.00%	0.00%
GARCH-T	0.04%	0.00%	0.00%	-	0.00%	0.00%	0.00%	7.69%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.01%	0.00%	0.00%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.01%	-	0.00%	0.00%	0.05%
GARCH-N-CMD	0.20%	0.00%	5.57%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.04%	0.00%	0.00%	7.69%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%	-

DOW	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	21.77%	0.31%	0.12%	0.11%	16.83%	0.29%	0.13%
Riskmetrics	0.00%	-	0.09%	0.00%	0.00%	0.00%	0.06%	0.00%	0.00%
GARCH-N	21.77%	0.09%	-	0.00%	0.00%	0.00%	28.12%	0.00%	0.00%
GARCH-T	0.31%	0.00%	0.00%	-	0.02%	38.52%	0.00%	94.17%	0.05%
GARCH-ST	0.12%	0.00%	0.00%	0.02%	-	53.36%	0.00%	0.01%	57.12%
GARCH-QML	0.11%	0.00%	0.00%	38.52%	53.36%	-	0.00%	45.02%	64.24%
GARCH-N-CMD	16.83%	0.06%	28.12%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.29%	0.00%	0.00%	94.17%	0.01%	45.02%	0.00%	-	0.00%
GARCH-ST-CMD	0.13%	0.00%	0.00%	0.05%	57.12%	64.24%	0.00%	0.00%	-

TOPIX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	93.24%	20.76%	12.20%	3.69%	92.72%	12.05%	4.05%
Riskmetrics	0.00%	-	0.12%	0.00%	0.00%	0.00%	0.15%	0.00%	0.00%
GARCH-N	93.24%	0.12%	-	0.00%	0.00%	0.00%	28.25%	0.00%	0.00%
GARCH-T	20.76%	0.00%	0.00%	-	1.06%	0.00%	0.00%	1.13%	0.00%
GARCH-ST	12.20%	0.00%	0.00%	1.06%	-	0.06%	0.00%	73.24%	0.82%
GARCH-QML	3.69%	0.00%	0.00%	0.00%	0.06%	-	0.00%	0.11%	82.21%
GARCH-N-CMD	92.72%	0.15%	28.25%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	12.05%	0.00%	0.00%	1.13%	73.24%	0.11%	0.00%	-	0.03%
GARCH-ST-CMD	4.05%	0.00%	0.00%	0.00%	0.82%	82.21%	0.00%	0.03%	-

WIG	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.10%	0.00%	60.67%	0.00%	0.39%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	40.75%	0.00%
GARCH-ST	0.00%	0.00%	0.10%	0.00%	-	0.00%	0.16%	0.00%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	60.67%	0.00%	0.16%	0.00%	-	0.00%	0.52%
GARCH-T-CMD	0.00%	0.00%	0.00%	40.75%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.39%	0.00%	0.00%	0.00%	0.52%	0.00%	-

Table 4: Diebold–Mariano test results calculated for 97.5% VaR for each individual index in considered time intervals. Null hypothesis is that the two forecasts are of equal accuracy and alternative hypothesis of different quality. Green fields indicate p-values greater than 5%.

Period I — Var 97.5% confidence level

ASX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.05%	0.00%	0.00%	0.00%	1.07%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.05%	0.00%	-	0.00%	0.00%	0.00%	27.22%	0.04%	0.01%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.81%	0.00%	74.92%	81.71%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	18.90%	0.00%	29.34%	62.28%
GARCH-QML	0.00%	0.00%	0.00%	0.81%	18.90%	-	0.00%	11.64%	28.62%
GARCH-N-CMD	1.07%	0.00%	27.22%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.04%	74.92%	29.34%	11.64%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.01%	81.71%	62.28%	28.62%	0.00%	0.00%	-

CAC	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.02%	0.00%	0.00%	0.00%	0.05%	0.01%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.02%	0.00%	-	0.00%	0.00%	0.00%	11.51%	30.89%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.01%	5.89%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.05%	0.00%	11.51%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.01%	0.00%	30.89%	0.01%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	5.89%	0.00%	0.00%	0.00%	0.00%	-

