Finance Research Letters 32 (2020) 101098

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Evaluation of volatility models for forecasting Value-at-Risk and Expected Shortfall in the Portuguese stock market

Nuno Sobreira[*,1], Rui Louro

ISEG-Lisbon School of Economics and Management, Universidade de Lisboa, Rua do Quelhas 6, Lisboa 1200-781, Portugal

A R T I C L E I N F O

Keywords:
VaR
Expected Shortfall
GARCH
Extreme Value Theory
Backtesting
Euronext Lisbon

JEL classification:
C22
C52
C53
C58
G31

A B S T R A C T

We run a forecasting competition of different methodologies to estimate Value-at-Risk (VaR) and Expected Shortfall (ES) with data on several stocks traded in the Euronext Lisbon stock exchange. The results are gauged using several backtesting procedures and compared with several loss functions. The asymmetric GARCH class with Extreme Value Theory generally performed better both for VaR and ES forecasting, especially, for more conservative coverage levels. Skewed distributions do not perform better than their conventional counterparts. The recommended sample size depends if the focus is on VaR or magnitude of the losses, although we find some superiority of larger sample sizes.

1. Introduction

Quantifying, forecasting and managing market risks are major concerns for financial institutions as exposure to extreme price movements in financial markets may cause sudden and significant losses.

In this context, the VaR became the most popular measure of market risk upon its introduction in the Basel I Accord amendment.

In spite of its establishment, VaR has been subject to criticism (Artzner et al., 1999; Yamai and Yoshiba, 2002; 2005) and the ES is becoming more important, in particular, given its recent role to obtain banks' capital charges (Bank for International Settlements (BIS), 2016).

The objective of this paper is to compare the performance of different methodologies to estimate the VaR and ES. It also intends to bring new insights about the methods used throughout this exercise.

We use Historical Simulation (HS), RiskMetrics, several models from the GARCH class with different distributions and the Extreme Value Theory approach (EVT) based on the Peak-over-Thresholds methodology (POT).

The data includes stock returns of companies traded in the Euronext Lisbon stock exchange that provide a representative set of the major companies in the Portuguese stock market. The results obtained for the VaR and ES are evaluated with backtesting procedures and compared using a loss function approach. We highlight that this paper compares the VaR and ES based on the stock returns of each company and not from a portfolio of these securities or a broad-based index such as PSI-20 as a proxy of the market portfolio.

* Corresponding author.

E-mail addresses: nsobreira@iseg.ulisboa.pt (N. Sobreira), rui.louro@phd.iseg.ulisboa.pt (R. Louro).

[1] We wish to thank CEMAPRE and the Department of Mathematics from ISEG. Financial support from FCT -*Fundação para a Ciência e Tecnologia* through national funds (CEMAPRE - UID/MULTI/00491/2013) is gratefully acknowledged.

https://doi.org/10.1016/j.frl.2019.01.010
Received 28 July 2018; Received in revised form 8 December 2018; Accepted 26 January 2019
Available online 29 January 2019
1544-6123/

This is motivated by the fact that we wanted to have a sufficiently large number of comparable series to enable a robust and exhaustive comparison of the competing methods while simultaneously focusing on the Portuguese stock market.

This paper makes the following contributions to previous empirical studies about VaR and ES. First, it uses a novel dataset that focuses on the Portuguese stock market throughout the current and last decade. The Portuguese stock market is a small but quite relevant and interesting case study given that the dataset covers a wide range of time varying economic conditions from the Global Financial Crisis, Eurozone debt crisis, the Troika bailout until the bailout exit. A detailed analysis of the results allows us to compare the performance of various methodologies under different economic conditions and with other financial markets. Second, we compare several methods to estimate the VaR and ES, at different coverage levels, with procedures both to validate and select the model in a unified dataset framework. Most of the literature considers only a subset of these possibilities. In fact, a substantial amount of it considers either model validation or model selection but not both and rely on VaR estimates alone discarding the ES analysis and the robustness of the results to different backtesting procedures.

