CIRRELT
1971 - 2021

CIRRELT-2021-04

The New International Regulation of Market Risk: Roles of VaR and CVaR in Model Validation

Samir Saissi Hassani
Georges Dionne

January 2021

Bureau de Montréal
Université de Montréal
C.P. 6 128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

The New International Regulation of Market Risk: Roles of VaR and CVaR in Model Validation

Samir Saissi Hassani[1], Georges Dionne[1,2,*]

[1] Canada Research Chair in Risk Management, HEC Montreal

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

Abstract. We model the new quantitative aspects of market risk management for banks that Basel established in 2016 and came into effect in January 2019. Market risk is measured by Conditional Value at Risk (CVaR) or Expected Shortfall at a confidence level of 97.5%. The regulatory backtest remains largely based on 99% VaR. As additional statistical procedures, in line with the Basel recommendations, supplementary VaR and CVaR backtests must be performed at different confidence levels. We apply these tests to various parametric distributions and use non-parametric measures of CVaR, including CVaR- and CVaR+ to supplement the modelling validation. Our data relate to a period of extreme market turbulence. After testing eight parametric distributions with these data, we find that the information obtained on their empirical performance is closely tied to the backtesting conclusions regarding the competing models.

Keywords: Basel III, VaR, CVaR, Expected Shortfall, backtesting, parametric model, non-parametric model, mixture of distributions, fat-tail distribution.

Acknowledgements. Research funded by the Canada Research Chair in Risk Management. We thank Claire Boisvert for her efficient assistance in the preparation of the manuscript. A preliminary version of this document was published in French in *Assurances et gestion des risques /* Insurance and Risk Management.

---

* Corresponding author: georges.dionne@hec.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2021

Introduction

In 2016, the Basel Committee decided that the market risk capital of banks should be calculated with CVaR or Expected Shortfall[1] at the 97.5% confidence level, while maintaining the backtesting of the models, as before, at 99% VaR (BCBS, 2016, 2019). This shift toward CVaR would be motivated by issues of consistency and the inadequacy of the risk coverage by the VaR, which has been noted over time. Market risk is now jointly managed by CVaR and VaR, at two different probabilities: $p = 2.5\,\%$ and $p = 1\,\%$ respectively.[2]

Further, Basel suggests adding statistical procedures to ensure the ex-post suitability of models (BCBS, 2016, page 82; BCBS, 2019, paragraph 32.13). We therefore perform four backtests in addition to the 1% VaR backtest, including two on VaR at $p = 2.5\,\%$ and $p = 5\,\%$ and two others on CVaR at $p = 2.5\,\%$ and $p = 5\,\%$. We use non-parametric measures, including CVaR- and CVaR+, to supplement the validation of the distributions used. The aim of this paper is to orchestrate all these aspects in a validation process compatible with the regulations in force.

We are working with data obtained from three risky stocks—IBM, General Electric and Walmart—, whose price fluctuations refute the usual assumptions of normality of returns during the period examined. The study period encompasses the extreme price fluctuations during the last economic recession in the United States (NBER,[3] December 2007 to June 2009) and the financial crisis of 2007-2009. We evaluate the behavior of VaR and CVaR using several parametric distributions to model returns: the normal distribution, Student's *t*, the EGB2 (Exponential GB2), SN2 (Skewed Normal Type 2), and SEP3 (Skewed Exponential Power Type 3).[4] We also construct homogeneous and heterogeneous mixtures of parametric densities. Eight models are analyzed in order to identify the distributions that best represent the data to manage the market risk contained therein.

[1] CVaR is also called Expected Shortfall in the literature. Both measures are equivalent with continuous distributions without jumps (Rockafellar and Uryasev, 2002). See also Dionne (2019).
[2] In this paper, we use the letter p to refer to the probability that the VaR is exceeded and 1-p for the corresponding confidence level. The *p*-value notation is for statistical tests.
[3] https://www.nber.org/cycles.html
[4] For a description of SN2 and SEP3, see Fernandez et al. (1995) and Rigby et al. (2014).

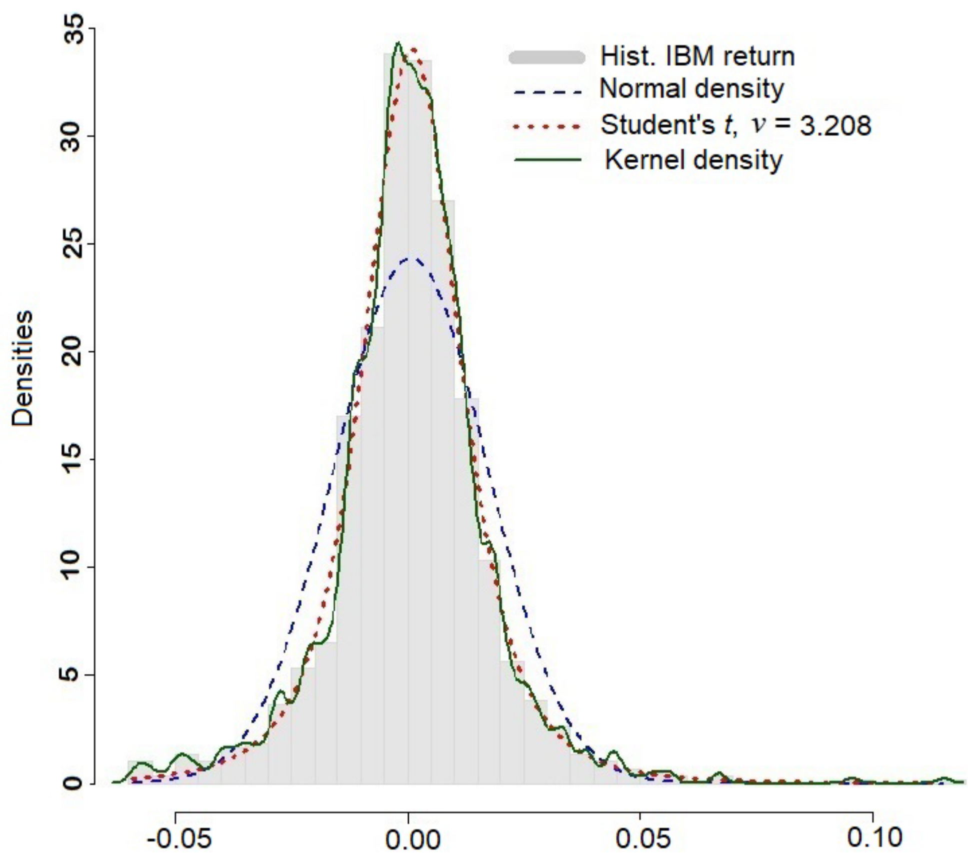
The analysis comprises three steps. First, the estimation of the models' parameters is validated by standard measures such as the AIC, BIC and Kolmogorov-Smirnov goodness-of-fit test. The second validation consists in comparing the kurtosis and asymmetry obtained from the parametric models with the same moments determined by a non-parametric approach of the data. The most important point in this step is to evaluate each model by comparing the value of its parametric CVaR against the non-parametric interval [$CVaR-_{np}$, $CVaR+_{np}$] which is computed from our sample of returns following Rockafellar and Uryasev (2002). Given that the three non-parametric measures of the sample obey the fundamental inequalities $CVaR-_{np} \leq CVaR_{np} \leq CVaR+_{np}$, we consider that a good model should also produce a CVaR that obeys the same framing: $CVaR-_{np} \leq CVaR_{Model} \leq CVaR+_{np}$. The third validation is the backtesting of the risk measures, which we carry out in compliance with the Basel regulations in force for market risk. We find that the results of the last two steps are strongly linked; failing to validate that a distribution properly fits the data would significantly affect the backtesting results.

Given that the calculations of parametric CVaR are much more complex than those of VaR, we define in detail each of the distributions or mixtures of distributions that we use in the eight models, together with the mathematical derivations of the corresponding CVaR. The mathematical developments are presented in the appendices. Appendix A1 describes the symbols of the different models. Appendix A2 shows the general expression of CVaR, and Appendix A3 outlines the general expression of CVaR from a mixture of distributions. Details of the statistical models and the backtesting procedure are also provided in the appendices.

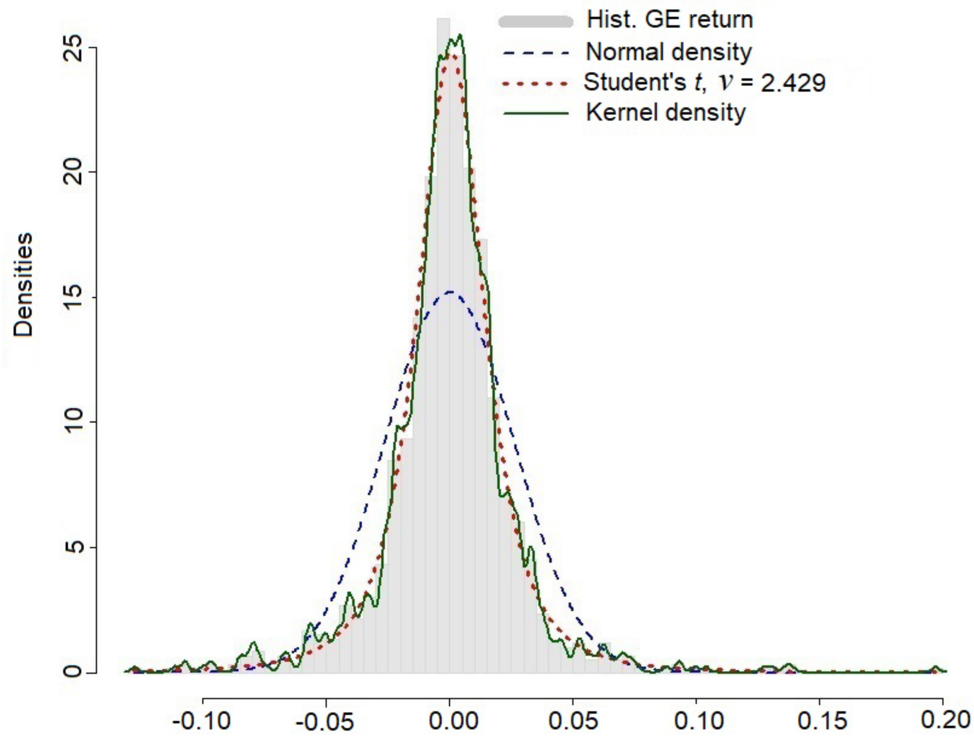
The following section presents the data used. Section 2 provides a preliminary analysis of the data. Section 3 is devoted to estimating the parameters of the eight competing models and empirically verifying their respective performances. Section 4 conducts backtesting of the models and the final section concludes the paper.

1. Data

First, we describe the data. The three risky stocks chosen are IBM, General Electric (GE) and Walmart (WM). The period consists of 1,200 days, from June 18, 2007 to March 20, 2012.[5] Returns are calculated by taking dividend payments into account. Figure 1 shows the distribution of returns for the securities over this period, which are far from normal. Fitting a Student's *t*-distribution to the returns, the estimated degrees of freedom (ν) are 3.2, 2.4 and 3.2 respectively, indicating the presence of a very fat-tail.



[5] The actual daily price extraction period is from June 15, 2007 to March 20, 2012, representing 1,201 days, which provides 1,200 daily returns from June 18, 2007 to March 20, 2012.



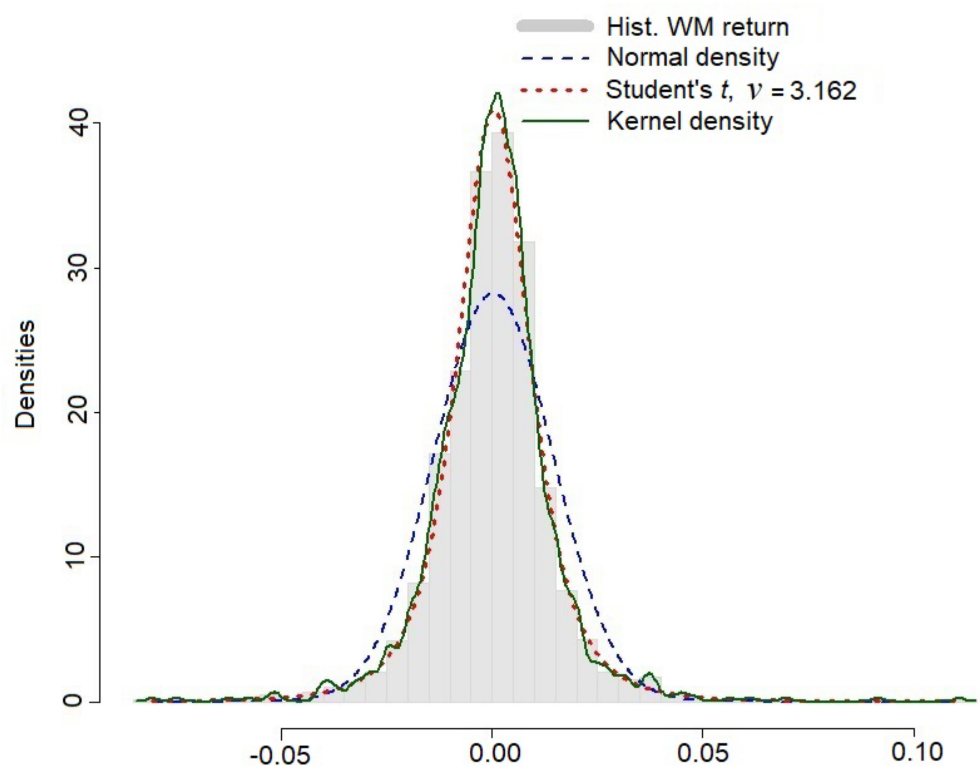


Figure 1: Histograms and densities of IBM, GE and WM stocks

Table 1 presents descriptive statistics including correlation matrix, variance-covariance matrix, and the first four nonparametric moments. The positive correlations are very strong during this period of financial crisis, at about 50%.

