



UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES

WORKING PAPERS

No. 12/2019 (297)

OLD-FASHIONED PARAMETRIC MODELS ARE STILL THE BEST. A COMPARISON OF VALUE-AT-RISK APPROACHES IN SEVERAL VOLATILITY STATES

MATEUSZ BUCZYŃSKI
MARCIN CHLEBUS

WARSAW 2019



Old-fashioned parametric models are still the best. A comparison of Value-at-Risk approaches in several volatility states.

Mateusz Buczyński*, Marcin Chlebus

University of Warsaw, Faculty of Economic Sciences

** Corresponding author: mateusz.buczynski@student.uw.edu.pl*

Abstract: Numerous advances in the modelling techniques of Value-at-Risk (VaR) have provided the financial institutions with a wide scope of market risk approaches. Yet it remains unknown which of the models should be used depending on the state of volatility. In this article we present the backtesting results for 1% and 2.5% VaR of six indexes from emerging and developed countries using several most known VaR models, among many: GARCH, EVT, CAViaR and FHS with multiple sets of parameters. The backtesting procedure has been based on the excess ratio, Kupiec and Christoffersen tests for multiple thresholds and cost functions. The added value of this article is that we have compared the models in four different scenarios, with different states of volatility in training and testing samples. The results indicate that the best of the models that is the least affected by changes in the volatility is GARCH(1,1) with standardized student's t-distribution. Non-parametric techniques (e.g. CAViaR with GARCH setup (see Engle and Manganelli, 2001) or FHS with skewed normal distribution) have very prominent results in testing periods with low volatility, but are relatively worse in the turbulent periods. We have also discussed an automatic method to setting a threshold of extreme distribution for EVT models, as well as several ensembling methods for VaR, among which minimum of best models has been proven to have very good results - in particular a minimum of GARCH(1,1) with standardized student's t-distribution and either EVT or CAViaR models.

Keywords: Value-at-Risk, GARCH, Extreme Value Theory, Filtered Historical Simulation, CAViaR, market risk, forecast comparison

JEL codes: G32, C52, C53, C58

1. Introduction

Every decision undertaken by any financial institution bears risks. These risks might have different sources, either market, credit or operational. Estimation of these three main types of risks is therefore essential to properly manage the institution. If we put aside the credit system, in most cases market risk becomes the most important risk factor in the business. According to many researchers, one of the most common market risk measures is Value-at-Risk (Mitra and Ji (2010); Jorion (2010)).

There are numerous approaches to Value-at-Risk estimation. However, the existence of the Basel III framework provides a set of regulations on how banking institutions should estimate risk measures. First of all, it is to the bank's preferences what statistical model it will use, but the Value-of-Risk should be calculated at the 99% confidence level, forecasts should be made for at least 10 days ahead, and the forecasting period should be at least a year long (250 observations).

Researchers outdo each other, providing results that support particular model's advantages (among many Danielsson and De Vries (2000); Angelidis, Benos, and Degiannakis (2007); Bao, Lee, and Saltoglu (2006); Huang (2009); Yu, Li, and Jin (2010); Ozun, Cifter, and Yilmazer (2010); Polanski and Stoja (2010); Gerlach, Chen, and Chan (2011)). But the truth is that there is no model that would be the best regardless of the market circumstances. If an individual model has been proven to have very satisfactory results under some set of conditions, it is highly probable that there exists another study considering the same model under different set of conditions and it proves the model's malfunctions. In such a case, Abad, Benito, and López (2014) have prepared a broad juxtaposition of the popular Value-at-Risk methodologies, for instance Filtered Historical Simulation (FHS), Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) models or Extreme Value Theory (EVT) models. Moreover, combined forecasts by using several procedures have been gaining a lot of attention recently with either mixed alternative risk models or model being dependent on the current volatility state (e.g. in McAleer, Jiménez-Martín, and Pérez-Amaral (2013)).

One of the most crucial factors that may influence the correctness of model's predictions are the dynamics of the volatility. As a matter of fact, in the last twenty years, we experienced a perfect moment to test models in such precedents – the global financial crisis. Another area that has an impact on the model's performance may be the country origin of the studied time series. Stock markets differ dramatically between states. Many researchers have been comparing the performance of the models before and after turbulent states (among others, in McAleer, Jiménez-Martín, and Pérez-Amaral (2013); Degiannakis, Floros, and Livada (2012); Chlebus (2016); Burchi and Martelli (2016)) and between different stock markets (among others in Žiković (2007); Djaković and Radiscaron (2010); Mutu, Balogh, and Moldovan (2011); Miletic and Miletic (2015); Buczyński and Chlebus (2018)). Most of them notice the advantages of GARCH models in periods of increased volatility, mentioning its conservatism and abrupt fit to the market situation. On the contrary, non-parametric and semiparametric methods are gaining a lot of attention lately, due to their very flexible fit to the data and low costs of usage.

For instance, Djaković and Radiscaron (2010) have studied performance of historical simulation models in different European countries. Their results indicate that model's risk estimates depend strongly on the market's volatility. According to them, Value-at-Risk models estimated based on the samples originating in developed countries may achieve better results than the ones based in

emerging markets. Another study conducted by Mutu, Balogh, and Moldovan (2011) proved that EVT and GARCH models perform better than other models in highly volatile periods. Buczyński and Chlebus (2018) also discovered that the GARCH models outperform semiparametric models in turbulent periods, but the discrepancies get smaller in tranquil periods. Additionally, they found that non-parametric methods do perform better in emerging markets. Degiannakis, Floros, and Livada (2012) performed a study on several Value-at-Risk models before and after crisis. GARCH models were performing in the pre-crisis period, but during the crisis only t-distributed models provided risk estimates of good quality. Conforming to them, these results apply to both developing and developed countries. In summary, Value-at-Risk estimation is a scientifically abundant field, however no study has compared the most popular models in many different aspects yet.

This study aims to estimate several different Value-at-Risk methods and to compare them depending on the time interval and stock exchange origin country. The aforementioned models include: GARCH(1,1) models with varied distributions (skewed normal, both standardized and non-standardized student's t and skewed student's t), Conditional Auto-Regressive Value at Risk (CAViaR) models (with all approaches presented in Engle and Manganelli (2004)), filtered historical simulation (FHS) with varied distributions (normal and skewed normal) and Extreme Value Theory (EVT) models with their own approach to threshold calculation. The models have been compared in four different time intervals: 2003-2007; mid 2004 – mid 2009; 2007-2011; 2008-2012. The comparison was based both on Kupiec and Christoffersen tests and on the analyses of Value-at-Risk cost functions.

The paper is organized as follows: the first section discusses in depth methodologies of all discussed models, the next section presents testing framework and in the third section empirical results of the estimation are presented.

2. Methodology

Let us bring the methodology of Value-at-Risk up front. The Value-at-Risk is defined as a maximum possible loss in a given time frame for a particular portfolio, at a given level of confidence in normal market conditions. Moreover, a mathematical equation to define Value-at-Risk could be presented as follows (Jorion (2010)):

$$P(r_t < VaR_\alpha(t) | \Omega_{t-1}) = \alpha \quad (1)$$

where:

r_t = financial return,

α = confidence level,

t = time interval,

Ω_{t-1} = information set available at time t-1.

One of the approaches to the Value-at-Risk estimation are GARCH models. Specifically in this paper the GARCH models under consideration are all GARCH(1,1) models, but with different assumptions about distribution of the underlying data: skewed normal, both standardized and non-standardized student's t and skewed student's t. GARCH(1,1) is one of the simplest of the whole GARCH family models and can be described by the two following equations:

$$\begin{aligned} r_t | \Omega_{t-1} &\sim IID(\mu_t, h_t), \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, \end{aligned} \quad (2)$$

where:

- r_t = financial return (a stochastic process),
- Ω_{t-1} = information set available at time $t - 1$,
- μ_t = mean of the process (hereby equal to 0),
- h_t = conditional variance,
- α_0 = constant,
- α_1 = parameter corresponding to the stochastic process' realization at time $t - 1$,
- ε_{t-1}^2 = squared random error in the period $t - 1$,
- β_1 = parameter corresponding to the conditional variance's realization at time $t - 1$.

Given these equations, we can define Value-at-Risk using a formula below (Angelidis, Benos, and Degiannakis 2004):

$$VaR_\alpha(t) = \hat{\mu}_t + q_\alpha \sqrt{\hat{h}_t}, \quad (3)$$

where:

- $VaR_\alpha(t)$ = Value-at-Risk at the α confidence level at time t ,
- $\hat{\mu}_t$ = forecast of the conditional mean at time t ,
- $\hat{h}_t$ = forecast of the conditional variance at the time t ,
- q_α = quantile of the assumed underlying distribution of the data.

Following Bollerslev and Wooldridge (1992), if the model is not conditionally normally distributed, but it specifies the first two conditional moments correctly, the estimates of the quasi likelihood function will be consistent and asymptotically normal. This makes it possible to use different than normal distributions of the underlying process. The reason for that is that most of the financial time series distributions tend to have fatter tails. Hence, the use of normal distribution may be not sufficient.

The usage of different distributions distorts the model's likelihood. While in case of normal distribution, one has to estimate only the parameters of the model. In case of other distributions, their specific parameters must be estimated as well, hence they are included in the likelihood functions. Therefore the assumption of the distribution of the underlying data bears the responsibility for the differences in volatility estimates in-between the models.

Another hereby studied approach to Value-at-Risk modeling are CAViaR models introduced by Engle and Manganelli (2004). The idea behind CAViaR models is to model the quantile of the distribution of the data directly instead of trying to model the whole distribution. Beta coefficients in the CAViaR models are estimated on the basis of Koenker and Basset (1978) regression quantile methodology, which minimalises the regression quantile objective function. The basic formula for the CAViaR model (with one lagged Value-at-Risk and one lagged observed value) can be expressed as:

$$VaR_\alpha(t) = \beta_0 + \beta_1 VaR_\alpha(t-1) + l(\beta_2, r_{t-1}, VaR_\alpha(t-1)) \quad (4)$$

where:

- r_t = rate of return,
- β_i = vector of parameters corresponding to lagged values,
- $l(.)$ = a function of a lagged values of observables and Value-at-Risk.

The point of using linking function within the CAViaR expression is to link the model outcome to the level of rates of return at time $t - 1$. In our research we include 3 various CAViaR model definitions, each of them using different linking function. All of them were provided in Engle and Manganelli (2004):

- Symmetric Absolute Value (SAV):

$$VaR_\alpha(t) = \beta_0 + \beta_1 VaR_\alpha(t-1) + \beta_2 |r_{t-1}|, \quad (5)$$

In this case, Value-at-Risk depends on its lag and lagged absolute rate of return, which implies symmetrical behavior regardless of gains or losses, focusing on the impact of changes;

- Indirect GARCH:

$$VaR_\alpha(t) = \sqrt{\beta_0 + \beta_1 VaR_\alpha(t-1) + \beta_2 (r_{t-1})^2}, \quad (6)$$

The indirect GARCH approach to CAViaR models is very similar to GARCH modeling. In fact, it would be correctly specified model if the underlying data were GARCH(1,1) process with IID distribution;

- Adaptive (AD):

$$VaR_\alpha(t) = VaR_\alpha(t-1) + \beta_0 ((1 + \exp(K(r_{t-1} + VaR_\alpha(t-1))))^{-1} - \alpha), \quad (7)$$

where:

K = positive finite number.

This version of CAViaR model increases Value-at-Risk estimates only on the occurrence of exceedance. In other case values should decrease slightly. Such model's structure is aimed at elimination of sequences of exceedances.

Yet another approach to Value-at-Risk estimation is Filtered Historical Simulation (FHS) introduced by Barone-Adesi, Giannopoulos, and Vosper (1999). It is based on the two most popular approaches – historical simulation and GARCH models. In FHS, the first step is to model the volatility and later to standardize the data using obtained estimates – filtering it in a way. Any GARCH model can serve as a filter here – in this paper GARCH(1,1) with normal and skewed normal distributions were used. Given equation 3, we can describe standardized residuals from the GARCH model as:

$$z_t = \frac{\varepsilon_t}{\sqrt{h_t}}. \quad (8)$$

The whole idea behind FHS is based on the Monte Carlo simulation on the IID returns obtained in 9. The simulation considers sampling from the standardized residuals a large number of times (Barone-Adesi, Giannopoulos, and Vosper (1999) suggested this number to be at least 10,000, hereby we sampled 20,000 times). Such a sample of simulated standardized residuals should be connected again with the data, therefore we generate the forecast of the volatility:

$$\sqrt{h_{t+1}} = \sqrt{\alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 h_t}, \quad (9)$$

and unstandardize simulated residuals by multiplication with it, constructing a vector of simulated returns. This is where historical simulation takes the steering wheel, because in the end, Value-at-Risk estimate equals to the assumed quantile of the simulated returns, as such:

$$VaR_\alpha(t+1) = \mu_{t+1} + q_\alpha \quad (10)$$

The last of the considered models are Extreme Value Theory models in the Peaks Over Threshold approach. The EVT is a branch of statistics that deals with estimating the probability of extreme events, thus it has a huge impact on Value-at-Risk. EVT has been long known technique, both as a statistical and risk management tool (McNeil (1999); Embrechts, Resnick, and Samorodnitsky (1999)).

EVT has two methods to estimate extreme events: Block Maxima (BM) and Peaks Over Threshold (POT). The BM method focuses on obtaining the asymptotic distribution of a series of maxima. The POT method helps in estimating tails of the distribution of the data, based on the exceedances of rate of return over a specified level. The POT method has been proven to be more effective whilst modeling limited data (McNeil, Frey, and Embrechts (2015)), which in case of financial data may not necessary be a problem, but when only rare events are included may the dataset become poorer.

These methods usually fit either generalized extreme value (GEV) distribution (suitable for BM) or generalized Pareto distribution (GPD), which is the case for the POT method. GPD can be given by:

$$G_{\xi, \sigma}(y) = \begin{cases} (1 + \frac{\xi}{\sigma}y)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - e^{-\frac{y}{\sigma}} & \text{if } \xi = 0 \end{cases} \quad (11)$$

where:

Introduction

ξ = shape parameter,
 σ = scale parameter.