The results show that the combination of asymmetric GARCH with EVT performs better both for VaR and ES forecasting. The evidence here is stronger for more conservative coverage levels. This result is robust for the majority of model validation backtests and loss functions considered. Moreover, skewed distributions generally do not perform better than their conventional counterparts. Finally, the sample size used clearly depends on the measure of interest as the VaR backtests and the loss function approach seem to favor smaller and larger samples, respectively. We also find that the methodologies considered benefit from larger sample sizes from a loss function perspective. Finally, we find evidence that factors which are generally set by regulators such as the coverage level or the backtesting strategy determine the likelihood of obtaining well calibrated VaR and ES models.

2. VaR and ES estimation methods

We consider the performance of HS, various parametric volatility models, RiskMetrics and the POT-EVT as these seem to be the most widely known approaches in the risk management community. The HS approach obtains the 1-day 100α% VaR, $VaR_{\alpha,t+1}$, as the negative quantile at level α of the log-returns, r_t. The 1-day 100α% ES, $ES_{\alpha,t+1}$, is obtained as the negative sample mean of those values of r_t lower than $-VaR_{\alpha,t+1}$. The remaining methods define the dynamics of r_t as:

$$r_t = \mu_t + u_t$$
$$u_t = \sigma_t \varepsilon_t,\ \varepsilon_t \overset{i.i.d.}{\sim} (0,1),\ \varepsilon_t \perp\!\!\!\perp u_{t-j},\ j \in \mathbb{N}$$
$$\sigma_t^2 = G(u_{t-1}^2, \ldots, u_{t-q}^2, \sigma_{t-1}^2, \ldots, \sigma_{t-p}^2) \tag{1}$$

where $\mu_t = E(r_t|\mathcal{F}_{t-1})$, $\sigma_t^2 = Var(r_t|\mathcal{F}_{t-1})$ with $\mathcal{F}_t = \{r_1, r_2, ..., r_t\}$, $\mathbb{N}$ denotes the Natural numbers and $G(\cdot)$ is a function relating σ_t^2 with lags of u_t^2 and of σ_t^2.

Under this framework, the 1-day 100α% VaR for a long position can be calculated as Tsay (2010):

$$VaR_{\alpha,t+1} = -(\mu_{t+1|t} + q_\alpha^\varepsilon \sigma_{t+1|t}) \tag{2}$$

where $\mu_{t+1|t} = E(r_{t+1}|\mathcal{F}_t)$, $\sigma_{t+1|t}^2 = Var(r_{t+1}|\mathcal{F}_t)$ and q_α^ε is the quantile at level α based on the distribution for ε_t. Likewise, the 1-day 100α% ES can be obtained as McNeil and Frey (2000):

$$ES_{\alpha,t+1} = -(\mu_{t+1|t} + \sigma_{t+1|t} E(\varepsilon_{t+1}|\varepsilon_{t+1} < q_\alpha^\varepsilon)) \tag{3}$$

As is readily seen, the calculation of $VaR_{\alpha,t+1}$ and $ES_{\alpha,t+1}$ requires values for $\mu_{t+1|t}$, $\sigma_{t+1|t}$, q_α^ε and $E(\varepsilon_{t+1}|\varepsilon_{t+1} < q_\alpha^\varepsilon)$.

We assume a zero conditional mean for the log returns ($\mu_t = 0$).[2] The remaining quantities were obtained from rolling $R-$sized window 1-step ahead forecasts according to each of the following possibilities:

- VaR and ES estimation methods: HS, RiskMetrics, parametric volatility time series models (GARCH, IGARCH, EGARCH, GJR-GARCH) and EVT (EVT GARCH, EVT IGARCH, EVT EGARCH, EVT GJR-GARCH)
- Distribution of ε_t: Normal, Student-t, GED, Skewed Normal, Skewed Student-t and skewed GED[3]
- R (rolling window size): 1000, 1500 and 2000
- α: 0.01 and 0.05