DAX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.04%	0.07%	1.14%	15.98%	0.02%	0.04%	1.53%
Riskmetrics	0.00%	-	1.28%	37.85%	0.00%	0.00%	1.72%	41.22%	0.00%
GARCH-N	0.04%	1.28%	-	11.04%	0.00%	0.00%	46.78%	9.13%	0.00%
GARCH-T	0.07%	37.85%	11.04%	-	0.00%	0.00%	17.20%	68.20%	0.00%
GARCH-ST	1.14%	0.00%	0.00%	0.00%	-	0.01%	0.00%	0.00%	36.53%
GARCH-QML	15.98%	0.00%	0.00%	0.00%	0.01%	-	0.00%	0.00%	0.09%
GARCH-N-CMD	0.02%	1.72%	46.78%	17.20%	0.00%	0.00%	-	11.73%	0.00%
GARCH-T-CMD	0.04%	41.22%	9.13%	68.20%	0.00%	0.00%	11.73%	-	0.00%
GARCH-ST-CMD	1.53%	0.00%	0.00%	0.00%	36.53%	0.09%	0.00%	0.00%	-

DOW	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.01%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.01%	0.00%	-	0.00%	0.00%	0.00%	0.73%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.29%	0.00%	0.04%	9.88%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.00%
GARCH-QML	0.00%	0.00%	0.00%	0.29%	0.00%	-	0.00%	15.40%	81.03%
GARCH-N-CMD	0.03%	0.00%	0.73%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.04%	0.00%	15.40%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	9.88%	0.00%	81.03%	0.00%	0.00%	-

TOPIX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	66.19%	31.26%	43.22%	40.12%	92.51%	58.44%	65.61%
Riskmetrics	0.00%	-	9.67%	3.22%	5.17%	4.48%	18.27%	8.71%	10.35%
GARCH-N	66.19%	9.67%	-	0.11%	4.78%	0.00%	5.91%	66.68%	94.63%
GARCH-T	31.26%	3.22%	0.11%	-	0.00%	32.24%	0.00%	1.61%	0.57%
GARCH-ST	43.22%	5.17%	4.78%	0.00%	-	79.99%	0.05%	21.37%	8.13%
GARCH-QML	40.12%	4.48%	0.00%	32.24%	79.99%	-	0.17%	39.84%	25.53%
GARCH-N-CMD	92.51%	18.27%	5.91%	0.00%	0.05%	0.17%	-	1.13%	5.49%
GARCH-T-CMD	58.44%	8.71%	66.68%	1.61%	21.37%	39.84%	1.13%	-	0.11%
GARCH-ST-CMD	65.61%	10.35%	94.63%	0.57%	8.13%	25.53%	5.49%	0.11%	-

WIG	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	5.95%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.20%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	52.71%	0.01%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.58%
GARCH-N-CMD	0.00%	0.00%	5.95%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.20%	52.71%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.01%	0.58%	0.00%	0.00%	-

Period II — Var 97.5% confidence level

ASX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	76.79%	5.87%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	11.10%	57.30%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	1.73%	0.00%	0.00%	36.50%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	1.73%	-	0.00%	0.00%	16.36%
GARCH-N-CMD	0.00%	0.00%	76.79%	11.10%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	5.87%	57.30%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	36.50%	16.36%	0.00%	0.00%	-

CAC	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	2.74%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	1.31%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	72.43%	0.00%	0.00%	0.09%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	72.43%	-	0.00%	0.00%	0.34%
GARCH-N-CMD	0.00%	0.00%	2.74%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	1.31%	0.00%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.09%	0.34%	0.00%	0.00%	-

DAX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	81.72%	0.00%	0.00%	1.48%	3.37%	0.00%
GARCH-T	0.00%	0.00%	81.72%	-	0.00%	0.00%	9.09%	0.06%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.78%	0.00%	1.57%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	1.48%	9.09%	0.78%	0.00%	-	96.43%	0.00%
GARCH-T-CMD	0.00%	0.00%	3.37%	0.06%	0.00%	0.00%	96.43%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	1.57%	0.00%	0.00%	0.00%	-

