VaR forecasts during the COVID-19 crisis: Comparing GARCH models performances

Abstract

The worldwide spread of COVID-19 was responsible for the second deepest fall in the American stock markets history when the share prices plummet more than 13% on March 13, 2020. The sharp downturn generated by the global pandemic put financial risk models to the test. Hence, we compare the performance of GARCH models to forecast 1%-VaR for the Dow Jones returns during the COVID-19 period. We consider the effects of extreme events by employing the wavelets-based technique for outliers' correction. From an initial set of 72 models, the Model Confidence Set procedure is implemented to select specifications with the higher predictive ability. Backtesting measures indicate that asymmetric models with skewed-heavy-tailed distributions are superior to monitor the risk. Moreover, correcting additive outliers in the series leads to a more accurate 1%-VaR coverage during the crisis.

Keywords: **Covid-19; GARCH; Model Confidence Set; outliers; Value-at-risk; wavelets.**

1 Introduction

Coronavirus or SARS-CoV2 was first reported late December 2019, in Wuhan City, China. Since then, it has caused the highly contagious COVID-19 disease that has spread all over the globe, with severe consequences for the worldwide economy (Goodell, 2020). Perceived as a real threat, the possibility of a prolonged slowdown increased the risk aversion in stock markets and initiated a broad assets sell-off. The Dow Jones Industrial Average index plummets 1,031.61 points, and the Cboe Volatility index (VIX) surged to its one-year highest level on February 24, 2020, with the investors' uncertainty rising. Stock prices continued on a downward trend as public health officials tried to find ways to contain the contagion of the virus, and governments were taking measures to cushion its impacts on the economy. The Dow Jones industrial average index tumbled officially into a Bear Market on March 11, 2020, after the World Health Organization (WHO) declared COVID-19 as a global pandemic. On March 13, 2020, the American market experienced its most profound drop since the "Black Monday" in 1987, with the Down Jones index falling more than 13%, whereas, on March 16, 2020, the VIX reached 82.69 points, surpassing its previous record of November 21, 2008. Figure 1 illustrates the increasing volatility as the COVID-19 went global after a relatively calm period for the financial markets. The upper panel depicts the Dow Jones Industrial Average returns from January 02, 2018, to May 11, 2020, and the bottom panel shows the VIX index for the same period.

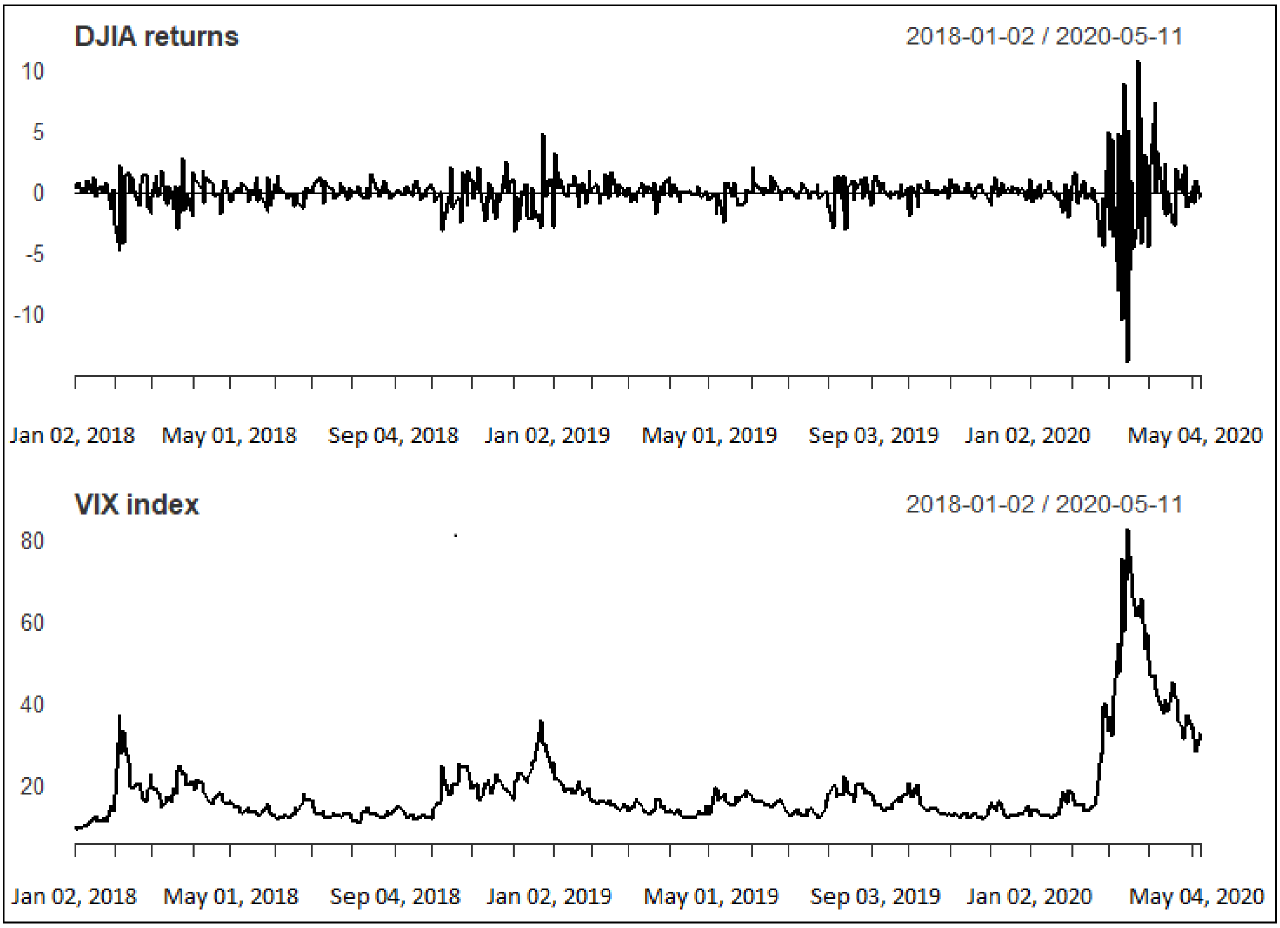

Figure 1: Dow Jones daily returns and VIX index, ranging from January 2, 2018, to May 11, 2020.

The rapid depreciation of stock prices in response to the Covid-19 outbreak poses new challenges for risk models used to monitor the capital exposure of financial institutions. As Jorion (2009) highlights for the Subprime period, the outset of a crisis comes to reinforce the importance of risk management. The effects of crises on stock markets are widely documented by the literature, with several studies investigating the behaviour of risk models under different markets conditions. For instance, Ho, Burridge, Cadle, & Theobald (2000) assess VaR forecasts for Asian stocks over a sequence of financial shocks that hit the markets in the second half of the nineties. Chen, Gerlach, Lin, & Lee (2012) examine the performance of parametric and Bayesian GARCH models for the Subprime period. Grané & Veiga (2014) calculate minimum capital risk requirements during financial crises by correcting the series for additive outliers effects through the wavelets-based technique. Bernardi & Catania (2016) compares the efficiency of GARCH and GAS models to produce VaR forecasts for the European Sovereign Debt period. Slim, Koubaa, & Bensaida (2017) perform VaR forecasts under different model distributions to evaluate long and short trading positions during and post-crisis. Song, Liu, & Sriboonchitta (2019) use Copulas models to estimate the VaR for emerging and developed markets, before and after highly volatile events.

A sudden shift on market conditions may impact on the performance of risk models with consequences for the capital exposure of financial institutions. Bao, Lee, & Saltoglu (2006) observe that most VaR models produce similar results under relatively stable periods, performing poorly during the crisis when the risk becomes more difficult to predict. As Dias (2013) points out, VaR models generally seem to behave adequately during non-crisis periods, but tend to underestimate the risk in the interim of a crisis, which means that regulatory capital may not cover large portfolio losses.

As in other highly volatile times, the outbreak of COVID-19 intensifies the necessity of evaluating risk measures to prevent investors from higher losses. Hence, the objective of this study is to compare the performance of GARCH-types modes to forecast the 1%-VaR for the Dow Jones Industrial Average returns during the initial days of the COVID-19 crisis. Forecasts are also appraised after correcting the presence of additive outliers in the series by employing the wavelets-based technique suggested by Grané & Veiga (2010). Considering models with and without outliers correction and six different distributions, the predictive ability of an initial set containing 72 specifications is tested through the Model Confidence Set procedure developed by Hansen, Lunde, & Nason (2011). Backtesting measures are also performed in order to assess the most accurate 1%-VaR forecasts for the period.

This study adds to the literature by expanding the outliers treatment previously performed by Grané & Veiga (2014) for more complex GARCH models. Withing the VaR framework, our results indicate that not only choosing the most appropriate specifications but also taking extreme observations into account can provide significant improvements for risk management, especially during more difficult markets conditions.

Besides this introduction, methodological procedures are discussed in section 2, data and results are presented in section 3, and conclusions in section 4.

2 Methodology

In this section, we discuss the methodology to assess the downside risk for the Dow Jones returns for the COVID-19 crisis. We start by estimating the conditional variance of returns through a variety of specifications derived from the classical ARCH/GARCH models of Engle (1982) and Bollerslev (1986). Thenceforth, the volatility expressed as conditional standard deviations of returns is used to calculate the risk measures, which are analysed through different approaches.

2.1 Models' setup

In order to deal with the autocorrelation of the series, the conditional mean of returns (r_t) is estimated as an AR(1) process, such that $r_t = \mu + \rho r_{t-1} + \varepsilon_{t-1}$. Thus, given ε_t the noise term, the GARCH(1,1) model is defined as follows:

$$\begin{aligned} \varepsilon_t =& \sigma_t z_t \\ \sigma_t^2 =& \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{1}$$

with $z_t \overset{i.i.d.}{\sim} N(0,1)$ for $t = (1, ..T)$ and conditional variance equals to σ_t^2. Parameters in equation 1 satisfies a stationary process such that $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$ and $(\alpha + \beta) < 1$, with the persistence of the conditional variance given by the sum of α and β.

Besides the standard GARCH(1,1), five other conditional heteroscedasticity specifications are performed to obtain the one-day ahead volatility of returns σ_t^2. The following equations defines the six models utilised in this study:

GARCH(1,1)	$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$	Bollerslev (1986)
iGARCH(1,1)	$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + (1-\alpha)\, \sigma_{t-1}^2$	Engle & Bollerslev (1986)
eGARCH(1,1)	$\log(\sigma_t^2) = \omega + \alpha z_{t-1}^2 + \gamma\left(\lvert z_{t-1}\rvert - E\lvert z_{t-1}\rvert\right) + \beta \log\left(\sigma_{t-1}^2\right)$	Nelson (1991)
gjrGARCH(1,1)	$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma I\left(\varepsilon_{t-1} < 0\right) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$	Glosten et al. (1993)
apARCH(1,1)	$\sigma_t^\delta = \omega + \alpha(\lvert\varepsilon_{t-1}\rvert - \gamma \varepsilon_{t-1})^\delta + \beta \sigma_t^\delta$	Ding, Granger, & Engle (1993)
csGARCH(1,1)	$\sigma_t^2 = q_t + \alpha\left(\varepsilon_{t-1}^2 - q_{t-1}\right) + \beta\left(\sigma_{t-1}^2 - q_{t-1}\right)$	Engle & Lee (1999)
	$q_t = \omega + \rho q_{t-1} + \phi\left(\varepsilon_{t-1}^2 - \sigma_{t-1}^2\right)$	

Despite the effectiveness of these models to describe the volatility process, some specifications may not capture the occurrence of extreme observations in the series, and so can be treated as additive outliers (Hotta & Tsay, 2012). Additive outliers are usually associated with unexpected events on stock markets (see Chatzikonstanti (2017)), whose effects lead to misspecified estimates and biased risk measures (Grané & Veiga, 2010). In this regard, we propose to tackle the issue by employing the wavelets-based technique for additive outliers proposed by Grané & Veiga (2010). The procedure follows a sequence of steps and uses the level wavelets coefficients and the detail wavelets coefficients to identify and correct outliers in financial returns.

Comparing daily observations with a determined threshold (k_1), discrepant values in GARCH residuals are firstly identified by applying the discrete wavelet transform (DWT) function. Defining k_1 as the 95^{th}-percentile of the distribution of the maximum of the first level of detail wavelets coefficients, each element in the series are located according to the following rule:

$$d_{Max} = Max_{1 \leq j \leq n/2} \left\{ \lvert d_j \rvert > k_1^{0.05} \right\} \tag{2}$$

where $dMax$ is an element belonging to the detail coefficients vector.

Sequentially, the procedure compares elements of the detail wavelet coefficients vector to the GARCH residuals to correct the magnitude of extreme observations and recompose the series through the application of the inverse discrete wavelet transform (IDWT) function. Although non-discrepant original values are not altered, the new residuals series is computed as the simple mean, without the influence of outliers at positions $2s$ and 2_{s-1}, with $\{2s, 2_{s-1}\} \in S$ where S is the vector with the outliers' location.

$$\bar{\varepsilon}_{n-2} = \frac{1}{n-2} \sum_{i \neq 2s, 2s-1}^{n} \varepsilon_i \tag{3}$$

with ε_i the residuals, and s the position of outliers in the series.

In our estimations, the threshold k_1^{α} is set 6.617, obtained from the generation of 1000 $i.i.d$ Student's-t random variables with 7 degrees of freedom, and α equals to 0.05. We use the first level of wavelets coefficients and detail wavelets coefficients for the outliers' identification (for a further discussion on the method see the original paper of Grané & Veiga (2010)).

Models with and without corrections for outliers are estimated under six innovations distributions; Normal (norm), Skewed Normal (snorm), Student's-t (std), Skewed Student's-t (sstd), Generalised Error Distribution (ged) and Skewed Generalised Error Distribution (sged). Hence, the initial set of candidate models totals 72 different specifications.

Conditional standard deviations computed from the models are used to predict one-day-ahead Value at risk (VaR) and Expected Shortfall (ES) at 1% significance level (α). VaR is calculated as follows:

$$VaR_{t+1}^{\alpha} = -Q_{\alpha} \hat{\sigma}_{t+1/t} \tag{4}$$

where Q_{α} is the α quantile for a given error distribution, and $\hat{\sigma}_{t+1/t}$ is the predicted conditional volatility obtained from the models presented before.

ES represents the expected loss conditioned to the loss that violates 1%-VaR, that is:

$$ES_{t+1}^{\alpha} = \frac{1}{\alpha} \int_{-\infty}^{VaR_{t+1}} z dF(\cdot) \tag{5}$$

where z is the standardised residuals and $F(\cdot)$ is the cumulative density function for a chosen distribution.

2.2 Risk model's evaluation

The first step of the analysis consists in employing the Model Confidence Set (MCS) procedure developed by Hansen et al. (2011) to define the most accurate models to forecast 1%-VaR. The MCS procedure is a sequence of statistical tests that compares and selects models containing "the best specifications" to explain a given phenomenon (Hansen et al., 2011). The selection is performed over an arbitrary loss function applied over VaR series. The asymmetric loss function suggested by González-Rivera et al. (2004) applied in this study is calculated as follows:

$$L\left(r_t, VaR_t^\alpha\right) = \left(\alpha - d_t^\alpha\right)\left(r_t - VaR_t^\alpha\right) \tag{6}$$

where r_t are the returns, the risk level $\alpha = 1\%$ and $d_t^\alpha \equiv 1(r_t < VaR_t^\alpha)$, whose (1-$\alpha$) is a weight that penalises more heavily observations when $r_t < VaR_t^\alpha$.

From an initial set of models M_0, the hypothesis of equal predictive ability (EPA) is tested sequentially, where the least accurate models are eliminated at each round. By comparing the average loss between pairs of models trough $T_{R,M}$ statistics, and the average loss between a single model and all elements in M_0 trough the $T_{Max,M}$ statistics, a final set of superior models is defined when the equal predictive ability hypothesis can no longer be rejected (see Bernardi & Catania (2016) for the computational implementation).

Given the superior set of models, the specifications are tested through the unconditional coverage (UC) test of Kupiec (1995), the conditional coverage test (CC) of Christoffersen (1998) and the dynamic quantile test (DQ) of Engle & Manganelli (2004). The general null hypothesis of these tests states the correct coverage of VaR forecasts on the left-tail of the distribution of returns at α significance level. Additionally, four backtesting measures are implemented to compare 1%-VaR forecasts, henceforward A/E, ADMean, ADMax and FZL.

1. A/E is the ratio of actual VaR violations to expected VaR violations:

$$A/E = \frac{\sum_{j=1}^{p} d_{t+j}}{\alpha\% P} \tag{7}$$

   with $d_{t+1} = 1(r_{t+1} < VaR_{t+1}^\alpha)$, where P is the number of observations of the out-of-sample. The closer A/E to the unity the more accurate the model.

2. ADMean and ADMax are respectively the mean and the maximum absolute deviations of the violations from the VaR forecasts, as proposed by McAleer & Da Veiga (2008). As a measure for the magnitude of losses, models with smaller ADMean and ADMax are preferable.

$$\begin{aligned} ADMean =& \frac{1}{T}\sum_{t+1}^{T}\left(|r_{t+1}| - \left|VaR_{t+1}^\alpha\right|\right) \\ ADMax =& Max\left(|r_{t+1}| - \left|VaR_{t+1}^\alpha\right|\right) \end{aligned} \tag{8}$$

   for $r_{t+1} < -VaR_{t+1}$

3. FZL is the joint VaR-ES loss function that combines both measures to assess the accuracy of the forecasts on the left-tail of returns distribution (see Fissler et al. (2016); Ardia et

al. (2018); Patton et al. (2019)). Models with smaller values imply higher precision to forecast the risk.

$$FZL_t^\alpha = \frac{1}{\alpha ES_t^\alpha} d_t \left(r_t - VaR_t^\alpha\right) + \frac{VaR_t^\alpha}{ES_t^\alpha} + \log\left(ES_t^\alpha\right) - 1 \quad (9)$$

for$ES_t^\alpha \leq VaR_t^\alpha < 0$

3 Data and results

The data[1] used in this study is represented by daily returns of the Dow Jones Industrial Average index calculated as $r_t = log(p_t/p_{t-1})$, where p_t is the adjusted closing price. Table 1 reports the descriptive statistics for the full sample and sub-samples periods under the analysis, spanning from January 27, 2010, to May 11, 2020. In-sample corresponds to 2500 observations used as estimation window. Out-of-sample forecasts are performed for the 90 first working days of 2020, which coincides with the COVID-19 outbreak.

Table 1: Descriptive statistics (x 100) for the Down Jones daily returns from January 27, 2010, to May 11, 2020.

	n obs	Mean	St.dev.	Min.	Q_1	Med.	Q_3	Max.	Skew.	Kurt.
full-sample	3107	0.0207	1.2792	-13.8418	-0.4008	0.0554	0.5337	10.7643	-0.4457	15.9643
in-sample	2500	0.0412	0.8866	-5.7061	-0.3258	0.0583	0.4837	4.8643	-0.4747	4.0615
out-of-sample	90	-0.1792	3.5117	-13.8418	-1.3963	0.0404	1.0996	10.7643	-0.4307	3.2406

Note: In-sample spans from January 27, 2010, to December 31, 2019 and out-of sample from January 2, 2020, to May 11 2020. Kurt is the excess kurtosis of log returns distributions.

The descriptive statistics of the in-sample period reflects ten years of an upward trend of the American stock markets. The low volatility of the period is observed by minimum and maximum returns between -5.70% and 4.86%, and an annualised standard deviation of 14.07% (0.8866 x$\sqrt{252}$)x100. Conversely, the outbreak of COVID-19 is responsible for an abrupt fall in the stocks prices and a rapid increase in the volatility. The out-of-sample annualised standard deviation for the period is 55,54% (3.5117 x$\sqrt{252}$)x100, with minimum returns and maximum returns reaching -13.84% and 10.76%, respectively. Whereas the distribution of returns is negatively skewed, the excess kurtosis evidences the presence of outliers in the series.

In order to evaluate the models' specifications described in the last section, one-day-ahead forecasts of the conditional standard deviation are performed in a rolling windows scheme. With an estimation sample of 2500 observations, daily 1% VaR and 1% ES are computed for the Dow Jones returns.

[1]the data was collected from the Thomsom Reuters Datastream

The model confidence set (MCS)[2] of Hansen et al. (2011) is initially applied to select the models with the equal predictive ability to forecast the 1%-VaR of the Dow Jones returns. By implementing the loss function proposed by González-Rivera et al. (2004), the final set is shown in table 2 and encompasses 18 out of 72 specifications. These models are ordered in the first and third columns according to their respective $T_{R,M}$ and $T_{Max,M}$ statistics. As it observed, all normal and normal-skewed models are eliminated, remaining only models that account for heady-tailed innovations. It is interesting to note that no GARCH(1,1) specifications are selected by the MCS procedure, which suggests that more complex models that take asymmetries of returns into account can be more accurate to predict the VaR during turbulent times. Furthermore, there are some indication that correcting additive outliers present in the series can improve VaR forecasts since 12 out of 18 models within the final set are estimated after the application of the wavelets-based technique. Both AR(1)-apARCH(1,1)-sged-wav and AR(1)-apARCH(1,1)-sstd-wav models are ranked in the first positions based on the $T_{R,M}$ and $T_{Max,M}$ statistics respectively. However, it is worth highlighting that the predictive ability to forecast the 1%-VaR cannot be rejected for any model belonging to the final set 2, so backtesting procedures are used alongside the MCS procedure to evaluate the accuracy of the models further.

Table 2: Model Confidence Set for 1%-VaR forecasts for the Dow Jones returns - from January 2, 2020, to May 11, 2020.

Models	$T_{R,M}$	Models	$T_{Max,M}$
AR(1)-apARCH(1,1)-sged-wav	-2.626	AR(1)-apARCH(1,1)-sstd-wav	-0.602
AR(1)-apARCH(1,1)-sstd-wav	-1.908	AR(1)-eGARCH(1,1)-sstd	0.602
AR(1)-gjrGARCH(1,1)-sged-wav	-0.827	AR(1)-eGARCH(1,1)-sstd-wav	0.778
AR(1)-eGARCH(1,1)-sstd	-0.665	AR(1)-apARCH(1,1)-sged-wav	0.816
AR(1)-gjrGARCH(1,1)-sstd-wav	-0.501	AR(1)-eGARCH(1,1)-sged	1.084
AR(1)-gjrGARCH(1,1)-std-wav	-0.273	AR(1)-gjrGARCH(1,1)-ged-wav	1.193
AR(1)-apARCH(1,1)-sged	-0.178	AR(1)-gjrGARCH(1,1)-std-wav	1.225
AR(1)-apARCH(1,1)-std-wav	-0.095	AR(1)-apARCH(1,1)-ged-wav	1.231
AR(1)-gjrGARCH(1,1)-ged-wav	-0.088	AR(1)-gjrGARCH(1,1)-sged-wav	1.247
AR(1)-apARCH(1,1)-sstd	0.136	AR(1)-apARCH(1,1)-std-wav	1.353
AR(1)-apARCH(1,1)-std	0.158	AR(1)-apARCH(1,1)-sstd	1.656
AR(1)-eGARCH(1,1)-sged	0.324	AR(1)-apARCH(1,1)-sged	1.662
AR(1)-eGARCH(1,1)-sstd-wav	0.411	AR(1)-apARCH(1,1)-std	1.967
AR(1)-iGARCH(1,1)-std-wav	0.430	AR(1)-iGARCH(1,1)-std-wav	1.988
AR(1)-iGARCH(1,1)-ged-wav	0.479	AR(1)-iGARCH(1,1)-ged-wav	2.025
AR(1)-apARCH(1,1)-ged-wav	0.489	AR(1)-apARCH(1,1)-ged	2.047
AR(1)-apARCH(1,1)-ged	0.539	AR(1)-gjrGARCH(1,1)-sstd-wav	2.052
AR(1)-iGARCH(1,1)-sstd-wav	0.554	AR(1)-iGARCH(1,1)-sstd-wav	2.831

Note: The MCS procedure is performed for under a 95% confidence level. The models are ordered according to their respective position within the superior set of models. $T_{R,M}$ is the statistics related to the relative average loss between pair of models. $T_{Max,M}$ is the statistics related to the average loss of a single model relative to the average losses of all models.

[2]The MCS procedure is performed through the R package MCS (Catania & Bernardi, 2017)

Tables 3, 4 and 5 show the p-values of the CC (Kupiec, 1995), UC (Christoffersen, 1998) and DQ (Engle & Manganelli, 2004) coverage tests. In line with the final set of models delivered by the MCS procedure, the highest number of rejections is primarily concentrated on models with normal innovations. In contrast, models whose innovations follow either a skewed Student's-t (*sstd*) or a skewed generalised error distribution (*sged*) seems to be more efficient to cover the left tail of the distributions of returns.

Table 3: Unconditional Coverage (UC) test. P-values for 1%-VaR forecasts for the Dow Jones returns.

	norm	std	ged	snorm	sstd	sged
AR(1)-GARCH(1,1)	0.000	0.016	0.016	0.003	0.315	0.080
AR(1)-iGARCH(1,1)	0.003	0.016	0.080	0.016	0.917	0.315
AR(1)-eGARCH(1,1)	0.000	0.016	0.016	0.003	0.315	0.080
AR(1)-gjrGARCH(1,1)	0.016	0.080	0.080	0.016	0.917	0.917
AR(1)-apARCH(1,1)	0.016	0.080	0.080	0.080	0.917	0.315
AR(1)-csGARCH(1,1)	0.003	0.016	0.016	0.003	0.080	0.080
AR(1)-GARCH(1,1)-wav	0.000	0.003	0.003	0.000	0.080	0.016
AR(1)-iGARCH(1,1)-wav	0.000	0.016	0.080	0.003	0.917	0.080
AR(1)-eGARCH(1,1)-wav	0.000	0.003	0.003	0.000	0.080	0.016
AR(1)-gjrGARCH(1,1)-wav	0.003	0.080	0.080	0.003	0.917	0.917
AR(1)-apARCH(1,1)-wav	0.003	0.016	0.016	0.016	0.917	0.080
AR(1)-csGARCH(1,1)-wav	0.000	0.003	0.003	0.000	0.016	0.016

Note: The table reports the p-values of the UC test of Kupiec (1995). 1%VaR forecasts are performed through a one-day-ahead rolling windows scheme for the COVID-19 period, ranging from January 2, 2020, to May 11, 2020.

Table 4: Conditional Coverage (CC) test. P-values for 1%-VaR forecasts for the Dow Jones returns.

	norm	std	ged	snorm	sstd	sged
AR(1)-GARCH(1,1)	0.001	0.045	0.045	0.008	0.577	0.194
AR(1)-iGARCH(1,1)	0.008	0.045	0.194	0.045	0.983	0.577
AR(1)-eGARCH(1,1)	0.001	0.045	0.045	0.008	0.577	0.194
AR(1)-gjrGARCH(1,1)	0.045	0.194	0.194	0.045	0.983	0.983
AR(1)-apARCH(1,1)	0.045	0.194	0.194	0.194	0.983	0.577
AR(1)-csARCH(1,1)	0.008	0.045	0.045	0.008	0.194	0.194
AR(1)-GARCH(1,1)-wav	0.000	0.008	0.008	0.001	0.194	0.045
AR(1)-iGARCH(1,1)-wav	0.001	0.045	0.194	0.008	0.983	0.194
AR(1)-eGARCH(1,1)-wav	0.000	0.008	0.008	0.001	0.194	0.045
AR(1)-gjrGARCH(1,1)-wav	0.008	0.194	0.194	0.008	0.983	0.983
AR(1)-apARCH(1,1)-wav	0.008	0.045	0.045	0.045	0.983	0.194
AR(1)-csARCH(1,1)-wav	0.000	0.008	0.008	0.000	0.045	0.045

Note: The table reports the p-values of the CC test of Christoffersen (1998). 1%VaR forecasts are performed through a one-day-ahead rolling windows scheme for the COVID-19 period, ranging from January 2, 2020, to May 11, 2020.

Table 5: Dynamic Quantile (DQ) test. P-values for 1%-VaR forecasts for the Dow Jones returns.

	norm	std	ged	snorm	sstd	sged
AR(1)-GARCH(1,1)	0.000	0.000	0.000	0.000	0.420	0.154
AR(1)-iGARCH(1,1)	0.000	0.000	0.000	0.000	0.990	0.416
AR(1)-eGARCH(1,1)	0.000	0.004	0.004	0.000	0.908	0.014
AR(1)-gjrGARCH(1,1)	0.000	0.203	0.204	0.000	0.996	0.996
AR(1)-apARCH(1,1)	0.000	0.248	0.249	0.247	0.998	0.602
AR(1)-csGARCH(1,1)	0.000	0.000	0.000	0.000	0.124	0.136
AR(1)-GARCH(1,1)-wav	0.000	0.000	0.000	0.000	0.146	0.000
AR(1)-iGARCH(1,1)-wav	0.000	0.000	0.000	0.000	0.987	0.147
AR(1)-eGARCH(1,1)-wav	0.000	0.000	0.000	0.000	0.000	0.000
AR(1)-gjrGARCH(1,1)-wav	0.000	0.226	0.227	0.000	0.994	0.994
AR(1)-apARCH(1,1)-wav	0.000	0.000	0.000	0.000	0.997	0.310
AR(1)-csGARCH(1,1)-wav	0.000	0.000	0.000	0.000	0.000	0.000

Note: The table reports the p-values of the DQ test of Engle & Manganelli (2004). 1%VaR forecasts are performed through a one-day-ahead rolling windows scheme for the COVID-19 period, ranging from January 2, 2020, to May 11, 2020.

Risk measures used to compare the models' performance are displayed in table 6. The models are divided into two groups, columns 1 to 5 refer to specifications with no corrections for additive outliers. In contrast, columns 6 to 10 contain models with series filtered through the wavelets-based technique. For each group, the best results for the corresponding measures are in boldface, with the overall best model indicated by $*$. Observing the proportion of actual to the expected VaR violations, the closest A/E ratio to the unity is 1.1, which is broadly produced by models with both *sstd* and *sged* innovations. Within the group of models with no corrections for additive outliers, AR(1)-eGARCH(1,1)-sstd model has the lowest average loss (FZL) and the smallest maximum absolute deviations (ADMax) of violations from 1%-VaR forecasts generally. However, iGARCH models with Student's-t innovations present the lowest ADMean for both cases. Based on these three measures, AR(1)-apARCH(1,1)-sstd-wav is the most efficient model to forecast the 1%-VaR, since it best accommodates the extremes returns observed during the highly volatile COVID-19 period. Besides the minimum A/E ratio, the model holds the lowest average loss according to the FZ function and also presents the second smallest ADmean measure overall.

Table 6: 1%-VaR Backtesting for Dow Jones returns.Out-of sample forecasts spanning from January 2, 2020, to May, 11 2020.

models	A/E	FZ	ADMean	ADMax	models	A/E	FZ	ADMean	ADMax
AR(1)-GARCH(1,1)-norm	6.667	3.336	0.798	1.955	AR(1)-GARCH(1,1)-norm-wav	7.778	3.545	0.984	1.955
AR(1)-GARCH(1,1)-std	4.444	2.553	0.608	1.701	AR(1)-GARCH(1,1)-std-wav	5.556	2.521	0.553	1.701
AR(1)-GARCH(1,1)-ged	4.444	2.631	0.643	1.725	AR(1)-GARCH(1,1)-ged-wav	5.556	2.619	0.616	1.725
AR(1)-GARCH(1,1)-snorm	5.556	2.932	0.723	1.823	AR(1)-GARCH(1,1)-snorm-wav	6.667	3.135	0.977	2.209
AR(1)-GARCH(1,1)-sstd	2.222	2.298	0.762	1.490	AR(1)-GARCH(1,1)-sstd-wav	3.333	2.272	0.651	1.490
AR(1)-GARCH(1,1)-sged	3.333	2.396	0.590	1.575	AR(1)-GARCH(1,1)-sged-wav	4.444	2.419	0.662	1.575
AR(1)-iGARCH(1,1)-norm	5.556	3.046	0.677	1.881	AR(1)-iGARCH(1,1)-norm-wav	6.667	3.098	0.733	1.881
AR(1)-iGARCH(1,1)-std	4.444	2.435	**0.473**	1.629	AR(1)-iGARCH(1,1)-std-wav	4.444	2.377	**0.474**	1.629
AR(1)-iGARCH(1,1)-ged	3.333	2.501	0.622	1.639	AR(1)-iGARCH(1,1)-ged-wav	3.333	2.444	0.624	1.639
AR(1)-iGARCH(1,1)-snorm	4.444	2.735	0.646	1.733	AR(1)-iGARCH(1,1)-snorm-wav	5.556	2.813	0.783	1.733
AR(1)-iGARCH(1,1)-sstd	**1.111**	2.265	1.407	1.407	AR(1)-iGARCH(1,1)-sstd-wav	**1.111**	2.204	1.407	1.407
AR(1)-iGARCH(1,1)-sged	2.222	2.337	0.749	1.481	AR(1)-iGARCH(1,1)-sged-wav	3.333	2.285	0.535	1.481
AR(1)-eGARCH(1,1)-norm	6.667	2.917	1.365	2.918	AR(1)-eGARCH(1,1)-norm-wav	7.778	3.331	1.591	3.535
AR(1)-eGARCH(1,1)-std	4.444	2.154	0.937	1.438	AR(1)-eGARCH(1,1)-std-wav	5.556	2.341	1.200	2.296
AR(1)-eGARCH(1,1)-ged	4.444	2.240	1.070	1.602	AR(1)-eGARCH(1,1)-ged-wav	5.556	2.476	1.349	2.512
AR(1)-eGARCH(1,1)-snorm	5.556	2.556	1.280	2.286	AR(1)-eGARCH(1,1)-snorm-wav	6.667	2.877	1.514	3.142
AR(1)-eGARCH(1,1)-sstd	2.222	**1.913**	0.886	**1.000**	AR(1)-eGARCH(1,1)-sstd-wav	3.333	2.003	1.096	1.517
AR(1)-eGARCH(1,1)-sged	3.333	2.024	0.903	1.164	AR(1)-eGARCH(1,1)-sged-wav	4.444	2.184	1.166	2.009
AR(1)-gjrGARCH(1,1)-norm	4.444	2.800	0.912	1.783	AR(1)-gjrGARCH(1,1)-norm-wav	5.556	2.806	0.885	1.783
AR(1)-gjrGARCH(1,1)-std	3.333	2.271	0.637	1.549	AR(1)-gjrGARCH(1,1)-std-wav	3.333	2.196	0.638	1.549
AR(1)-gjrGARCH(1,1)-ged	3.333	2.315	0.683	1.562	AR(1)-gjrGARCH(1,1)-ged-wav	3.333	2.241	0.684	1.562
AR(1)-gjrGARCH(1,1)-snorm	4.444	2.515	0.716	1.648	AR(1)-gjrGARCH(1,1)-snorm-wav	5.556	2.457	0.605	1.648
AR(1)-gjrGARCH(1,1)-sstd	**1.111**	2.160	1.320	1.320	AR(1)-gjrGARCH(1,1)-sstd-wav	**1.111**	2.083	1.320	1.320
AR(1)-gjrGARCH(1,1)-sged	**1.111**	2.196	1.399	1.399	AR(1)-gjrGARCH(1,1)-sged-wav	**1.111**	2.119	1.399	1.399
AR(1)-apARCH(1,1)-norm	4.444	2.518	0.818	1.545	AR(1)-apARCH(1,1)-norm-wav	5.556	2.712	1.097	2.205
AR(1)-apARCH(1,1)-std	3.333	2.088	0.516	1.278	AR(1)-apARCH(1,1)-std-wav	4.444	2.084	0.565	1.278
AR(1)-apARCH(1,1)-ged	3.333	2.129	0.582	1.298	AR(1)-apARCH(1,1)-ged-wav	4.444	2.154	0.673	1.298
AR(1)-apARCH(1,1)-snorm	3.333	2.276	0.843	1.395	AR(1)-apARCH(1,1)-snorm-wav	4.444	2.404	1.072	1.750
AR(1)-apARCH(1,1)-sstd	**1.111**	1.952	1.021	1.021	**AR(1)-apARCH(1,1)-sstd-wav***	**1.111**	**1.880**	1.021	**1.021**
AR(1)-apARCH(1,1)-sged	2.222	1.988	0.577	1.119	AR(1)-apARCH(1,1)-sged-wav	3.333	1.954	0.505	1.119
AR(1)-csGARCH(1,1)-norm	5.556	3.440	0.949	2.003	AR(1)-csGARCH(1,1)-norm-wav	7.778	3.673	1.018	2.003
AR(1)-csGARCH(1,1)-std	4.444	2.747	0.749	1.800	AR(1)-csGARCH(1,1)-std-wav	5.556	2.738	0.742	1.800
AR(1)-csGARCH(1,1)-ged	4.444	2.766	0.722	1.793	AR(1)-csGARCH(1,1)-ged-wav	5.556	2.759	0.714	1.793
AR(1)-csGARCH(1,1)-snorm	5.556	3.051	0.754	1.883	AR(1)-csGARCH(1,1)-snorm-wav	7.778	3.288	0.896	2.302
AR(1)-csGARCH(1,1)-sstd	3.333	2.440	0.646	1.616	AR(1)-csGARCH(1,1)-sstd-wav	4.444	2.460	0.733	1.616
AR(1)-csGARCH(1,1)-sged	3.333	2.512	0.705	1.654	AR(1)-csGARCH(1,1)-sged-wav	4.444	2.557	0.816	1.654

Note: A/E is the proportion of actual to the expected 1%-VaR violations. FZ is the average of the joint VaR-ES loss function firstly suggested by Fissler et al. (2016). ADMean and ADMax are the mean and maximum absolute deviation measures suggested by McAleer & Da Veiga (2008). The measures of the best models are in boldface and the superior model overall is indicated by *.

These findings are illustrated in figure 2, where the 1%-VaR coverage of the standard AR(1)-GARCH(1,1)-std is contrasted with the respective measure produced by the AR(1)-apARCH(1,1)-sstd-wav model. Red points in the figures indicate VaR violations when the actual returns are lower than the 1%-VaR forecasts. It is observed that the AR(1)-GARCH(1,1)-std fails in four situations, two of them before the first shock of the COVID-19 on the stock markets. The single violation produced by the AR(1)-apARCH(1,1)-sstd-wav is seen on February 24, 2020, when the Down Jones index posts a sharp drop of 3.1% as American Centers for Disease Control and Prevention releases public health responses to the COVID-19 disease. The AR(1)-apARCH(1,1)-sstd-wav captures the deepest daily fall in more than tens years on March 8, 2020, and all other subsequent losses of the Down Jones' shares during the period. Moreover, the AR(1)-apARCH(1,1)-sstd-wav shows more coverage precision than its counterpart, since the AR(1)-GARCH(1,1)-std overestimates the risk after its fourth violation, with 1%VaR going bellow -20%. Therefore, our results are in line with those obtained by Grané & Veiga (2014), in the sense that treating outliers in the series through the wavelets-based technique helps to improve risk measures. On the other hand, different from their findings, the application of the procedure to more complex specifications and error distributions delivers more accurate VaR forecasts for the period under analysis.

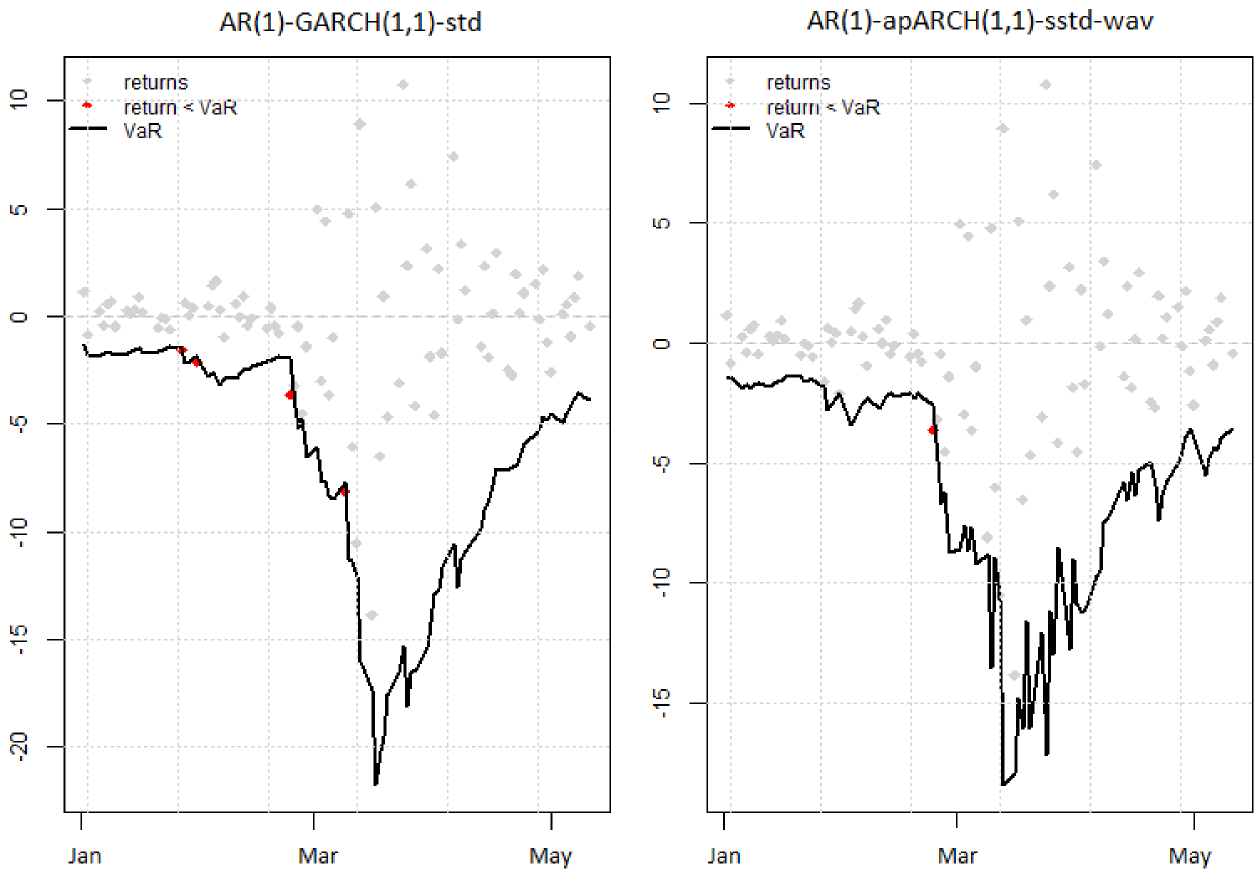

Figure 2: 1%VaR forecasts (line) ranging from January 2, 2020, to May 11, 2020. Gray points indicate realised log returns, and red points are 1%VaR violations. Graphs obtained trough the R package rugarch (Ghalanos, 2020)

4 Conclusion

This study compares the performance of several GARCH-type models to forecast the 1%-VaR for the Dow Jones Index Average returns during the period of the Covid-19 crisis. Conditional volatility of returns are estimated through the combination of six models specifications and six innovations distributions. Estimates are also carried out after identifying and correcting additive outliers present in the series by applying the wavelets-based technique of Grané & Veiga (2010). The Model Confidence Set procedure of Hansen et al. (2011) is used to select the most accurate specifications, and 1%-VaR forecasts are evaluated through backtesting measures.

We observe that the superior set of models selected by the MCS procedure presents lower mean losses (FZ measure) and reduced maximum absolute deviations of VaR violations (AD-Max). These models generally allow for heavy-tailed and skewed innovations distributions. By comparing the results of models with and without corrections for additive outliers, the application of the wavelet technique seems to improve VaR forecasts in a considerable number of cases. Furthermore, models with more complex specifications deliver better performance for the highly volatile period of the Covid-19 crisis. Since it naturally accommodates the leverage effect of financial returns into its specification, the asymmetric power ARCH model, in the form AR(1)-apARCH-sstd-wav, produces the most accurate 1%-VaR forecasts for the Dow Jones returns overall.

Our findings reveal that, besides asymmetric models specifications and distributions, considering the effects of outliers in the series is crucial to managing the risk during turbulent times. As future work, these analyses can be performed for the post-COVID-19-crisis. Moreover, VaR forecasts for the same period can be investigated by introducing exogenous regressors into the estimates in a high-frequency approach as in Bernardi, Catania, & Petrella (2017).

References

Ardia, D., Bluteau, K., Boudt, K., & Catania, L. (2018). Forecasting risk with markov-switching garch models: A large-scale performance study. *International Journal of Forecasting*, *34*(4), 733–747.

Bao, Y., Lee, T.-H., & Saltoglu, B. (2006). Evaluating predictive performance of value-at-risk models in emerging markets: a reality check. *Journal of forecasting*, *25*(2), 101–128.

Bernardi, M., & Catania, L. (2016). Comparison of value-at-risk models using the mcs approach. *Computational Statistics*, *31*(2), 579–608.

Bernardi, M., Catania, L., & Petrella, L. (2017). Are news important to predict the value-at-risk? *The European Journal of Finance*, *23*(6), 535–572.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Catania, L., & Bernardi, M. (2017). Mcs: Model confidence set procedure [Computer software manual]. (R package version 0.1.3)

Chatzikonstanti, V. (2017). Breaks and outliers when modelling the volatility of the us stock market. *Applied Economics*, *49*(46), 4704–4717.

Chen, C. W., Gerlach, R., Lin, E. M., & Lee, W. (2012). Bayesian forecasting for financial risk management, pre and post the global financial crisis. *Journal of Forecasting*, *31*(8), 661–687.

Christoffersen. (1998). Evaluating interval forecasts. *International economic review*, 841–862.

Dias, A. (2013). Market capitalization and value-at-risk. *Journal of Banking & Finance*, *37*(12), 5248–5260.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, *1*(1), 83–106.

Engle. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, *5*(1), 1–50.

Engle, & Lee, G. (1999). A long-run and short-run component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497.

Engle, & Manganelli, S. (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, *22*(4), 367–381.

Fissler, T., Ziegel, J. F., et al. (2016). Higher order elicitability and osband's principle. *The Annals of Statistics*, *44*(4), 1680–1707.

Ghalanos, A. (2020). rugarch: Univariate garch models. [Computer software manual]. (R package version 1.4-2.)

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779–1801.

González-Rivera, G., Lee, T.-H., & Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of forecasting*, *20*(4), 629–645.

Goodell, J. W. (2020). Covid-19 and finance: Agendas for future research. *Finance Research Letters*, 101512.

Grané, A., & Veiga, H. (2010). Wavelet-based detection of outliers in financial time series. *Computational Statistics & Data Analysis*, *54*(11), 2580–2593.

Grané, A., & Veiga, H. (2014). Outliers, garch-type models and risk measures: A comparison of several approaches. *Journal of Empirical Finance*, *26*, 26–40.

Hansen, Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497.

Ho, L.-C., Burridge, P., Cadle, J., & Theobald, M. (2000). Value-at-risk: Applying the extreme value approach to asian markets in the recent financial turmoil. *Pacific-Basin Finance Journal*, *8*(2), 249–275.

Hotta, L. K., & Tsay, R. S. (2012). Outliers in garch processes. *Economic time series: modeling and seasonality*, 337–358.

Jorion, P. (2009). Risk management lessons from the credit crisis. *European Financial Management*, *15*(5), 923–933.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives*, *3*(2).

McAleer, M., & Da Veiga, B. (2008). Single-index and portfolio models for forecasting value-at-risk thresholds. *Journal of Forecasting*, *27*(3), 217–235.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370.

Patton, A. J., Ziegel, J. F., & Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of econometrics*, *211*(2), 388–413.

Slim, S., Koubaa, Y., & Bensaida, A. (2017). Value-at-risk under lévy garch models: Evidence from global stock markets. *Journal of International Financial Markets, Institutions and Money*, *46*, 30–53.

Song, Q., Liu, J., & Sriboonchitta, S. (2019). Risk measurement of stock markets in brics, g7, and g20: Vine copulas versus factor copulas. *Mathematics*, *7*(3), 274.