Each possible combination was assessed as follows: We estimate the unknown parameters with the first R days of the sample and corresponding forecasts necessary to obtain (2) and (3). Then, the fixed $R-$days sample is rolled forward and this procedure is repeated until the last R observations. This produces a time series of 1-day VaR and ES estimates, $\{\widehat{VaR}_{\alpha,t}\}_{t=R+1}^T$ and $\{\widehat{ES}_{\alpha,t}\}_{t=R+1}^T$, which are validated with backtesting procedures and ranked with the loss function approach. We consider the following VaR backtests: POF (Kupiec, 1995), TUFF (Kupiec, 1995), IND (Christofferssen, 1998), DBI (Christofferssen and Pelletier, 2004), CC (Christofferssen, 1998) and DQ (Engle and Manganelli, 2004). As for ES backtesting, four tests were used: MF (McNeil and Frey, 2000), AS1, AS2 (Acerbi and Szekely, 2014) and EKT (Emmer et al., 2015). A comprehensive review of several VaR and ES

[2] Previous articles reported a negligible or a negative impact of assuming an ARMA specification for μ_t (see,for example, Angelidis et al., 2004; Braione and Scholtes, 2016). We also ran this experiment and found lower differences between the average proportion of realized and expected exceptions with $\mu_t = 0$.

[3] We use the skewed distributions proposed by Fernandez and Steel (1998).

backtesting procedures may be found in Jorcano (2017) and Novales and Garcia-Jorcano (2018).

As for the loss function approach, we obtain results for the daily capital charges criterion from McAleer et al. (2013) (DCC) and the loss functions surveyed by Abad et al. (2015) (AML).[4]

We obtain log-returns of several stocks of companies indexed in the Euronext Lisbon stock exchange between 3 January 2005, or the earliest available date, and 27 June 2017.[5] Information for these companies along with the descriptive statistics can be found in Table 1.

Table 1
Company description and descriptive statistics for the daily log returns.

Sector	Company	Max.	Min.	Mean	Standard deviation	Skewness	Kurtosis
Consumer Staples	Jerónimo Martins	0.10	-0.17	7×10^{-4}	0.02	-0.54	10.11
	SONAE	0.21	-0.13	1×10^{-4}	0.02	0.28	10.16
Financials	Millenium BCP	0.24	-0.16	-0.001	0.03	0.28	8.43
	Banco Santander	0.19	-0.20	0.00	0.02	-0.13	10.22
	Banco BPI	0.24	-0.14	-3×10^{-4}	0.02	0.65	11.94
Industrials	Altri, SGPS	0.17	-0.16	0.001	0.02	0.35	9.46
	The Navigator Company	0.09	-0.09	4×10^{-4}	0.02	-0.08	6.43
	Cimpor	0.28	-0.30	-8×10^{-4}	0.02	0.80	24.15
	Mota-Engil	0.18	-0.21	1×10^{-4}	0.02	0.20	8.63
Telecommunications	Pharol	0.22	-0.30	-0.001	0.03	0.29	17.64
	NOS, SGPS	0.12	-0.11	-2×10^{-4}	0.02	0.06	8.10
Energy	GALP	0.22	-0.14	3×10^{-4}	0.02	0.46	12.23
Utilities	EDP	0.12	-0.19	1×10^{-4}	0.02	-0.42	12.92
	EDP Renováveis	0.13	-0.13	0.00	0.03	0.07	7.34
	REN	0.12	-0.13	-1×10^{-4}	0.01	-0.20	13.72

The selection was made to encompass various sectors of activity and simultaneously to include stock returns with a negligible number of zeros.

All computations were done in the R environment using *rugarch* (Ghalanos, 2015) with procedures included in this package and code developed by the authors.

3. Empirical results

For each of the 15 firms included in the analysis, we obtain VaR and ES estimates for the periods after the sample defined by the rolling window. Overall, the results included 2250 time series of VaR and ES estimates for each α which will be used to assess the performance of the various methods considered in this paper. Naturally, we need to devise a strategy to summarize this information such that the different procedures are evaluated and compared in a straightforward manner.

Previous studies reported the number and proportion of exceptions in the hold-out sample and compared them to the predefined coverage level. They also reported the value of the backtesting statistics along with the corresponding p-values and the value of the loss functions. However, here we would need to report an enormous amount of results and tables. Therefore, we proceed by computing and comparing the proportion of VaR and ES series where the null hypothesis of each backtesting statistic is not rejected for each possible variant. Therefore, a percentage of non-rejections of the null hypothesis for each backtest is presented for each estimation approach, GARCH model, distribution of ε_t, rolling window size and coverage level. This provides an indication of the likelihood of each option providing a model with adequate VaR and ES forecasts according to each backtesting procedure. Moreover, we report the combination that yields the best results for each company under the loss function approach.