An excess distribution over a given threshold u is expressed by:

$$F_u(y) = P(X - u \leq y - X > u) = \frac{F(y+u) - F(u)}{1 - F(u)} = \frac{F(y) - F(u)}{1 - F(u)}, \quad (12)$$

where $0 < y < x_F - u$ and:

F_u = conditional excess distribution function,

$y = x - u$,

Therefore we may assume that there exists such extreme distribution F and for some threshold u , $F_u(x) = G_{\xi, \sigma}(x)$ and $\xi \in \mathbb{R}, \sigma > 0$. For any $x \geq u$:

$$\begin{aligned}\hat{F}(x) &= (X > u)P(X > x - X > u) \\ &= \hat{F}(u)P(X - u > x - u - X > u) \\ &= \hat{F}(u)\hat{F}_u(x - u) \\ &= \hat{F}(u)\left(1 + \frac{\xi}{\sigma}(x - u)\right)^{-\frac{1}{\xi}}.\end{aligned}\tag{13}$$

If we inverse the formula in 14, we will obtain a quantile function, which is exactly what we need to calculate excess Value-at-Risk:

$$Va\hat{R}_{u\alpha}(t) = q_\alpha(F) = u + \frac{\sigma}{\xi} \left(\left(\frac{1 - \alpha}{F(u)} \right)^{-\xi} - 1 \right).\tag{14}$$

We may replace $F(u)$ with the relation of the observations below the threshold to the number of observations $((N - N_u)/N)$ to obtain an estimator for tail probability (Smith (1987)):

$$Va\hat{R}_{u\alpha}(t) = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{N}{N_u} \alpha \right)^{-\xi} - 1 \right).\tag{15}$$

where:

N_u = number of observations exceeding threshold u ,

N = number of observations.

Such theoretical distribution could be fit to the original data, have proposed a dynamic approach to EVT Value-at-Risk estimation. Prior to the GED fitting, we could estimate the current market volatility using GARCH models and fit the theoretical extreme distribution to the residuals of such model. Therefore the final Value-at-Risk one day ahead forecast can be expressed by:

$$VaR_\alpha(t+1) = \mu_{t+1} + Va\hat{R}_{u\alpha}(t+1)\sqrt{h_{t+1}}\tag{16}$$

One unknown remains – how to set the value of threshold u . Usually two options are used: mean excess plots, which represent the empirical mean excess function over the threshold value u ; Hill estimators, which are conditional maximum likelihood estimators. Both the techniques rely on subjective choice by an expert knowledge. However, choosing the threshold each time the estimation window moves would be a huge effort. To address this issue, we decided to set the threshold automatically – at threshold u equal to a certain quantile q_k of the standardized residuals, where $k < \alpha$, hence creating EVT k -quantile POT models. In this study, we consider the values of $k \in \{0.90, 0.95\}$.

If there are numerous approaches to Value-at-Risk estimation, that perform the best in different market states, would it not be more efficient to combine the forecasts to perhaps obtain better

estimates? According to Armstrong (2001), combining the forecast may provide results that outrun single models due to reduced inconsistency in estimates. The first to introduce a concept of combined forecast were Bates and Granger (1969) who proposed an unbiased forecast as a linear combination of two forecast tests, where weights are estimated on the past errors of original forecasts. On the other hand, Diebold (1988) argues that single model might be better off in Value-at-Risk estimation than the combination of several models that are based on incorrect assumptions.

More recent studies shed some more light on the combined forecasts in Value-at-Risk estimation. The simplest approach to combined forecasts used in the literature is averaging – a simple arithmetic mean of the sets of separate models' forecasts. Such an estimate could be written as:

$$VaR_{\alpha}(t) = \frac{\sum_{k=1}^K VaR_{\alpha,k}(t)}{K} \quad (17)$$

where:

K = number of averaged Value-at-Risk approaches.

Another approach would be to select the minimum or maximum (depending on the conservatism of the approach) Value-at-Risk estimate among the chosen models:

$$VaR_{\alpha}(t) = \min\{VaR_{\alpha,1}(t), \dots, VaR_{\alpha,K}(t)\}. \quad (18)$$

3. Backtesting Framework

An advantage of Value-at-Risk estimates is that its performance is easily comparable between the models because as an output they provide values that can be assessed a posteriori. At the moment, the two most popular approaches are used in Value-at-Risk evaluation – statistical tests of adequacy and loss functions.

There are several statistical accuracy tests based on hypothesis testing. One of the most common is unconditional coverage test developed by Kupiec (1995) (also known as the Kupiec test). To derive its formula let us start with the definition of exceedance as an indicator function:

$$I_t = \begin{cases} 1 & \text{if } r_t < VaR_{\alpha}(t) \\ 0 & \text{if } r_t \geq VaR_{\alpha}(t) \end{cases} \quad (19)$$

where:

r_t = rate of return realization at time t ,

$VaR_{\alpha}(t)$ = Value-at-Risk forecast for time t .

Kupiec (1995) has proven that assuming the probability of an exceedance is constant, then overall the number of exceptions over testing sample $N_e = \sum_{t=1}^N I_t$ is a Bernoulli process (it comes from binomial distribution $B(N, \alpha)$, where N is number of observations). Moreover, an excess ratio can be defined as:

$$\hat{\alpha} = \frac{\sum_{t=1}^N I_t}{N} \quad (20)$$

where:

$\hat{\alpha}$ = observed excess ratio,
 N = number of observations.

An unconditional coverage relates to the inference that accurate Value-at-Risk model should produce forecasts with $\hat{\alpha}$ equal to α which is Kupiec test's null hypothesis. Its statistic is a likelihood ratio test, which follows asymptotic $\chi^2(1)$ distribution and can be written as:

$$LR_{UC} = 2(\log(\hat{\alpha}^{N_e}(1 - \hat{\alpha})^{N - N_e}) - \log(\alpha^{N_e}(1 - \alpha)^{N - N_e})), \quad (21)$$

where:

$\hat{\alpha}$ = observed excess ratio,
 α = assumed confidence level.

Proof of Kupiec (1995) is also a foundation of the Basel Traffic Light test, which depending on a particular cumulative binomial probability percentages, assigns one of three color zones to the analyzed Value-at-Risk measure. Consequently, all models with the number of exceedances are lower than 95% of cumulative binomial distribution (for specific N and α) are assigned to the green zone, meaning that these models are accurate. Above 95% is the beginning of the yellow zone, which raise questions in regard of accuracy, but they do not perform badly themselves. Conclusively, all models whose number of exceedances fall above 99.99% are assigned to the red zone and assumed to be incorrect.

Another approach to statistical hypotheses testing has been proposed by Christoffersen (1998), who established the conditional coverage test (also known as the Christoffersen test). It extends the unconditional coverage test by verifying serial independence of the Value-at-Risk exceedances. The statistic for Christoffersen test follows asymptotic $\chi^2(2)$ distribution and can be expressed by:

$$LR_{CC} = LR_{UC} + LR_{IND}, \quad (22)$$

where:

R_{IND} — likelihood ratio statistic for the hypothesis of serial independence against first-order Markov dependence, expressed by:

$$2 \left(\log((1 - \pi_{01})^{N_{00}} \pi_{01}^{N_{01}} (1 - \pi_{11})^{N_{10}} \pi_{11}^{N_{11}}) - \log((1 - \hat{\alpha})^{N_{00} + N_{01}} \hat{\alpha}^{N_{00} + N_{01}}) \right) \quad (23)$$

where:

N_{ij} — number of observations in state j after having been in state i in the previous period (0 – not an exceedance, 1 – exceedance),

$\pi_{01} = \frac{N_{01}}{N_{01} + N_{00}}$,

$\pi_{11} = \frac{N_{11}}{N_{11} + N_{10}}$.

One of the most serious flaws of the aforementioned accuracy test is the lack of ability to assess the Value-at-Risk measure given many confidence levels. To overcome this issue, Pérignon and Smith (2008) proposed a multivariate unconditional test, which is also a likelihood test but can test more than one coverage rates at once. The test itself comes from $\chi^2(j)$ distribution, where j is the number of tested coverage rates. For a particular number of confidence levels j , we can build j disjoint intervals of Value-at-Risk estimates: $\{(-\infty; VaR_{\alpha_0}], (VaR_{\alpha_0}; VaR_{\alpha_1}], \dots, (VaR_{\alpha_j}; \infty)\}$.

We can assign the probability of falling into each of these intervals by subtracting confidence level of a particular interval and confidence level of a previous interval, as such: $\{\theta_0 = \alpha_0, \theta_1 = \alpha_1 - \alpha_0, \dots, \theta_j = \alpha_j - \alpha_{j-1}\}$. Moreover an indicator function for a particular interval can be expressed by:

$$J_{j,t} = I_t(\alpha_j) - I_t(\alpha_{j-1}) = \begin{cases} 1 & \text{if } VaR_{\alpha-1}(t) < r_t < VaR_{\alpha}(t) \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

where:

r_t = rate of return realization at time t ,
 $VaR_{\alpha}(t)$ = Value-at-Risk forecast for time t .

Therefore, the formal derivation of the Pérignon and Smith test is given by:

$$LR_{MUC} = 2 \left(\sum_{i=0}^j N_i \log\left(\frac{N_i}{N}\right) - \sum_{i=0}^j N_i \log(\theta_i) \right) \quad (25)$$

where:

N_i = sum of indicator function of a particular interval $J_{i,t}$,
 θ_i = theoretical probability of falling into a particular interval.

The second approach to Value-at-Risk assessment is cost functions, which calculate the probable losses conditional of using a particular model. The first researcher to develop this area of evaluation was Lopez (1998), who defined cost function as a tool to measure the magnitude of a loss in case an exceedance occurred. The model cost function he proposed can be expressed by:

$$CF_t = \begin{cases} g(r_t, VaR_{\alpha}(t)) & \text{if } r_t < VaR_{\alpha}(t) \\ 0 & \text{if } r_t \geq VaR_{\alpha}(t) \end{cases} \quad (26)$$

where:

$g(\cdot)$ = penalty function for Value-at-Risk exceedance.

Researchers have proposed many forms of penalty function. For example Abad, Benito, and Lopez (2015) proposed the following cost function:

$$C_{AB_t} = \begin{cases} |r_t - VaR_{\alpha}(t)| & \text{if } r_t < VaR_{\alpha}(t) \\ 0 & \text{if } r_t \geq VaR_{\alpha}(t) \end{cases} \quad (27)$$

The overall score for a model is calculated as a sum of cost function values for all forecasts divided by number of exceedances N_e , $\frac{\sum_{t=1}^N CF_t}{N_e}$.

Some researchers argue that cost functions should also include an opportunity cost as a measure of the model's usage for cases when Value-at-Risk estimate is not exceeded. For example Caporin (2008) suggested following cost function:

$$C_{C_t} = \begin{cases} |r_t - VaR_{\alpha}(t)| & \text{if } r_t < VaR_{\alpha}(t) \\ |r_t - VaR_{\alpha}(t)| & \text{if } r_t \geq VaR_{\alpha}(t) \end{cases}, \quad (28)$$

which is in fact Abad, Benito and Lopez's cost function expanded to every forecast. Such a function takes into account the magnitude of the models, which tend to overestimate Value-at-Risk

for the low number of exceedances. Another function of that type was proposed by Chlebus (2016), who extends the concept of cost functions distinguishing penalty for opportunity cost between the cases of positive and negative rate of return. Moreover, in case of Value-at-Risk exceedance, he proposes not to measure the magnitude of loss above the reserve, but the magnitude of whole loss. The higher the value is the more conservative the model tends to be. His approach is as follows:

$$C_{EC_t} = \begin{cases} |VaR_\alpha(t)| & \text{if } r_t \geq VaR_\alpha(t) \text{ and } r_t \geq 0 \\ |r_t - VaR_\alpha(t)| & \text{if } r_t \geq VaR_\alpha(t) \text{ and } r_t < 0 \\ |r_t| & \text{if } r_t < VaR_\alpha(t) \text{ and } r_t < 0 \end{cases} \quad (29)$$

Last of the cost functions that we have considered in the research was proposed by Amendola and Candila (2016). They have introduced it in a way that the overestimated Value-at-Risk models are less penalized than the models that underestimate the risk. Such approach can very clearly distinguish between the conservative and liberal models. Let $X = \frac{\hat{\alpha} - \alpha}{\alpha}$ be the relative deviation from the assumed level of excess ratio. Then:

$$P = \begin{cases} e^X & \text{if } X > 0 \\ 1 & \text{if } X \leq 0 \end{cases} \quad (30)$$

where:

P = penalizing quantity.

P is later included in the cost function and it penalizes more the higher the observed excess ratio $\hat{\alpha}$ becomes:

$$C_{A_t} = \begin{cases} P(1 + |r_t - VaR_\alpha(t)|) & \text{if } r_t \leq VaR_\alpha(t) \\ |r_t - VaR_\alpha(t)| & \text{if } r_t > VaR_\alpha(t) \end{cases} \quad (31)$$

In the three aforementioned cases, the overall score is equal to the sum of the cost functions for each forecast divided by the number of forecasts N , $\frac{\sum_{t=1}^N CF_t}{N}$.

4. Data

Comparison and assessment of different Value-at-Risk techniques have been performed for log rate of return from six separate indexes: three from developed and three from developing countries. The indexes included WIG (Poland), BUX (Bulgaria), RTS (Russia), DJI (United States of America), LSE (Great Britain) and NKX (Japan). Next, over the last 20 years, four distinct periods were selected, each of them lasting 5 years and characterized with different volatility details, including division into learning and testing samples, where testing sample always covers almost the entire last year of the period (250 observations). Period I includes years 2003 – 2007, period II lasts from July 2004 to June 2009, period 3 includes years 2007 – 2011 and period IV includes years 2008 – 2012. Periods were selected so that all of the combinations of turbulent and tranquil volatility measures were covered, hence periods I and II build the model on the tranquil in-sample, whereas periods III and IV are built over the sample of rather high volatility. Periods I and IV use low volatility samples to test the model, while for periods II and III forecasts were built for turbulent out-of-sample.