DOW	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.01%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	15.96%	4.22%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	13.34%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	90.25%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	15.96%	0.00%	0.00%	0.00%	-	11.93%	0.00%
GARCH-T-CMD	0.00%	0.00%	4.22%	13.34%	0.00%	0.00%	11.93%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	90.25%	0.00%	0.00%	0.00%	-

TOPIX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.01%	0.02%	84.53%	81.51%	0.07%	0.41%
GARCH-T	0.00%	0.00%	0.01%	-	0.00%	0.02%	0.02%	12.75%	1.54%
GARCH-ST	0.00%	0.00%	0.02%	0.00%	-	0.06%	0.07%	45.09%	7.20%
GARCH-QML	0.00%	0.00%	84.53%	0.02%	0.06%	-	97.81%	0.11%	0.43%
GARCH-N-CMD	0.00%	0.00%	81.51%	0.02%	0.07%	97.81%	-	0.05%	0.26%
GARCH-T-CMD	0.00%	0.00%	0.07%	12.75%	45.09%	0.11%	0.05%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.41%	1.54%	7.20%	0.43%	0.26%	0.00%	-

WIG	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	3.60%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	42.35%	0.00%	0.00%	0.01%
GARCH-N	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.64%	0.00%	0.00%
GARCH-T	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	0.62%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	95.55%	0.00%
GARCH-QML	3.60%	42.35%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	0.64%	0.00%	0.00%	0.00%	-	0.00%	0.00%
GARCH-T-CMD	0.00%	0.00%	0.00%	0.62%	95.55%	0.00%	0.00%	-	0.00%
GARCH-ST-CMD	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-

Period III	Var 97.5% confidence level								
ASX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	0.58%	3.49%	0.00%	13.50%	0.55%	3.99%
GARCH-T	0.00%	0.00%	0.58%	-	0.00%	0.00%	1.94%	99.63%	0.00%
GARCH-ST	0.00%	0.00%	3.49%	0.00%	-	0.00%	8.79%	0.00%	45.26%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.01%	13.50%	1.94%	8.79%	0.00%	-	1.87%	9.75%
GARCH-T-CMD	0.00%	0.00%	0.55%	99.63%	0.00%	0.00%	1.87%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	3.99%	0.00%	45.26%	0.00%	9.75%	0.00%	-

CAC	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.02%	0.01%	0.00%	0.00%	0.02%	0.01%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.02%	0.00%	-	5.00%	0.00%	0.00%	16.70%	4.80%	0.00%
GARCH-T	0.01%	0.00%	5.00%	-	0.00%	0.00%	20.13%	90.33%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	1.08%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.02%	0.00%	16.70%	20.13%	0.00%	0.00%	-	19.86%	0.00%
GARCH-T-CMD	0.01%	0.00%	4.80%	90.33%	0.00%	0.00%	19.86%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	1.08%	0.00%	0.00%	0.00%	-

DAX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.09%	0.08%	0.04%	0.03%	0.07%	0.08%	0.03%
Riskmetrics	0.00%	-	0.16%	0.14%	0.06%	0.05%	0.12%	0.13%	0.06%
GARCH-N	0.09%	0.16%	-	37.31%	13.67%	1.08%	32.84%	48.13%	13.36%
GARCH-T	0.08%	0.14%	37.31%	-	1.41%	55.97%	69.18%	60.53%	1.15%
GARCH-ST	0.04%	0.06%	13.67%	1.41%	-	69.18%	31.88%	0.10%	90.07%
GARCH-QML	0.03%	0.05%	1.08%	55.97%	69.18%	-	19.87%	47.40%	71.18%
GARCH-N-CMD	0.07%	0.12%	32.84%	69.18%	31.88%	19.87%	-	79.26%	32.02%
GARCH-T-CMD	0.08%	0.13%	48.13%	60.53%	0.10%	47.40%	79.26%	-	0.00%
GARCH-ST-CMD	0.03%	0.06%	13.36%	1.15%	90.07%	71.18%	32.02%	0.00%	-

