Economics Bulletin

Volume 36, Issue 1

On the existence of an optimal estimation window for risk measures

Marcelo Brutti Righi
Federal University of Rio Grande do Sul

Paulo Sergio Ceretta
Federal University of Santa Maria

Abstract

We investigate whether there can exist an optimal estimation window for financial risk measures. Accordingly, we propose a procedure that achieves optimal estimation window by minimizing estimation bias. Using results from a Monte Carlo simulation for Value at Risk and Expected Shortfall in distinct scenarios, we conclude that the optimal length for the estimation window is not random but has very clear patterns. Our findings can contribute to the literature, as studies have typically neglected the estimation window choice or relied on arbitrary choices.

Citation: Marcelo Brutti Righi and Paulo Sergio Ceretta, (2016) "On the existence of an optimal estimation window for risk measures", *Economics Bulletin*, Volume 36, Issue 1, pages 1-9
Contact: Marcelo Brutti Righi - marcelobrutti@hotmail.com, Paulo Sergio Ceretta - ceretta10@gmail.com.
Submitted: September 12, 2015. **Published:** January 26, 2016.

1. Introduction

Estimating risk measures is now a standard approach in the financial field. A risk manager should consider the available information in order to forecast for the next period. Although many studies focus on both the introduction and comparison of estimation techniques, the role of information amounts is still rather neglected. If little information is considered for forecasting the next period, there is a possibility that market fundamentals will be ignored with too much reliance on short term adjustments, which tend to be volatile. However, if too much past information is used, current facts that could improve forecasting are given low importance. Thus, correctly balancing such a trade-off is crucial for correct risk measurement.

Accordingly, the following question naturally arises: is there an optimal amount of past information to use in forecasting risk measures? It is difficult to know, as there are many variables in the entire process. Nevertheless, evidence can be obtained. In that sense, the objective of this note is to show that the optimal amount is not random, thus opening the way for studies seeking to compare and obtain the number of past data in the same vein that occurs for distinct quantile levels or forecasting horizons. To that end, we present a procedure based on minimizing estimation bias. Results from Monte Carlo simulations sustain our conclusions.

2. Proposed procedure and simulation details

In this study, we focus on the risk measures most often used by both academic researchers and industry practitioners: Value at Risk (VaR) and Expected Shortfall (ES). Let *X* be the random payoff of a financial position with distribution function *F*. At the significance level $\alpha \in (0,1)$, VaR represents the lost on *X* that is only overcome with probability α, in other words, the quantile of F, i.e., $VaR^{\alpha} = -\inf\ \{q : F(q) \geq \alpha\} = -q_{\alpha}(X)$[1]. Despite its simplicity and wide use, VaR does not consider potential losses beyond the quantile level and lacks theoretical properties[2]. ES does not suffer from such drawbacks[3], as it is the expected value of a loss once it overcomes the VaR, i.e., $ES^{\alpha} = -E[X|X < -VaR^{\alpha}]$.

These risk measures are typically estimated using the last *N* observations. Thus, let $\widehat{VaR}_N^{\alpha}$ and $\widehat{ES}_N^{\alpha}$, respectively, be estimated VaR and ES at the significance level *α* based on the last *N* observations. A risk manager has the set of possible choices for *N* in the form of $\mathcal{N} = \{N_1, N_2, \cdots, N_k\}$, with $N_1 < N_2 < \cdots < N_k$. More specifically, the choice is for a $N_i, i = 1,2,\cdots,k$ between two extreme options, such as $N_{min} \leq N_i \leq N_{max}$. Most researchers and risk managers consider – for example, for daily estimation – estimation windows from one to eight years, i.e., $250 \leq N_i \leq 2000$. With that in mind, we propose to consider that the optimal choice for the estimation window is the one that minimizes the bias from the true risk measure value. In this note, formulations (1) and (2) mathematically define it for VaR and ES, respectively.