The choice of the corresponding periods has been done under two assumptions. First of all (as far as it is possible) the periods were supposed to last five full years from 1st of January to

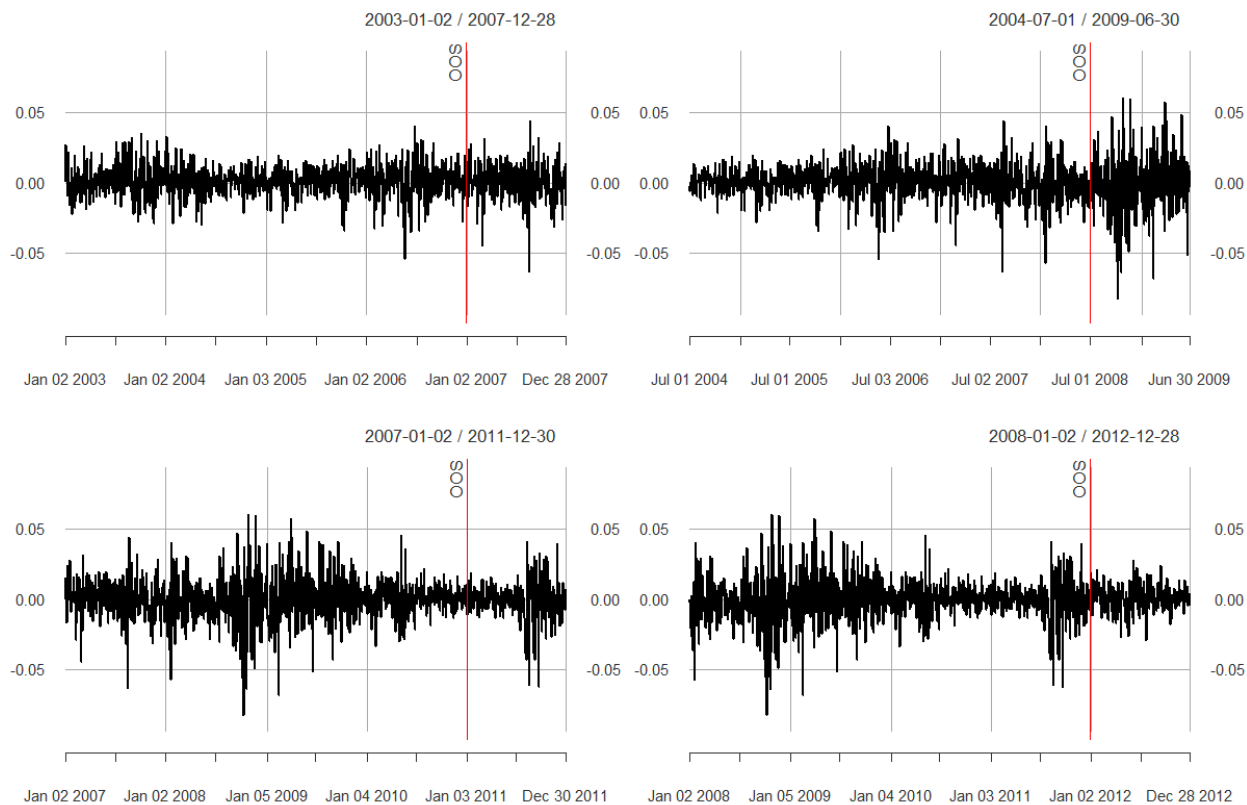


Figure 1: Rate of returns for WIG with illustratory division into four considered time frames

31st of December, with the last year (more or less, it is always 250 observations) being the testing one. Second of all the tranquil and volatile periods were supposed to have high relative volatility. In the end, four aforementioned periods were formed that fit these rules the best for each of the considered indices (with one exception of period II, which starts in July of 2004, so that the most of the volatility that was an effect of the world crisis is captured in the testing sample).

All of the considered approaches were built to satisfy Basel III requirements, therefore using the rolling-window method one-day before the Value-at-Risk estimates were calculated. For confidence level, two values were selected: 99% and 97.5% (99% is established by the current legal framework (and therefore reported in hereby work), but is going to be replaced with 97.5% in the forthcoming updates and a will change to expected shortfall as the market risk indicator). The number of observations in the learning sample differs by several days in each period and for each index, but the testing sample always covered 250 observations. To sum up, the evaluation will be considered for each approach, for six different indexes in four different time periods at two separate confidence levels. All of the calculations were performed in R.

5. Results

5.1. Single models forecasts

Based on the results of statistical tests for both considered confidence levels, the best models include FHS with skewed normal distribution, GARCH models with both skewed student's t and standardized student's t distributions, the EVT 90-th quantile model and the CAViaR GARCH model. Each of these models has scored the least in terms of rejections of null hypotheses (p -value that was considered was equal to 0.01). On the other hand, the worst model appears to be CAViaR AD, for which the null hypotheses were rejected the most number of times. The remaining models have fairly good results with only a few rejections that appear mostly in the IV studied period for both studied confidence levels.

If we add information about the number of exceedances to the information the above situation is further clarified. In terms of exceedances, the best model is GARCH with standardized student's t distribution – the mean number of exceedances is the lowest among all models for the Value-at-Risk at 1% and 2.5% confidence levels, being nearly equal to the assumed confidence levels in the IV period (only 4 yellow zones and no red zones in terms of Basel Traffic Light test.) Similarly the GARCH with skewed student's t distribution does not achieve any exceedances for the Value-at-Risk at 1% confidence level, but the mean number of exceedances is slightly higher throughout all the studied periods (6 yellow zones). Although the CAViaR GARCH model has a lower excess ratio for the Value-at-Risk at 1% confidence level (mean of 3.67 exceedances, twice in the yellow zone), there occurred several more exceedances for the Value-at-Risk at 2.5% confidence level, in addition the mean numbers of exceedances are also higher than in case of GARCH with standardized student's t distribution. However, it must be noted that in case of II and III periods where models have been tested on relatively volatile samples, the CAViaR GARCH model has been proven to achieve better results than GARCH with skewed student's t distribution. In conclusion, these models almost always received the green light in terms of Basel Traffic Light test, with just a few exceptions.

The remaining models in terms of Value-at-Risk at 1% confidence level have mostly achieved 2 to 3 exceedances in the relatively low volatile periods and as much as 4 or 5 exceedances in the turbulent periods and in terms of Value-at-Risk at 2.5% confidence level 1 or 2 exceedances, regardless of the studied period. Moreover, in terms of CAViaR models, the exceedances were so large (10 – 12 exceedances at 1% confidence level; 20 – 24 exceedances at 2.5% confidence level) that models CAViaR AD earned a red light in the Basel Traffic Light test. The CAViaR in the adaptive approach was not able to match the market conditions in most cases, once again confirming its worst position in this study.

When it comes to the cost of the models, the cheapest in terms of Excessive cost function is FHS with skewed normal distribution, followed by CAViaR GARCH. The most expensive one in terms of all costs is the best models so far – GARCH with standardized student's t distribution. A fairly cheaper solution is its 'sibling' model – GARCH with skewed student's t distribution. Considering all models, the cheapest one appears to be GARCH with skewed normal distribution, but its relatively high number of exceedances increases the cost of usage in terms of bad forecasts. Despite the prevalence with historical simulation, CAViaR AD is one of the most expensive models due to its propensity to exceed.

When comparing the model's correctness in the space of different indices, we have not found any differences between a particular model's number of exceedances for indices from emerging

and developed countries. There are some slight deviations in case of mean number of exceedances for these groups, but it seems that all models are affected. This may impose that the hypothesis about model's differing quality of outcome depending on time series country of origin is highly affected by underlying data, and such deviations should not be attributed to the level of country's development.

The level of volatility in the in- and out-of-sample does influence the model's predictions. Many of the models that performed well in the calm periods are not able to estimate correctly in case of highly volatile periods. These results are in accordance with the literature – parametric models are able to estimate forecasts that would obtain low number of exceedances even in case of high volatility. Semi- or non-parametric models are able to forecast accurately when there are no uncommon events, even more accurately than parametric models, but if they occur, it is likely that parametric model would outdo them in terms of banking system's needs.

In conclusion, the best models that we propose are:

- parametric: GARCH models with skewed student's t and standardized student's t distributions;
- semi- or non-parametric: FHS with skewed normal distribution, EVT model with a threshold set to 90-th quantile and CAViaR model with GARCH-like specification.

Table 1: Benchmarking results for each of the considered Value-at-Risk models for 1% confidence level in each considered time frame

		PERIOD I						PERIOD II						PERIOD III						PERIOD IV					
		No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴
CAVIAR SAV	WIG	4.00	1.60	1.20	3.23	2.76	0.06	5.00	2.00	0.98	5.65	4.97	0.11	3.00	1.20	0.91	3.34	2.95	0.05	2.00	0.80	0.31	2.71	2.34	0.04
	BUX	5.00	2.00	0.63	2.97	2.55	0.08	4.00	1.60	1.37	6.74	5.76	0.10	5.00	2.00	0.55	4.42	3.93	0.10	1.00	0.40	0.06	3.94	3.46	0.04
	RTS	1.00	0.40	3.21	4.55	3.98	0.05	2.00	0.80	2.84	12.22	10.81	0.13	6.00	2.40	2.05	5.53	4.93	0.15	3.00	1.20	0.44	4.77	4.24	0.06
	EC	3.33	1.33	1.68	3.58	3.10	0.07	3.67	1.47	1.73	8.20	7.18	0.11	4.67	1.87	1.17	4.43	3.94	0.10	2.00	0.80	0.27	3.81	3.34	0.05
	DJI	5.00	2.00	0.98	6.17	5.35	0.12	2.00	0.80	1.64	3.32	2.86	0.04	1.00	0.40	0.15	2.40	2.12	0.03	9.00	3.60	0.71	2.05	1.78	0.51
	LSE	5.00	2.00	0.72	2.65	2.29	0.08	5.00	2.00	1.36	7.72	6.83	0.13	3.00	1.20	2.32	4.13	3.62	0.06	0.00	0.00	–	2.96	2.61	0.03
	NKX	4.00	1.60	0.74	2.87	2.51	0.06	5.00	2.00	1.02	6.19	5.33	0.12	4.00	1.60	1.50	3.86	3.45	0.07	0.00	0.00	–	3.18	2.73	0.03
	DC	4.67	1.87	0.81	3.89	3.38	0.09	4.00	1.60	1.34	5.74	5.01	0.10	2.67	1.07	1.32	3.47	3.06	0.05	3.00	1.20	–	2.73	2.37	0.19
	AC	4.00	1.60	1.25	3.74	3.24	0.08	3.83	1.53	1.54	6.97	6.09	0.10	3.67	1.47	1.25	3.95	3.50	0.08	2.50	1.00	–	3.27	2.86	0.12
CAVIAR GARCH	WIG	4.00	1.60	1.18	3.36	2.88	0.06	6.00	2.40	0.69	5.71	5.05	0.15	2.00	0.80	1.12	3.43	3.03	0.04	2.00	0.80	0.38	2.69	2.32	0.03
	BUX	6.00	2.40	0.65	2.89	2.48	0.13	4.00	1.60	0.95	6.91	5.95	0.10	4.00	1.60	0.92	4.32	3.81	0.07	0.00	0.00	–	3.84	3.36	0.04
	RTS	1.00	0.40	3.12	4.35	3.79	0.05	1.00	0.40	5.80	12.29	10.86	0.13	6.00	2.40	1.99	5.68	5.08	0.16	2.00	0.80	0.46	4.73	4.18	0.06
	EC	3.67	1.47	1.65	3.54	3.05	0.08	3.67	1.47	2.48	8.30	7.29	0.13	4.00	1.60	1.34	4.48	3.97	0.09	1.33	0.53	–	3.75	3.29	0.04
	DJI	4.00	1.60	1.26	6.04	5.22	0.09	3.00	1.20	1.02	3.29	2.84	0.05	0.00	0.00	–	2.33	2.04	0.02	7.00	2.80	0.80	2.13	1.84	0.19
	LSE	4.00	1.60	0.93	2.65	2.28	0.06	5.00	2.00	1.24	7.71	6.83	0.13	2.00	0.80	3.12	4.49	3.96	0.05	1.00	0.40	0.53	2.91	2.56	0.03
	NKX	4.00	1.60	0.80	2.87	2.51	0.06	3.00	1.20	1.33	6.27	5.38	0.08	2.00	0.80	2.95	3.75	3.31	0.05	1.00	0.40	0.13	3.14	2.70	0.04
	DC	4.00	1.60	1.00	3.85	3.34	0.07	3.67	1.47	1.20	5.76	5.02	0.09	1.33	0.53	–	3.52	3.10	0.04	3.00	1.20	0.48	2.73	2.37	0.09
	AC	3.83	1.53	1.32	3.69	3.19	0.07	3.67	1.47	1.84	7.03	6.15	0.11	2.67	1.07	–	4.00	3.54	0.07	2.17	0.87	–	3.24	2.83	0.06
CAVIAR AD	WIG	2.00	0.80	1.85	3.75	3.25	0.05	4.00	1.60	1.35	5.54	4.85	0.08	9.00	3.60	1.11	4.28	3.96	0.53	0.00	0.00	–	3.60	3.21	0.04
	BUX	1.00	0.40	0.30	4.17	3.72	0.05	10.00	4.00	3.05	6.65	5.75	0.89	4.00	1.60	2.38	5.75	5.22	0.09	3.00	1.20	0.27	3.01	2.55	0.04
	RTS	1.00	0.40	0.72	5.47	4.91	0.06	11.00	4.40	3.28	10.24	9.10	1.46	7.00	2.80	2.81	8.61	8.02	0.26	3.00	1.20	0.85	4.39	3.86	0.06
	EC	1.33	0.53	0.96	4.46	3.96	0.05	8.33	3.33	2.56	7.47	6.57	0.81	6.67	2.67	2.10	6.21	5.73	0.29	2.00	0.80	–	3.67	3.21	0.05
	DJI	11.00	4.40	1.88	5.15	4.43	1.39	3.00	1.20	0.63	5.42	4.98	0.07	1.00	0.40	0.50	2.50	2.22	0.03	9.00	3.60	0.62	1.93	1.66	0.51
	LSE	7.00	2.80	0.86	2.29	1.93	0.19	12.00	4.80	1.79	5.75	4.99	2.24	4.00	1.60	1.70	6.08	5.58	0.09	0.00	0.00	–	2.81	2.46	0.03
	NKX	4.00	1.60	1.14	2.78	2.42	0.06	7.00	2.80	2.98	6.57	5.78	0.24	10.00	4.00	1.00	4.86	4.49	0.86	12.00	4.80	0.42	1.94	1.57	2.17
	DC	7.33	2.93	1.29	3.41	2.93	0.55	7.33	2.93	1.80	5.91	5.25	0.85	5.00	2.00	1.07	4.48	4.10	0.33	7.00	2.80	–	2.23	1.90	0.90
	AC	4.33	1.73	1.12	3.93	3.45	0.30	7.83	3.13	2.18	6.69	5.91	0.83	5.83	2.33	1.58	5.35	4.92	0.31	4.50	1.80	–	2.95	2.55	0.47
FHS NORM	WIG	4.00	1.60	1.06	3.52	3.04	0.06	5.00	2.00	0.90	5.59	4.91	0.11	3.00	1.20	0.75	3.36	2.97	0.05	2.00	0.80	0.29	2.71	2.34	0.04
	BUX	4.00	1.60	0.76	2.94	2.52	0.06	5.00	2.00	1.22	6.61	5.67	0.12	4.00	1.60	0.93	4.31	3.80	0.07	0.00	0.00	–	3.59	3.10	0.04
	RTS	1.00	0.40	3.02	4.41	3.85	0.05	5.00	2.00	1.90	11.14	9.83	0.17	4.00	1.60	3.07	5.79	5.16	0.09	2.00	0.80	0.02	4.71	4.16	0.06
	EC	3.00	1.20	1.61	3.62	3.14	0.06	5.00	2.00	1.34	7.78	6.80	0.13	3.67	1.47	1.58	4.49	3.98	0.07	1.33	0.53	–	3.67	3.20	0.04
	DJI	3.00	1.20	0.87	5.96	5.12	0.07	3.00	1.20	1.03	3.38	2.93	0.05	2.00	0.80	0.11	2.28	2.01	0.03	10.00	4.00	0.54	2.01	1.74	0.83
	LSE	6.00	2.40	0.75	2.50	2.15	0.12	5.00	2.00	1.38	7.03	6.14	0.13	2.00	0.80	2.07	4.52	4.00	0.05	1.00	0.40	0.11	2.97	2.62	0.03
	NKX	4.00	1.60	1.09	2.83	2.47	0.06	2.00	0.80	1.85	6.55	5.64	0.07	3.00	1.20	2.65	3.52	3.09	0.05	2.00	0.80	0.09	2.97	2.55	0.04
	DC	4.33	1.73	0.91	3.76	3.25	0.08	3.33	1.33	1.42	5.65	4.91	0.08	2.33	0.93	1.61	3.44	3.03	0.04	4.33	1.73	0.25	2.65	2.31	0.30
	AC	3.67	1.47	1.26	3.69	3.19	0.07	4.17	1.67	1.38	6.72	5.86	0.11	3.00	1.20	1.60	3.96	3.50	0.06	2.83	1.13	–	3.16	2.75	0.17
FHS SNORM	WIG	4.00	1.60	1.06	3.49	3.02	0.06	4.00	1.60	1.10	5.52	4.83	0.08	3.00	1.20	0.74	3.36	2.97	0.05	2.00	0.80	0.32	2.72	2.35	0.04
	BUX	4.00	1.60	0.70	2.95	2.53	0.06	6.00	2.40	1.07	6.60	5.68	0.16	4.00	1.60	0.90	4.30	3.79	0.07	1.00	0.40	0.08	3.60	3.12	0.04
	RTS	1.00	0.40	2.76	4.45	3.89	0.05	5.00	2.00	1.93	11.06	9.75	0.17	4.00	1.60	3.11	5.80	5.16	0.09	1.00	0.40	0.06	4.71	4.15	0.05
	EC	3.00	1.20	1.51	3.63	3.15	0.06	5.00	2.00	1.37	7.73	6.75	0.14	3.67	1.47	1.59	4.49	3.98	0.07	1.33	0.53	0.15	3.68	3.21	0.04
	DJI	2.00	0.80	1.21	6.01	5.15	0.07	3.00	1.20	1.01	3.40	2.95	0.05	2.00	0.80	0.11	2.28	2.01	0.03	7.00	2.80	0.70	2.06	1.77	0.19
	LSE	6.00	2.40	0.78	2.49	2.14	0.12	5.00	2.00	1.36	7.00	6.11	0.12	2.00	0.80	2.04	4.54	4.02	0.05	1.00	0.40	0.12	2.96	2.62	0.03
	NKX	4.00	1.60	1.06	2.83	2.47	0.06	3.00	1.20	1.36	6.57	5.71	0.08	3.00	1.20	2.69	3.55	3.12	0.05	1.00	0.40	0.11	2.98	2.55	0.03
	DC	4.00	1.60	1.02	3.78	3.26	0.08	3.67	1.47	1.24	5.66	4.92	0.08	2.33	0.93	1.61	3.46	3.05	0.04	3.00	1.20	0.31	2.67	2.31	0.09
	AC	3.50	1.40	1.26	3.71	3.20	0.07	4.33	1.73	1.31	6.69	5.84	0.11	3.00	1.20	1.60	3.97	3.51	0.06	2.17	0.87	0.23	3.17	2.76	0.06