DOW	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.01%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.01%	-	0.01%	0.01%	0.00%	0.00%	0.00%	0.01%	0.00%
GARCH-N	0.00%	0.01%	-	93.49%	0.00%	0.00%	7.04%	1.23%	0.00%
GARCH-T	0.01%	0.01%	93.49%	-	0.00%	0.00%	29.76%	3.65%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%	61.76%
GARCH-QML	0.00%	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.00%	0.00%
GARCH-N-CMD	0.00%	0.00%	7.04%	29.76%	0.00%	0.00%	-	0.24%	0.00%
GARCH-T-CMD	0.00%	0.01%	1.23%	3.65%	0.00%	0.00%	0.24%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	61.76%	0.00%	0.00%	0.00%	-

TOPIX	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	1.13%	0.46%	0.35%	2.79%	1.46%	0.30%	0.09%
Riskmetrics	0.00%	-	0.15%	0.04%	0.04%	0.44%	0.18%	0.03%	0.01%
GARCH-N	1.13%	0.15%	-	39.64%	0.00%	2.18%	32.15%	1.63%	0.00%
GARCH-T	0.46%	0.04%	39.64%	-	2.79%	0.45%	8.25%	5.66%	0.01%
GARCH-ST	0.35%	0.04%	0.00%	2.79%	-	0.00%	0.34%	64.70%	3.18%
GARCH-QML	2.79%	0.44%	2.18%	0.45%	0.00%	-	29.81%	0.04%	0.00%
GARCH-N-CMD	1.46%	0.18%	32.15%	8.25%	0.34%	29.81%	-	0.69%	0.00%
GARCH-T-CMD	0.30%	0.03%	1.63%	5.66%	64.70%	0.04%	0.69%	-	0.10%
GARCH-ST-CMD	0.09%	0.01%	0.00%	0.01%	3.18%	0.00%	0.00%	0.10%	-

WIG	Historical	Riskmetrics	GARCH-N	GARCH-T	GARCH-ST	GARCH-QML	GARCH-N-CMD	GARCH-T-CMD	GARCH-ST-CMD
Historical	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Riskmetrics	0.00%	-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GARCH-N	0.00%	0.00%	-	2.12%	0.00%	0.44%	73.84%	1.25%	0.00%
GARCH-T	0.00%	0.00%	2.12%	-	0.00%	0.04%	9.50%	16.24%	0.00%
GARCH-ST	0.00%	0.00%	0.00%	0.00%	-	0.00%	0.05%	0.00%	0.01%
GARCH-QML	0.00%	0.00%	0.44%	0.04%	0.00%	-	2.18%	0.02%	0.00%
GARCH-N-CMD	0.00%	0.00%	73.84%	9.50%	0.05%	2.18%	-	6.95%	0.01%
GARCH-T-CMD	0.00%	0.00%	1.25%	16.24%	0.00%	0.02%	6.95%	-	0.00%
GARCH-ST-CMD	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.01%	0.00%	-

6 CONCLUSIONS

Findings of the research presented in this study, evaluating the effectiveness of Value-at-Risk forecasts for six utility indices from developed markets under different volatility regimes can be decomposed into three main sections:

The first one analyzes the overall behavior of all modes in each testing period and tries to answer the question of what is the preferred environment for them to accurately assess the market risk of an underlying instrument. From the empirical results one can state that VaR forecasts are the most accurate and resilient when surrounding market moves are of moderate nature. In testing period two, where based on the level of VIX index, in-sample period has medium volatility, and out-of-sample time interval is calm, all models scored well in terms of Kupiec, Christoffersen, and Dynamic Quantile tests at both 99% and 97.5% confidence levels of VaR calculations. On the other hand, all models do not perform well in a highly turbulent era of Covid-19 pandemic (Period III). Most of the approaches to VaR estimation are underestimating the downside market risk, leading to the highest number of exceedances for both confidence intervals of calculations. These results are in line with expectations and with the well-described regularities in the literature on the subject.