Table 1 Matrices of correlations, variance-covariances and nonparametric moments

	IBM	General Electric	Walmart
IBM	1		
General Electric	0.567592	1	
Walmart	0.491463	0.430835	1
IBM	0.02688%	0.02443%	0.01138%
General Electric	0.02443%	0.06894%	0.01598%
Walmart	0.01138%	0.01598%	0.01995%
Mean	0.07580%	-0.00286%	0.03472%
Variance	0.02688%	0.06894%	0.01995%
Skewness	0.27190	0.35375	0.35429
Kurtosis	7.43415	9.95718	10.68244

The average daily returns are practically nil. The skewness coefficients are positive, showing that all three distributions are pulled toward positive returns. This is surprising because one might have expected to see a shift toward the left tail of losses during this time period. The most important point is that the kurtosis coefficients are very large, about three times larger than the kurtosis of a normal distribution. This confirms the very large tail thickness reported in the previous paragraphs.

2. Preliminary data analysis

We begin by calculating the optimal weights of a portfolio that minimizes the relative VaR at the 95% confidence level (p = 5%) under the constraint that the weights sum to 1. We assume the returns to be normally distributed for the moment so minimizing the relative VaR is equivalent to minimizing the CVaR plus the mean of the portfolio returns. Moreover the optimal weights are independent of the chosen p, as we will see later. We also assume that these weights will remain optimal for all distributions studied in the next sections to obtain comparable backtesting results between the different models. In other words, we suppose there is separation between portfolio optimization decision and model backtesting as it is often observed in many financial institutions.

Under the normal assumption, relative VaR of the portfolio is written as:

$$\mathrm{VaR}_r = -\sigma_{\mathrm{portfolio}} \times \Phi_0^{-1}(p) = -\sqrt{\beta^T \Sigma \beta} \times q_0 > 0 \tag{1}$$

where $\Phi_0^{-1}(\cdot)$ is the inverse of the cumulative function of $N(0,1)$ evaluated at p, β is the vector of security weights, β^T is the transpose of β, Σ is the variance-covariance matrix of security returns and q_0 is the quantile of $N(0,1)$ relative to p. The VaR expression is positive since $q_0 < 0$ in left tail. On the other hand, since q_0 depends only on p, equation (1) is minimised on the term $\sqrt{\beta^T \Sigma \beta}$ alone. Thus, the optimal weights are independent of the chosen p. The Excel file[6] gives the results of Table 2 in percentages for p = 5%. The table shows that the VaR of the optimal portfolio are lower than those of the weighted assets, which is consistent with the diversification principle.

Table 2
Optimal portfolio: Relative VaR[1] minimization (normal model)

		IBM	General Electric	Walmart	Total	
Weight		0.3889444	-0.0465131	0.6575686	1.00000	
Portfolio	p = 5%	$q_0 = -1.64485$				
Mean	0.05244%	0.07580%	-0.00286%	0.03472%		
Variance	0.01680%					
Standard deviation	1.29631%	1.63961%	2.62559%	1.41254%		
Skewness	0.3588781					
Kurtosis	9.8157828					
	Portfolio				Weighted sum	Difference
VaRa	2.07980%	2.62112%	4.32157%	2.28871%	2.72546%	-0.64566%
VaRr	2.13224%	2.69692%	4.31871%	2.32343%	2.77764%	-0.64540%

[1] VaRa and VaRr refer to absolute VaR and relative VaR respectively.

We now compute the CVaR of the optimal portfolio. With the normal assumption, CVaR is written according to equation (2) (based on equation A6 in the appendix):

[6] The Excel file is available on the Canada Research Chair website at https://chairegestiondesrisques.hec.ca/en/seminars-and-publications/book-wiley/

$$\begin{aligned} \mathrm{CVaR} &= -\mu_{\mathrm{portfolio}} + \sigma_{\mathrm{portfolio}} \frac{\phi_0(q_0)}{p} \\ &= -\mu_{\mathrm{portfolio}} + \sqrt{\beta^T \Sigma \beta} \times \frac{\phi_0(q_0)}{p} \end{aligned} \tag{2}$$

where $\phi_0(\cdot)$ is the density function of $N(0,1)$. Table 3 shows the results for the optimal portfolio. From equation (2), we can see that $\mathrm{CVaR} + \mu_{\mathrm{portfolio}}$ is minimized on $\sqrt{\beta^T \Sigma \beta}$ alone, since $\phi_0(q_0)/p$ is constant. Thus, the optimal weights are the same as minimizing relative VaR. The portfolio's CVaR is also naturally lower than that of the weighted assets, which means that the basic principle of diversification is followed. In the next section we continue to work with p = 5%. At the end of the section we will discuss the effect of this choice with respect to p = 2.5% and p = 1%.

We want to calculate the VaR, CVaR, CVaR- and CVaR+ measures for the sample data at p = 5% following Rockafellar and Uryasev (2002). With our notation, we can write:

$$\mathrm{CVaR}^+ = E\{X_t | X_t < -\mathrm{VaR}\} \text{ and } \mathrm{CVaR}^- = E\{X_t | X_t \leq -\mathrm{VaR}\} \tag{3}$$

where the vector $\{X_t\}_{t=1}^{t=1200}$ denotes the portfolio returns. Note the minus sign in front of VaR since $\mathrm{VaR} > 0$ and $X_t < 0$ in the left tail of the distribution. The data $\{X_t\}$ can be seen as a discrete finite sample drawn from an unknown distribution. Therefore, a non-parametric estimate of equations in (3) can be obtained from the historical simulation method by writing:

$$\mathrm{CVaR}_{np}^+ = \frac{1}{N_T^+} \sum_{t=1}^{T} X_t \times (X_t < -\mathrm{VaR}_{np}) \text{ and } \mathrm{CVaR}_{np}^- = \frac{1}{N_T^-} \sum_{t=1}^{T} X_t \times (X_t \leq -\mathrm{VaR}_{np}) \tag{4}$$

where $N_T^+ = \sum_{t=1}^{T} 1 \times (X_t < -\mathrm{VaR}_{np})$ and $N_T^- = \sum_{t=1}^{T} 1 \times (X_t \leq -\mathrm{VaR}_{np})$.

The results are presented in Table 4. CVaR_{np} is not shown in the table because CVaR_{np} is equal to $\mathrm{CVaR+}_{np}$. This last fact is also verified at p = 2.5% and p = 1%.[7] The relative difference

[7] The reason is that 1,200 × 5 = 60, 1,200 × 2.5% = 30 and 1,200 × 1% = 12 are all integers. Therefore, there are no split atoms over VaR.

between CVaR$-_{np}$ and CVaR+$_{np}$ $(2.97795\% - 2.96269\%)/2.96269\% = 0.51\%$ is extremely small. This adds a significant selective requirement in identifying a distribution or mixture of continuous parametric distributions whose CVaR must be framed by CVaR- and CVaR+.[8]

Table 3
Optimal portfolio : CVaR (normal model)

		IBM	General Electric	Walmart	Total	
Weight		0.3889444	-0.0465131	0.6575686	1.00000	

	Portfolio				Weighted sum	Difference
CVaR	2.62147%	2.62112%	4.32157%	2.28871%	2.72546%	-0.10399%
$\mu_{portfolio}$+CVaR	2.67392%	2.69692%	4.31871%	2.32343%	2.77764%	-0.10372%

Figure 2 clearly shows that a normal density would not be appropriate to represent the optimal portfolio data. On the other hand, Student's *t* would not be sharp enough and does not keep enough mass around the mode, as would be required by the kernel density plot of the data. These remarks would rather suggest a Laplace distribution.

Table 4
Non-parametric[1] measurements of VaR, CVaR- and CVaR+

		IBM	General Electric	Walmart		
Weight B		0.3889444	-0.0465131	0.6575686	1.00000	
p = 5%	$q_0 = -1.64485$					
	Portfolio		Order: 61		Weighted sum	Difference
VaR$_{np}$	2.04736%	2.69619%	4.12926%	2.02620%	2.57310%	-0.52573%
CVaR$-_{np}$	2.96269%	3.80857%	6.39598%	3.19753%	3.88142%	-0.91873%
CVaR+$_{np}$	2.97795%	3.82711%	6.43375%	3.21706%	3.90322%	-0.92528%

[1] np: non-parametric.

[8] Assuming that there is some discontinuity in the distribution of returns around VaR$_{np}$.

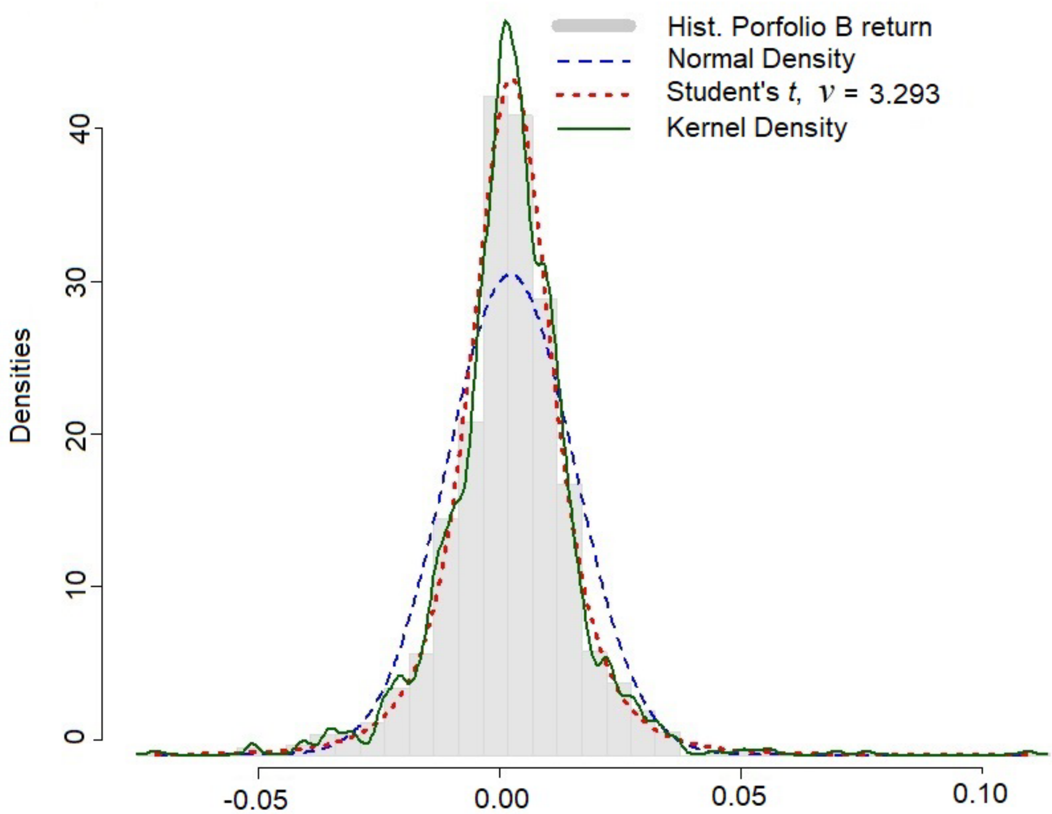


Figure 2: Histogram and densities of optimal portfolio

3. Estimation of the parametric distributions

From now on, we are using the weights of optimal portfolio in Section 2 as the reference portfolio. The VaR calculated from the parametric models will be the absolute VaR relative to 0, like the CVaR. The models are denoted as M1 to M8. Complete estimations of model coefficients are presented in the following tables and in tables A2 and A3 in the appendix.

For comparison, we start with the M1 model, assuming that the data follow a normal distribution (see definitions and expressions in Appendix A4). Model M1 is denoted 1:NO, in that it consists of a single normal distribution. The results are presented in Table 5. VaR of this model is higher than VaR_{np} (non-parametric VaR), whereas CVaR is much lower than the two non-parametric CVaR. Unsurprisingly, the Kolmogorov-Smirnov (KS) test rejects this model (*p*-value = 0.0015 < 10%). The reported asymmetry and kurtosis values, 0.3589 and 9.8158, correspond to the empirical moments of portfolio returns (Table 2). They are very different compared to those of the normal distribution.