Table 1: Benchmarking results for each of the considered Value-at-Risk models for 1% confidence level in each considered time frame
(continued)

		PERIOD I						PERIOD II						PERIOD III						PERIOD IV					
		No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴
EVT95	WIG	4.00	1.60	1.17	3.43	2.96	0.06	4.00	1.60	1.16	5.51	4.82	0.08	4.00	1.60	0.62	3.32	2.94	0.06	2.00	0.80	0.38	2.64	2.26	0.03
	BUX	4.00	1.60	0.85	2.80	2.37	0.06	6.00	2.40	1.14	6.48	5.55	0.16	4.00	1.60	0.86	4.33	3.83	0.07	1.00	0.40	0.02	3.59	3.11	0.04
	RTS	1.00	0.40	2.89	4.38	3.82	0.05	5.00	2.00	2.13	10.83	9.51	0.16	4.00	1.60	3.07	5.85	5.22	0.09	3.00	1.20	0.05	4.68	4.16	0.06
	EC	3.00	1.20	1.64	3.54	3.05	0.06	5.00	2.00	1.48	7.61	6.63	0.14	4.00	1.60	1.52	4.50	3.99	0.07	2.00	0.80	0.15	3.64	3.18	0.05
	DJI	2.00	0.80	1.35	5.94	5.09	0.07	3.00	1.20	0.99	3.38	2.94	0.05	2.00	0.80	0.13	2.29	2.01	0.03	9.00	3.60	0.63	1.97	1.70	0.51
	LSE	5.00	2.00	0.86	2.54	2.18	0.08	5.00	2.00	1.28	7.01	6.12	0.12	2.00	0.80	2.26	4.41	3.88	0.05	1.00	0.40	0.14	2.88	2.54	0.03
	NKX	4.00	1.60	0.96	2.99	2.63	0.06	2.00	0.80	1.90	6.58	5.67	0.07	3.00	1.20	2.54	3.66	3.23	0.05	0.00	0.00	–	3.11	2.67	0.03
	DC	3.67	1.47	1.06	3.83	3.30	0.07	3.33	1.33	1.39	5.66	4.91	0.08	2.33	0.93	1.64	3.45	3.04	0.04	3.33	1.33	–	2.66	2.30	0.19
	AC	3.33	1.33	1.35	3.68	3.17	0.06	4.17	1.67	1.43	6.63	5.77	0.11	3.17	1.27	1.58	3.98	3.52	0.06	2.67	1.07	–	3.15	2.74	0.12
EVT90	WIG	5.00	2.00	1.00	3.34	2.87	0.09	6.00	2.40	0.86	5.44	4.78	0.15	4.00	1.60	0.58	3.35	2.97	0.06	2.00	0.80	0.30	2.70	2.33	0.04
	BUX	4.00	1.60	0.82	2.83	2.40	0.06	6.00	2.40	1.29	6.33	5.40	0.16	4.00	1.60	0.85	4.33	3.82	0.07	1.00	0.40	0.00	3.59	3.12	0.04
	RTS	1.00	0.40	2.93	4.44	3.88	0.05	6.00	2.40	1.85	10.64	9.35	0.21	4.00	1.60	3.22	5.68	5.05	0.09	3.00	1.20	0.10	4.64	4.12	0.06
	EC	3.33	1.33	1.58	3.54	3.05	0.06	6.00	2.40	1.33	7.47	6.51	0.17	4.00	1.60	1.55	4.45	3.95	0.07	2.00	0.80	0.13	3.65	3.19	0.05
	DJI	2.00	0.80	1.27	6.05	5.20	0.07	3.00	1.20	0.98	3.40	2.95	0.05	2.00	0.80	0.11	2.29	2.02	0.03	8.00	3.20	0.66	2.02	1.74	0.31
	LSE	5.00	2.00	0.84	2.56	2.19	0.08	5.00	2.00	1.28	7.00	6.12	0.12	3.00	1.20	1.53	4.33	3.82	0.06	1.00	0.40	0.13	2.87	2.53	0.03
	NKX	4.00	1.60	1.00	2.94	2.58	0.06	2.00	0.80	1.98	6.65	5.74	0.07	3.00	1.20	2.39	3.77	3.34	0.05	0.00	0.00	–	3.17	2.72	0.03
	DC	3.67	1.47	1.04	3.85	3.32	0.07	3.33	1.33	1.42	5.68	4.94	0.08	2.67	1.07	1.34	3.46	3.06	0.05	3.00	1.20	–	2.69	2.33	0.13
	AC	3.50	1.40	1.31	3.69	3.19	0.07	4.67	1.87	1.37	6.58	5.72	0.13	3.33	1.33	1.45	3.96	3.50	0.06	2.50	1.00	–	3.17	2.76	0.09
GARCH SNORM	WIG	6.00	2.40	0.92	3.22	2.77	0.13	7.00	2.80	1.10	4.96	4.29	0.22	5.00	2.00	0.77	3.02	2.65	0.08	3.00	1.20	0.41	2.44	2.08	0.04
	BUX	4.00	1.60	0.72	2.91	2.49	0.06	7.00	2.80	1.24	6.11	5.19	0.23	5.00	2.00	1.10	3.96	3.45	0.09	1.00	0.40	0.25	3.28	2.80	0.04
	RTS	1.00	0.40	3.18	3.97	3.41	0.04	9.00	3.60	2.24	9.04	7.78	0.59	5.00	2.00	2.93	5.10	4.47	0.11	5.00	2.00	0.45	4.14	3.64	0.10
	EC	3.67	1.47	1.61	3.37	2.89	0.08	7.67	3.07	1.53	6.70	5.76	0.35	5.00	2.00	1.60	4.02	3.53	0.10	3.00	1.20	0.37	3.29	2.84	0.06
	DJI	4.00	1.60	0.89	5.57	4.76	0.08	4.00	1.60	0.97	3.08	2.63	0.06	3.00	1.20	0.30	2.05	1.78	0.04	9.00	3.60	0.62	1.98	1.70	0.51
	LSE	6.00	2.40	0.87	2.38	2.02	0.12	6.00	2.40	1.49	6.35	5.46	0.16	4.00	1.60	1.64	3.76	3.25	0.07	3.00	1.20	0.20	2.47	2.15	0.04
	NKX	5.00	2.00	0.93	2.76	2.41	0.08	3.00	1.20	1.86	6.17	5.30	0.08	3.00	1.20	2.71	3.53	3.10	0.05	1.00	0.40	0.13	2.97	2.53	0.03
	DC	5.00	2.00	0.90	3.57	3.06	0.10	4.33	1.73	1.44	5.20	4.46	0.10	3.33	1.33	1.55	3.11	2.71	0.05	4.33	1.73	0.31	2.48	2.13	0.19
	AC	4.33	1.73	1.25	3.47	2.98	0.09	6.00	2.40	1.48	5.95	5.11	0.22	4.17	1.67	1.57	3.57	3.12	0.07	3.67	1.47	0.34	2.88	2.49	0.13
GARCH-ST	WIG	4.00	1.60	1.12	3.49	3.01	0.06	7.00	2.80	0.90	5.20	4.54	0.22	4.00	1.60	0.73	3.25	2.86	0.06	2.00	0.80	0.33	2.65	2.28	0.03
	BUX	4.00	1.60	0.57	3.08	2.65	0.06	5.00	2.00	1.34	6.44	5.50	0.12	5.00	2.00	0.85	4.17	3.67	0.10	1.00	0.40	0.18	3.42	2.95	0.04
	RTS	1.00	0.40	2.26	4.93	4.37	0.05	4.00	1.60	2.08	11.33	9.99	0.14	4.00	1.60	3.07	5.81	5.17	0.09	2.00	0.80	0.16	4.61	4.07	0.05
	EC	3.00	1.20	1.32	3.83	3.35	0.06	5.33	2.13	1.44	7.66	6.68	0.16	4.33	1.73	1.55	4.41	3.90	0.08	1.67	0.67	0.22	3.56	3.10	0.04
	DJI	2.00	0.80	0.94	6.31	5.46	0.07	3.00	1.20	0.94	3.44	3.00	0.05	3.00	1.20	0.09	2.30	2.03	0.04	7.00	2.80	0.60	2.15	1.87	0.19
	LSE	5.00	2.00	0.81	2.57	2.21	0.08	5.00	2.00	1.40	6.80	5.91	0.12	3.00	1.20	1.84	4.13	3.62	0.06	1.00	0.40	0.17	2.76	2.41	0.03
	NKX	4.00	1.60	1.08	2.89	2.53	0.06	3.00	1.20	1.70	6.51	5.64	0.08	3.00	1.20	2.68	3.66	3.23	0.05	1.00	0.40	0.03	3.05	2.61	0.03
	DC	3.67	1.47	0.95	3.92	3.40	0.07	3.67	1.47	1.35	5.59	4.85	0.08	3.00	1.20	1.54	3.36	2.96	0.05	3.00	1.20	0.27	2.65	2.30	0.09
	AC	3.33	1.33	1.13	3.88	3.37	0.06	4.50	1.80	1.39	6.62	5.76	0.12	3.67	1.47	1.54	3.88	3.43	0.07	2.33	0.93	0.25	3.11	2.70	0.06
GARCH-T	WIG	6.00	2.40	0.89	3.24	2.79	0.13	7.00	2.80	1.11	4.96	4.29	0.22	4.00	1.60	0.81	3.15	2.77	0.06	3.00	1.20	0.30	2.55	2.19	0.04
	BUX	4.00	1.60	0.65	2.98	2.56	0.06	6.00	2.40	1.21	6.34	5.42	0.16	5.00	2.00	0.80	4.21	3.71	0.10	1.00	0.40	0.16	3.45	2.97	0.04
	RTS	1.00	0.40	3.21	4.04	3.48	0.04	7.00	2.80	2.08	10.02	8.74	0.27	4.00	1.60	3.43	5.37	4.73	0.08	4.00	1.60	0.32	4.32	3.81	0.07
	EC	3.67	1.47	1.58	3.42	2.94	0.08	6.67	2.67	1.47	7.11	6.15	0.22	4.33	1.73	1.68	4.24	3.74	0.08	2.67	1.07	0.26	3.44	2.99	0.05
	DJI	2.00	0.80	1.31	5.87	5.01	0.07	3.00	1.20	1.17	3.20	2.75	0.05	3.00	1.20	0.22	2.14	1.87	0.04	8.00	3.20	0.63	2.04	1.76	0.31
	LSE	6.00	2.40	0.84	2.40	2.04	0.12	6.00	2.40	1.45	6.39	5.51	0.16	3.00	1.20	2.09	3.92	3.41	0.05	2.00	0.80	0.20	2.60	2.26	0.03
	NKX	5.00	2.00	1.01	2.71	2.36	0.08	3.00	1.20	2.05	6.16	5.28	0.08	3.00	1.20	2.92	3.48	3.05	0.05	2.00	0.80	0.09	2.91	2.48	0.04
	DC	4.33	1.73	1.05	3.66	3.14	0.09	4.00	1.60	1.55	5.25	4.51	0.10	3.00	1.20	1.75	3.18	2.78	0.05	4.00	1.60	0.31	2.51	2.17	0.13
	AC	4.00	1.60	1.32	3.54	3.04	0.08	5.33	2.13	1.51	6.18	5.33	0.16	3.67	1.47	1.71	3.71	3.26	0.06	3.33	1.33	0.28	2.98	2.58	0.09

Table 1: Benchmarking results for each of the considered Value-at-Risk models for 1% confidence level in each considered time frame (continued)

		PERIOD I						PERIOD II						PERIOD III						PERIOD IV					
		No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴
GARCH-T STAND	WIG	3.00	1.20	1.29	3.67	3.18	0.05	4.00	1.60	1.15	5.57	4.88	0.08	3.00	1.20	0.51	3.59	3.21	0.05	1.00	0.40	0.14	2.93	2.55	0.03
	BUX	3.00	1.20	0.63	3.18	2.75	0.05	5.00	2.00	0.91	6.86	5.93	0.12	2.00	0.80	0.69	4.80	4.26	0.06	0.00	0.00	–	3.96	3.48	0.04
	RTS	1.00	0.40	2.33	5.08	4.52	0.05	2.00	0.80	2.17	12.91	11.51	0.14	3.00	1.20	3.34	6.59	5.94	0.08	0.00	0.00	–	5.35	4.77	0.05
	EC	2.33	0.93	1.42	3.98	3.49	0.05	3.67	1.47	1.41	8.45	7.44	0.12	2.67	1.07	1.51	4.99	4.47	0.06	0.33	0.13	–	4.08	3.60	0.04
	DJI	2.00	0.80	0.56	6.94	6.09	0.08	2.00	0.80	0.88	3.81	3.36	0.05	1.00	0.40	0.07	2.58	2.29	0.03	6.00	2.40	0.60	2.26	1.97	0.12
	LSE	5.00	2.00	0.72	2.67	2.31	0.08	5.00	2.00	1.08	7.29	6.41	0.13	2.00	0.80	2.05	4.75	4.23	0.06	0.00	0.00	–	3.17	2.82	0.03
	NKX	4.00	1.60	1.00	3.00	2.65	0.06	3.00	1.20	1.39	6.81	5.95	0.08	3.00	1.20	2.59	3.77	3.34	0.05	0.00	0.00	–	3.11	2.66	0.03
	DC	3.67	1.47	0.76	4.20	3.68	0.07	3.33	1.33	1.12	5.97	5.24	0.09	2.00	0.80	1.57	3.70	3.29	0.05	2.00	0.80	–	2.84	2.48	0.06
	AC	3.00	1.20	1.09	4.09	3.58	0.06	3.50	1.40	1.26	7.21	6.34	0.10	2.33	0.93	1.54	4.35	3.88	0.05	1.17	0.47	–	3.46	3.04	0.05

Note:

This table above presents the number of exceedances, excess ratio, Caporin, Abad & Benito's and Chlebus' cost functions, divided into the analysed periods and models for each analysed index. The analysis was performed at 250 Value-at-Risk forecasts generated by each model.

¹ Abad & Benito's cost function

² Caporin's cost function

³ Chlebus' cost function

⁴ Amendola & Candila' cost function

Table 2: Statistical benchmarking results for the considered Value-at-Risk models at 1% confidence level in each considered time frame

		PERIOD I			PERIOD II			PERIOD III			PERIOD IV		
		UC test ¹	CC test ²	MV UC test ³	UC test ¹	CC test ²	MV UC test ³	UC test ¹	CC test ²	MV UC test ³	UC test ¹	CC test ²	MV UC test ³
CAVIAR SAV	WIG	0.38	0.64	0.63	0.16	0.34	0.35	0.76	0.06	0.95	0.74	0.93	0.94
	BUX	0.16	0.34	0.31	0.38	0.64	0.21	0.16	0.34	0.09	0.28	0.55	0.51
	RTS	0.28	0.55	0.13	0.74	0.93	0.03	0.06	0.15	0.04	0.76	0.92	0.58
	DJI	0.16	0.34	0.37	0.74	0.93	0.87	0.28	0.55	0.55	0	0	0
	LSE	0.16	0.34	0.35	0.16	0.34	0.35	0.76	0.92	0.23	0.02	0.08	0.05
	NKX	0.38	0.64	0.56	0.16	0.34	0.06	0.38	0.09	0.02	0.02	0.08	0.07
CAVIAR GARCH	WIG	0.38	0.64	0.63	0.06	0.15	0.11	0.74	0.93	0.87	0.74	0.93	0.94
	BUX	0.06	0.15	0.11	0.38	0.64	0.56	0.38	0.64	0.16	0.02	0.08	0.08
	RTS	0.28	0.55	0.13	0.28	0.55	0.01	0.06	0.15	0.11	0.74	0.93	0.87
	DJI	0.38	0.64	0.56	0.76	0.92	0.23	0.02	0.08	0.08	0.02	0.05	0.05
	LSE	0.38	0.64	0.56	0.16	0.34	0.35	0.74	0.93	0.94	0.28	0.55	0.01
	NKX	0.38	0.64	0.67	0.76	0.92	0.3	0.74	0.02	0.87	0.28	0.55	0.33
CAVIAR AD	WIG	0.74	0.93	0.94	0.38	0.64	0.1	0	0	0	0.02	0.08	0.08
	BUX	0.28	0.55	0.13	0	0	0	0.38	0.64	0.63	0.76	0.92	0.58
	RTS	0.28	0.55	0.13	0	0	0	0.02	0.05	0.02	0.76	0.92	0.79
	DJI	0	0	0	0.76	0.92	0.06	0.28	0.55	0.55	0	0	0
	LSE	0.02	0.03	0	0	0	0	0.38	0.64	0.21	0.02	0.08	0.03
	NKX	0.38	0.64	0.04	0.02	0.05	0.01	0	0	0	0	0	0
FHS NORM	WIG	0.38	0.64	0.63	0.16	0.34	0.35	0.76	0.06	0.88	0.74	0.93	0.78
	BUX	0.38	0.64	0.42	0.16	0.34	0.02	0.38	0.64	0.16	0.02	0.08	0.08
	RTS	0.28	0.55	0.13	0.16	0.34	0.06	0.38	0.64	0.63	0.74	0.93	0.87
	DJI	0.76	0.92	0.79	0.76	0.92	0.88	0.74	0.93	0.87	0	0	0
	LSE	0.06	0.15	0.14	0.16	0.34	0.35	0.74	0.93	0.57	0.28	0.55	0.33
	NKX	0.38	0.64	0.56	0.74	0.93	0.53	0.76	0.06	0.58	0.74	0.93	0.22
FHS SNORM	WIG	0.38	0.64	0.63	0.38	0.64	0.67	0.76	0.06	0.88	0.74	0.93	0.94
	BUX	0.38	0.64	0.63	0.06	0.15	0.03	0.38	0.64	0.16	0.28	0.55	0.51
	RTS	0.28	0.55	0.13	0.16	0.34	0.12	0.38	0.64	0.63	0.28	0.55	0.55
	DJI	0.74	0.93	0.15	0.76	0.92	0.88	0.74	0.93	0.94	0.02	0.05	0.02
	LSE	0.06	0.15	0.14	0.16	0.34	0.35	0.74	0.93	0.57	0.28	0.55	0.51
	NKX	0.38	0.64	0.67	0.76	0.92	0.79	0.76	0.06	0.58	0.28	0.55	0.33
EVT95	WIG	0.38	0.64	0.63	0.38	0.64	0.67	0.38	0.09	0.16	0.74	0.93	0.94
	BUX	0.38	0.64	0.42	0.06	0.15	0.05	0.38	0.64	0.16	0.28	0.55	0.33
	RTS	0.28	0.55	0.13	0.16	0.34	0.12	0.38	0.64	0.63	0.76	0.92	0.58
	DJI	0.74	0.93	0.15	0.76	0.92	0.88	0.74	0.93	0.87	0	0	0
	LSE	0.16	0.34	0.21	0.16	0.34	0.35	0.74	0.93	0.57	0.28	0.55	0.33
	NKX	0.38	0.64	0.67	0.74	0.93	0.53	0.76	0.06	0.58	0.02	0.08	0.07
EVT90	WIG	0.16	0.34	0.35	0.06	0.15	0.11	0.38	0.09	0.16	0.74	0.93	0.87
	BUX	0.38	0.64	0.42	0.06	0.15	0.03	0.38	0.64	0.16	0.28	0.55	0.33
	RTS	0.28	0.55	0.13	0.06	0.15	0.09	0.38	0.64	0.63	0.76	0.92	0.58
	DJI	0.74	0.93	0.3	0.76	0.92	0.88	0.74	0.93	0.87	0.01	0.02	0.01
	LSE	0.16	0.34	0.21	0.16	0.34	0.35	0.76	0.92	0.58	0.28	0.55	0.33
	NKX	0.38	0.64	0.56	0.74	0.93	0.53	0.76	0.06	0.58	0.02	0.08	0.07
GARCH SNORM	WIG	0.06	0.15	0.04	0.02	0.05	0.02	0.16	0.08	0.23	0.76	0.92	0.88
	BUX	0.38	0.64	0.02	0.02	0.05	0.06	0.16	0.34	0.01	0.28	0.55	0.33
	RTS	0.28	0.55	0.33	0	0	0	0.16	0.34	0.35	0.16	0.34	0.01
	DJI	0.38	0.64	0.38	0.38	0.64	0.38	0.76	0.92	0.88	0	0	0
	LSE	0.06	0.15	0.05	0.06	0.15	0.11	0.38	0.64	0.63	0.76	0.92	0.58
	NKX	0.16	0.34	0.37	0.76	0.92	0.53	0.76	0.06	0.58	0.28	0.55	0.33

Table 2: Statistical benchmarking results for the considered Value-at-Risk models at 1% confidence level in each considered time frame (continued)

		PERIOD I			PERIOD II			PERIOD III			PERIOD IV		
		UC test ¹	CC test ²	MV UC test ³	UC test ¹	CC test ²	MV UC test ³	UC test ¹	CC test ²	MV UC test ³	UC test ¹	CC test ²	MV UC test ³
GARCH-ST	WIG	0.38	0.64	0.63	0.02	0.05	0.04	0.38	0.09	0.63	0.74	0.93	0.94
	BUX	0.38	0.64	0.02	0.16	0.34	0.21	0.16	0.34	0.09	0.28	0.55	0.33
	RTS	0.28	0.55	0.01	0.38	0.64	0.21	0.38	0.64	0.63	0.74	0.93	0.87
	DJI	0.74	0.93	0.78	0.76	0.92	0.95	0.76	0.92	0.88	0.02	0.05	0.06
	LSE	0.16	0.34	0.37	0.16	0.34	0.35	0.76	0.92	0.88	0.28	0.55	0.55
	NKX	0.38	0.64	0.56	0.76	0.92	0.79	0.76	0.06	0.58	0.28	0.55	0.33
GARCH-T	WIG	0.06	0.15	0.11	0.02	0.05	0.06	0.38	0.09	0.63	0.76	0.92	0.88
	BUX	0.38	0.64	0.02	0.06	0.15	0.14	0.16	0.34	0.01	0.28	0.55	0.33
	RTS	0.28	0.55	0.33	0.02	0.05	0.02	0.38	0.64	0.56	0.38	0.64	0.16
	DJI	0.74	0.93	0.15	0.76	0.92	0.3	0.76	0.92	0.95	0.01	0.02	0.02
	LSE	0.06	0.15	0.03	0.06	0.15	0.16	0.76	0.92	0.79	0.74	0.93	0.87
	NKX	0.16	0.34	0.12	0.76	0.92	0.02	0.76	0.06	0.88	0.74	0.93	0.22
GARCH-T STAND	WIG	0.76	0.92	0.95	0.38	0.64	0.67	0.76	0.06	0.58	0.28	0.55	0.51
	BUX	0.76	0.92	0.23	0.16	0.34	0.23	0.74	0.93	0.87	0.02	0.08	0.02
	RTS	0.28	0.55	0.01	0.74	0.93	0.78	0.76	0.92	0.58	0.02	0.08	0.07
	DJI	0.74	0.93	0.87	0.74	0.93	0.87	0.28	0.55	0.33	0.06	0.15	0.17
	LSE	0.16	0.34	0.09	0.16	0.34	0.23	0.74	0.93	0.57	0.02	0.08	0.07
	NKX	0.38	0.64	0.63	0.76	0.92	0.58	0.76	0.06	0.58	0.02	0.08	0.07

Note:

The table above presents the results of formal tests: Kupiec (unconditional coverage), Christoffersen (conditional coverage), and Multivariate Unconditional test for each analysed index divided into models in all analysed periods. Green fields indicate p-values greater than 5%. EC, DC and AC stand for emerging, developed and all countries. Source - own calculations.

¹ p-value of the unconditional coverage test

² p-value of the conditional coverage test

³ p-value of the multivariate unconditional coverage test for 1% and 2.5% VaR

5.2. Mixed Forecasts

Based on the findings from the previous section, a set of the best models have been identified and mixed forecasts have been calculated in the following way:

1. The overall number of ‘the best’ models is equal to 5, hence sets of combinations of from 2 to 5 models have been created. This kind of method produces 26 mixed forecasts of different models: 10 of 2 models, 10 of 3 models, 5 of 4 models and 1 of 5 models.
2. The next step considered using aforementioned techniques for mixed forecasts: minimum and arithmetic mean of the generated sets. This produces 52 different Value-at-Risk forecasts.

Logically, the minimum technique will produce the most conservative forecasts with high costs. Value-at-Risk forecasts for both considered confidence levels in the minimum variant is the most interesting case. We have either 0 or 1 exceedance, and there is very little space to propose a better model. In terms of cost, these combinations of models are very costly, however the lowest costs are achieved by combinations of FHS with skewed normal distribution and EVT with 90% quantile, models that also achieved low costs in the single results. Considering the financial institution’s tendency towards protection, such forecast combinations may be a very interesting approach to Value-at-Risk estimation, given that the costs of using such methods increase relatively slower than the increase in model’s ‘faultlessness’. In addition, the cost ratio of the ensembled model to its components is at maximum equal to 2 - 3, which given its non-defectiveness might be a very superior option for banking institutions. One of the best models produced in that step is a combination of GARCH with standardized student’s t distribution and one of the two semi-parametrical models: either FHS with skewed normal distribution or CAViaR with GARCH-like specification.

For the averaging ensembling method, all the combined forecasts are fairly similar in their exceedances and the costs. One of the best combinations is the mean of CAViaR GARCH, GARCH with standardized student’s t distribution and either FHS with skewed normal distribution or EVT with 90% quantile. These forecasts obtained the lowest number of yellow lights in terms of Basel Traffic Lights test, with circa 3 – 4 exceedances in all the observed periods. However, none of these mean combined forecasts have obtained results that would be better than the original single models. The best model here was a combination of GARCH models with skewed and standardized student’s t distributions.

In conclusion, ensembling might provide a very good model, assuming that we use the minimum method - in such a case the combined predictive power of the best models is able to nearly perfectly forecast the volatile periods while simultaneously being relatively cheap in relation to the component models.

Table 3: Benchmarking results for exemplary Value-at-Risk models combined by minimum ensembling method in each considered time frame

		PERIOD I						PERIOD II						PERIOD III						PERIOD IV					
		No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴
minimum CAVIAR GARCH GARCH-T STAND	WIG	1.00	0.40	1.68	4.68	4.17	0.05	0.00	0.00	-	8.38	7.60	0.08	0.00	0.00	-	7.46	7.03	0.07	0.00	0.00	-	3.54	3.15	0.04
	BUX	0.00	0.00	-	4.52	4.05	0.05	0.00	0.00	-	16.51	15.45	0.17	0.00	0.00	-	7.15	6.56	0.07	0.00	0.00	-	4.62	4.13	0.05
	RTS	0.00	0.00	-	7.52	6.94	0.08	0.00	0.00	-	27.17	25.68	0.27	2.00	0.80	0.95	10.09	9.46	0.11	0.00	0.00	-	5.79	5.21	0.06
	EC	0.33	0.13	-	5.57	5.05	0.06	0.00	0.00	-	17.36	16.24	0.17	0.67	0.27	-	8.23	7.68	0.08	0.00	0.00	-	4.65	4.17	0.05
	DJI	0.00	0.00	-	4.78	4.36	0.05	0.00	0.00	-	14.48	13.50	0.14	0.00	0.00	-	11.04	10.48	0.11	0.00	0.00	-	3.72	3.37	0.04
	LSE	1.00	0.40	0.71	4.82	4.44	0.05	0.00	0.00	-	18.56	17.61	0.19	0.00	0.00	-	12.87	12.38	0.13	0.00	0.00	-	3.89	3.44	0.04
	NKX	0.00	0.00	-	11.81	10.92	0.12	0.00	0.00	-	7.54	7.05	0.08	0.00	0.00	-	2.98	2.69	0.03	1.00	0.40	0.15	3.23	2.90	0.04
	DC	0.33	0.13	-	7.14	6.57	0.07	0.00	0.00	-	13.53	12.72	0.14	0.00	0.00	-	8.96	8.52	0.09	0.33	0.13	-	3.61	3.24	0.04
	AC	0.33	0.13	-	6.35	5.81	0.06	0.00	0.00	-	15.44	14.48	0.15	0.33	0.13	-	8.60	8.10	0.09	0.17	0.07	-	4.13	3.70	0.04
minimum FHS SNORM EVT90	WIG	2.00	0.80	0.98	4.48	3.99	0.05	1.00	0.40	0.22	7.95	7.20	0.08	0.00	0.00	-	6.37	5.93	0.06	0.00	0.00	-	3.29	2.90	0.03
	BUX	2.00	0.80	0.47	3.88	3.44	0.05	0.00	0.00	-	14.46	13.39	0.14	1.00	0.40	0.42	6.46	5.90	0.07	0.00	0.00	-	4.16	3.68	0.04
	RTS	1.00	0.40	0.49	6.23	5.67	0.07	0.00	0.00	-	21.68	20.19	0.22	2.00	0.80	1.75	9.31	8.67	0.10	1.00	0.40	0.06	5.09	4.53	0.05
	EC	1.67	0.67	0.64	4.86	4.37	0.06	0.33	0.13	-	14.70	13.60	0.15	1.00	0.40	-	7.38	6.83	0.08	0.33	0.13	-	4.18	3.70	0.04
	DJI	1.00	0.40	0.05	4.22	3.82	0.05	0.00	0.00	-	13.85	12.86	0.14	0.00	0.00	-	11.40	10.84	0.11	0.00	0.00	-	3.50	3.15	0.04
	LSE	1.00	0.40	0.98	4.55	4.17	0.05	0.00	0.00	-	15.38	14.43	0.15	0.00	0.00	-	11.25	10.77	0.11	0.00	0.00	-	3.63	3.18	0.04
	NKX	0.00	0.00	-	10.17	9.27	0.10	0.00	0.00	-	7.33	6.85	0.07	0.00	0.00	-	2.83	2.54	0.03	5.00	2.00	0.25	2.63	2.34	0.08
	DC	0.67	0.27	-	6.31	5.75	0.07	0.00	0.00	-	12.19	11.38	0.12	0.00	0.00	-	8.49	8.05	0.08	1.67	0.67	-	3.25	2.89	0.05
	AC	1.17	0.47	-	5.59	5.06	0.06	0.17	0.07	-	13.44	12.49	0.14	0.50	0.20	-	7.93	7.44	0.08	1.00	0.40	-	3.72	3.30	0.05
minimum FHS SNORM GARCH-ST	WIG	2.00	0.80	0.98	4.48	3.99	0.05	1.00	0.40	0.22	7.95	7.20	0.08	1.00	0.40	0.15	6.00	5.59	0.06	0.00	0.00	-	3.14	2.75	0.03
	BUX	2.00	0.80	0.07	4.27	3.84	0.05	0.00	0.00	-	14.83	13.76	0.15	1.00	0.40	0.55	6.33	5.77	0.07	0.00	0.00	-	3.96	3.47	0.04
	RTS	1.00	0.40	0.28	6.43	5.88	0.07	0.00	0.00	-	21.68	20.19	0.22	2.00	0.80	1.83	9.22	8.59	0.10	1.00	0.40	0.06	5.09	4.53	0.05
	EC	1.67	0.67	0.44	5.06	4.57	0.06	0.33	0.13	-	14.82	13.72	0.15	1.33	0.53	0.84	7.19	6.65	0.08	0.33	0.13	-	4.06	3.58	0.04
	DJI	0.00	0.00	-	4.27	3.86	0.04	0.00	0.00	-	13.85	12.86	0.14	0.00	0.00	-	11.40	10.84	0.11	0.00	0.00	-	3.50	3.15	0.04
	LSE	1.00	0.40	1.12	4.41	4.03	0.05	0.00	0.00	-	15.00	14.05	0.15	1.00	0.40	0.12	10.96	10.52	0.11	0.00	0.00	-	3.60	3.15	0.04
	NKX	0.00	0.00	-	10.57	9.67	0.11	0.00	0.00	-	7.33	6.85	0.07	0.00	0.00	-	2.83	2.54	0.03	2.00	0.80	0.30	2.84	2.52	0.04
	DC	0.33	0.13	-	6.42	5.85	0.07	0.00	0.00	-	12.06	11.25	0.12	0.33	0.13	-	8.39	7.96	0.09	0.67	0.27	-	3.31	2.94	0.04
	AC	1.00	0.40	-	5.74	5.21	0.06	0.17	0.07	-	13.44	12.49	0.14	0.83	0.33	-	7.79	7.31	0.08	0.50	0.20	-	3.69	3.26	0.04
minimum EVT90 GARCH-T STAND	WIG	1.00	0.40	1.68	4.68	4.17	0.05	1.00	0.40	0.28	7.89	7.15	0.08	0.00	0.00	-	6.77	6.34	0.07	0.00	0.00	-	3.54	3.15	0.04
	BUX	0.00	0.00	-	4.50	4.03	0.04	0.00	0.00	-	15.87	14.80	0.16	0.00	0.00	-	7.15	6.56	0.07	0.00	0.00	-	4.62	4.13	0.05
	RTS	0.00	0.00	-	7.14	6.57	0.07	0.00	0.00	-	25.46	23.97	0.25	2.00	0.80	0.95	10.09	9.46	0.11	0.00	0.00	-	5.79	5.21	0.06
	EC	0.33	0.13	-	5.44	4.92	0.06	0.33	0.13	-	16.41	15.31	0.17	0.67	0.27	-	8.00	7.45	0.08	0.00	0.00	-	4.65	4.17	0.05
	DJI	0.00	0.00	-	4.50	4.09	0.04	0.00	0.00	-	14.48	13.50	0.14	0.00	0.00	-	11.04	10.48	0.11	0.00	0.00	-	3.72	3.37	0.04
	LSE	1.00	0.40	0.98	4.55	4.17	0.05	0.00	0.00	-	15.38	14.43	0.15	0.00	0.00	-	11.25	10.77	0.11	0.00	0.00	-	3.63	3.18	0.04
	NKX	0.00	0.00	-	11.69	10.80	0.12	0.00	0.00	-	7.54	7.05	0.08	0.00	0.00	-	2.98	2.69	0.03	1.00	0.40	0.32	3.06	2.73	0.03
	DC	0.33	0.13	-	6.92	6.35	0.07	0.00	0.00	-	12.47	11.66	0.12	0.00	0.00	-	8.43	7.98	0.08	0.33	0.13	-	3.47	3.09	0.04
	AC	0.33	0.13	-	6.18	5.64	0.06	0.17	0.07	-	14.44	13.48	0.15	0.33	0.13	-	8.21	7.72	0.08	0.17	0.07	-	4.06	3.63	0.04

Note:

The table above presents the results of ensembling using minimum of up to 5 models. Only two the best and two the worst models are presented. The total number of models created was 26.

¹ Abad & Benito's cost function² Caporin's cost function³ Chlebus' cost function⁴ Amendola & Candila' cost function

Table 4: Benchmarking results for exemplary Value-at-Risk models combined by averaging ensembling method in each considered time frame