After describing the general behavior of the models, a vital question is the individual performance of each Value-at-Risk assessment technique. At a 1% significance level of forecast, in Period I, characterized by low and increasing volatility, the model that works best is the GARCH(1,1) with skewed t-student's distribution and exogenous regressors in the form of lagged returns on commodity futures. The model gives no reason to reject the null hypotheses of two coverage tests for all indices, and Dynamic Quantile test for most of them. For the less conservative VaR predictions at a 2.5% significance interval, the GARCH(1,1) with skewed student's t distribution scored the highest in accuracy and independency tests. However, Diebold-Mariano trial indicates that at the significance level of 95%, both the GARCH-ST and the GARCH-ST-CMD produce the same quality of forecasts. In Period II (medium, decreasing volatility), the GARCH(1,1) with standard normal distribution should be distinguished from all VaR 99% forecasts, and in the context of VaR 97.5%, GARCH-ST-CMD is on the top of the podium again. In the most turbulent period of the coronavirus epidemic rollout, and therefore the highest volatility, the GARCH(1,1) with QML estimator has the best performance (99% VaR) as it is able to accurately predict VaR levels for 5 out of 6 indices. For the less conservative forecast (97.5% VaR), GARCH-ST-CMD seems to provide the highest backtesting quality. For all indices and time periods, the Riskmetrics® model obtains the worst results, as it is based on the assumption of stock returns being normally distributed which with regard to the data used in this comparison is not valid. Overall, the best performance can be attributed to the GARCH-ST-CMD approach. Due to its form, this model is able to include some stylized facts about empirical data such as fat tail or skewness, and exogenous regressors are stabilizing the behavior of the model in volatile times.

Finally, this paper creates an opportunity to compare the performance of the GARCH models that are enriched with additional variables of commodity prices with those that are not. In periods and confidence intervals where the GARCH-ST-CMD has the best performance, Diebold-Mariano test exhibits a fact, that for Australian and Japan markets mainly, both GARCH-ST-CMD and GARCH-ST provide researchers with the same accuracy of forecasts. It can be concluded that the model selection should be made individually for each instrument, as more complex models do not always bring the improvement of results.

As a further development of the analysis conducted in this paper, authors recommend enriching it with additional models, especially semi-parametric techniques which have recently gained considerable popularity such as the CaViAR (Conditional Autoregressive Value at Risk), the FHS (Filtered Historical Simulation) or models based on the EVT (Extreme Value Theory). Moreover, the number of tested market volatility scenarios can be inflated to provide more versatile benchmarking.

REFERENCES

Manganelli, S. and Engle, R. (2001). A detailed comparison of value at risk estimates. *European Central Bank Working Paper No. 75.*

Linsmeier, T. and Pearson, N. (2000). Value at Risk, *Financial Analysts Journal, 56:2, 47-67.*

Sansa, N. (2020). The Impact of the COVID - 19 on the Financial Markets: Evidence from China and USA. *Electronic Research Journal of Social Sciences and Humanities Vol 2: Issue II.*

Whaley, R. (1995). Predicting stock market volatility: A new measure. *The Journal of Futures Markets No.3, 265-302.*

Kim, B., Kim, H and Lee, B. (2015). Spillover effects of the U.S. financial crisis on financial markets in emerging Asian countries. *International Review of Economics & Finance Vol. 39, September 2015, 192-210.*

Sharma, M. (2012). Evaluation of Basel III revision of quantitative standards for implementation of internal models for market risk. *IIMB Management Review Vol. 24, Issue 4, December 2012, 234-244.*

Buczyński, M. and Chlebus, M. (2020). Size does matter. A study on the required window size for optimal quality market risk models. *Faculty of Economic Sciences, University of Warsaw Working Papers 9/2020 (315).*

Buczyński, M. and Chlebus, M. (2019). Old-fashioned parametric models are still the best. a comparison of value-at-risk approaches in several volatility states. *Faculty of Economic Sciences, University of Warsaw Working Papers, 12(297).*