3.1. Estimation approach

One of the objectives of this work is to ascertain which approach performs best. Table 2 shows the summary results.

For the more liberal coverage level ($\alpha = 0.05$), RiskMetrics and HS have the worst performance for the majority of the backtests. In fact, RiskMetrics performs quite poorly with regard to conditional coverage and ES whereas HS fails both in POF and conditional coverage. Moreover, in general, the EVT approach dominates the other approaches. When a more conservative coverage level is considered ($\alpha = 0.01$), the EVT approach continues to show the best performance overall, outperforming more substantially the parametric and RiskMetrics approaches.

In general, we obtain a higher percentage of adequate VaR and ES methods for more liberal coverage levels. The exceptions are the HS and EVT approaches, which show improved results with several VaR backtesting procedures for a more conservative α.

Finally, the number of VaR and ES models that are deemed adequate greatly depends on the backtesting procedure considered.

[4] The VaR and ES estimation methods, the backtests and the loss functions are detailed in the supplementary appendix.

[5] We collected data from the Lisbon stock exchange website https://www.bolsadelisboa.com.pt/cotacoes/accoes-lisboa.

Table 2
Percentage of non-rejections of the null hypothesis of the backtesting procedures per VaR/ES estimation approach.

	$\alpha = 0.01$				$\alpha = 0.05$			
	HS	Parametric	EVT	RiskMetrics	HS	Parametric	EVT	RiskMetrics
POF	73%	49%	88%	9%	51%	87%	86%	100%
TUFF	89%	87%	90%	84%	100%	100%	100%	93%
IND	69%	88%	92%	76%	42%	76%	76%	47%
DBI	69%	93%	92%	84%	33%	92%	95%	87%
CC	62%	53%	85%	2%	33%	70%	72%	58%
DQ	44%	38%	68%	0%	13%	67%	58%	24%
MF	93%	33%	97%	18%	100%	21%	95%	0%
AS1	—	46%	47%	0%	—	48%	50%	0%
AS2	—	60%	50%	0%	—	79%	79%	49%
EKT	4%	1%	16%	0%	2%	11%	19%	0%

Unconditional coverage is validated for a generous percentage of models, followed by the assumption of independence of exceptions. The conditional coverage, and especially the ES approval, are more demanding assumptions and are rejected for more models. In fact, the performance of the models considered according to the EKT backtest is quite poor as the null is rejected for most of the ES models.

3.2. GARCH model

Tables 3 and 4 show the percentage of non-rejections of the null by GARCH model type.

We begin by analyzing the results of the parametric approach. With $\alpha = 0.05$, the best performing model is the EGARCH. The worst performing models are the IGARCH and GARCH for VaR and ES, respectively. However, for $\alpha = 0.01$, the GARCH and IGARCH become competitive and no parametric model seems to uniformly provide better results.

Table 3
Percentage of non-rejection of the null hypothesis of the backtesting procedures for the parametric approach per GARCH model type.

	Parametric approach							
	$\alpha = 0.01$				$\alpha = 0.05$			
	GARCH	IGARCH	EGARCH	GJR-GARCH	GARCH	IGARCH	EGARCH	GJR-GARCH
POF	44%	60%	51%	41%	90%	78%	90%	89%
TUFF	88%	88%	87%	86%	100%	100%	100%	100%
IND	86%	84%	94%	90%	74%	69%	83%	77%
DBI	94%	93%	91%	92%	91%	87%	97%	95%
CC	47%	55%	62%	49%	68%	60%	78%	74%
DQ	34%	38%	48%	32%	58%	50%	85%	74%
MF	29%	40%	37%	28%	17%	19%	28%	20%
AS1	43%	45%	49%	46%	49%	50%	50%	44%
AS2	55%	65%	62%	59%	73%	84%	81%	76%
EKT	1%	1%	3%	1%	8%	9%	17%	11%

Table 4
Percentage of non-rejection of the null hypothesis of the backtesting procedures for the EVT approach per GARCH model type.

	EVT approach							
	$\alpha = 0.01$				$\alpha = 0.05$			
	GARCH	IGARCH	EGARCH	GJR-GARCH	GARCH	IGARCH	EGARCH	GJR-GARCH
POF	86%	87%	91%	90%	85%	86%	84%	90%
TUFF	91%	91%	89%	89%	100%	100%	100%	100%
IND	91%	83%	96%	97%	74%	69%	83%	78%
DBI	92%	93%	94%	91%	92%	91%	100%	96%
CC	83%	76%	91%	89%	68%	66%	79%	77%
DQ	65%	60%	73%	72%	57%	46%	70%	60%
MF	98%	97%	96%	95%	93%	94%	97%	96%
AS1	44%	47%	53%	43%	49%	52%	54%	44%
AS2	47%	53%	54%	43%	73%	83%	81%	77%
EKT	15%	16%	14%	21%	17%	16%	25%	19%

Now let's compare the different parametric volatility models under the EVT approach. Here, the models with asymmetric effects perform better over the considered symmetric GARCH for $\alpha = 0.05$. When $\alpha = 0.01$, the EGARCH and GJR-GARCH models continue to perform better for most of the backtesting procedures but the differences are much less substantial.

3.3. Distribution of ε_t

Another important variant for VaR and ES forecasting is the distribution of ε_t whose choice remains under discussion: a number of studies concluded that complex distributional schemes have superior VaR performance[6] but sometimes simpler distributions reveal themselves competitive.[7] The results across different distributions are shown in Tables 5 and 6.

Various aspects are worth highlighting. First, the performance of VaR and ES models with skewed distributions is clearly not significantly different from their symmetric counterparts with this dataset. Hence, there seems to be no advantage in using skewed distributions. Second, although the performance across different distributions may be seen as somewhat similar for some backtests, we observe some superiority of the symmetric and skewed Student-t both for VaR and ES, especially for more conservative coverage levels. Finally, as regards the EVT approach, McNeil and Frey (2000) proposed using the Normal Distribution and, in general, we find no particular advantage in using a different distribution in the first step. The results across different statistical distributions are quite similar for most backtests.

Table 5
Percentage of non-rejections of the null hypothesis of the backtesting procedures for the parametric approach per statistical distribution of ε_t.

	Parametric approach											
	$\alpha = 0.01$						$\alpha = 0.05$					
	$N(0, 1)$	t	GED	Sk.$N(0, 1)$	Sk.t	Sk.GED	$N(0, 1)$	t	GED	Sk.$N(0, 1)$	Sk.t	Sk.GED
POF	49%	53%	46%	48%	54%	43%	87%	86%	89%	86%	85%	89%
TUFF	86%	90%	86%	85%	90%	87%	100%	100%	100%	100%	100%	100%
IND	88%	88%	88%	89%	88%	89%	73%	81%	75%	74%	80%	72%
DBI	90%	95%	93%	91%	93%	93%	92%	93%	92%	92%	92%	93%
CC	52%	61%	48%	51%	61%	48%	67%	72%	72%	67%	72%	71%
DQ	38%	46%	33%	36%	46%	30%	69%	71%	63%	68%	72%	58%
MF	32%	38%	31%	31%	38%	30%	19%	26%	19%	21%	27%	14%
AS1	7%	80%	52%	4%	81%	49%	1%	81%	64%	2%	83%	60%
AS2	17%	85%	81%	13%	82%	82%	70%	79%	84%	72%	78%	89%
EKT	1%	1%	1%	2%	2%	1%	12%	13%	11%	11%	13%	8%

Table 6
Percentage of non-rejections of the null hypothesis of the backtesting procedures for the EVT approach per statistical distribution of ε_t.

	EVT approach											
	$\alpha = 0.01$						$\alpha = 0.05$					
	$N(0, 1)$	t	GED	Sk.$N(0, 1)$	Sk.t	Sk.GED	$N(0, 1)$	t	GED	Sk.$N(0, 1)$	Sk.t	Sk.GED
POF	84%	91%	90%	84%	91%	91%	82%	87%	87%	83%	88%	89%
TUFF	91%	90%	89%	90%	90%	89%	100%	100%	100%	100%	100%	100%
IND	91%	93%	90%	92%	93%	90%	74%	78%	77%	75%	77%	74%
DBI	94%	90%	93%	95%	89%	92%	93%	96%	96%	93%	95%	96%
CC	84%	86%	86%	83%	86%	84%	68%	76%	74%	69%	74%	74%
DQ	68%	69%	67%	67%	69%	65%	59%	61%	56%	59%	58%	56%
MF	98%	97%	96%	97%	96%	96%	93%	96%	95%	94%	97%	94%
AS1	1%	83%	56%	0%	86%	56%	1%	81%	71%	1%	81%	64%
AS2	5%	77%	69%	1%	77%	69%	73%	76%	88%	71%	72%	92%
EKT	16%	17%	14%	18%	17%	17%	19%	19%	20%	18%	19%	19%

[6] See, for example, (Mittnik and Paolella (2000); Brooks and Persand (2003); Giot and Laurent (2003, 2004); Kuester et al. (2006); Bali et al. (2008); Braione and Scholtes (2016); BenSaïda et al. (2018)).

[7] See, for example, (Bams et al. (2005); Hansen and Lunde (2005); Slim et al. (2017)).

3.4. Rolling window size

Another active stream of research in this literature is the size of the sample used to estimate the VaR and ES models. The evidence is mixed, with some articles recommending larger sample sizes[8] and others smaller[9] ones.[10] Table 7 shows the percentage of non-rejections of the null of the various backtests for different rolling window sizes (1000, 1500 and 2000).

Table 7
Percentage of non-rejections of the null hypothesis of the backtesting procedures per rolling window size.

	HS approach						Parametric approach						EVT approach					
	$\alpha = 0.01$			$\alpha = 0.05$			$\alpha = 0.01$			$\alpha = 0.05$			$\alpha = 0.01$			$\alpha = 0.05$		
	1000	1500	2000	1000	1500	2000	1000	1500	2000	1000	1500	2000	1000	1500	2000	1000	1500	2000
POF	73%	80%	67%	60%	60%	33%	44%	47%	55%	86%	88%	86%	89%	91%	85%	90%	83%	86%
TUFF	87%	93%	87%	100%	100%	100%	89%	88%	85%	100%	100%	100%	93%	91%	86%	100%	100%	100%
IND	73%	67%	67%	40%	53%	33%	90%	90%	86%	66%	78%	83%	95%	92%	88%	71%	81%	76%
DBI	73%	73%	60%	33%	27%	40%	93%	96%	89%	94%	93%	90%	89%	96%	92%	96%	95%	93%
CC	67%	60%	60%	40%	33%	27%	54%	53%	54%	68%	71%	71%	87%	85%	83%	73%	74%	71%
DQ	40%	40%	53%	20%	7%	13%	41%	40%	33%	66%	70%	65%	70%	70%	62%	58%	60%	56%
MF	93%	93%	93%	100%	100%	100%	33%	29%	38%	19%	21%	23%	94%	98%	98%	96%	95%	94%
AS1	—	—	—	—	—	—	47%	46%	43%	48%	48%	49%	46%	48%	46%	51%	49%	50%
AS2	—	—	—	—	—	—	60%	63%	57%	79%	80%	76%	52%	54%	42%	81%	79%	76%
EKT	7%	7%	0%	0%	0%	7%	1%	2%	1%	8%	13%	13%	19%	18%	13%	19%	19%	19%

For the parametric approach, the performance of all rolling window sizes is rather similar, although the results are slightly favorable to VaR models with a smaller sample size for $\alpha = 0.01$. The ES performance is clearly sensitive to the backtesting procedure. As regards to the HS and EVT approaches the results are somewhat inconclusive for the case $\alpha = 0.05$ with some tests pointing towards smaller sample sizes and others larger ones. However, when a more conservative confidence level is considered, the results seem to favor again the use of models with smaller window sizes.

3.5. Loss function

Table 8 shows the methods that achieved the minimum value of the DCC for each company and α as we give more weight to the results of this criterion.[11]

Table 8
Best performing models according to the DCC criterion.

Company	$\alpha = 0.01$	$\alpha = 0.05$
Altri, SGPS	EVT GJR-GARCH $N(0, 1)$ 1000	HS 1500
Millenium BCP	EVT EGARCH Sk.t 2000	EGARCH Sk.t 2000
Banco BPI	EVT EGARCH Sk.t 1500	EGARCH Sk.t 2000
Cimpor	EVT EGARCH GED 2000	EGARCH Sk.t 1500
EDP	EVT EGARCH GED 1500	EGARCH t 1500
EDP Renovaveis	EVT GJR-GARCH Sk.t 2000	EGARCH t 2000
GALP	GJR-GARCH $N(0, 1)$ 1000	EVT GJR-GARCH $N(0, 1)$ 1500
Jeronimo Martins	EGARCH $N(0, 1)$ 1500	EVT EGARCH Sk.$N(0, 1)$ 1000
Mota-Engil	EVT EGARCH Sk.$N(0, 1)$ 2000	EVT EGARCH Sk.GED 1000
The Navigator Comp.	EGARCH $N(0, 1)$ 1000	EVT EGARCH Sk.$N(0, 1)$ 1000
NOS, SGPS	GJR-GARCH Sk.$N(0, 1)$ 1000	EVT IGARCH Sk.GED 1000
Pharol	EVT EGARCH $N(0, 1)$ 2000	EGARCH Sk.t 2000
REN	GARCH $N(0, 1)$ 2000	GARCH t 2000
Banco Santander	EGARCH $N(0, 1)$ 1500	EGARCH t 1000
SONAE	GJR-GARCH Sk.$N(0, 1)$ 1500	EGARCH Sk.t 1000

We start by looking at the results of the more conservative $\alpha = 0.01$. Here, the EVT approach was selected more often according to DCC. This result is reinforced considering all tables being the preferred approach for a good amount of loss functions. Hence, the EVT

[8] See Danıelsson (2002), among others.
[9] See, for example, Hoppe (1998).
[10] See Angelidis et al. (2004), for additional references on this issue.
[11] Tables with the results for all the loss functions are displayed in the supplementary appendix.

approach not only delivers a higher proportion of adequate models, according to the backtests, but also dominates the remaining approaches in terms of model selection. As regards to model type, the loss function is minimized for the majority of the cases by the EGARCH (9 for DCC and 42% overall) and GJR-GARCH (5 for DCC and 43% overall). As for the distribution of ε_t, the most commonly selected are the GED and skewed Student-t. Finally, a rolling window size of 2000 observations (6 for DCC and 44% overall) was selected more often.

When we analyze the more liberal $\alpha = 0.05$, some differences are worth noting. The EVT approach still performs better for many companies and across different loss functions. However, the parametric approach has a slight advantage as it appears in 51% of the cases. The EGARCH collects most of the preferences (11 for DCC and 55% overall) and the GJR-GARCH is still preferred to models without asymmetric effects albeit by a smaller margin. As regards to the preferred distribution of ε_t, the conventional and Skewed Student-t are now selected more often both according to DCC and to the AML loss functions. Another point worth mentioning is that there is no clear advantage of using skewed distributions as the results are reasonably evenly split.

Finally, the 2000 sample rolling window size continues to be selected as the most favorable case (48% overall). Hence, here we find evidence that these market risk models may benefit from using larger sample sizes.

4. Conclusions

This paper implements an exhaustive VaR and ES analysis to the Portuguese stock market with stocks of companies that represent the most relevant sectors of activity in the Portuguese economy. Hence, we have focused on the VaR and ES evaluation based on the returns of individual securities. As a future line of research, it would be interesting to build various portfolios based on different asset allocation strategies and compare their performance in terms of VaR and ES. Moreover, it would be useful to extend our analysis with additional VaR and ES forecasting methods and performance measures.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.frl.2019.01.010.

References

Abad, P., Muela, S., Lopez, C., 2015. The role of the loss function in value-at-risk comparisons. J. Risk Model Validation 9, 1–19.
Acerbi, C., Szekely, B., 2014. Baktesting Expected Shortfall. Risk Mag. 27, 76–81.
Angelidis, T., Benos, A., Degiannakis, S., 2004. The use of GARCH models in VaR estimation. Stat. Methodol. 1, 105–128.
Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Math. Finance 9 (3), 203–228.
Bali, T.G., Mo, H., Tang, Y., 2008. The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. J. Bank. Finance 32 (2), 269–282.
Bams, D., Lehnert, T., Wolff, C., 2005. An evaluation framework for alternative VaR-models. J. Int. Money Finance 24, 944–958.
Bank for International Settlements (BIS), 2016. Minimum capital requirements for market risk.
BenSaïda, A., Boubaker, S., Nguyen, D.K., Slim, S., 2018. Value-at-Risk under market shifts through highly flexible models. J. Forecast. 37, 790–804.
Braione, M., Scholtes, N., 2016. Forecasting Value-at-Risk under Different Distributional Assumptions. Econometrics 4 (1), 3–30.
Brooks, C., Persand, G., 2003. The Effect of Asymmetries on Stock Index Return Value-at-Risk Estimates. J. Risk Finance 4 (2), 29–42.
Christofferssen, P., 1998. Evaluating Interval Forecasts. Int. Econ. Rev. 39 (4), 841–862.
Christofferssen, P., Pelletier, P., 2004. Backtesting Value-at-Risk: A Duration-Based approach. J. Financ. Econom. 2 (1), 84–108.
Danıelsson, J., 2002. The emperor has no clothes: limits to risk modelling. J. Bank. Finance 26 (7), 1273–1296.
Emmer, S., Kratz, M., Tasche, D., 2015. What is the best risk measure in practice? a comparison of standard measures. J. Risk 18 (2), 31–60.
Engle, R., Manganelli, S., 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. J. Bus. Econ. Stat. 22, 367–381.
Fernandez, C., Steel, M.F.J., 1998. On bayesian modeling of fat tails and skewness. J. Am. Stat. Assoc. 93 (441), 359–371.
Ghalanos, A., 2015. Introduction to the Rugarch package (version 1.3-1).
Giot, P., Laurent, S., 2003. Value-at-Risk for Long and Short Trading Positions. J. Appl. Econom. 18 (6), 641–663.
Giot, P., Laurent, S., 2004. Modelling daily Value-at-Risk using realized volatility and ARCH type models. J. Empirical Finance 11 (3), 379–398.
Hansen, P.R., Lunde, A., 2005. A forecast comparison of volatility models: does anything beat a GARCH (1,1)? J. Appl. Econom. 20 (7), 873–889.
Hoppe, R., 1998. VaR and the unreal world. Risk Mag. 11, 45–50.
Jorcano, L.G., 2017. Sample size, skewness and leverage effects in Value-at-Risk and Expected Shortfall estimation. Universidad Complutense de Madrid, Madrid, Spain Ph.D thesis.
Kuester, K., Mittnik, S., Paolella, M.S., 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. J. Financ. Econom. 4 (1), 53–89.
Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. J. Derivatives 3 (2), 73–84.
McAleer, M., Jiménez-Martín, J.-Á., Pérez-Amaral, T., 2013. Has the Basel Accord improved risk management during the global financial crisis? North Am. J. Econ. Finance 26, 250–265.
McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. J. Empirical Finance 7, 271–300.
Mittnik, S., Paolella, M.S., 2000. Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates. J. Forecast. 19, 313–333.
Novales, A., Garcia-Jorcano, L., 2018. Backtesting extreme value theory models of expected shortfall. Quant. Finance 1–27.
Slim, S., Koubaa, Y., Bensaida, A., 2017. Value-at-Risk under Lévy GARCH models: evidence from global stock markets. J. Int. Financ. Markets Inst. Money 46, 30–53.
Tsay, R., 2010. Analysis of Financial Time Series, Third. J. Wiley and Sons, Hoboken.
Yamai, Y., Yoshiba, T., 2002. Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization. Monet. Economic Stud. 20 (1), 87–121.
Yamai, Y., Yoshiba, T., 2005. Value-at-risk versus expected shortfall: a practical perspective. J. Bank. Finance 29 (4), 997–1015.