		PERIOD I						PERIOD II						PERIOD III						PERIOD IV						
		No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	No. exceedances	Excess ratio (%)	CAB (%) ¹	CC (%) ²	CEC (%) ³	CA ⁴	
mean	CAVIAR GARCH FHS SNORM	WIG	8.00	3.20	0.91	2.67	2.23	0.32	9.00	3.60	1.42	4.34	3.71	0.53	11.00	4.40	1.43	2.78	2.47	1.36	5.00	2.00	0.57	2.10	1.75	0.08
		BUX	6.00	2.40	1.07	2.23	1.82	0.12	11.00	4.40	2.72	5.47	4.63	1.41	6.00	2.40	1.90	3.32	2.81	0.13	3.00	1.20	0.34	2.71	2.26	0.04
		RTS	4.00	1.60	0.90	3.44	2.91	0.06	9.00	3.60	3.66	8.96	7.79	0.59	9.00	3.60	1.98	4.40	3.84	0.54	5.00	2.00	0.83	3.67	3.16	0.09
		EC	6.00	2.40	0.96	2.78	2.32	0.17	9.67	3.87	2.60	6.25	5.38	0.84	8.67	3.47	1.77	3.50	3.04	0.68	4.33	1.73	0.58	2.83	2.39	0.07
		DJI	11.00	4.40	0.70	2.16	1.84	1.35	10.00	4.00	1.66	5.64	4.87	0.87	8.00	3.20	1.53	3.41	2.96	0.33	2.00	0.80	0.07	2.34	2.01	0.03
		LSE	9.00	3.60	0.69	2.33	2.01	0.51	10.00	4.00	2.76	5.52	4.79	0.88	3.00	1.20	4.02	3.11	2.66	0.05	2.00	0.80	0.48	2.54	2.11	0.03
		NKX	10.00	4.00	1.40	4.92	4.23	0.86	6.00	2.40	1.48	2.85	2.43	0.13	3.00	1.20	0.33	1.90	1.64	0.03	17.00	6.80	0.61	1.67	1.43	22.61
		DC	10.00	4.00	0.93	3.14	2.69	0.91	8.67	3.47	1.97	4.67	4.03	0.63	4.67	1.87	1.96	2.81	2.42	0.14	7.00	2.80	0.39	2.18	1.85	7.56
AC	8.00	3.20	0.94	2.96	2.51	0.54	9.17	3.67	2.28	5.46	4.70	0.74	6.67	2.67	1.87	3.15	2.73	0.41	5.67	2.27	0.48	2.50	2.12	3.81		
mean	CAVIAR GARCH GARCH-T STAND	WIG	5.00	2.00	1.29	2.82	2.35	0.08	8.00	3.20	1.50	4.43	3.79	0.34	10.00	4.00	1.42	2.92	2.60	0.84	4.00	1.60	0.57	2.23	1.88	0.05
		BUX	6.00	2.40	0.86	2.43	2.02	0.12	11.00	4.40	2.44	5.72	4.90	1.41	6.00	2.40	1.61	3.59	3.09	0.13	2.00	0.80	0.23	2.94	2.48	0.04
		RTS	2.00	0.80	1.56	3.65	3.11	0.04	7.00	2.80	3.59	9.80	8.59	0.27	6.00	2.40	2.54	4.76	4.16	0.15	3.00	1.20	0.95	3.97	3.44	0.05
		EC	4.33	1.73	1.24	2.97	2.49	0.08	8.67	3.47	2.51	6.65	5.76	0.67	7.33	2.93	1.86	3.76	3.28	0.37	3.00	1.20	0.58	3.05	2.60	0.05
		DJI	10.00	4.00	0.72	2.21	1.89	0.83	7.00	2.80	2.12	5.83	5.00	0.23	6.00	2.40	1.92	3.51	3.03	0.13	1.00	0.40	0.02	2.46	2.12	0.03
		LSE	9.00	3.60	0.63	2.38	2.07	0.51	10.00	4.00	2.64	5.63	4.90	0.88	3.00	1.20	4.00	3.12	2.68	0.05	2.00	0.80	0.48	2.54	2.11	0.03
		NKX	6.00	2.40	1.79	5.29	4.52	0.15	6.00	2.40	1.38	2.94	2.52	0.13	3.00	1.20	0.26	1.98	1.71	0.03	13.00	5.20	0.64	1.78	1.53	3.51
		DC	8.33	3.33	1.05	3.29	2.83	0.50	7.67	3.07	2.05	4.80	4.14	0.41	4.00	1.60	2.06	2.87	2.47	0.07	5.33	2.13	0.38	2.26	1.92	1.19
AC	6.33	2.53	1.14	3.13	2.66	0.29	8.17	3.27	2.28	5.72	4.95	0.54	5.67	2.27	1.96	3.31	2.88	0.22	4.17	1.67	0.48	2.66	2.26	0.62		
mean	GARCH-ST GARCH-T STAND	WIG	5.00	2.00	1.23	2.88	2.41	0.08	8.00	3.20	1.58	4.36	3.72	0.34	11.00	4.40	1.40	2.82	2.51	1.36	4.00	1.60	0.55	2.25	1.89	0.05
		BUX	5.00	2.00	0.85	2.58	2.17	0.08	11.00	4.40	2.58	5.59	4.76	1.41	6.00	2.40	1.58	3.62	3.12	0.13	2.00	0.80	0.22	2.95	2.49	0.04
		RTS	1.00	0.40	2.93	3.80	3.24	0.04	9.00	3.60	3.34	9.26	8.11	0.59	5.00	2.00	2.98	4.81	4.19	0.10	4.00	1.60	0.81	3.87	3.35	0.07
		EC	3.67	1.47	1.67	3.09	2.61	0.07	9.33	3.73	2.50	6.40	5.53	0.78	7.33	2.93	1.99	3.75	3.27	0.53	3.33	1.33	0.53	3.02	2.58	0.05
		DJI	11.00	4.40	0.73	2.14	1.82	1.35	10.00	4.00	1.61	5.68	4.92	0.87	6.00	2.40	1.92	3.51	3.03	0.13	3.00	1.20	0.08	2.30	1.98	0.04
		LSE	9.00	3.60	0.63	2.39	2.07	0.51	10.00	4.00	2.73	5.55	4.81	0.88	3.00	1.20	4.07	3.06	2.61	0.05	2.00	0.80	0.48	2.54	2.12	0.03
		NKX	6.00	2.40	1.83	5.25	4.48	0.15	6.00	2.40	1.43	2.89	2.47	0.13	3.00	1.20	0.33	1.91	1.64	0.03	12.00	4.80	0.66	1.82	1.56	2.18
		DC	8.67	3.47	1.06	3.26	2.79	0.67	8.67	3.47	1.93	4.70	4.07	0.63	4.00	1.60	2.11	2.83	2.43	0.07	5.67	2.27	0.40	2.22	1.88	0.75
AC	6.17	2.47	1.37	3.17	2.70	0.37	9.00	3.60	2.21	5.55	4.80	0.70	5.67	2.27	2.05	3.29	2.85	0.30	4.50	1.80	0.47	2.62	2.23	0.40		
mean	FHS SNORM EVT90 GARCH-ST	WIG	8.00	3.20	0.90	2.68	2.24	0.32	9.00	3.60	1.48	4.28	3.66	0.53	12.00	4.80	1.40	2.71	2.40	2.20	4.00	1.60	0.67	2.13	1.78	0.05
		BUX	6.00	2.40	0.93	2.36	1.95	0.12	11.00	4.40	2.89	5.31	4.46	1.41	6.00	2.40	1.84	3.37	2.86	0.13	3.00	1.20	0.31	2.75	2.29	0.04
		RTS	2.00	0.80	1.69	3.53	2.98	0.04	9.00	3.60	4.31	8.36	7.17	0.59	8.00	3.20	2.21	4.42	3.85	0.34	5.00	2.00	0.95	3.56	3.05	0.09
		EC	5.33	2.13	1.18	2.86	2.39	0.16	9.67	3.87	2.89	5.98	5.10	0.84	8.67	3.47	1.82	3.50	3.04	0.89	4.00	1.60	0.64	2.81	2.37	0.06
		DJI	11.00	4.40	0.79	2.08	1.76	1.35	10.00	4.00	1.82	5.49	4.72	0.87	8.00	3.20	1.51	3.43	2.98	0.33	3.00	1.20	0.17	2.21	1.89	0.04
		LSE	9.00	3.60	0.70	2.32	2.01	0.51	11.00	4.40	2.57	5.46	4.75	1.41	3.00	1.20	4.06	3.06	2.62	0.05	2.00	0.80	0.47	2.55	2.12	0.03
		NKX	11.00	4.40	1.35	4.85	4.17	1.38	6.00	2.40	1.51	2.82	2.40	0.13	4.00	1.60	0.30	1.85	1.59	0.05	17.00	6.80	0.59	1.68	1.45	22.61
		DC	10.33	4.13	0.94	3.08	2.64	1.08	9.00	3.60	1.96	4.59	3.96	0.80	5.00	2.00	1.95	2.78	2.40	0.14	7.33	2.93	0.41	2.15	1.82	7.56
AC	7.83	3.13	1.06	2.97	2.52	0.62	9.33	3.73	2.43	5.29	4.53	0.82	6.83	2.73	1.89	3.14	2.72	0.52	5.67	2.27	0.53	2.48	2.10	3.81		

Note:

The table above presents the results of ensembling using average of up to 5 models. Only two the best and two the worst models are presented. The total number of models created was 26.

¹ Abad & Benito's cost function² Caporin's cost function³ Chlebus' cost function⁴ Amendola & Candila' cost function

Table 5: Ratio of the ensembled and component models for exemplary Value-at-Risk models combined by minimum ensembling method in each considered time frame

			PERIOD I					PERIOD II					PERIOD III					PERIOD IV				
			No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴	No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴	No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴	No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴
minimum CAVIAR GARCH GARCH-T STAND	CAVIAR GARCH	EC	11.00	–	0.63	0.60	1.39	–	–	0.48	0.45	0.73	6.00	–	0.54	0.52	1.06	–	–	0.81	0.79	0.92
		DC	12.00	–	0.54	0.51	0.93	–	–	0.43	0.39	0.63	–	–	0.39	0.36	0.45	9.00	–	0.75	0.73	2.32
		AC	11.50	–	0.58	0.55	1.13	–	–	0.46	0.42	0.69	8.00	–	0.47	0.44	0.75	13.00	–	0.78	0.76	1.54
	GARCH-T STAND	EC	7.00	–	0.71	0.69	0.89	–	–	0.49	0.46	0.66	4.00	–	0.61	0.58	0.73	–	–	0.88	0.86	0.91
		DC	11.00	–	0.59	0.56	1.00	–	–	0.44	0.41	0.63	–	–	0.41	0.39	0.51	6.00	–	0.79	0.77	1.63
		AC	9.00	–	0.64	0.62	0.95	–	–	0.47	0.44	0.65	7.00	–	0.51	0.48	0.62	7.00	–	0.84	0.82	1.23
minimum FHS SNORM EVT90	FHS SNORM	EC	1.80	2.34	0.75	0.72	1.03	15	–	0.53	0.50	0.93	3.67	–	0.61	0.58	0.89	4.00	–	0.88	0.87	0.98
		DC	6.00	–	0.60	0.57	1.26	–	–	0.46	0.43	0.69	–	–	0.41	0.38	0.53	1.80	–	0.82	0.80	1.70
		AC	3.00	–	0.66	0.63	1.16	26	–	0.50	0.47	0.82	6.00	–	0.50	0.47	0.70	2.17	–	0.85	0.84	1.37
	EVT90	EC	2.00	2.46	0.73	0.70	1.17	18	–	0.51	0.48	1.17	4.00	–	0.60	0.58	0.95	6.00	–	0.87	0.86	1.05
		DC	5.50	–	0.61	0.58	1.05	–	–	0.47	0.43	0.68	–	–	0.41	0.38	0.55	1.80	–	0.83	0.81	2.47
		AC	3.00	–	0.66	0.63	1.11	28	–	0.49	0.46	0.95	6.67	–	0.50	0.47	0.74	2.50	–	0.85	0.84	1.82
minimum FHS SNORM GARCH-ST	FHS SNORM	EC	1.80	3.41	0.72	0.69	1.00	15	–	0.52	0.49	0.92	2.75	1.88	0.62	0.60	0.90	4.00	–	0.91	0.90	1.00
		DC	12.00	–	0.59	0.56	1.27	–	–	0.47	0.44	0.70	7.00	–	0.41	0.38	0.53	4.50	–	0.81	0.79	2.41
		AC	3.50	–	0.65	0.61	1.14	26	–	0.50	0.47	0.82	3.60	–	0.51	0.48	0.70	4.33	–	0.86	0.85	1.65
	GARCH-ST	EC	1.80	2.98	0.76	0.73	1.03	16	–	0.52	0.49	1.08	3.25	1.84	0.61	0.59	1.06	5.00	–	0.88	0.86	1.01
		DC	11.00	–	0.61	0.58	1.07	–	–	0.46	0.43	0.70	9.00	–	0.40	0.37	0.57	4.50	–	0.80	0.78	2.40
		AC	3.33	–	0.68	0.65	1.05	27	–	0.49	0.46	0.91	4.40	–	0.50	0.47	0.80	4.67	–	0.84	0.83	1.65
minimum EVT90 GARCH-T STAND	EVT90	EC	10.00	–	0.65	0.62	1.16	18	–	0.46	0.43	1.05	6.00	–	0.56	0.53	0.89	–	–	0.78	0.77	0.97
		DC	11.00	–	0.56	0.52	0.98	–	–	0.46	0.42	0.66	–	–	0.41	0.38	0.56	9.00	–	0.77	0.75	3.47
		AC	10.50	–	0.60	0.57	1.06	28	–	0.46	0.42	0.88	10.00	–	0.48	0.45	0.72	15.00	–	0.78	0.76	2.06
	GARCH-T STAND	EC	7.00	–	0.73	0.71	0.91	11	–	0.51	0.49	0.70	4.00	–	0.62	0.60	0.75	–	–	0.88	0.86	0.91
		DC	11.00	–	0.61	0.58	1.03	–	–	0.48	0.45	0.69	–	–	0.44	0.41	0.54	6.00	–	0.82	0.80	1.69
		AC	9.00	–	0.66	0.64	0.98	21	–	0.50	0.47	0.69	7.00	–	0.53	0.50	0.65	7.00	–	0.85	0.84	1.25

¹ Abad & Benito's cost function² Caporin's cost function³ Chlebus' cost function⁴ Amendola & Candila' cost function

Table 6: Ratio of the ensembled and component models for exemplary Value-at-Risk models combined by averaging ensembling method in each considered time frame

			PERIOD I					PERIOD II					PERIOD III					PERIOD IV				
			No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴	No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴	No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴	No. exceedances	CAB ¹	CC ²	CEC ³	CA ⁴
mean CAVIAR GARCH FHS SNORM	CAVIAR GARCH	EC	0.61	1.72	1.27	1.31	0.47	0.38	0.95	1.33	1.35	0.15	0.46	0.76	1.28	1.31	0.13	0.31	–	1.33	1.37	0.62
		DC	0.40	1.07	1.23	1.24	0.07	0.42	0.61	1.23	1.24	0.14	0.29	–	1.25	1.28	0.30	0.43	1.25	1.25	1.28	0.01
		AC	0.48	1.40	1.25	1.27	0.14	0.40	0.80	1.29	1.31	0.14	0.40	–	1.27	1.30	0.16	0.38	–	1.29	1.33	0.02
	FHS SNORM	EC	0.50	1.57	1.31	1.36	0.34	0.52	0.53	1.24	1.26	0.16	0.42	0.90	1.28	1.31	0.10	0.31	0.26	1.30	1.34	0.61
		DC	0.40	1.09	1.20	1.21	0.09	0.42	0.63	1.21	1.22	0.14	0.50	0.82	1.23	1.26	0.33	0.43	0.80	1.22	1.25	0.01
		AC	0.44	1.34	1.25	1.28	0.13	0.47	0.57	1.23	1.24	0.15	0.45	0.86	1.26	1.29	0.14	0.38	0.48	1.27	1.30	0.02
mean CAVIAR GARCH GARCH-T STAND	CAVIAR GARCH	EC	0.85	1.33	1.19	1.22	0.95	0.42	0.99	1.25	1.27	0.19	0.55	0.72	1.19	1.21	0.24	0.44	–	1.23	1.26	0.90
		DC	0.48	0.95	1.17	1.18	0.14	0.48	0.59	1.20	1.21	0.21	0.33	–	1.23	1.25	0.57	0.56	1.27	1.21	1.23	0.07
		AC	0.61	1.16	1.18	1.20	0.25	0.45	0.81	1.23	1.24	0.20	0.47	–	1.21	1.23	0.29	0.52	–	1.22	1.25	0.10
	GARCH-T STAND	EC	0.54	1.15	1.34	1.40	0.61	0.42	0.56	1.27	1.29	0.17	0.36	0.82	1.33	1.36	0.17	0.11	–	1.34	1.38	0.88
		DC	0.44	0.73	1.28	1.30	0.15	0.43	0.55	1.24	1.26	0.21	0.50	0.76	1.29	1.33	0.64	0.38	–	1.26	1.29	0.05
		AC	0.47	0.95	1.31	1.35	0.21	0.43	0.55	1.26	1.28	0.18	0.41	0.79	1.31	1.35	0.24	0.28	–	1.30	1.35	0.08
mean GARCH-ST GARCH-T STAND	GARCH-ST	EC	0.82	0.79	1.24	1.28	0.86	0.57	0.58	1.20	1.21	0.21	0.59	0.78	1.18	1.19	0.15	0.50	0.43	1.18	1.20	0.81
		DC	0.42	0.89	1.20	1.22	0.10	0.42	0.70	1.19	1.19	0.13	0.75	0.73	1.19	1.22	0.68	0.53	0.66	1.19	1.22	0.11
		AC	0.54	0.83	1.22	1.25	0.17	0.50	0.63	1.19	1.20	0.17	0.65	0.75	1.18	1.20	0.22	0.52	0.53	1.19	1.21	0.16
	GARCH-T STAND	EC	0.64	0.85	1.29	1.34	0.74	0.39	0.56	1.32	1.35	0.15	0.36	0.76	1.33	1.37	0.12	0.10	–	1.35	1.40	0.80
		DC	0.42	0.72	1.29	1.32	0.11	0.38	0.58	1.27	1.29	0.14	0.50	0.75	1.31	1.35	0.65	0.35	–	1.28	1.32	0.08
		AC	0.49	0.80	1.29	1.33	0.17	0.39	0.57	1.30	1.32	0.14	0.41	0.75	1.32	1.36	0.18	0.26	–	1.32	1.36	0.13
mean FHS SNORM EVT90 GARCH-ST	FHS SNORM	EC	0.56	1.28	1.27	1.32	0.36	0.52	0.47	1.29	1.32	0.16	0.42	0.87	1.28	1.31	0.08	0.33	0.24	1.31	1.35	0.69
		DC	0.39	1.08	1.23	1.23	0.08	0.41	0.63	1.23	1.24	0.11	0.47	0.82	1.24	1.27	0.32	0.41	0.76	1.24	1.27	0.01
		AC	0.45	1.19	1.25	1.27	0.11	0.46	0.54	1.27	1.29	0.14	0.44	0.85	1.26	1.29	0.11	0.38	0.44	1.28	1.32	0.02
	EVT90	EC	0.62	1.35	1.24	1.28	0.40	0.62	0.46	1.25	1.28	0.21	0.46	0.85	1.27	1.30	0.08	0.50	0.21	1.30	1.35	0.74
		DC	0.35	1.10	1.25	1.26	0.06	0.37	0.72	1.24	1.25	0.10	0.53	0.69	1.24	1.28	0.34	0.41	–	1.25	1.28	0.02
		AC	0.45	1.23	1.24	1.27	0.11	0.50	0.57	1.24	1.26	0.16	0.49	0.77	1.26	1.29	0.12	0.44	–	1.28	1.32	0.02
	GARCH-ST	EC	0.56	1.12	1.34	1.40	0.37	0.55	0.50	1.28	1.31	0.19	0.50	0.85	1.26	1.29	0.09	0.42	0.35	1.27	1.31	0.69
		DC	0.35	1.00	1.27	1.29	0.06	0.41	0.69	1.22	1.23	0.10	0.60	0.79	1.21	1.24	0.35	0.41	0.66	1.23	1.26	0.01
		AC	0.43	1.07	1.31	1.34	0.10	0.48	0.57	1.25	1.27	0.15	0.54	0.82	1.24	1.26	0.13	0.41	0.47	1.25	1.29	0.02