$$N_{optimal} = \inf\left(N_i \in \underset{N_i \in \mathcal{N}}{\operatorname{argmin}}\left|\widehat{VaR}_{N_i}^{\alpha} - VaR^{\alpha}\right|\right) \quad (1)$$

$$N_{optimal} = \inf\left(\underset{N_i \in \mathcal{N}}{\operatorname{argmin}}\left|\widehat{ES}_{N_i}^{\alpha} - ES^{\alpha}\right|\right) \quad (2)$$

This minimization procedure is based on absolute deviation, but other functional forms, such as least squares, can be used. Nevertheless, readers can note that for this specific case, both approaches would tend to give the same solution. Moreover, absolute deviation is linked more to distance, which is the definition of bias, beyond the fact that it does not leverage discrepancies. We consider the infimum for the case of ties because of parsimony, as one can use less data to obtain the same results.

As true VaR^{α} and ES^{α} are not observable, it is impossible to solve the problem for empirical data, but it is possible to consider simulated data where one knows the true value for risk measures. If one solves (1) or (2) for *j* samples with the same data generation process and no clear optimum exists, then $N_J = \{N^1_{optimal}, \cdots, N^j_{optimal}\} \sim \mathcal{U}(N_{min}, N_{max})$. In other words, the optimal estimation windows would assume any value between the minimum and maximum candidates with uniform (or, at least, very similar) probability.

To verify whether there is any pattern distinct from the uniformity for optimal values for *N*, we perform a Monte Carlo simulation study. To that end, we consider that returns *X*, drawn from AR (1) – GARCH (1,1) models[4], conform (3).

$$X_T = 0.50X_{T-1} + \varepsilon_T,\ \varepsilon_T = \sigma_T z_T,\ z_T \sim t_v,$$
$$\sigma_T^2 = \sigma^2(1 - 0.10 - 0.85) + 0.10\varepsilon_{T-1}^2 + 0.85\sigma_T^2. \quad (3)$$

Where, X_T, σ_T^2, ε_T and z_T are for period *T*, respectively, return, conditional variance, innovation on the expectation and a *v* degrees of freedom student white noise with $E[z_T] = 0$ and $E[(z_T)^2] = 1$. σ^2 is the unconditional variance. We consider four scenarios to contemplate the presence ($v = 6$) or not (($v = \infty$, i.e., Normal distribution) of extreme returns, as well as periods of low ($\sigma = 0.0125$) and high ($\sigma = 0.022$) volatility. The parameters have been chosen to match those obtained for daily returns of the S&P 500 index before and during the sub-prime crisis[5]. Under this specification the true values for the risk measures are $VaR_T^{\alpha} = -(0.50X_{T-1} + \sigma_T t_v^{-1}(\alpha))$ and $ES_T^{\alpha} = -\left(0.50X_{T-1} + \sigma_T\left(\alpha^{-1}\int_0^{\alpha} t_v^{-1}(s)\mathrm{d}s\right)\right)$. We choose $N_{min} = 250$ and $N_{max} = 2000$, around one and eight years, as this is the range of values that is typically is used in studies about risk estimation.

We simulate 10,000 samples with length 2001 (N_{max} plus 1 observation for the forecasting) for each scenario, and compute VaR and ES considering the Historical Simulation (HS) method. This non-parametric empirical method does not have assumptions about data and is the most commonly used in both academic studies and the financial industry[6]. Let *E* be the empirical distribution of returns $\{X\}_{2000-N+1}^{2000}$, then HS estimators are $\widehat{VaR}_N^{\alpha} = -E^{-1}(\alpha)$ and $\widehat{ES}_N^{\alpha} = -(N\alpha)^{-1}\sum_{i=1}^{N}\left(\{X\}_{2000-N+1}^{2000} * \mathbf{1}_{\{X\}_{2000-N+1}^{2000} < -\widehat{VaR}_N^{\alpha}}\right)$, where $\mathbf{1}_p$ is the indicator function that assumes value 1 if *p* is true and 0 otherwise. We compute, for each sample $\widehat{VaR}_N^{\alpha}$ and $\widehat{ES}_N^{\alpha}$ for $N_{min} = 250$, $N_{max} = 2{,}000$ and $N_{optimal}$ (solving (1) with VaR_{2001}^{α} and (2) with ES_{2001}^{α} as true values, respectively for VaR and ES. We consider 1% and 5% as values for *α*.