Table 5
Model M1 = 1:NO

M1 parameters		Normal distribution (1:NO)			p = 5%		Order[1]: 61
μ	0.0005244		Quantiles, coefficients and probabilities				
σ	0.0129631		q_0	μ	$\Phi^0[q_0]$	$-\sigma$	$\phi^0[q_0]$
			-1.644854	0.000524	0.050000	-0.012963	0.10313564
						q	$F_{Model}[q]$
VaR	2.07980%	VaR_{np}: 2.04736%				-2.07980%	0.050000
CVaR	2.62147%	$CVaR-/+_{np}$: [2.96269%, 2.97795%]					

	Model	Data				
Skewness	0.0000	0.3589	AIC	-7,021.1012	KS stat.	0.0775
Kurtosis	3.0000	9.8158	BIC	-7,010.9210	KS *p*-value	0.0015

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.

We now turn to the Student's *t*-distribution (M2 = 1:T; see definitions and expressions in Appendix A5). The estimated degree of freedom parameter is $\nu = 3,28871$. The kurtosis is thus undefined $(\text{since } \nu < 4)$. This time, the VaR < VaR_{np} and the CVaR > $\left[CVaR -_{np},\ CVaR +_{np}\right]$, which is the opposite situation compared to 1:NO. Further, the *p*-value = 0.1285 > 10% means that the KS test does not reject this model. The AIC and BIC criteria improve compared to the normal M1 model. Moreover, these AIC and BIC values are the smallest of all the models. Still, the fact that the CVaR > $\left[CVaR -_{np},\ CVaR +_{np}\right]$ is problematic. The reason for this is probably related to the fact that Student's *t* allows to account for tail thickness, but its kurtosis is undefined. In addition, Student's *t* does not capture the asymmetry of the data.

Table 6
Model M2 = 1:T

M2 parameters		Student's *t* distribution (1:T)		p = 5%		Order[1]: 61
μ	0.0006974		Quantiles, coefficients and probabilities			
σ	0.0085310	q_0	μ	$F^0_{T,\nu}\left[q_0\right]$	$-\sigma$	$\text{Tail}^0_{T,\nu}\left[q_0\right]$
ν	3.2887197	-2.271479	0.000697	0.050000	-0.008531	0.180675
					q	$F_{\text{Model}}\left[q\right]$
VaR	1.86806%	VaR_{np}: 2.04736%			-1.86806%	0.050000
CVaR	3.01294%	$CVaR\text{-/+}_{np}$: [2.96269%, 2.97795%]				

	Model	Data				
Skewness	0.0000	0.3589	AIC	-7,249.1447	KS stat.	0.0483
Kurtosis	Indefinite	9.8158	BIC	-7,233.8745	KS *p*-value	0.1285

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.

We now move on to the M3 model, using the EGB2 distribution (CVaR calculations are made by numerical integrals because the analytical expression is not available; see definitions and expressions in Appendix A6). This model provides one more parameter. Indeed, the parameters ν and τ characterize both the tail thickness and the asymmetry of the distribution. The distribution is skewed negatively or positively, or is symmetric when $\nu < \tau$, $\nu > \tau$, or $\nu = \tau$ respectively. As for the thickness of the tail, the smaller the ν, the thicker the left tail (all other parameters kept equal). Estimation of the M3 model gives $\tau = 0.1652.$ and $\nu = 0.1587$. Given that ν is very close to $\tau,$ we have a slight negative skewness = –0.081 (Table 7), which is not compatible with the nonparametric skewness coefficient of the data = 0.359.

Table 7
Model M3 = 1:EGB2

M3 parameters		Exponential GB2 distribution (1:EGB2)		p = 5%	Order[1]:61
μ	0.0008884				
σ	0.0014108				
ν	0.1587161			q	$F_{Model}[q]$
τ	0.1652522			-2.00674%	0.050000
VaR	2.00674%	VaR_{np}: 2.04736%			
CVaR	2.89562%	CVaR-/+$_{np}$: [2.96269%, 2.97795%]			
	Model	Data			
Skewness	-0.0813	0.3589	AIC -7,243.1659	KS stat.	0.0375
Kurtosis	5.8076	9.8158	BIC -7,222.8056	KS *p*-value	0.3490

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.

VaR of M3 is the closest to VaR_{np} so far. However, $CVaR = 2.89562\,\% < \left[CVaR -_{np},\ CVaR +_{np}\right]$. The kurtosis of 5.8 is still insufficient compared with 9.8. Despite the great flexibility mentioned in the literature regarding the four-parameter EGB2, these results seem to indicate that a single parametric distribution would not be sufficient to properly identify the risks inherent in our data, despite the fact that the KS test does not reject this model (*p*-value = 0.3490).

A last word concerning the values $\nu = 0.1587$ and $\tau = 0.1652$. Given that the parameters ν and τ are very small and near zero, we know the lemma 2 of Caivano and Harvey (2014), which says that the EGB2 tends toward a Laplace density when $\nu \approx \tau \approx 0$. This directly corroborates the observation of the sharp mode of the kernel density plotted in Figure 2. We will observe this convergence toward a Laplace distribution below. We now move on to mixed distributions.

We estimate VaR and CVaR of the M4 model constructed with a mixture of two normal distributions (2:NO) using the expressions given in appendices A2, A3 and A4. The quantile q_m at the degree of confidence $(1-p)$ of a mixture of densities is obtained by a numerical method. VaR is equal to $-q_m$. CVaR is calculated using the value of q_m. The results of the 2:NO model, presented in Table 8, show that we may be on the right track with a mixture of densities. $\text{VaR} = 1.95397\,\%$ and $\text{CVaR} = 3.11363\,\%$ clearly approach the non-parametric measurements compared with those obtained for the 1:NO distribution. Kurtosis = 6.7 also improves. The KS test gives a *p*-value = 0.2181, which is comfortably above 10%. However, the CVaR of the 2:NO model is still far from the range $\left[\text{CVaR}-_{\text{np}},\ \text{CVaR}+_{\text{np}}\right]$.

Table 8
Model M4 = 2:NO

M4 Parameters			Mixture of 2 normal distributions (2:NO)			p = 5%		Order[1]: 61
Distribution 1	μ_1	-0.0004845						
	σ_1	0.0226636			Quantiles, coefficients and probabilities			
Distribution 2	μ_2	0.0008151	Density	q_0	μ	$\Phi^0[q_0]$	$-\sigma$	$\phi^0[q_0]$
	σ_2	0.0082545	1	-0.840784	-0.000484	0.200234	-0.022664	0.280159
	c_1	0.2231962	2	-2.465888	0.000815	0.006834	-0.008255	0.019078
							q_m	$F_{\text{Model}}[q_m]$
							-1.95397%	0.050000
VaR[2]		1.95397%	VaR_{np}: 2.04736%					
CVaR		3.11363%	$\text{CVaR-/+}_{\text{np}}$: [2.96269%, 2.97795%]					

	Model	Data				
Skewness	-0.1386	0.3589	AIC	-7,228.0889	KS stat.	0.0433
Kurtosis	6.6789	9.8158	BIC	-7,202.6385	KS *p*-value	0.2180

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.
[2] Obtained numerically from the Excel solver by minimizing $\left(F_m\left(q_m\right)-p\right)^2$.

The estimation of the mixture of two Student's *t* distributions (2:T, see definitions and expressions in appendices A2, A3 and A5) presented in Table 9 demonstrates a very large parameter for the degree of freedom of the first Student’s *t* $\nu_1 = 23{,}642.3\ > 30$, clearly indicating that the first Student's *t* is practically a normal distribution. The second distribution with a degree of freedom $\nu_2 = 6.4162 > 4$ allows the mixture to now have a well-defined kurtosis of 8.4, close to the kurtosis of data of 9.8. We have a *p*-value equal to 0.1100, which is at the limit of rejection at 10%. The BIC = -7,207.65 is worse than that of 1:T and 1:EGB2. VaR of 2.02945% is very close to the non-parametric distribution, but $\text{CVaR} = 3.04198\,\% > \left[\text{CVaR}-_{\text{np}},\ \text{CVaR}+_{\text{np}}\right]$. Note also that the asymmetry coefficient of 2:T = -0.15 is negative while the non-parametric = 0.36 > 0. This suggests that the asymmetry in the data should be better integrated into the modeling. This 2:T mixture appears to be an improvement, but remains insufficient because it does not seem to allow the asymmetry to be modeled directly.

Table 9
Model M5 = 2:T

M5 parameters		Mixture of 2 Student's *t* distributions (2:T)			p = 5%		Order[1]: 61
Distribution 1	μ_1	-0.0012920					
	σ_1	0.0066854			Quantiles, coefficients and probabilities		
	ν_1	23642.31	Density	q_0	μ	$F^0_{T,\nu}\left[q_0\right]$	$-\sigma$	$\text{Tail}^0_{T,\nu}\left[q_0\right]$
Distribution 2	μ_2	-0.0004740	1	-3.228904	0.001292	0.000622	-0.006685	-0.002176
	σ_2	0.0140598	2	-1.409730	-0.000474	0.102601	-0.014060	0.218922
	ν_2	6.4162601						
	c_1	0.5158049					q_m	$F_{\text{Model}}\left[q_m\right]$
							-2.02945%	0.050000
VaR[2]		2.02945%	VaR$_{np}$: 2.04736%					
CVaR		3.04198%	CVaR -/+$_{np}$: [2.96269%, 2.97795%]					

	Model	Data				
Skewness	-0.1544	0.3589	AIC	-7,243.2842	KS stat.	0.0500
Kurtosis	8.3993	9.8158	BIC	-7,207.6537	KS *p*-value	0.1100

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.
[2] Obtained numerically from the Excel solver by minimizing $\left(F_m\left(q_m\right)-p\right)^2$.

Before exploring the addition of a parameter capturing asymmetry, we want to examine what happens for a mixture of three normal densities. Model M6 is constructed with a 3:NO mixture. In Table 10, the *p*-value of the KS test is 0.2280 > 10%. Moreover, M6 is the first model whose kurtosis of 9.4 is almost identical to the non-parametric distribution. This time, the asymmetry coefficient is positive, as is the non-parametric distribution.

VaR (3:NO) = 2.03847% is almost identical to the non-parametric distribution. As for the CVaR (3:NO) = 3.00452%, it is the closest to the interval $\left[\text{CVaR}-_{np},\ \text{CVaR}+_{np}\right]$ thus far. This model appears to be better suited to the data. We will come back to this finding when we perform the backtests of the models.

Table 10
Model M6 = 3:NO

M6 parameters		Mixture of 3 normal distributions (3:NO)				p = 5%		Order[1]: 61
Distribution 1	μ_1	-0.0004753						
	σ_1	0.0150441			Quantiles, coefficients and probabilities			
Distribution 2	μ_2	0.0043390	Density	q_0	μ	$\Phi^0[q_0]$	$-\sigma$	$\phi^0[q_0]$
	σ_2	0.0376531	1	-1.323399	-0.000475	0.092851	-0.015044	0.166189
Distribution 3	μ_3	0.0011752	2	-0.656617	0.004339	0.255714	-0.037653	0.321579
	σ_3	0.0065771	3	-3.278018	0.001175	0.000523	-0.006577	0.001852
	c_1	0.4433715						
	c_2	0.0334707					q_m	$F_{Model}[q_m]$
							-2.03847%	0.050000
VaR[2]		2.03847%	VaR_{np}: 2.04736%					
CVaR		3.00452%	$CVaR_{-/+np}$: [2.96269%, 2.97795%]					

	Model	Data				
Skewness	0.1224	0.3589	AIC	-7,244.1307	K-S stat.	0.0425
Kurtosis	9.43212	9.8158	BIC	-7,203.4101	K-S *p*-value	0.2280

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.
[2] Obtained numerically from the Excel solver by minimizing $\left(F_m(q_m)-p\right)^2$.

To advance in the modeling, we now explore the effect of adding an asymmetry parameter as an enhancement to the previous 3:NO model. The SN2 density (Skewed Normal type 2, Fernandez et al., 1995, appendix A7) allows this. We inject an asymmetry parameter in two normal densities and keep the third one as is. The mixture of this model M7 becomes 2:SN2 + 1:NO.

The 2:SN2+1:NO model includes 10 parameters. The effect of capturing asymmetry is clear in all the results presented in Table 11. The asymmetry coefficient is closest to the non-parametric one, and the kurtosis of M7 is even slightly higher than that of the non-parametric distribution. The *p*-value of the KS test is 0.1980 > 10%. The VaR = 2.05018% $\approx$ the VaRnp. The CVaR = 2.98338% is almost stuck to the upper bound of the CVaR interval. We probably have a serious candidate to represent the risks of the data, even if the CVaR is not quite framed by the

$\left[\text{CVaR}-_{np},\ \text{CVaR}+_{np}\right]$. Finally, the parameter $\nu_1 = 0.8831 < 1$ confirms the capture of some degree of asymmetry for the first SN2 density. The second density, which has $\nu_2 = 0.9939 \approx 1$ degenerates to a simple normal distribution. A mixture of 1:SN2 + 2:NO would probably have been sufficient, while saving a parameter for the estimation.