¹ Abad & Benito's cost function² Caporin's cost function³ Chlebus' cost function⁴ Amendola & Candila' cost function

6. Conclusion

Given all the results, we have identified the five best models that provide Value-at-Risk forecasts with good quality. These models are: CAViaR with Indirect GARCH approach, the GARCH model with both standardized student's t and skewed student's t distributions, FHS with skewed normal distribution and EVT with 90% quantile. All these models provided sufficiently good forecasts for both 1% and 2.5% confidence levels that only in very rare conditions obtained as much as yellow light of Basel Traffic Light test.

Moreover, these models might be divided into conservative and liberal groups. We would like to propose the GARCH model with standardized student's t distribution as the most conservative one. Another model that can be put in the same category as the former is the GARCH model with skewed student's t distribution, which was just a little bit less accurate in its predictions, but its costs were much lower. Remaining models are in our sense liberal ones due to their relatively very low costs. It should be noted that these models are not in any sense bad quality, but almost always had worse results in terms of excess ratios, which might be an incentive for the financial institutions to use the more accurate ones. The best models that we have selected in this study cover both the parametric and semi-/non-parametric groups of models. Parametric models include GARCH models with skewed student's t and standardized student's t distributions, which due to their conservatism are able to obtain the lowest excess ratios. Non- and semi-parametric models include FHS with skewed normal distribution, the EVT model with a threshold set to 90-th quantile and CAViaR model with GARCH-like specification, which are more liberal than the former group, but the lack of quick reaction to high and spontaneous increases in the volatility.

In this study, we also researched the quality of forecast combinations, and we have found out that the ensembling method that uses a minimum of n model's forecasts can provide forecasts that are nearly perfect (0 – 1 exceedances) with a relatively small increase in the cost functions values related to its component models. The model built using that technique is not extremely expensive, which combining with the previous conclusion might be interesting for financial institutions looking for more conservative ways to forecast market risk. The best combination of the models that achieved circa 1/3 exceedance on average was a combination of the GARCH(1,1) model with standardized student's t distribution and either FHS with skewed normal distribution or the GARCH-like specified CAViaR model.

Furthermore, we did not find any models that would over- or underestimate Value-at-Risk depending on the underlying time series country of origin, but we confirmed that level of volatility does have an impact on this.

References

- Abad, Pilar, Sonia Benito, and Carmen Lopez. 2015. "The Role of the Loss Function in Value-at-Risk Comparisons." Journal Article. *Journal of Risk Model Validation* 9: 1–19. <https://doi.org/10.21314/JRMV.2015.132>.
- Abad, Pilar, Sonia Benito, and Carmen López. 2014. "A Comprehensive Review of Value at Risk Methodologies." Journal Article. *The Spanish Review of Financial Economics* 12 (1): 15–32. <https://doi.org/10.1016/j.srfe.2013.06.001>.
- Amendola, A. and V. Candila. 2016. "valuation of volatility predictions in a VaR framework." Journal Article. *Quantitative Finance* 16 (5): 695–709. <http://dx.doi.org/10.1080/14697688.2015.1062122>.
- Angelidis, Timotheos, Alexandros Benos, and Stavros Degiannakis. 2004. "The Use of Garch Models in Var Estimation." Journal Article. *Statistical Methodology* 1 (1): 105–28. <https://doi.org/10.1016/j.stamet.2004.06.001>.

2004.08.004.

———. 2007. “A Robust Var Model Under Different Time Periods and Weighting Schemes.” Journal Article. *Review of Quantitative Finance Accounting* 28 (2): 187–201. <https://doi.org/10.1007/s11156-006-0010-y>.

Armstrong, J. Scott. 2001. “Evaluating Forecasting Methods.” Book Section. In *Principles of Forecasting: A Handbook for Researchers and Practitioners*, edited by J. Scott Armstrong, 443–72. Boston, MA: Springer US. https://doi.org/10.1007/978-0-306-47630-3_20.

Bao, Yong, Tae-Hwy Lee, and Burak Saltoglu. 2006. “Evaluating Predictive Performance of Value-at-Risk Models in Emerging Markets: A Reality Check.” Journal Article. *Journal of Forecasting* 25 (2): 101–28. <https://doi.org/doi:10.1002/for.977>.

Barone-Adesi, Giovanni, Kostas Giannopoulos, and Les Vosper. 1999. “VaR Without Correlations for Portfolios of Derivative Securities.” Journal Article. *The Journal of Futures Markets* 19 (5): 583–602. [https://doi.org/doi:10.1002/\(SICI\)1096-9934\(199908\)19:5<583::AID-FUT5>3.0.CO;2-S](https://doi.org/doi:10.1002/(SICI)1096-9934(199908)19:5<583::AID-FUT5>3.0.CO;2-S).

Bates, J. M., and C. W. J. Granger. 1969. “The Combination of Forecasts.” Journal Article. *Journal of the Operational Research Society* 20 (4): 451–68. <https://doi.org/10.1057/jors.1969.103>.

Bollerslev, Tim, and Jeffrey M. Wooldridge. 1992. “Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances.” Journal Article. *Econometric Reviews* 11 (2): 143–72. <https://doi.org/10.1080/07474939208800229>.

Buczyński, Mateusz, and Marcin Chlebus. 2018. “Comparison of Semi-Parametric and Benchmark Value-at-Risk Models in Several Time Periods with Different Volatility Levels.” Journal Article. *E-Finanse* 14: 67–82. <https://e-finanse.com/current-issue/?number=57&id=407>.

Burchi, A, and D Martelli. 2016. “Measuring Market Risk in the Light of Basel Iii: New Evidence from Frontier Markets.” Book Section. In *Handbook of Frontier Markets*, 99–122. Elsevier.

Caporin, Massimiliano. 2008. “Evaluating Value-at-Risk Measures in Presence of Long Memory Conditional Volatility.” Journal Article. *Journal of Risk* 10 (3). <https://doi.org/10.21314/JOR.2008.172>.

Chlebus, Marcin. 2016. “EWS-Garch: New Regime Switching Approach to Forecast Value-at-Risk.” Journal Article. *Central European Economic Journal* 3 (50): 1–25.

Christoffersen, Peter F. 1998. “Evaluating Interval Forecasts.” Journal Article. *International Economic Review* 39 (4): 841–62. <https://doi.org/10.2307/2527341>.

Danielsson, Jon, and Casper G. De Vries. 2000. “Value-at-Risk and Extreme Returns.” Journal Article. *Annales d'Économie et de Statistique*, no. 60: 239–70. <https://doi.org/10.2307/20076262>.

Degiannakis, Stavros, Christos Floros, and Alexandra Livada. 2012. “Evaluating Value-at-Risk Models Before and After the Financial Crisis of 2008: International Evidence.” Journal Article. *Managerial Finance* 38 (4): 436–52.

Diebold, Francis X. 1988. “Serial Correlation and the Combination of Forecasts.” Journal Article. *Journal of Business & Economic Statistics* 6 (1): 105–11. <https://doi.org/10.2307/1391423>.

Djaković, Vladimir, and Slobodan Radiscaron. 2010. “Application of Var in Emerging Markets: A Case of Selected Central and Eastern European Countries.” Journal Article. *African Journal of Business Management* 4 (17): 3666–80.

Embrechts, Paul, Sidney I. Resnick, and Gennady Samorodnitsky. 1999. “Extreme Value Theory as a Risk Management Tool.” Journal Article. *North American Actuarial Journal* 3 (2): 30–41. <https://doi.org/10.1080/10920277.1999.10595797>.

- Engle, Robert F., and Simone Manganelli. 2004. "CAViaR." Journal Article. *Journal of Business & Economic Statistics* 22 (4): 367–81. <https://doi.org/10.1198/073500104000000370>.
- Gerlach, Richard H., Cathy W. S. Chen, and Nancy Y. C. Chan. 2011. "Bayesian Time-Varying Quantile Forecasting for Value-at-Risk in Financial Markets." Journal Article. *Journal of Business & Economic Statistics* 29 (4): 481–92. <https://doi.org/10.1198/jbes.2010.08203>.
- Huang, Alex YiHou. 2009. "A Value-at-Risk Approach with Kernel Estimator." Journal Article. *Applied Financial Economics* 19 (5): 379–95. <https://doi.org/10.1080/09603100701857906>.
- Jorion, Philippe. 2010. "Risk Management." Journal Article. *Annual Review of Financial Economics* 2 (1): 347–65. <https://EconPapers.repec.org/RePEc:anr:refeco:v:2:y:2010:p:347-365>.
- Koenker, Roger, and Gilbert Bassett, Jr. 1978. "Regression Quantiles." Journal Article. *Econometrica* 46 (1): 33–50. <https://EconPapers.repec.org/RePEc:anr:refeco:v:2:y:2010:p:347-365>.
- Kupiec, Paul H. 1995. "Techniques for Verifying the Accuracy of Risk Measurement Models." Journal Article. *The Journal of Derivatives* 3 (2): 73–84. <https://doi.org/10.3905/jod.1995.407942>.
- Lopez, Jose A. 1998. "Methods for Evaluating Value-at-Risk Estimates." Journal Article. *Economic Policy Review* 4 (3). <https://doi.org/http://dx.doi.org/10.2139/ssrn.1029673>.
- McAleer, Michael, Juan-Ángel Jiménez-Martín, and Teodosio Pérez-Amaral. 2013. "Has the Basel Accord Improved Risk Management During the Global Financial Crisis?" Journal Article. *The North American Journal of Economics and Finance* 26: 250–65.
- McNeil, Alexander J. 1999. "Extreme Value Theory for Risk Managers." Journal Article. *Departement Mathematik ETH Zentrum*.
- McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools Revised Edition*. Book. Economics Books. Princeton University Press. <https://doi.org/DOI>:
- Miletic, Mirjana, and Sinisa Miletic. 2015. "Performance of Value at Risk Models in the Midst of the Global Financial Crisis in Selected Cee Emerging Capital Markets." Journal Article. *Economic Research-Ekonomska Istraživanja* 28 (1): 132–66. <https://doi.org/10.1080/1331677X.2015.1028243>.
- Mitra, Sovan, and Tong Ji. 2010. "Risk Measures in Quantitative Finance." Journal Article. *International Journal of Business Continuity and Risk Management* 1 (2): 125–35. <https://EconPapers.repec.org/RePEc:ids:ijbcrm:v:1:y:2010:i:2:p:125-135>.
- Mutu, S, P Balogh, and D Moldovan. 2011. "The Efficiency of Value at Risk Models on Central and Eastern European Stock Markets." Journal Article. *International Journal of Mathematics and Computers in Simulation* 5 (2): 110–17.
- Ozun, Alper, Atilla Cifter, and Sait Yilmazer. 2010. "Filtered Extreme-value Theory for Value-at-risk Estimation: Evidence from Turkey." Journal Article 11 (2): 164–79. <https://doi.org/doi:10.1108/15265941011025189>.
- Pérignon, Christophe, and Daniel R Smith. 2008. "A New Approach to Comparing Var Estimation Methods." Journal Article 16 (2): 54–66. <https://doi.org/10.3905/JOD.2008.16.2.054> %J The Journal of Derivatives.
- Polanski, Arnold, and Evarist Stoja. 2010. "Incorporating Higher Moments into Value-at-Risk Forecasting." Journal Article. *Journal of Forecasting* 29 (6): 523–35. <https://doi.org/doi:10.1002/for.1155>.
- Smith, Richard L. 1987. "Estimating Tails of Probability Distributions." Journal Article. *Ann. Statist.* 15 (3): 1174–1207. <https://doi.org/10.1214/aos/1176350499>.

- Yu, Philip L. H., Wai Keung Li, and Shusong Jin. 2010. "On Some Models for Value-at-Risk." Journal Article. *Econometric Reviews* 29 (5-6): 622–41. <https://doi.org/10.1080/07474938.2010.481972>.
- Žiković, Saša. 2007. "Testing Popular Var Models in Eu New Member and Candidate States." Journal Article. *Zbornik Radova Ekonomskog Fakulteta U Rijeci: Časopis Za Ekonomsku Teoriju I Praksu* 25 (2): 325–46.



UNIVERSITY OF WARSAW

FACULTY OF ECONOMIC SCIENCES

44/50 DŁUGA ST.

00-241 WARSAW

WWW.WNE.UW.EDU.PL