3. Simulation results

The results from our Monte Carlo simulation are presented in Table 1 and Figs. 1 to 4. The results from Table 1 indicate that the HS estimator produces relevant bias and variability, overestimating risk, especially in periods that are more turbulent. The exception is for Normal innovations with low volatility, which underestimates risk. Such deficiencies are in accordance with the points raised by Pritsker (2006). Nevertheless, considering the optimal estimation window reduces both bias and variability.

More specifically on the optimal estimation window, the results in Figs. 1 to 4 indicate that a common pattern is identified in most scenarios and significance levels. The optimal estimation window has more probability of occurring between 250 and 500 days (around 1 and 2 years), with some significant probability, except for the Normal distribution with low volatility around the maximum possible of 2,000 days (8 years)[7]. In some cases, as for scenarios of high volatility, there is also relevant probability for estimation windows between 750 and 1,000 days (3 and 4 years). In such cases, discrepancy is small most likely because estimation consistency is partially lost on turbulent periods. In all situations, the empirical distribution of the optimal lengths differs significantly from a Uniform distribution[8]. It is worth mentioning that the results are relatively homogenous for both VaR and ES at 1% and 5% significance levels.

Such results are in partial discordance with studies that argue in favor of larger estimation windows to improve risk forecasting, as Kuester et al. (2006) and Alexander and Sheedy (2008) for VaR, as well as Wong et al. (2012) and Righi and Ceretta (2015) for ES. This outcome can be linked to the fact that these types of studies typically rely on an arbitrary amount of past data, and even when more candidates for the estimation window are used, the comparison is very limited to specific lengths (and not to an entire interval of possible lengths as we do in our simulation exercise) and there is no consensus. Of course, we are not saying here that this is the optimal solution for everyone who uses empirical risk estimation, but it is very strong evidence that an optimal estimation window can exist. This phenomenon is not well investigated in the current literature.

4. Conclusion

In this note, we conduct a Monte Carlo simulation to show that the optimal amount of past information in risk measures forecasting is not random and can directly affect the quality of forecasting. To that end, we propose a procedure that chooses the optimal estimation window by minimizing estimation bias. Our results, which are obtained for VaR and ES under distinct scenarios and quantiles, indicate that the optimal estimation windows are not uniformly distributed, and that most probability is for the interval between 1 and 2 years (for daily forecasting). Our focus here is not to say what the optimum is, because we only consider one estimation model (HS) and a limited number of possibilities, but indicate that such an optimum can exist. The literature must start to pursue it rather than place trust in very arbitrary choices.

[1] See Duffie and Pan (1997) for a review on VaR.
[2] VaR is not coherent in the sense of Artzner et al. (1999) as it does not have the subadditivity property that guaranteed risk diversification.
[3] ES is coherent, as explained in Acerbi and Tasche (2002).
[4] This data generation process is often considered for finance because it contemplates stylized facts of daily financial returns, such as volatility clusters and heavy tails.
[5] This is a choice of the authors because this index is one of the most representative and is usually considered in simulation studies (see Christoffersen and Gonçalves (2005) for instance).
[6] Pérignon and Smith (2010) indicate that 76% of financial institutions that disclose their VaR methodology use HS for estimation.
[7] Perhaps if a larger value for N_{max} is considered, such probability around 2,000 could be dispersed.
[8] We conduct usual chi-squared tests for the null hypothesis of Uniform distribution.

References

Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. J. Bank. Financ. 26, 1487–1503.

Alexander, C., Sheedy, E., 2008. Developing a stress testing framework based on market risk models. J. Bank. Financ. 32, 2220–2236.

Artzner, P., Delbaen, F., Eber, J.-M., Heath D., 1999. Coherent measures of risk. Math. Financ. 9, 203–228.

Christoffersen, P., Gonçalves, S., 2005. Estimation risk in financial risk management. J. Risk. 7, 1–28.

Duffie, D., Pan, J.,1997. An overview of value at risk. J. Deriv. 4, 7–49.

Kuester, K., Mittnik S., Paolella, M. S., 2006. Value-at-risk prediction: A comparison of alternative strategies. J. Financ. Economet. 4, 53–89.

Pérignon, C., Smith, D. R., 2010. The level and quality of Value-at-Risk disclosure by commercial banks. J. Bank. Financ. 34, 362–377.

Pritsker, M., 2006. The hidden dangers of historical simulation. J. Bank. Financ. 30, 561–582.

Righi, M. B., Ceretta, P. S., 2015. A comparison of expected shortfall estimation models. J. Econ. Bus. 78, 14–47.

Wong, W. K., Fan, G., Zeng, Y., 2012. Capturing tail risks beyond VaR. Rev. Pac. Basin Finan. Mark. Pol. 15, 1250015.

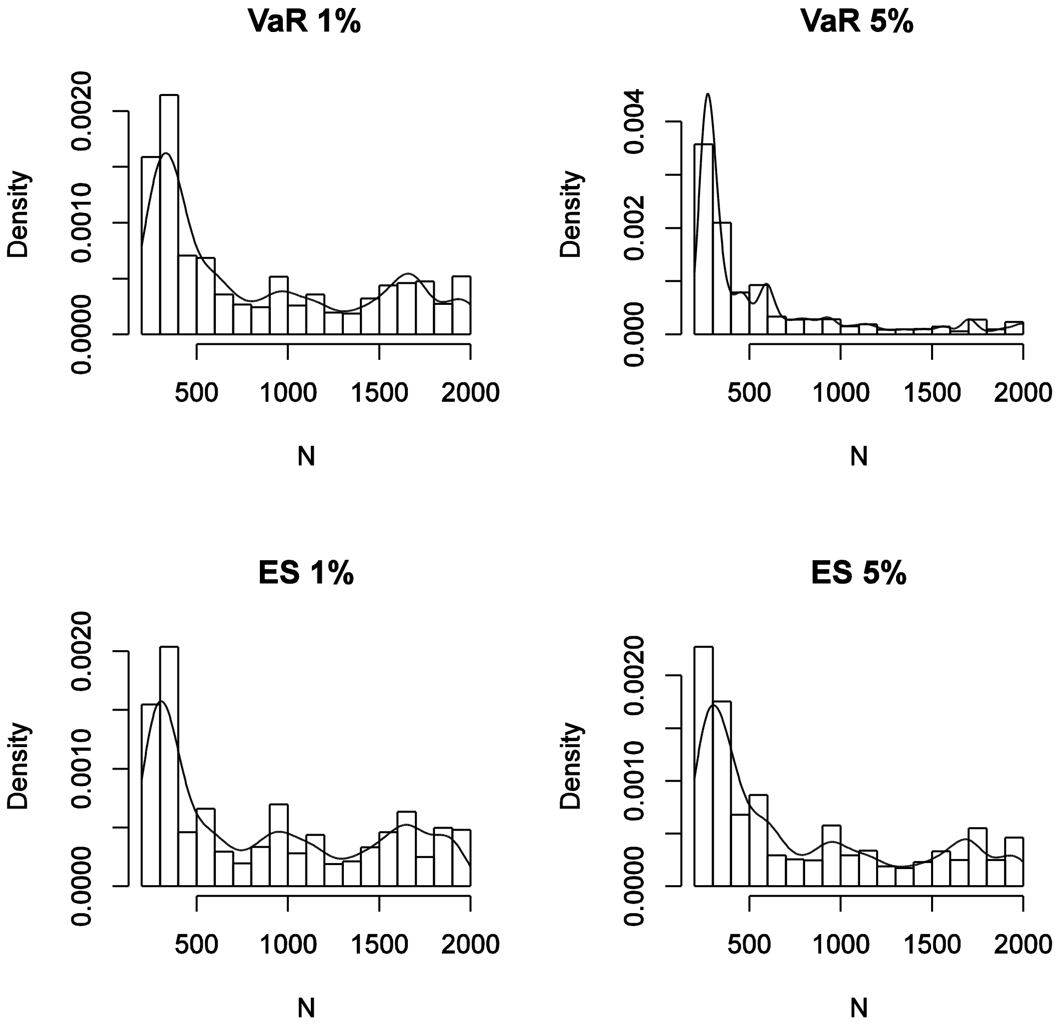

Fig. 1. Histograms and densities of VaR and ES optimal *N* obtained through Monte Carlo simulation under Normal GARCH with low volatility.

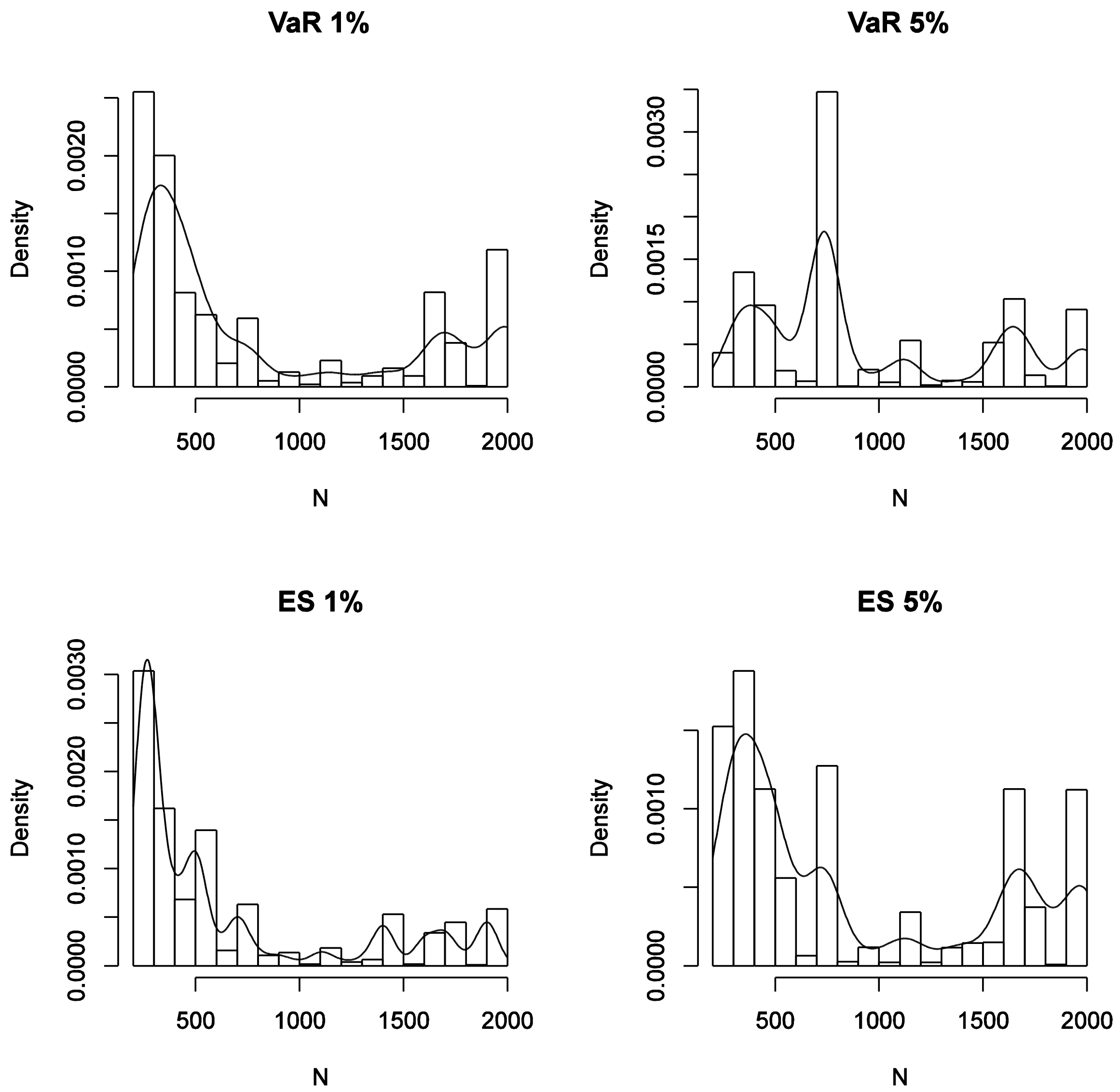

Fig. 2. Histograms and densities of VaR and ES optimal *N* obtained through Monte Carlo simulation under Normal GARCH with high volatility.

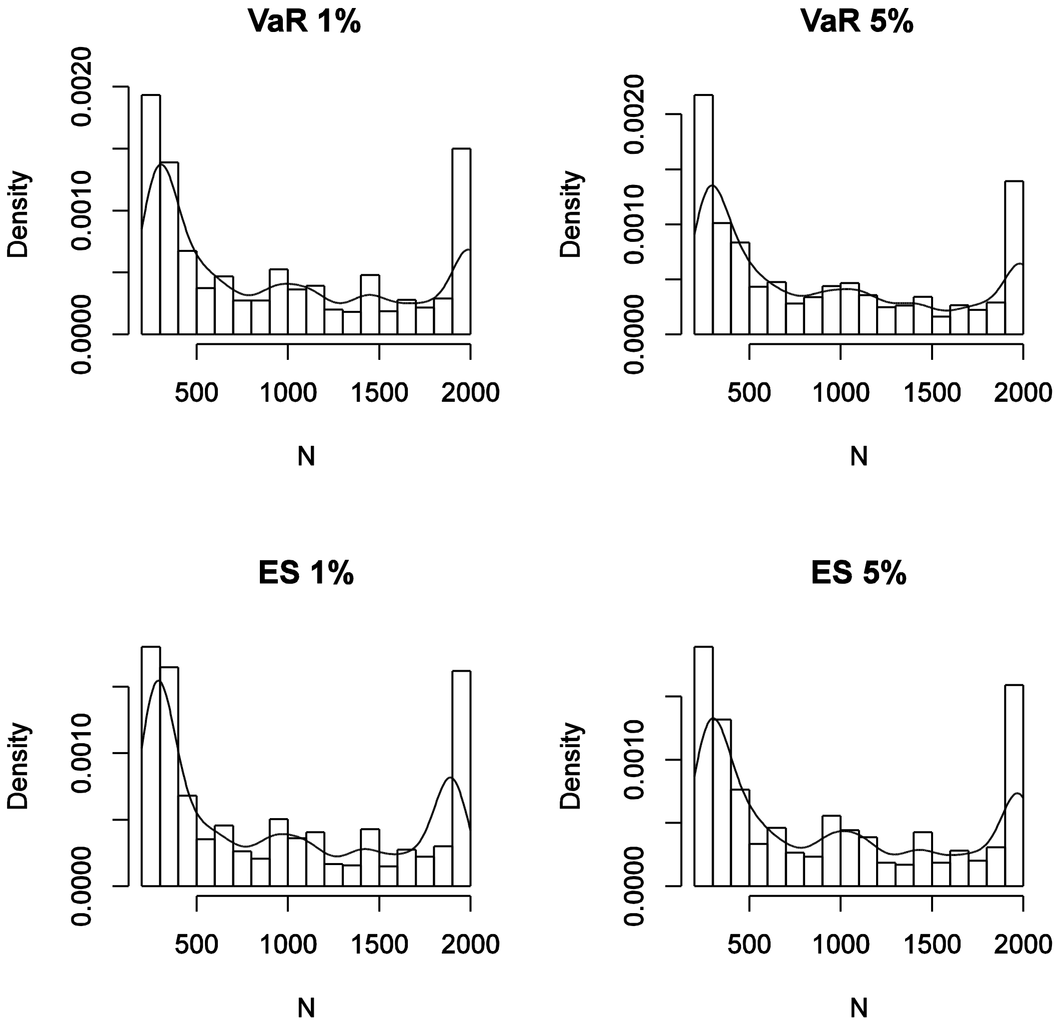

Fig. 3. Histograms and densities of VaR and ES optimal *N* obtained through Monte Carlo simulation under Student's *t* GARCH with low volatility.

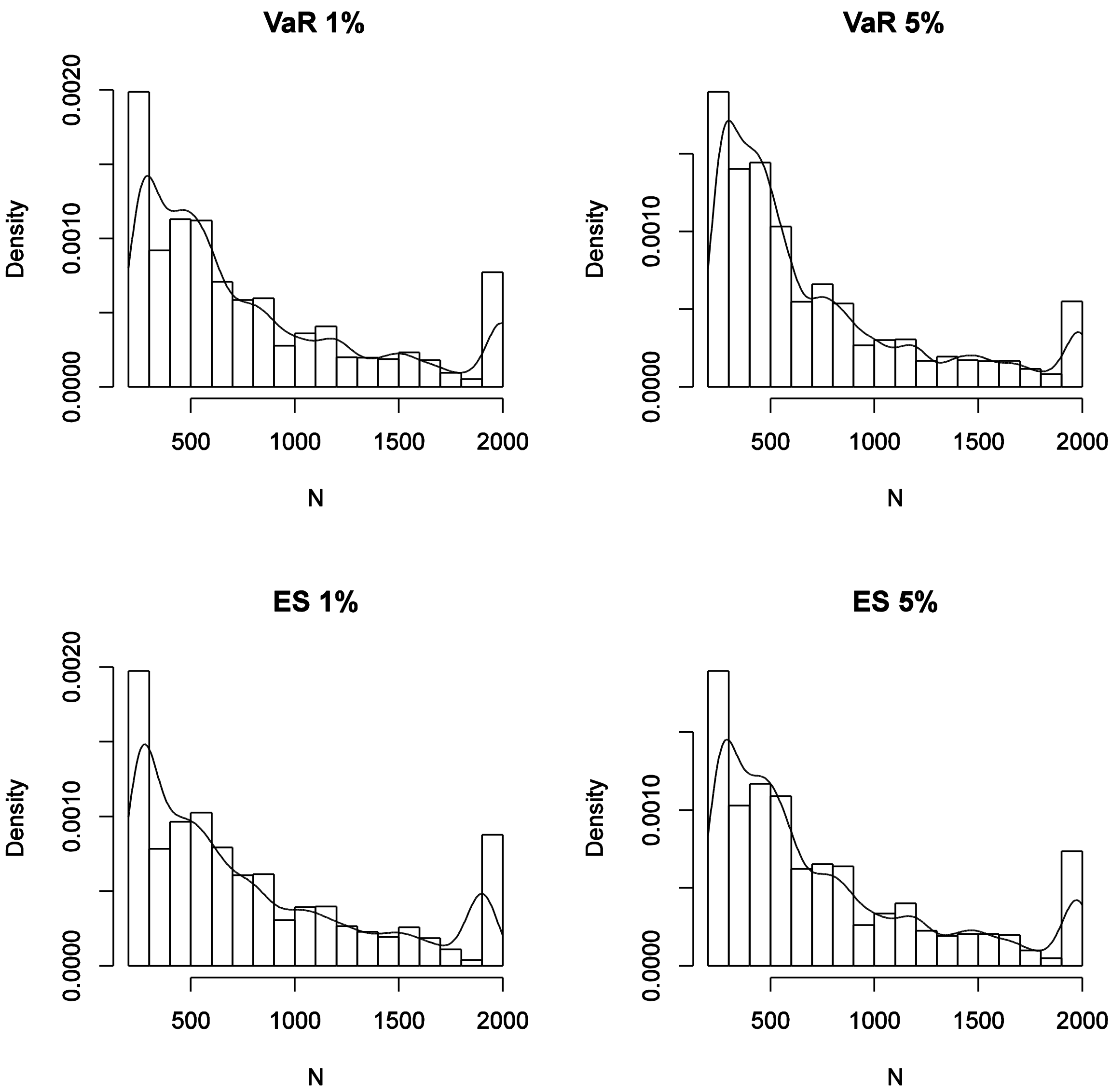

Fig. 4. Histograms and densities of VaR and ES optimal *N* obtained through Monte Carlo simulation under Student's *t* GARCH with high volatility.

Table 1 . Bias, Root Mean Squared Error and optimal *N* obtained in the Monte Carlo Simulations.

Normal GARCH						
	Low Volatility			**High Volatility**		
	Bias (%)	RMSE (%)	Mean N_J	Bias (%)	RMSE (%)	Mean N_J
$\widehat{VaR}^{1\%}_{250}$	-0.0492	0.1464	250	0.5755	0.3596	250
$\widehat{VaR}^{1\%}_{2000}$	-0.0581	0.1722	2000	0.5855	0.3008	2000
$\widehat{VaR}^{1\%}_{Optimal}$	-0.0413	0.1201	847	0.4689	0.2739	816
$\widehat{VaR}^{5\%}_{250}$	-0.0952	0.1754	250	0.7767	0.3566	250
$\widehat{VaR}^{5\%}_{2000}$	-0.1784	0.1486	2000	0.6074	0.3077	2000
$\widehat{VaR}^{5\%}_{Optimal}$	-0.0975	0.1262	563	0.5816	0.2836	943
$\widehat{ES}^{1\%}_{250}$	-0.0491	0.2002	250	0.4910	0.3625	250
$\widehat{ES}^{1\%}_{2000}$	0.0023	0.1899	2000	0.6155	0.3188	2000
$\widehat{ES}^{1\%}_{Optimal}$	-0.0215	0.1194	865	0.4187	0.2757	655
$\widehat{ES}^{5\%}_{250}$	-0.0584	0.1843	250	0.6457	0.3517	250
$\widehat{ES}^{5\%}_{2000}$	-0.0922	0.1646	2000	0.6208	0.3029	2000
$\widehat{ES}^{5\%}_{Optimal}$	-0.0580	0.1221	786	0.5246	0.2778	882
Student's *t* GARCH						
	Low Volatility			**High Volatility**		
	Bias (%)	RMSE (%)	Mean N_J	Bias (%)	RMSE (%)	Mean N_J
$\widehat{VaR}^{1\%}_{250}$	0.7535	1.4633	250	0.7383	1.2357	250
$\widehat{VaR}^{1\%}_{2000}$	0.9379	2.3075	2000	1.4199	2.8599	2000
$\widehat{VaR}^{1\%}_{Optimal}$	0.3937	0.7748	941	0.4142	0.6474	770
$\widehat{VaR}^{5\%}_{250}$	0.4386	1.1167	250	0.6868	1.0167	250
$\widehat{VaR}^{5\%}_{2000}$	0.2087	0.7295	2000	0.8489	1.4105	2000
$\widehat{VaR}^{5\%}_{Optimal}$	0.1301	0.4097	920	0.3372	0.4742	711
$\widehat{ES}^{1\%}_{250}$	0.8477	1.6805	250	0.6567	1.1969	250
$\widehat{ES}^{1\%}_{2000}$	1.4043	3.2037	2000	1.8530	3.5519	2000
$\widehat{ES}^{1\%}_{Optimal}$	0.5327	1.0453	904	0.4279	0.7581	768
$\widehat{ES}^{5\%}_{250}$	0.6738	1.3607	250	0.7337	1.1448	250
$\widehat{ES}^{5\%}_{2000}$	0.7537	1.7613	2000	1.3449	2.4114	2000
$\widehat{ES}^{5\%}_{Optimal}$	0.3361	0.6915	947	0.4085	0.6030	759