Table 11
Model M7 = 2:SN2 + 1:NO

M7 Parameters		Mixture of 2 SN2 + 1 normal (2:SN2+1:NO)					p = 5%	Order[1]: 61
Distribution 1	μ_1	0.00259297						
	σ_1	0.01467882						
	ν_1	0.88306710			Quantiles, coefficients and probabilities			
Distribution 2	μ_2	0.00092266	Density	q_0	μ	$F_i^0[q_0]$	$-\sigma$	$\text{Tail}_{f_i}^0[q_0]$
	σ_2	0.00638969	1	-1.573342	0.002593	0.092550	-0.014679	0.193377
	ν_2	0.99395520	2	-3.352978	0.000923	0.000433	-0.006390	0.001564
Distribution 3	μ_3	0.00918332	3	-0.763277	0.009183	0.222649	-0.038892	0.298127
	σ_3	0.03889170						
	c_1	0.47293330						
	c_2	0.50005730					q_m	$F_{Model}[q_m]$
							-2.05018%	0.050000
VaR[2]		2.05018%		VaR_{np}: 2.04736%				
CVaR		2.98338%		CVaR-/+_{np}: [2.96269%, 2.97795%]				

	Model	Data				
Skewness	0.2433	0.3589	AIC	-7,241.4466	KS stat.	0.0442
Kurtosis	9.8409	9.8158	BIC	-7,190.5458	KS *p*-value	0.1980

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.
[2] Obtained numerically from the Excel solver by minimizing $\left(F_m(q_m)-p\right)^2$.

In addition to the direct parameterization of the asymmetry, we also want to capture the tail thickness. Fernandez et al. (1995) propose the SEP3 density (Skewed Exponential Power type 3; see also Rigby et al., 2014). We wish to reduce the number of parameters at the same time. The M8 model is constructed with the 2:SEP3 mixture (mixture of two SEP3 distributions, see definitions and calculations in appendices A3, A8 and A9).

First, the *p*-value of the KS test in Table 12 is the largest. The asymmetry coefficient is very small and positive. The kurtosis is large, but smaller than in the previous model. Given the values of AIC and BIC, the model fits the data better than the previous model. The VaR(2:SEP3) = 1.99295% is a little far from the non-parametric distribution, but most importantly, the VaR = 2.97397% falls within the range $\left[\text{CVaR}-_{\text{np}},\ \text{CVaR}+_{\text{np}}\right]$ for this 2:SEP3 model despite the narrow interval.

Note the asymmetry parameters $\nu_1 = 1.0315 \approx 1$ and $\nu_2 = 0.6137 < 1.$ We find ourselves in the same configuration as the previous model, with a density that captures the asymmetry. Tail thickness parameters are equal to $\tau_1 = 0.9599 \approx 1$ and $\tau_2 = 2.1084 \approx 2.$ We therefore have a first SEP3 distribution that is practically a Laplace distribution ($\nu = 1$, $\tau = 1$). The second SEP3 degenerates into an asymmetric $(\nu \neq 1)$ normal ($\tau = 2$), which is finally an SN2. In this case, a Laplace mixture added to an SN2 would probably have suited the data. This result directly corroborates similar findings in the recent market risk literature highlighting the mixture qualities of a Laplace and a Gaussian distribution (see Haas et al., 2006; Haas, 2009; Broda and Paolella, 2011; Miao et al., 2016; Taylor, 2019).

Table 12
Model M8 = 2:SEP3

M8 Parameters			Mixture of 2 SEP3 distributions (2:SEP3)				p = 5%	Order[1]: 61
Distribution 1	μ_1	-0.0007520						
	σ_1	0.0004529						
	ν_1	1.0315089						
	τ_1	0.9598700			Quantiles, coefficients and probabilities			
Distribution 2	μ_2	0.0075456	Density	q_0	μ	$F_i^0[q_0]$	$-\sigma$	$\mathrm{Tail}_{f_i}^0[q_0]$
	σ_2	0.0065018	1	-4.234276	-0.000752	0.066185	-0.004529	0.425876
	ν_2	0.6137048	2	-4.225762	0.007546	0.004189	-0.006502	0.019551
	τ_2	2.1083901						
	c_1	0.7389303					q_m	$F_{Model}[q_m]$
							-1.99295%	0.050000
VaR[2]		1.99295%	VaR_{np}: 2.04736%					
CVaR		2.97397%	$CVaR{-}/{+}_{np}$: [2.96269%, 2.97795%]					

	Model	Data				
Skewness	0.0051	0.3589	AIC	-7,245.9268	KS stat.	0.0392
Kurtosis	7.1752	9.8158	BIC	-7,200.1161	KS *p*-value	0.3040

[1] Non-parametric VaR is equal to the 61st smallest value in the sample.
[2] Obtained numerically from the Excel solver by minimizing $\left(F_m\left(q_m\right)-p\right)^2$.

Before moving on to the VaR and CVaR backtesting step, note that all eight models maintain the same behaviors at probabilities p = 2.5% and 1%. However, at p = 1%, CVaR of models M7 and M8 are close to the upper bound of $\left[\text{CVaR}-_{np},\ \text{CVaR}+_{np}\right]$ rather than being in that interval (see Table A4 in Appendix A11). This percentile is actually too far down the tail of losses for CVaR to be accurate. This should not pose a problem under current regulatory requirements, given that Basel requires backtest on VaR rather than CVaR at this 1% percentile and CVaR at 2.5 % is the measure of market risk.

To conclude this section, Figure 3 graphically summarizes the VaR and CVaR behavior of models 2:SEP3 and 2:SN2 + 1:NO, as well as 1:NO and 1:T in the left tail of portfolio returns.

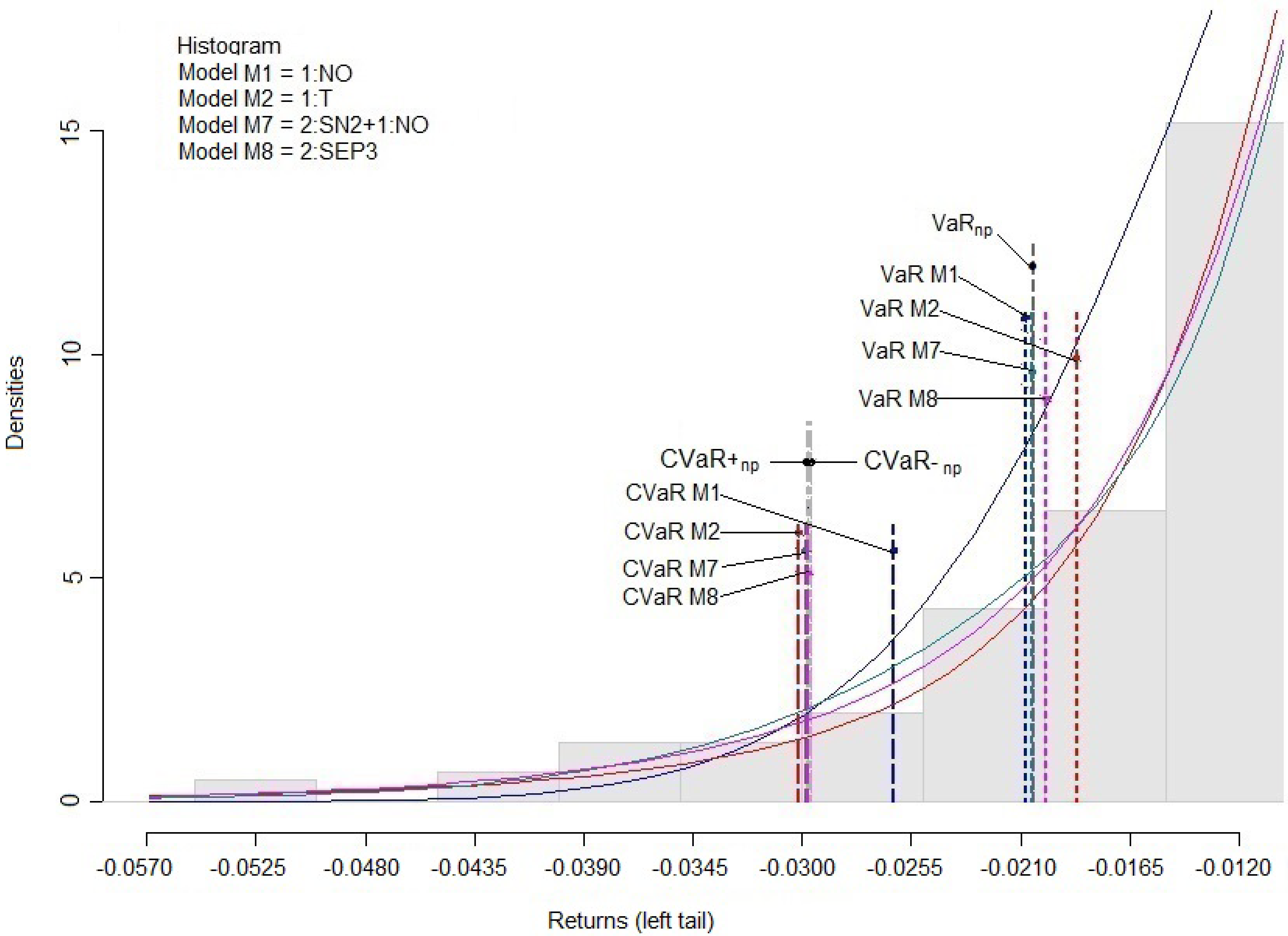


Figure 3: VaR and CVaR plots of selected models in the left tail of returns

4. Backtesting of VaR and CVaR in compliance with the Basel regulations in force

4.1 Validation methodology for VaR and CVaR models

The VaRs of the different models will be validated by three backtests. The uc backtest validates the frequency of hits unconditionally (Kupiec 1995). Second, Christoffersen's (1998) cc backtest is conditional on inter-hit independence. The last test is the DQ backtest of Engle and Manganelli (2004). DQ is used in parallel with cc to detect both consecutive exceptions and those spaced with a lag of up to about a week with daily data. Christofferson's cc test detects successive exceedances with a lag of only one day. If there are clusters of more or less closely spaced exceedances with lags greater than one day, cc does not detect them but DQ does.

As for model CVaR, we apply the Z_{ES} backtest of Acerbi and Szekely (2017), and the backtest of Righi and Ceretta (2015), which will now be called RC. We also show the results of the backtests Z_1 and Z_2 for information purposes only.[9]

Here we deploy five backtests to validate the VaR and CVaR of competing models in order to: (i) satisfy the regulatory requirement to perform the 1% VaR backtest (BCBS, 2016, page 77; BCBS, 2019, paragraph 32.5); (ii) as a complement, validate the 2.5% CVaR and the 2.5% VaR; (iii) as another complement, validate the 5% CVaR and the 5% VaR.

Currently, none of the four backtests in points (ii) and (iii) is explicitly required to validate the banks' overall market risk coverage. However, we propose them as part of the Basel recommendation to foresee additional statistical tests with varying degrees of confidence to support model accuracy (BCBS, 2016, page 82; BCBS, 2019, paragraph 32.13).

It is thus natural to consider adding validation of the risk measures of (ii) at 2.5% given that 2.5% CVaR determines the coverage. The 5% backtests of (iii) would be of less importance, but should help confirm the robustness of the models. Note that the five backtests are carried out as out-of-sample tests, the approach of which is set out in Appendix A10.

[9]Although they are currently quite popular in the literature, these backtests have some problems as reported in the literature that prevent us from drawing conclusions based on their results.

4.2 Backtest results of the VaR and CVaR models

Backtest results for the eight models are presented in Table 13. Unsurprisingly, the normal model 1:NO is rejected because of its 1% VaR, 2.5% VaR and 2.5% CVaR. However, we did not expect that the Student 1:T model would be rejected for similar reasons. The *p*-values are larger, but remain <10%, which is the critical rejection threshold.

The backtests of 1% VaR of EGB2 have higher *p*-values, but still below 10%. One might be tempted not to reject 1% VaR, especially because 2.5% VaR behaves rather well, with *p*-values of the uc, cc and DQ backtests all >10% (0.1552, 0.2875 and 0.1131 respectively). In contrast, 2.5% CVaR is rejected by Z_{ES} *p*-value = 0.00 < 10%) and by RC (*p*-value = 0.0124 < 10%). This model is a concrete example where 2.5% CVaR does not pass the backtest, while VaR performs relatively well for the 1% regulatory backtest. It also does so at 2.5%, which is an additional validation.

The next case illustrates the opposite situation. With the 2:NO model, 2.5% VaR is rejected by the uc and DQ backtests (0.0783 < 10% and 0.0898 < 10%). In contrast, 2.5% CVaR is well validated by Z_{ES} and RC (*p*-value = 0.3516 > 10% and *p*-value = 0.3396 > 10%). Model 2:T replicates almost the same behavior. This model is certainly an improvement over the 2:NO model, but remains insufficient for the data (as is the case with 1:T versus 1:NO).

The 3:NO mixture, despite its eight parameters, is inferior to the previous models, including model 2:NO, which has only five parameters. Yet 3:NO seemed to perform well in the discussion of Table 10 in the previous section (see Section 3). This confirms the merits of injecting additional parameters to capture the asymmetry in the data.

More specifically, model 2:SN2+1:NO, which includes two parameters for asymmetry, seems to fit better for 1% VaR. The *p*-values of uc, cc and DQ are 0.2695, 0.4378 and 0.3496 respectively. VaR at 2.5% also seems to perform well for uc (*p*-value = 0.2819) and cc (*p*-value = 0.4004), except for the independence of hits, DQ test *p*-value is 0.0685. CVaR at 2.5% is not rejected according to Z_{ES} (*p*-value = 0.1516 > 10%), but is rejected for RC (*p*-value = 0.0516 < 10%).

A word about criterion (iii). The uc backtest validates VaR at 5% (*p*-value = 0.4338) and the backtest Z_{ES} validates 5% CVaR (*p*-value = 0.4020), but RC rejects it (*p*-value = 0.0596). To summarize, the 2:SN2+1:NO mixture shows a clear improvement over previous mixtures, with the injection of the two asymmetry parameters. However, the improvement is not yet sufficient to model the risks of the data effectively.

Now we come to the backtests for 2:SEP3. Clearly, 1% VaR is validated given the respective *p*-values of uc, cc and DQ (0.2695 > 10%, 0.4378 > 10% and 0.0994 ≈ 10%). VaR at 2.5% is also validated according to uc (*p*-value = 0.2819), cc (*p*-value = 0.4004) and DQ (*p*-value = 0.1200). The 2.5% CVaR is comfortably validated by both Z_{ES} (*p*-value = 0.5000 > 10%) and RC (*p*-value = 0.3168 > 10%). As for criterion (iii), the uc backtest can be considered to validate VaR at 5% (*p*-value = 0.0954 ≈ 10%). The backtests Z_{ES} and RC validate 5% CVaR with comfortable *p*-values (*p*-value = 0.8412 and 0.3776 respectively). In conclusion, the 2:SEP3 mixture, which captures both asymmetry and tail thickness, appears to have superior abilities to model the risks incorporated in our data. These results directly confirm the conclusions of recent work on the superiority of a mixture of a normal distribution and a Laplace distribution, as seen previously.

Conclusion

This paper presented a framework for validating market risk models. The approach jointly deploys CVaR and VaR backtests, in compliance with international regulations in force (coverage with 2.5% CVaR and required backtest on 1% VaR). Further, given the use of actual data that cover a period of extreme market turbulence, the assumption of normality of returns is definitively outdated. Identifying a parametric model entails comparing the magnitudes resulting from the calculations using the model parameters with the equivalent magnitudes estimated in a non-parametric distribution. The keystone of this article is the specification of the framework of CVaR of the model to be evaluated by the interval [$CVaR-_{np}$, $CVaR+_{np}$], which appears to be an important criterion for evaluating models and is very closely linked to the conclusions of the backtests of the models. As seen in the different estimates, nonparametric kurtosis and asymmetry also help guide the research approach to determine the direction in which to move forward.

Further, this research is an exercise in the actual implementation of VaR and CVaR backtesting when choosing a parametric model that can manage the market risk embedded in the data. The identification of the 2:SEP3 mixture, which seems to work well with our data, is not a coincidence. In fact, the mixing of a normal distribution with a Laplace distribution directly corroborates the conclusions of the recent literature, which positions this mixture as the natural replacement for normal distribution for market risk (see Haas et al., 2006; Haas, 2009; Broda and Paolella, 2011; Miao et al., 2016; Taylor, 2019).

Table 13
Out-of-sample backtests of VaR and CVaR

p	#Hits	Model	— uc —		cc	DQ	— Z_{ES} —		— RC —		— Z_1 —		— Z_2 —	
			Stat	p-value	p-value	p-value	Stat	p-value	Stat	p-value	Stat	p-value	Stat	p-value
0.050	71	1:NO	2.010	0.1563	0.0326	0.0012	-0.257	0.0000	-0.915	0.0000	-0.184	0.0000	-0.401	0.0024
0.025	47	1:NO	8.450	0.0037	0.0105	0.0035	-0.379	0.0000	-0.611	0.0000	-0.182	0.0000	-0.852	0.0000
0.010	26	1:NO	12.372	0.0004	0.0012	0.0000	-0.649	0.0000	-0.535	0.0000	-0.229	0.0000	-1.663	0.0000
0.050	78	1:T	5.215	0.0224	0.0033	0.0001	-0.152	0.0388	-0.193	0.3012	-0.051	0.2876	-0.366	0.0160
0.025	45	1:T	6.685	0.0097	0.0343	0.0163	-0.168	0.0724	0.010	0.9056	-0.028	0.5724	-0.542	0.0148
0.010	20	1:T	4.487	0.0342	0.0756	0.0303	-0.229	0.0860	-0.058	0.4304	-0.062	0.4144	-0.771	0.0236
0.050	70	1:EGB2	1.669	0.1964	0.0037	0.0000	-0.144	0.0008	-0.464	0.0056	-0.087	0.0220	-0.268	0.0028
0.025	38	1:EGB2	2.020	0.1552	0.2875	0.1131	-0.185	0.0000	-0.294	0.0124	-0.102	0.0340	-0.396	0.0040
0.010	19	1:EGB2	3.504	0.0612	0.1277	0.0791	-0.292	0.0008	-0.206	0.0072	-0.122	0.0724	-0.777	0.0024
0.050	74	2:NO	3.211	0.0731	0.0004	0.0000	-0.143	0.4956	-0.213	0.4496	-0.058	0.1892	-0.305	0.5700
0.025	40	2:NO	3.100	0.0783	0.1817	0.0898	-0.193	0.3516	-0.139	0.3396	-0.076	0.0492	-0.435	0.5140
0.010	17	2:NO	1.864	0.1722	0.3084	0.0880	-0.306	0.1024	-0.293	0.0064	-0.170	0.0173	-0.657	0.4044
0.050	74	2:T	3.211	0.0731	0.0004	0.0000	-0.147	0.2604	-0.212	0.2796	-0.063	0.2544	-0.311	0.2528
0.025	40	2:T	3.100	0.0783	0.1817	0.0670	-0.204	0.2508	-0.146	0.2540	-0.081	0.1184	-0.441	0.2500
0.010	18	2:T	2.627	0.1051	0.2044	0.0532	-0.312	0.2108	-0.174	0.0836	-0.140	0.0984	-0.709	0.2436
0.050	73	3:NO	2.782	0.0954	0.0013	0.0000	-0.135	0.1836	-0.384	0.1432	-0.060	0.1356	-0.254	0.2996
0.025	42	3:NO	4.387	0.0362	0.1017	0.0267	-0.175	0.1204	-0.220	0.1428	-0.089	0.0336	-0.343	0.2676
0.010	19	3:NO	3.504	0.0612	0.1023	0.0126	-0.283	0.0544	-0.306	0.0052	-0.152	0.0096	-0.633	0.1688
0.050	66	2:SN2 + 1:NO	0.613	0.4338	0.0086	0.0000	-0.105	0.4020	-0.410	0.0596	-0.068	0.1316	-0.175	0.6520
0.025	36	2:SN2 + 1:NO	1.158	0.2819	0.4004	0.0685	-0.137	0.1516	-0.277	0.0516	-0.079	0.1140	-0.295	0.4116
0.010	16	2:SN2 + 1:NO	1.219	0.2695	0.4378	0.3496	-0.219	0.0296	-0.236	0.0144	-0.129	0.0441	-0.505	0.2004
0.050	73	2:SEP3	2.782	0.0954	0.0301	0.0006	-0.111	0.8412	-0.153	0.3776	-0.037	0.3308	-0.262	0.8640
0.025	36	2:SEP3	1.158	0.2819	0.4004	0.1200	-0.127	0.5000	-0.131	0.3168	-0.064	0.1699	-0.277	0.9912
0.010	16	2:SEP3	1.219	0.2695	0.4378	0.0994	-0.198	0.1516	-0.206	0.0388	-0.101	0.1097	-0.468	0.6016

References

Acerbi, C. and Szekely, B. (2014). Back-testing expected shortfall. *Risk*, 27(11):76-81.

Acerbi, C. and Szekely, B. (2017). General properties of backtestable statistics. Available at SSRN: https://ssrn.com/abstract=2905109 or http://dx.doi.org/10.2139/ssrn.2905109.

Basel Committee on Banking Supervision (BCBS) (2016). Minimum capital requirements for market risk, publication no 352. *Bank For International Settlements (BIS)*, (Jan-2016):1-92.

Basel Committee on Banking Supervision (BCBS) (2019). Minimum capital requirements for market risk, publication no 457. *Bank For International Settlements (BIS)*, (Jan-2019):1-136.

Broda, S.A. and Paolella, M.S. (2011). Expected shortfall for distributions in finance. *In Statistical Tools for Finance and Insurance*, pages 57-99. Springer.

Caivano, M. and Harvey, A. (2014). Time-series models with an EGB2 conditional distribution. *Journal of Time Series Analysis*, 35(6): 558-571.

Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4): 841-862.

Cummins, J.D., Dionne, G., McDonald, J.B. and Pritchett, B.M. (1990). Applications of the GB2 family of distributions in modeling insurance loss processes. *Insurance: Mathematics and Economics*, 9(4):257-272.

Dionne, G. (2019). Corporate risk management: Theories and applications. John Wiley, 384 pages.

Dionne, G. and Saissi Hassani, S. (2017). Hidden Markov regimes in operational loss data: Application to the recent financial crisis. *Journal of Operational Risk*, 12(1): 23-51.

Efron, B. and Tibshirani, R.J. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC press, 456 p.

Engle, R.F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, 22(4):367-381.

Fernandez, C., Osiewalski, J. and Steel, M.F. (1995). Modeling and inference with ↑-spherical distributions. *Journal of the American Statistical Association*, 90(432):1331-1340.

Haas, M. (2009). Modelling skewness and kurtosis with the skewed Gauss-Laplace sum distribution. *Applied Economics Letters*, 16(12): 1277-1283.

Haas, M., Mittnik, S. and Paolella, M.S. (2006). Modelling and predicting market risk with laplace-gaussian mixture distributions. *Applied Financial Economics*, 16(15): 1145-1162.

Kerman, S.C. and McDonald, J.B. (2015). Skewness-kurtosis bounds for egb1, egb2, and special cases. *Communications in Statistics-Theory and Methods*, 44(18):3857-3864.

Kupiec, P.H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2):73-84.

McDonald, J.B. (2008). Some generalized functions for the size distribution of income. Dans: *Modeling Income Distributions and Lorenz Curves*, Springer, p. 37-55.

McDonald, J.B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3):647-663.

McDonald, J.B. and Michelfelder, R.A. (2016). Partially adaptive and robust estimation of asset models: accommodating skewness and kurtosis in returns. *Journal of Mathematical Finance*, 7(1): 219.

McDonald, J.B. and Xu, Y.J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics*, 66(1-2): 133-152.

Miao, D.W.C., Lee, H.C. and Chen, H. (2016). A standardized normal-Laplace mixture distribution fitted to symmetric implied volatility smiles. *Communications in Statistics-Simulation and Computation* 45(4): 1249-1267.

Rigby, B., Stasinopoulos, M., Heller, G. and Voudouris, V. (2014). The distribution toolbox of GAMLSS. *gamlss.org*.

Righi, M. and Ceretta, P.S. (2015). A comparison of expected shortfall estimation models. *Journal of Economics and Business*, 78: 14-47.

Rockafellar, R.T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance, 26*(7): 1443-1471.

Taylor, J.W. (2019). Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric laplace distribution. *Journal of Business and Economic Statistics* 37 (1): 121-133.

Theodossiou, P. (2018). Risk measures for investment values and returns based on skewed-heavy tailed distributions: Analytical derivations and comparison. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3194196.

Appendices

A1. Estimated models

The appendices present the mathematical developments of the equations retained and the tables of results of parameter estimation for different statistical distributions of returns. Given that these developments are algebraic, the signs of the final expressions of VaR and CVaR should be reversed to obtain positive measures. Table A1 presents the symbols of the estimated models.

Table A.1
Model Symbol Definitions

Model	Symbol	Description of the model
M1	1:NO	Normal distribution
M2	1:T	Student's *t* distribution
M3	1:EGB2	Exponential GB2 distribution
M4	2:NO	Mixture of 2 normal distributions
M5	2:T	Mixture of 2 Student's *t* distributions
M6	3:NO	Mixture of 3 normal distributions
M7	2:SN2+1:NO	Mixture of 2 SN2 + 1 normal distributions
M8	2:SEP3	Mixture of 2 SEP3 distributions

Let's start by deriving the general formulas of CVaR for a statistical distribution (Appendix A2) and for a mixture of distributions (Appendix A3).

A2. Expression of CVaR

Expression of the density and cumulative function of a reduced distribution

We are interested in the family of location-scale parametric distributions $\overline{F}$ having a location parameter μ and a scale parameter σ. If $F \in \overline{F}$ and $y \sim F$, then the reduced variable $z = (y-\mu)/\sigma$ follows the distribution F^0 defined with equality:

$$F(y) = F^0(z)\,.$$

F^0 is said to be a reduced cumulative function of F. The reduced density $f^0(\cdot)$ is related to the density $f(\cdot)$ by writing:

$$f(y) = \frac{f^0(z)}{\sigma} \tag{A0}$$

All densities in this document belong to $\overline{F}$, including the normal distribution and Student's *t*. The location and scale parameters coincide with the mean and standard deviation of the normal distribution. This is not always the case for the other distributions $\in \overline{F}$.

General expression of CVaR

We note $q < 0$ the quantile of VaR corresponding to the degree of confidence (1 – p). As in the study by Broda and Paolella (2011), the tail quantity of a density $f(\cdot)$ at point *x* is defined by: $\text{tail}_f(x) \overset{\text{def}}{=} \int_{-\infty}^{x} t\, f(t)\,dt.$ We develop the expression of CVaR using its definition:

$$\text{CVaR}_f = E\left[y \middle| y \le q\right] = \frac{1}{F(q)} \int_{-\infty}^{q} y\, f(y)\,dy \tag{A1}$$

$$= \frac{1}{p}\left\{\int_{-\infty}^{\frac{q-\mu}{\sigma}} (\mu + \sigma z)\frac{f^0(z)}{\sigma}\sigma d(z)\right\} \tag{A2}$$

$$= \frac{1}{p}\left\{\int_{-\infty}^{\frac{q-\mu}{\sigma}} \mu\, f^0(z)dz + \sigma\int_{-\infty}^{\frac{q-\mu}{\sigma}} z\, f^0(z)d(z)\right\} \tag{A3}$$

$$= \frac{1}{p}\left\{\mu\, F^0\left(\frac{q-\mu}{\sigma}\right) + \sigma \text{Tail}_{f^0}\left(\frac{q-\mu}{\sigma}\right)\right\} \tag{A4}$$

Equation (A2) is obtained by using (A0) after a change of variable $z = (y-\mu)/\sigma$, or $y = \mu + \sigma z$, where $dy = \sigma dz$. Equations (A3) and (A4) come from algebraic calculations on the previous line (A2). Note that there are two parts in formula (A4): the first one is μ/p times the centered reduced cumulative $F^0(\cdot)$ evaluated at the centered reduced quantity $(q-\mu)/\sigma$. The second one is σ/p times the tail of f^0, also evaluated at $(q-\mu)/\sigma$.

A last remark is that we have of course $F^0\left((q-\mu)/\sigma\right)=p$, which would simplify the expression (A4). Even so, we will leave the expression as it is so that it will be of the same form as for mixtures of distributions where there will indeed be several cumulatives $F_i^0(\cdot)$, for which $F_i^0\left((q-\mu)/\sigma\right)\neq p$.

A3. CVaR of a mixture of distributions

Let $m(\cdot)$ be a mixture of *n* densities $f_i(\cdot), i=1,\ldots,n.$ Each density $f_i\in\overline{F}$ has a parameter of location μ_i and scale σ_i. The mixture density $f_m(\cdot)$ and its distribution $F_m(\cdot)$ are written as:

$$f_m(y)=\sum_1^n c_i\ f_i(y),\qquad F_m(y)=\sum_1^n c_i\ F_i(y)$$

where c_i is a probability, to be estimated, regarding the weight of density $f_i(\cdot)$. The sum of the c_i is equal to 1.

Let q_m be the quantile corresponding to VaR of the mixture at the confidence level $(1-p)$. We denote $f_i^0(\cdot)$ and $F_i^0(\cdot)$ as the reduced density and the reduced cumulative of the *i*th density. The expression of the CVaR_m is developed by taking the sum (Σ) out of the integral:

$$\begin{aligned}\text{CVaR}_m &= E\left[y\,|\,y\le q_m\right]=\frac{1}{p}\int_{-\infty}^{q_m} y\,f_m(y)\,dy\\ &=\frac{1}{p}\sum_1^n c_i\int_{-\infty}^{q_m} y\,f_i(y)\,d(y)\\ &=\frac{1}{p}\sum_1^n c_i\int_{-\infty}^{\frac{q_m-\mu_i}{\sigma_i}}(\mu_i+\sigma_i z_i)\frac{f_i^0(z_i)}{\sigma_i}\sigma_i d(z_i)\\ &=\frac{1}{p}\sum_1^n c_i\left\{\mu_i F_i^0\left(\frac{q_m-\mu_i}{\sigma_i}\right)+\sigma_i \text{Tail}_{f_i^0}\left(\frac{q_m-\mu_i}{\sigma_i}\right)\right\}\end{aligned}$$

or, in vector form, convenient for calculations:

$$\mathrm{CVaR}_m = \frac{1}{p}\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}^T \times \left[\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \times \begin{pmatrix} F_1^0\left(\dfrac{q_m-\mu_1}{\sigma_1}\right) \\ \vdots \\ F_n^0\left(\dfrac{q_m-\mu_n}{\sigma_n}\right) \end{pmatrix} + \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{pmatrix} \times \begin{pmatrix} \mathrm{Tail}_{f_1}^0\left(\dfrac{q_m-\mu_1}{\sigma_1}\right) \\ \vdots \\ \mathrm{Tail}_{f_n}^0\left(\dfrac{q_m-\mu_n}{\sigma_n}\right) \end{pmatrix} \right].$$

In the general case, q_m is found numerically as a solution to the equation $F_m(q_m) - p = 0$. The Excel file allows this. Note that for a distribution *i*, $F_i^0\left((q_m-\mu_i)/\sigma_i\right) \neq p$. The only case where there is equality is when the distribution F_i^0 is unique (no mixture).

A4. Expression of CVaR of a normal distribution

The density $\phi_{\mu,\sigma}(\cdot)$ of a normal distribution $N(\mu,\sigma)$ is:

$$\phi_{\mu,\sigma}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right).$$

We denote both $\phi_0(\cdot)$ and $\Phi_0(\cdot)$ as the density and the cumulative of the standard normal distribution $N(0,1)$. It is easy to show that:

$$\frac{\partial}{\partial x}\phi_0(x) = -x\phi_0(x). \qquad \text{(A5a)}$$

For $y \sim N(\mu,\sigma)$, the quantile q of VaR at the confidence level $(1-p)$ is found by $P(y = \mu + \sigma z \leq q) = p$, hence $\mathrm{VaR}_{p,\mu,\sigma}$ is:

$$q = \mu + \sigma\Phi_0^{-1}(p).$$

Further, with the definition of tail and with the help of equation (A5a) , we find:

$$\mathrm{Tail}_{\phi_0}(x) = \int_{-\infty}^{x} z\,\phi_0(z)\,dz = -\phi_0(x). \qquad \text{(A5)}$$

We apply (A4) and (A5) to obtain:

$$\mathrm{CVaR}_{\phi,\mu,\sigma} = \frac{1}{p}\left\{\mu\Phi_0\left(\frac{q-\mu}{\sigma}\right) - \sigma\phi_0\left(\frac{q-\mu}{\sigma}\right)\right\}. \tag{A6}$$

A5. Expression of the CVaR of the Student's *t* distribution

The density $f_{T,\mu,\sigma,\nu}(\cdot)$ of the Student's *t* of parameters μ (location), σ (scale) and ν (degrees of freedom) is written as:

$$f_{T,\mu,\sigma,\nu}(y) = \frac{A}{\sigma}\left(1 + \left(\frac{y-\mu}{\sigma}\right)^2 \frac{1}{\nu}\right)^{-b}$$

where $A = \left[\sqrt{\nu} \times B(1/2, \nu/2)\right]^{-1}$ and $b = (\nu+1)/2$. $B(\cdot)$ is the beta function.[10] The reduced functions are noted $f^0_{T,\nu}(\cdot)$ and $F^0_{T,\nu}(\cdot)$. We determine *q* from the VaR of $y \sim t(\mu,\sigma,\nu)$ to the degree of confidence $(1-p)$:

$$P(y \le q) = P(\mu + \sigma z \le q) \Rightarrow F^0_{T,\nu}\left(\frac{q-\mu}{\sigma}\right) = p$$

$$q = \mu + \sigma F^{0-1}_{T,\nu}(p)$$

where $F^{0-1}_{T,\nu}(\cdot)$ is the quantile (or inverse) function of $F^0_{T,\nu}(\cdot)$. The tail at point *x* is by definition:

$$\mathrm{Tail}_{f^0_{T,\nu}}(x) = \int_{-\infty}^{x} z\, f^0_{T,\nu}(z)\,dz = A\int_{-\infty}^{x} z\left(1 + z^2/\nu\right)^{-b} dz. \tag{A7}$$

We change the variable $u = z^2/\nu$, hence $zdz = \nu du/2$. The integral of equation (A7) becomes:

[10] There is another way to write the constant A with the gamma function $\Gamma(\cdot)$ instead of the beta function.

$$\mathrm{Tail}_{f^0_{T,\nu}}(x) = A\frac{\nu}{2}\int_{-\infty}^{\frac{x^2}{\nu}}(1+u)^{-b}\,du$$

$$= \frac{A}{2}\frac{\nu}{-b+1}\Big[(1+u)^{-b+1}\Big]_{-\infty}^{\frac{x^2}{\nu}} = \frac{\nu}{2(1-b)}\left(1+\frac{x^2}{\nu}\right)\times A\left(1+\frac{x^2}{\nu}\right)^{-b} \qquad \text{(A8)}$$

$$= -\frac{\nu+x^2}{\nu-1}\times f^0_{T,\nu}(x). \qquad \text{(A9)}$$

In equation (A8), we replace b with its value $(\nu+1)/2$. The final expression of the tail is simplified in (A9). In order to be valid we need to have $\nu > 1$. We now apply (A9) in (A4) to find:

$$\mathrm{CVaR}_{f_{T,\mu,\sigma,\nu}} = \frac{1}{p}\left\{\mu F^0_{T,\nu}\left(\frac{q-\mu}{\sigma}\right) - \sigma\frac{\nu+\left(\frac{q-\mu}{\sigma}\right)^2}{\nu-1} f^0_{T,\nu}\left(\frac{q-\mu}{\sigma}\right)\right\}.$$

Important: In Excel, the functions related to Student's *t* distribution consider the degree of freedom ν to be an integer. Therefore, calculations cannot be made in standard form, and an additional module is required. The XRealStats.xlam module is used. It must be downloaded from their website[11], placed in the C:/TP5 directory and activated to use the functions that allow calculations with ν∈R. The cumulative and density functions are called by T_DIST. The inverse of the cumulative function is T_INV.

A6. The EGB2 distribution: Exponential GB2

The EGB2 (Exponential Generalized Beta type 2) density has four parameters and is written, according to Kerman and McDonald (2015), for $y \in R$:

$$f(y|\mu,\sigma,\nu,\tau) = \frac{e^{\nu z}}{|\sigma|\times B(\nu,\tau)\left(1+e^z\right)^{\nu+\tau}}$$

where $z = (y-\mu)/\sigma$, $\mu,\sigma \in R$, $\nu,\tau > 0$. $B(\cdot)$ is the standard beta function. The density of the GB2 was originally proposed by McDonald (1984).

[11] http://www.real-statistics.com/free-download/

The parameters ν and τ characterize both tail thickness and the asymmetry of the distribution. The distribution has a negative or positive asymmetry, or is symmetrical when $\nu < \tau$, $\nu > \tau$ or $\nu = \tau$ respectively. As for the tail thickness, the smaller the ν, the thicker the left tail (all other parameters being equal).

The EGB2 includes many parametric distributions as special cases. Specifically, when $\nu \approx \tau \to +\infty$, the distribution converges to the normal. In practice, this convergence can be considered to have been reached when $\nu \approx \tau > 15$. When $\nu = \tau = 1$, EGB2 becomes a logistic distribution. Further, lemma 2 of Caivano and Harvey (2014) shows that EGB2 tends toward a Laplace density when $\nu \approx \tau \approx 0$. Other interesting special cases of EGB2 and GB2 are presented by Kerman and McDonald (2015), McDonald (2008) and McDonald and Xu (1995).

Cummins, Dionne, McDonald, and Pritchett (1990) applies the GB2 to compute reinsurance premiums and quantiles for the distribution of total insurance losses. EGB2 is increasingly used in finance, as the studies by Caivano and Harvey (2014), McDonald and Michelfelder (2016), and Theodossiou (2018) exemplify and in operational risk management (Dionne and Saissi Hassani, 2017).

A7. The *Skewed Normal* Type 2 distribution: SN2

The definition of the density of Skewed Normal Type 2 (SN2) by Fernandez et al. (1995) for $y \in R$ can be written as:

$$f_{SN2,\mu,\sigma,\nu}(y) = \frac{2\nu}{\sigma\sqrt{2\pi}\left(1+\nu^2\right)}\left\{\exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 \nu^2\right) I_{(y<\mu)} + \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 \frac{1}{\nu^2}\right) I_{(y\geq\mu)}\right\} \quad \text{(A10)}$$

where $\mu \in R$, $\sigma > 0, \nu > 0$. If $\nu < 1$, asymmetry is to the left (negative returns); if $\nu > 1$, asymmetry is positive. When $\nu = 1$, we return to a normal (symmetrical) distribution. This density is also useful to compute capital in operational risk management, as in the study by Dionne and Saissi Hassani (2017). A random variable $y \sim F_{SN2,\mu,\sigma,\nu} \Rightarrow z = (y-\mu)/\sigma \sim F^0_{SN2,\nu}$. For $z < 0$, only the left side of the equation (A10) is non-zero. The reduced density is then written as:

$$f^0_{SN2,\nu}(z) = \frac{2\nu}{1+\nu^2}\phi_0(z \times \nu).$$

The cumulative at point $z<0$ is written as:

$$F_{SN2,\nu}^{0}(z)=\int_{-\infty}^{z} f_{SN2,\nu}^{0}(t)\,dt=\frac{2}{1+\nu^2}\,\Phi_0(z\times\nu).$$

The functions $\phi_0(\cdot)$ and $\Phi_0(\cdot)$ designate the cumulative and the reduced centered normal density $N(0,1)$. The previous equation allows us to find the expression of the VaR at the confidence level $(1-p)$:

$$P(y\le q)=P\left(z\le\frac{q-\mu}{\sigma}\right)\Rightarrow F_{SN2,\nu}^{0}\left(\frac{q-\mu}{\sigma}\right)=p$$

$$\frac{2}{1+\nu^2}\Phi_0\left(\frac{q-\mu}{\sigma}\times\nu\right)=p$$

$$q=\mu+\sigma\frac{1}{\nu}\Phi_0^{-1}\left(p\,\frac{1+\nu^2}{2}\right). \tag{A11}$$

The expression (A11) is valid only if $p(1+\nu^2)/2\le 1,$ otherwise $\Phi^{-1}(\cdot)$ would not be defined. This requires that $\nu\le\sqrt{2/p-1}.$

The expression of the tail is developed as follows:

$$\mathrm{Tail}_{SN2,\nu}(x)=\int_{-\infty}^{x} z\,f^{0}(z)\,dz=\int_{-\infty}^{x}\frac{2\nu}{1+\nu^2}\,\phi_0(z\times\nu)\,dz$$

$$=\frac{2\nu}{1+\nu^2}\int_{-\infty}^{x\times\nu}\frac{u}{\nu}\phi_0(u)\frac{du}{\nu}=\frac{2}{\nu\left(1+\nu^2\right)}\Big[-\phi_0(u)\Big]_{-\infty}^{x\times\nu}$$

$$=-\frac{2}{\nu\left(1+\nu^2\right)}\,\phi_0(x\times\nu). \tag{A12}$$

Equation (A12) is obtained by changing the variable $u=z\times\nu$ and using equation (A5a). Equations (A11) and (A12) in (A4) give the expression of CVaR:

$$\mathrm{CVaR}_{SN2,\mu,\sigma,\nu}=\frac{1}{p}\left\{\frac{2}{1+\nu^2}\left[\mu\Phi_0\left(\frac{q-\mu}{\sigma}\nu\right)-\sigma\frac{1}{\nu}\phi_0\left(\frac{q-\mu}{\sigma}\nu\right)\right]\right\}.$$

Again, when $\nu=1$ we find the CVaR of $N(\mu,\sigma)$.

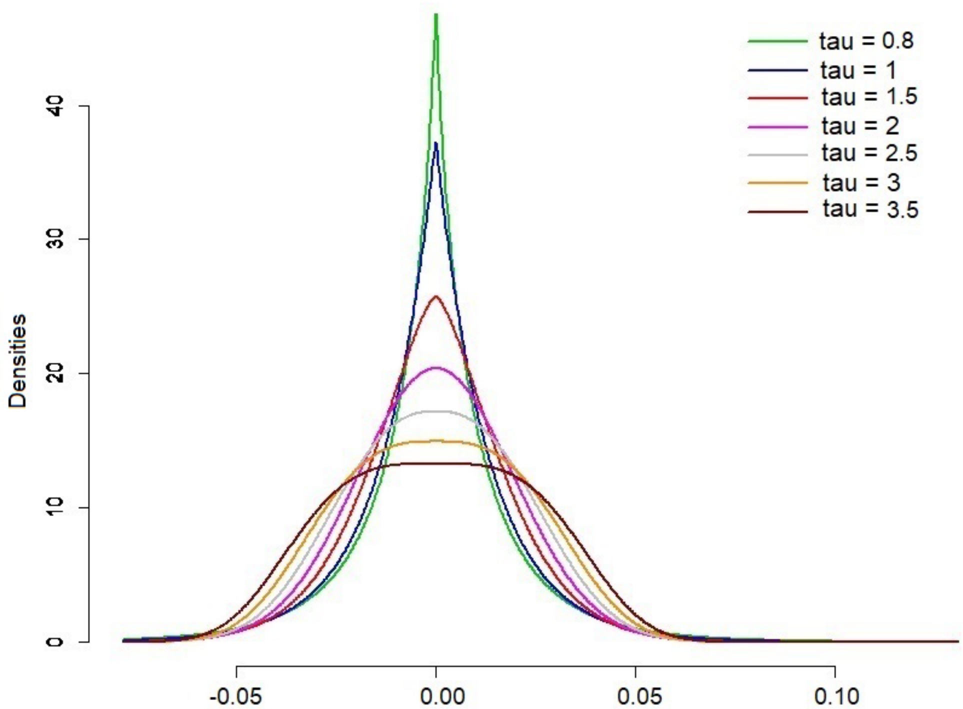
A8. The Skewed Exponential Power type 3 Distribution: SEP3

Fernandez et al. (1995) defined and named this distribution. SEP3 refers to the classification proposed by Rigby et al. (2014). The density of SEP3 is written as:

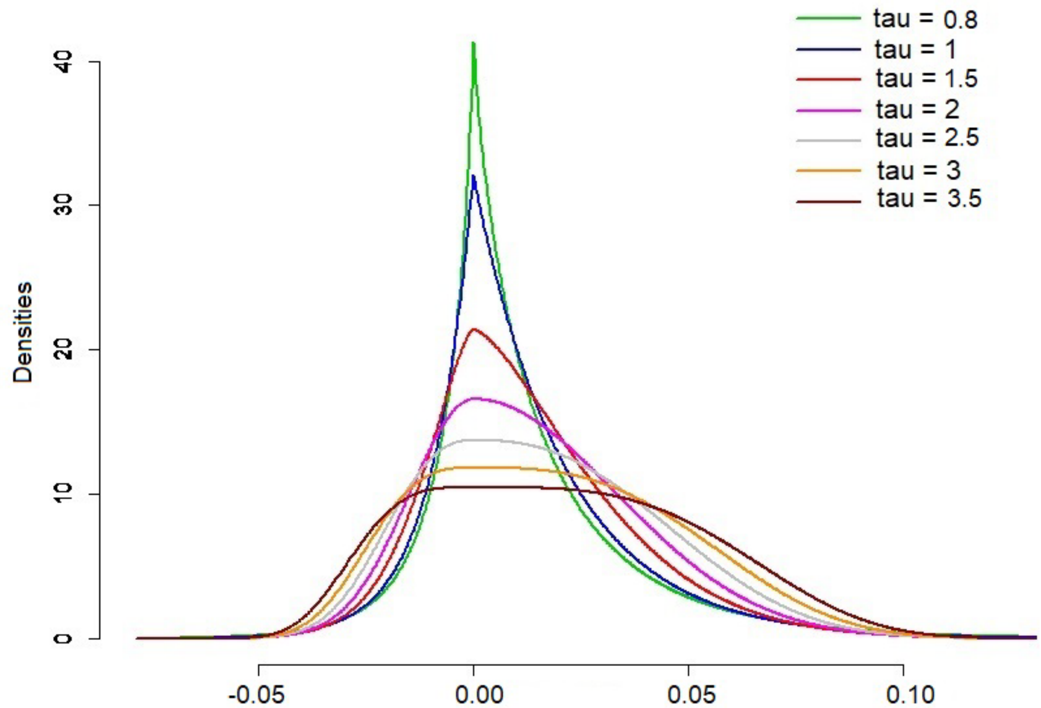
$$f_{SEP3,\mu,\sigma,\nu,\tau}(y)=\frac{c}{\sigma}\left\{\exp\left(-\frac{1}{2}\left|\frac{y-\mu}{\sigma}\nu\right|^{\tau}\right)I_{(y<\mu)}+\exp\left(-\frac{1}{2}\left|\frac{y-\mu}{\sigma}\frac{1}{\nu}\right|^{\tau}\right)I_{(y\geq\mu)}\right\}$$

where $c=\nu\times\tau\times\left[\left(1+\nu^2\right)2^{1/\tau}\Gamma\left(1/\tau\right)\right]^{-1}$ and where $\mu\in R,\sigma>0,\nu\in R,\tau>0$. They are respectively the parameters of location, scale, asymmetry, and tail thickness. SEP3 has as special cases the SN2 when $\tau=2$ and a Laplace distribution (asymmetric version) when $\tau=1$. Note that other names exist in the literature to designate distributions comparable to SEP3, such as AP (Asymmetric Power) and AEP (Asymmetric Exponential Power).

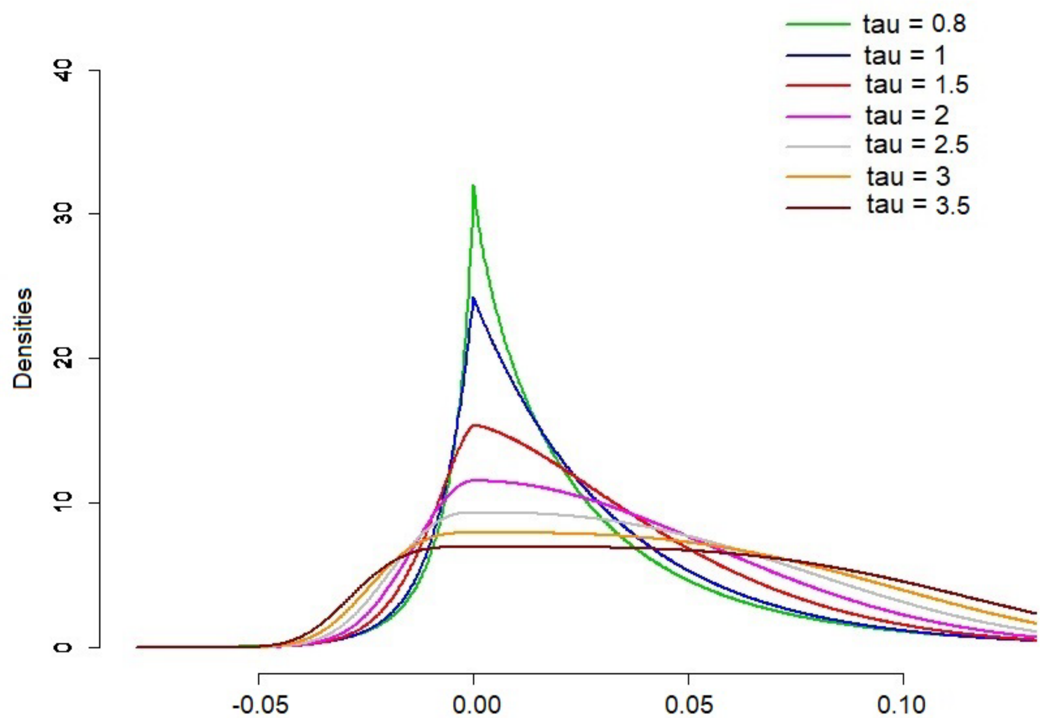
SEP3 can be leptokurtic when $\tau<2$ or platykurtic when $\tau>2$ (see Figure A1). VaR and CVaR calculations use gamma functions and the gamma distribution, as shown in the next section.



$\nu=1$



$\nu = 1.5$



$\nu = 2$

Figure A1: Plots of SEP3 with different values of τ and ν

A9. Expression of VaR and CVaR with SEP3

As we did for SN2, we develop the expression of reduced cumulative of SEP3 for $z<0$ (left tail) by writing:

$$F^0_{SEP3,\nu,\tau}(z)=\int_{-\infty}^{z}\frac{\tau\nu}{(1+\nu^2)2^{1/\tau}\Gamma(1/\tau)}\times\exp\left(-\frac{1}{2}\left|w\nu\right|^{\tau}\right)dw$$

$$=\frac{2^{1/\tau}}{\nu\tau}\frac{\tau\nu}{(1+\nu^2)2^{1/\tau}\Gamma(1/\tau)}\int_{(z\nu)^{\tau}/2}^{+\infty}u^{1/\tau-1}e^{-u}\,du \qquad \text{(A13)}$$

$$=\frac{1}{(1+\nu^2)\Gamma(1/\tau)}\int_{(z\nu)^{\tau}/2}^{+\infty}u^{1/\tau-1}e^{-u}\,du. \qquad \text{(A14)}$$

Equation (A13) is immediate after the change of variable $u=(-w\nu)^{\tau}/2$ and by positing $s=-z>0$. Note that the inside of the integral $u^{1/\tau-1}e^{-u}\,du$ is reminiscent of the gamma function. We need the complete gamma function $\Gamma(\cdot)$ and its incomplete version $\gamma(\cdot,\cdot)$, which are defined by:

$$\gamma(a,r)=\int_0^r t^{a-1}e^{-t}dt \qquad a>0,\, r>0$$

$$\Gamma(a)=\int_0^{+\infty} t^{a-1}e^{-t}dt \qquad a>0\,.$$

Parameter *a* is for the shape of these functions. It is easy to see that $\Gamma(a)=\gamma(a,+\infty)$. We also have a distribution that bears the same name, i.e. gamma,[12] whose cumulative parameter shape = a (and scale$=1$ because it is standardized) evaluated at the point $x>0$. It is written as $G_a(x)=\left[\Gamma(a)\right]^{-1}\gamma(a,x)$. The calculation of $F^0_{SEP3,\nu,\tau}(z)$ can be obtained from equality (A14):

[12] Under the same "gamma" designation, three entities can be distinguished: the function $\Gamma(.)$ (complete from 0 to $+\infty$) and the incomplete function (its integral stops at a point $r<+\infty$). The third entity is the gamma distribution with two parameters: shape and scale.

$$F^0_{SEP3,\nu,\tau}(z) = \frac{1}{(1+\nu^2)\Gamma(1/\tau)} \int_{(z\nu)^\tau/2}^{+\infty} u^{1/\tau-1} e^{-u}\, du$$

$$= \frac{1}{(1+\nu^2)\Gamma(1/\tau)} \left\{ \int_0^{+\infty} - \int_0^{(z\nu)^\tau/2} \right\} = \frac{1}{(1+\nu^2)} \frac{\Gamma(1/\tau) - \gamma\left(1/\tau, (z\nu)^\tau/2\right)}{\Gamma(1/\tau)} \qquad \text{(A15)}$$

$$= \frac{1}{1+\nu^2} \left(1 - G_{1/\tau}\left(\frac{|z\nu|^\tau}{2} \right) \right) \qquad \text{(A16)}$$

Equality (A15) is a cut-off of the integral's bounds that allows to find the gamma functions. To save space, we have not inserted the complete mathematical expressions of the two integrals in (A15), which are the same as in the previous equation. The expression is simplified by using the cumulative $G_{1/\tau}(\text{shape} = 1/\tau \text{ and scale} = 1)$. By inverting (A16), the quantile of VaR at the degree of confidence $(1-p)$ is immediate:

$$F^0_{SEP3,\nu,\tau}\left(\frac{q-\mu}{\sigma} \right) = p$$

$$q = \mu + \sigma \times \frac{\left[2 \times G^{-1}_{1/\tau}\left(1 - p(1+\nu^2) \right) \right]^{1/\tau}}{\nu}$$

The calculation of the tail of SEP3 is similar to that done for the cumulative, but with a shape $2/\tau$ parameter for $x < 0$:

$$\text{Tail}_{SEP3,\nu,\tau}(x) = \int_{-\infty}^{x} z f^0(z)\, dz = \int_{-\infty}^{x} c \times z \times \exp\left(-\frac{1}{2}|z\nu|^\tau \right) dz$$

$$= \frac{-2^{1/\tau}}{\nu(1+\nu^2)\Gamma(1/\tau)} \int_{(-x\nu)^\tau/2}^{+\infty} u^{2/\tau-1} e^{-u}\, du$$

$$= \frac{-2^{1/\tau}}{\nu(1+\nu^2)\Gamma(1/\tau)} \Gamma(2/\tau) \left(1 - G_{2/\tau}\left(\frac{|x\nu|^\tau}{2} \right) \right). \qquad \text{(A17)}$$

Finally, by putting (A16) and (A17) in (A4) we find:

$$\mathrm{CVaR}_{\mathrm{SEP3},\mu,\sigma,\nu,\tau}=\frac{1}{p}\left\{\frac{1}{1+\nu^2}\left[\mu\times\left(1-G_{1/\tau}\left(\frac{\left|\frac{q-\mu}{\sigma}\nu\right|^{\tau}}{2}\right)\right)-\sigma\times\frac{2^{1/\tau}}{\nu}\frac{\Gamma(2/\tau)}{\Gamma(1/\tau)}\left(1-G_{2/\tau}\left(\frac{\left|\frac{q-\mu}{\sigma}\nu\right|^{\tau}}{2}\right)\right)\right]\right\}.$$

Remember that $G_{n/\tau}(x)$ is the cumulative gamma distribution of shape $= n/\tau$ and scale = 1 evaluated at point *x*. When $\tau = 2$, we return to SN2. If $\tau = 2$, and $\nu = 1$, we get a normal distribution. The gamma distribution and the complete gamma function exist in standard Excel.

A10. CVaR backtest notations and expressions

The backtests performed in this paper are out of sample. The series of 1,200 daily returns is named $\{X_t\}_{t=1}^{t=1200}$. We have eight models $M_i, i = 1...8.$ For model M and day *t*, we take the 250 returns preceding this day to estimate the vector of parameters of model M which we note as θ_t. Based on this vector θ_t, we calculate the measures $\mathrm{VaR}_{p,t}$ and $\mathrm{CVaR}_{p,t}$ relative to the degree of confidence $(1-p)$. We recall here that $\mathrm{VaR}_{p,t} > 0$ and $\mathrm{CVaR}_{p,t} > 0$ for all *t* by convention. For each model, we will have built two series of size 1,200 each: $\{\mathrm{VaR}_{p,t}\}_{t=1}^{t=1200}$ and $\{\mathrm{CVaR}_{p,t}\}_{t=1}^{t=1200}$.

The first backtest is Z_{ES}, introduced by Acerbi and Szekely (2017). The expression of its statistic is:

$$Z_{ES}(X_t)=\frac{1}{T}\sum_{t=1}^{T}\frac{p\times\left(\mathrm{CVaR}_{p,t}-\mathrm{VaR}_{p,t}\right)+\left(X_t+\mathrm{VaR}_{p,t}\right)\left(X_t+\mathrm{VaR}_{p,t}\right)<0}{p\times\mathrm{CVaR}_{p,t}}. \qquad \text{(A18)}$$

The null hypothesis H_0 of the test Z_{ES} is that CVaR is appropriate and, in this case, the statistic $Z_{ES}(X_t)$ must be statistically zero. The alternative hypothesis H_1 is under- or overestimated: $H_0: Z_{ES}(X_t)=0; H_1: Z_{ES}(X_t)\neq 0.$ Here is the procedure for calculating the distribution of the null hypothesis. For each day *t*, we draw N random values using the M model with the parameters θ_t. Taking N = 5,000, for example, the draws generate a matrix $\{Y_t^n\}$ of

1,200 columns and 5,000 rows. Applying equation (A18) and replacing X_t with Y_t^n we calculate the series of 5,000 H_0 values, $\left\{Z_{ES}\left(Y_t^n\right)\right\}_{n=1}^{n=5000}$. The *p*-value of the test Z_{ES} is then equal to $\min\left[\Pr\left(Z_{ES}(Y) < Z_{ES}(X)\right), \Pr\left(Z_{ES}(Y) > Z_{ES}(X)\right)\right]$.

The second backtest is denoted RC and is proposed by Righi and Ceretta (2015). Its statistic is defined by the expression:

$$RC\left(X_t\right) = \frac{1}{T}\sum_{t=1}^{T}\frac{\left(X_t + CVaR_{p,t}\right)\times\left(X_t + VaR_{p,t}\right) < 0}{SD_{p,t}} \qquad (A19)$$

where $SD_{p,t} = \sqrt{\operatorname{variance}\left(X_t \times \left(X_t + VaR_{p,t}\right) < 0\right)}$ is the standard deviation of X_t those that exceed the VaR. In the standard version of Righi and Ceretta (2015), the *p*-value is obtained by bootstrapping according to Efron and Tibshirani (1994). Here, we will obtain it instead by following exactly the same construction as for Z_{ES}.

Finally, and for information purposes only, the Z_1 and Z_2 statistics are defined by:

$$Z_1 = \frac{\Sigma_{t=1}^{T}\frac{X_t}{CVaR_{p,t}}\times\left(X_t + VaR_{p,t}\right) < 0}{\Sigma_{t=1}^{T} 1\times\left(X_t + VaR_{p,t}\right) < 0} + 1$$

$$Z_2 = \frac{1}{T\times p}\sum_{t=1}^{T}\frac{X_t\times\left(X_t + VaR_{p,t}\right) < 0}{CVaR_{p,t}} + 1.$$

A11. Model estimation and parametric and non-parametric VaR and CVaR calculations

The estimated parameters of the distributions are given in the following tables.

Table A.2
Model estimation - Panel A

	1:NO	1:T	1:EGB2	2:NO	2:T
μ_1	0.0005244	0.0006974**	0.0008884*	−0.0004845	0.0012920**
	(0.0003741)	(0.0002977)	(0.0004982)	(0.0015691)	(0.0005553)
σ_1	0.0129631***	0.0085310***	0.0014108**	0.0226636***	0.0066854***
	(0.0002645)	(0.0003410)	(0.0006812)	(0.0018632)	(0.0009171)
ν_1		3.2887197***	0.1587161**		23,642.31***
		(0.3809600)	(0.0796200)		(0.0000001)
τ_1			0.1652522*		
			(0.0851634)		
μ_2				0.0008151**	−0.0004740
				(0.0003448)	(0.0008931)
σ_2				0.0082545***	0.0140598***
				(0.0005136)	(0.0025828)
ν_2					6.4162601**
					(2.5707612)
τ_2					
c_1				0.2231962***	0.5158049***
				(0.0497856)	(0.1538992)
No of params	2	3	4	5	7
LogLik	3,512.5506	3,627.5723	3,625.5829	3,619.0444	3,628.6421
AIC	−7,021.1012	−7,249.1446	−7,243.1659	−7,228.0889	−7,243.2842
BIC	−7,010.9210	−7,233.8744	−7,222.8056	−7,202.6385	−7,207.6536
KS (*p*-value)	0.0015	0.1285	0.3490	0.2180	0.1100
No of obs.	1,200	1,200	1,200	1,200	1,200

$^{***}p < 0.01$ $^{**}p < 0.05$ $^{*}p < 0.1$

Table A.3
Model Estimation - Panel B

	3:NO	2:SN2 + 1:NO	2:SEP3
μ_1	−0.0004753	0.0025930	−0.0007520***
	(0.0009649)	(0.0065819)	(0.0001560)
σ_1	0.0150441***	0.0146788***	0.0045291***
	(0.0022041)	(0.0020169)	(0.0014071)
ν_1		0.8830671***	1.0315089***
		(0.2038416)	(0.0376383)
τ_1			0.9598700***
			(0.1180946)
μ_2	0.0043390	0.0009227	0.0075456**
	(0.0098212)	(0.0013866)	(0.0032033)
σ_2	0.0376531***	0.0063897***	0.0065018**
	(0.0101797)	(0.0012556)	(0.0025539)
ν_2		0.9939552***	0.6137048***
		(0.2581567)	(0.2182171)
τ_2			2.1083901*
			(1.1436395)
μ_3	0.0011752***	0.0091833	
	(0.0004491)	(0.0149751)	
σ_3	0.0065771***	0.0388917***	
	(0.0008483)	(0.0081479)	
c_1	0.4433715***	0.4729333***	0.7389303***
	(0.1089861)	(0.1594929)	(0.1181466)
c_2	0.0334707	0.5000573***	
	(0.0303812)	(0.1734363)	
No. of params	8	10	9
LogLik	3,630.0653	3,630.7232	3,631.9633
AIC	−7,244.1307	−7,241.4465	−7,245.9267
BIC	−7,203.4101	−7,190.5457	−7,200.1160
KS (*p*-value)	0.2280	0.1980	0.3040
No of Obs.	1,200	1,200	1,200

***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

Table A.4
Calculation and comparison of CVaRs

p	Densities Mixtures	VaR (in %)	CVaR-/CVaR+ CVaR (in%)	Mean (in%)	Variance (in%)	Skewness	Kurtosis
0.050	nparam	2.04736	2.96269/2.97795	0.0524	0.0168	0.3589	9.8158
0.050	1:NO	2.07980	2.62147	0.0524	0.0168	0.0000	3.0000
0.050	1:T	1.86805	3.01294	0.0697	0.0186	0.0000	
0.050	1:EGB2	2.00674	2.89562	0.0525	0.0157	-0.0813	5.8076
0.050	2:NO	1.95397	3.11363	0.0525	0.0168	-0.1386	6.6789
0.050	2:T	2.02945	3.04197	0.0437	0.0163	-0.1544	8.3993
0.050	3:NO	2.03846	3.00451	0.0549	0.0172	0.1224	9.4321
0.050	2:SN2 + 1:NO	2.05018	2.98338	0.0524	0.0168	0.2433	9.8409
0.050	2:SEP3	1.99293	2.97395	0.0544	0.0163	0.0051	7.1752
0.025	nparam	2.54290	3.63040/3.66665				
0.025	1:NO	2.48828	2.97808				
0.025	1:T	2.51522	3.87890				
0.025	1:EGB2	2.62287	3.51175				
0.025	2:NO	2.81354	3.90424				
0.025	2:T	2.71654	3.74976				
0.025	3:NO	2.66598	3.68928				
0.025	2:SN2 + 1:NO	2.68920	3.62898				
0.025	2:SEP3	2.66110	3.66159				
0.010	nparam	3.59575	4.44800/4.51902				
0.010	1:NO	2.96323	3.40250				
0.010	1:T	3.54473	5.29712				
0.010	1:EGB2	3.43734	4.32622				
0.010	2:NO	3.89559	4.82632				
0.010	2:T	3.62577	4.72258				
0.010	3:NO	3.47885	4.71115				
0.010	2:SN2 + 1:NO	3.47913	4.53241				
0.010	2:SEP3	3.57259	4.58396				