Abad, P. and Benito, S. (2013). A detailed comparison of value at risk estimates. *Mathematics and Computers in Simulation, 94:258–276.*

Totić, S., Bulajićm M., and Vlastelica, T. (2011). Empirical comparison of conventional methods and extreme value theory approach in value-at-risk assessment. *African Journal of Business Management Vol. 5(33), pp. 12810-12818.*

Ünal, G. (2011). Value-at-risk forecasts: a comparison analysis of extreme-value versus classical approaches. *The Journal of Risk Model Validation (59–76) Volume 5/Number 3, Fall 2011.*

Wang, Y., Jiang, Y. and He, G. (2019). Using Artificial Neural Network to Predict Value-at-Risk of S&P 500 Market Index withExternal Information. *2019 IEEE International Conference on Big Data, Cloud Computing, Data Science & Engineering (BCD), 2019, pp. 90-93.*

Abad, P., Benito, S., and López, C. (2014). A comprehensive review of value at risk methodologies. *The Spanish Review of Financial Economics, 12(1):15–32.*

Bryson, J. and Vanchan, V. (2020). Covid-19 and alternative conceptualization of value and risk in GPN research. *Tijdschrift voor Economische en Sociale Geografie – 2020, Vol. 111, No. 3, pp. 530–542.*

Kambouroudis, D. and McMillan, D. (2016). Does VIX or volume improve GARCH volatility forecasts?. *Applied Economics, 48:13, 1210-1228*.

CBOE. Cboe White Paper: Cboe Volatility Index. *Accessed 20.08.2021.*

Jorion, P. (2001). Value at Risk: The New Benchmark for Managing Financial Risk. *2nd Edition, McGraw-Hill, United States of America.*

Parameswaran, S. (2011). Fundamentals of Financial Instruments. *John Wiley & Sons (Asia) Pte. Ltd.*, *283-290.*

Butler, J. and Schachter, B. (1996). Estimating Value-at-Risk with a precision measure by combining kernel estimation with historical simulation. *1996 Chicago Fed Bank Structure Conference.*

Down, K. (2002). Measuring Market Risk. *John Wiley & Sons, Chichester.*

Richardson, M. and Smith, T. (1993). A Test for Multivariate Normality in Stock Returns. *The Journal of Business , Apr., 1993, Vol. 66, No. 2 (Apr., 1993), pp. 295-321.*

Morgan Guaranty Trust Company of New York (1996). RiskMetrics Technical Document - Fourth Edition.

Orhan, M. and Köksal, B. (2011). A comparison of GARCH models for VaR estimation.

Expert Systems with Applications Volume 39, Issue 3, 15 February 2012, Pages 3582-3592

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics 31, 307—327.*

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

Angelidis, T., Benos, A., and Degiannakis, S. (2004). The use of garch models in var estimation. *Statistical Methodology, 1, 105–128.*

Nana, G., Korn, R., Erlwein-Sayer, C. (2013). GARCH-extended models: theoretical properties and applications. *Cornell University website, Accessed 07.09.2021.*

Duro, G. (2020). Modelling Daily Volatility with External Regressors. *Instituto Universitario De Lisboa repository, Accessed 18.09.2021.*

Han, H., Kristensen, D. (2015). Asymptotic Theory for the QMLE in GARCH-X Models with Stationary and Non-Stationary Covariates. *CREATES Research Paper 2012-25.*

Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal Article. The Journal of Derivatives 3 (2): 73–84.*

Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review. 39(4), 841–862.*

Niepola, O. (2009). Backtesting Value-at-Risk Models. *Helisinki School of Economics*.

Markov chain Definition. *US English by Oxford Dictionaries. English. Retrieved 2017-12-14. Accessed 05.10.2021*.

Engle, R., Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles*. Journal of Business & Economic Statistics 22, no. 4: 367-381*.

Geman, D., Geman, H., Taleb, N. (2014). Tail Risk Constraints and Maximum Entropy. *Entropy 17 (6), 3724-3737, 2015.*

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl