The Role of the Skewed Distributions in the Framework of Extreme Value Theory (EVT)

Sonia Benito Muela[1], Carmen López-Martín[2] & M ª Ángeles Navarro[3]

[1] Faculty of Economics and Business Administration, National Distance Education University (UNED), Madrid, Spain

[2] School of Law and Social Sciences, University Carlos III of Madrid, Madrid, Spain

[3] PhD. Student of the Faculty of Economics and Business Administration, National Distance Education University (UNED), Madrid, Spain

Correspondence: Sonia Benito Muela, Faculty of Economics and Business Administration, National Distance Education University (UNED), Senda del Rey, 11, 28040 Madrid, Spain.

Abstract

In this paper, we analyze the role of the heavy tail and skewed distribution in market risk estimation (Value at Risk (VaR)). In particular, we are interested in knowing if in the framework of the conditional extreme value theory, the estimation of the volatility model below heavy tail and skewed distribution contributes to improve the VaR estimation respect to these obtained from a symmetric distribution. The study has been carried out for six individual assets belonging to the digital sector: ADP, Amazon, Cerner, Apple, Microsoft and Telefonica. The analysis period runs from January 1st, 2008 to the end of December 2013. Although the evidence found is a little bit weak, the results obtained seem to indicate that the heavy tail and skewed distribution outperforms the symmetric distribution both in terms of accuracy VaR estimations as in terms of firm's loss function. Furthermore, the market risk capital requirements fixed on the base of the VaR estimations are also lowest below a skewed distribution.

JEL Classifications: C15, C22, C52, G17, G32.

Keywords: extreme value theory, loss functions, skewness distributions, value at risk

1. Introduction

A context of risk is one in which we do not know with certainty the consequences associated with a decision. The only thing that we know is possible outcomes associated with it and the likelihood of achieving such results. In the financial field, the notion of risk implies that we know the various yields can potentially get to make an investment and also know the probability of achieving such results. This allows us to estimate the average expected yield and the possible diversion 'above' or 'below' the average value, that is, the risk. The most popular and traditional risk measure is volatility (variance). In fact, traditional financial theory defines risk as the dispersion of returns due to movements in financial variables.

Another way of measuring risk, which is the most commonly used at present, is to evaluate the losses that may occur when the price of the asset that makes up the portfolio goes down. This is what Value at Risk (VaR) does. The Value at Risk of a portfolio indicates the maximum amount that an investor may lose over a given time horizon and with a given probability. In this case, the concept of risk is associated with the danger of losses.

According to Jorion (2001), "VaR measure is defined as the worst expected loss over a given horizon under normal market conditions at a given level of confidence". For the estimation of VaR measure different methodologies have been developed. These methodologies can be classified in three groups: (i) parametric methods (as Riskmetrics); (ii) non-parametric method (as historical simulation) and (iii) semiparametric method (as extreme value theory, filter historical simulation and CaViar Method)[1].

Among all of them, one of the most used by the financial institution is parametric method. This method assumes

---

[1] See Abad, Benito and López-Martín (2014) for a review of those methodologies.

that financial returns follow a known distribution. Thus, below this method, the VaR of a portfolio at $(1 - \alpha)$ confidence level is calculated as the product of the conditional standard deviation of the return portfolio by the α quantile (k_α) of the assumed distribution[2]. The mainly assumed distribution is the normal distribution. Below this assumption, k_α is the α quantile of the standard normal.

The major drawbacks of this method are related to the normal distribution assumption for financial returns. Empirical evidence shows that financial returns do not follow normal distribution. The skewness coefficient is in most cases negative and statistically significant, implying that the financial returns distribution is skewed to the left. This result is not in accord with the properties of a normal distribution, which is symmetric. In addition to this, empirical distribution of financial return has been documented to exhibit significantly excessive kurtosis (fat tails and peakness) (see Bollerslev, 1987). Consequently, the size of the actual losses is much higher than that predicted by a normal distribution.

Taking this into account, the research in the framework of parametric method has focus on investigating other density functions which capture the skew and kurtosis of financial returns. In this line, Abad, Benito, López-Martín and Sánchez (2016), Chen, Gerlach, Lin and Lee (2012), Polansky and Stoja (2010), Bali and Theodossiou (2008), Ausin and Galeano (2007), Zhang and Cheng (2005) among others, show that, in the context of parametric method, assuming fat tail and skewness distributions improve the performance of this model in VaR estimation.

As we will see later, below the conditional extreme value theory (which is one of the most successful method in estimating VaR), the value at risk of a portfolio at $(1 - \alpha)\%$ confidence level is calculated as the product of the conditional standard deviation of the return portfolio by the α (k_α) quantile of Pareto generalize distribution (PGD). Traditionally, the conditional standard deviation of the return portfolio is estimated by assuming a symmetric distribution for the financial return. So, in the same line that aforementioned paragraphs in this paper, we analyze if, in the framework of the method based on the conditional extreme value theory, the estimation of the volatility model below a fat tail and skewness distribution contributes to improve the results in VaR estimation.

The rest of the paper is as follow. In section 2 we present the volatility models used in the empirical application. In section 3 we describe the VaR measure, the methodology used in this paper to calculate this measure and the backtesting techniques. The results of the empirical application are presented in section 4. Section 5 includes the main conclusion.

2. Volatility Models

According to Engle (1982), Bollerslev (1986) and Nelson (1991), the financial returns are characterized by a variance changing over time, alternating periods of low volatility followed by periods of high volatility. In other words, as noted by Mandelbrot (1963), "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". This effect, known as *volatility clustering*, indicates the presence of conditional heteroscedasticity in the return series and the need to model the behaviour of the conditional variance. Furthermore, financial returns are subject to leverage (Black, 1976). This means that volatility tends to be higher after negative returns, this is typically attributed to leverage (hence the name).

To capture cluster in volatility Engle (1982) proposed Autoregressive Conditional Heteroscedasticity (ARCH), which featured a variance that does not remain fixed but rather varies throughout a period. Bollerslev (1986) extended the ARCH model into the Generalized ARCH (GARCH). The GARCH model captures volatility clustering but does not capture the leverage effect. In this model the positive surprises have the same effect on volatility than negative surprises. To capture the leverage effect have been developed asymmetric GARCH models as for instance the APARCH model (Ding, Granger & Engle, 1993) or the EGARCH model proposed by Nelson (1991).

On the other hand, it is well known that, long memory property observed in the volatility of the financial series cannot be captured by traditional GARCH models, so it is necessary to model a long-term component. To capture this characteristic, Engle and Lee (1999) proposed a model in which the variance is decomposed into a component of long-term and short-term component. The main role of short-term component is to pick up the temporary increase in the variance after an impact on the price. While the persistence or long-term memory indicates how long the shock takes to be absorbed.

Creal, Koopman and Lucas (2008, 2011) and Harvey and Chakravarty (2008) propose a model that integrates

---

[2]This is true when we assume that the portfolio return mean is zero.

three characteristics aforementioned: cluster in volatility, leverage effect and long memory. The model proposed by these authors is known Beta-t-EGARCH model. In this model, the short-term component responds in the same manner as in the traditional GARCH. Nevertheless, the long-term component is not sensitive to extreme observations, as it does in the standard GARCH model. That is, in the standard GARCH, the existence of an outlier in yields has a persistent effect on volatility, which increases the variance. In contrast, the Beta-t-EGARCH model corrects volatility in case of existence of an outlier, absorbing the effect and returning to previous levels of volatility.

Formally let $r_1, r_2, \ldots, r_n$ be a sequence of independent and identically distributed random variables representing financial returns. Assume that $\{r_t\}$ follows a stochastic process:

$$r_t = exp(\lambda_{t|t-1})(\varepsilon_t), \quad \varepsilon_t \sim t(0, \sigma_\varepsilon, v) \qquad\qquad v > 2 \qquad\qquad (1)$$

where the conditional error ε_t is distributed as a t-Student with zero mean, unconditional standard deviation σ_ε and degrees of freedom parameter v. $\lambda_{t|t-1}$ is the conditional scale or volatility, that does not need to be equal to the conditional standard deviation. The conditional standard deviation is obtained as $\sigma_\varepsilon \, exp\, (\lambda_{t|t-1})^3$.

In the Beta-t-EGARCH model, in (1) is defined as:

$$\lambda_{t|t-1} = \delta + \lambda_{1,t|t-1}^+ + \lambda_{2,t|t-1}^+ \qquad\qquad (2)$$

being the long-term component is $\lambda_{1,t|t-1}^+ = \varphi_1 \lambda_{1,t-1|t-2}^+ + k_1 u_{t-1}$ and the short-term component is $\lambda_{2,t|t-1}^+ = \varphi_2 \lambda_{2,t-1|t-2}^+ + k_2 u_{t-1}$. The leverage effect may be introduced into the model using the sign of the observations. Thus, the short component with leverage effect is as follow:

$$\lambda_{2,t|t-1}^+ = \varphi_2 \lambda_{2,t-1|t-2}^+ + k_2 u_{t-1} + k^*(-r_{t-1})(u_{t-1} + 1) \qquad\qquad (3)$$

and u_t is the score conditional which is given by:

$$u_t = \frac{(v+1)(r_t)^2}{v exp(2\lambda_{t|t-1}) + (r_t)^2} - 1 \qquad\qquad -1 \leq u_t \leq v, \; v > 0 \qquad\qquad (4)$$

Taking the signs of minus, r_{t-1} means that the parameter of k^* is normally non-negative for stock returns as in the GARCH model, the long-term component $\lambda_{1,t|t-1}^+$, has φ_1 close to one or even set equal to one, while the short-term component $\lambda_{2,t|t-1}^+$, will typically have a higher k combined with a lower φ. The model is not identifiable if the $\varphi_1 = \varphi_2$. Imposing the constraint $0 < \varphi_2 < \varphi_1 < 1$ ensures identifiability and stationarity.

Finally, the empirical literature has shown that financial returns also exhibit skewness and fat-tailed distributions. To capture these features, Harvey and Sucarrat (2014) extended the Beta-t-EGARCH model, combining the skewness of the conditional distribution with a leverage effect in the dynamic of the scale. This model is known as Beta-skewness-t-EGARCH.

Skewness is introduced into the Beta-t-EGARCH model using the method proposed by Fernandez and Steel (1998) (see Harvey and Sucarrat (2014) for more details of this method). Thus, in equation (1), they assume that the conditional error ε_t is distributed as a skewed t-student with mean μ_ε, scale σ_ε, degrees of freedom parameter v and skewness parameter γ^4.

$$r_t = exp(\lambda_{t|t-1})(\varepsilon_t - \mu_\varepsilon), \varepsilon_t \sim st(\mu_\varepsilon, \sigma_\varepsilon, v, \gamma) \qquad v > 2, \; \gamma \epsilon (0, \infty) \qquad\qquad (5)$$

where $\lambda_{t|t-1}$ is given by:

$$\lambda_{t|t-1} = \delta + \lambda_{1,t|t-1}^+ + \lambda_{2,t|t-1}^+ \qquad\qquad (6)$$

The long-term component is $\lambda_{1,t|t-1}^+ = \varphi_1 \lambda_{1,t-1|t-2}^+ + k_1 u_{t-1}$ and the short-term component is $\lambda_{2,t|t-1}^+ = \varphi_2 \lambda_{1,t-1|t-2}^+ + k_2 u_{t-1} + k^*(-(r_{t-1}))(u_{t-1} + 1)$. Again, only the short-term component has a leverage effect. In

---

[3]The specification for the standard GARCH model is as follows: $r_t = \sigma_t \varepsilon_t$, $\varepsilon_t \sim$ IID $(0, \sigma_\varepsilon^2)$, where σ_t^2 is the scale o volatility, which is modeled $\sigma_t^2 = \delta + \varphi_1 \sigma_{t-1}^2 + k_1 \varepsilon_{t-1}^2$ with $\sigma_\varepsilon^2 = 1$.

[4]The conditional error ε_t is an uncentred (i.e., mean not necessarily equal to zero) Skewed t variable with v degrees of freedom, skewness parameter γ. A centred and symmetric t-distribution variable with mean zero is obtained when $\gamma = 1$, in which $\mu_\varepsilon = 0$, whereas a left-skewed (right-skewed) t-variable is obtained when $\gamma < 1$, $(\gamma > 1)$. More details on the distribution can be found in Harvey and Sucarrat (2013) and Sucarrat (2013).

this model the conditional score is given by:

$$u_t = \frac{(v+1)(r_t + \mu_\varepsilon exp(\lambda_{t|t-1}))(r_t)}{v\gamma^{2(r_t + \mu_\varepsilon exp(\lambda_{t|t-1}))} exp(2\lambda_{t|t-1}) + (r_t + \mu_\varepsilon exp(\lambda_{t|t-1}))^2} \tag{7}$$

3. VaR Methodology

Let $r_1, r_2, \ldots \ldots r_n$ be identically distributed independent random variables representing the financial returns. Use $F(r)$ to denote the cumulative distribution function $F(r) = Pr(r_t < r|\Omega_{t-1})$ conditioned to an information set, Ω_{t-1}, available at time t-1. Assume that $\{r_t\}$ follows the stochastic process:

$$r_t = \mu + \sigma_t \varepsilon_t \qquad \varepsilon_t \sim iid(0,1) \tag{8}$$

being μ the mean of returns; $\sigma_t^2 = E(\varepsilon_t^2|\Omega_{t-1})$ and ε_t has the conditional distribution function $G(\varepsilon)$, $G(\varepsilon) = Pr(\varepsilon_t < \varepsilon |\Omega_{t-1})$.

The VaR at a given level of confidence $1 - \alpha$, denoted by $VaR(\alpha)$, is defined as the α quantile of the probability distribution of financial returns.

$$F(VaR(\alpha)) = Pr(r_t < VaR(\alpha)) = \alpha \tag{9}$$

There are two ways to estimate this quantile: (1) inverting the distribution function of financial returns, $F(r)$, or (2) inverting the distribution function of innovations, $G(\varepsilon)$. The latter case will also require estimating the standard deviation of returns.

$$VaR(\alpha) = F^{-1}(\alpha) = \mu + \sigma_t G^{-1}(\alpha) \tag{10}$$

Thus, the VaR estimation involves the specifications of the distribution function of financial returns, $F(r)$, or the distribution function of innovations, $G(\varepsilon)$, along with the standard deviation of returns, σ_t.

The historical simulation method, Monte Carlo simulation and unconditional approach based on the Extreme Value Theory focus on the estimation of $F(r)$, while Parametric Method and the Conditional Extreme Value Theory estimate $G(\varepsilon)$.

Below, we will describe the Conditional Extreme Value Theory, which have been used in this study to compute VaR.

3.1 Conditional Extreme Value Theory

The Extreme Value Theory (EVT) approach focuses on the limiting distribution of extreme returns observed over a long time period, which is essentially independent of the distribution of the returns themselves.

There are two methods based on EVT: a) Block Maxima Model (BMM) proposed by McNeil (1998) and b) Peaks Over Threshold Model (POT). This second model is generally considered to be the most useful for practical applications, due to the more efficient use of the data on extreme values.

Within the POT models framework, we can distinguish two types of analysis: (i) the Semi-Parametric Model built around Hill estimator and (ii) and the fully parametric models based on the generalized Pareto distribution. The latter method is commonly most used in practice.

Below, the fundamental theory of this approach is described, considering both unconditional and conditional.

Let $r_1, r_2, \ldots \ldots r_n$ be a random sequence of observations representing the financial returns. Given a threshold denoted by u, we will be interested in excess losses over the threshold[5] denoted by $y_1, y_2, y_3, \ldots y_{N_u}$, where $y_i = r_i - u$ and N_u are the number of sample data greater than u.

Then, for instance, if the threshold is equal to 1.5 %, we are left with all returns lower to 1.5 %.

The Extreme Value Theory assumes that the distribution of excess losses above the threshold follows a Generalized Pareto distribution (GPD) given for the expression:

$$G_{k,\xi}(y) = 1 - [1 + \frac{k}{\xi}y]^{-1/k} \tag{11}$$

being k y ξ the parameters of the distribution[6].

It can be shown that under this assumption, the α percentile of the distribution, as the VaR can be estimated as:

---

[5]The most common is to use the 10% percentile as the threshold level.

[6]These parameters can be estimated by maximum likelihood.

$$VaR(\alpha) = G_{k,\xi}^{-1}(\alpha) = u + \frac{\xi}{k}\left[\left[\frac{n}{N_u}(1-\alpha)\right]^{-\xi} - 1\right] \qquad r < u \qquad (12)$$

where n represents the number of sample data.

The extreme value method described in the preceding paragraphs does not consider the level of volatility. This method is known as *Unconditional Extreme Value Theory.*

Since financial returns are variables that are characterized by heteroscedasticity, McNeil and Frey (2000) proposed a new methodology, called *Conditional Extreme Value Theory*, for estimating VaR that combines extreme value theory with models of volatility.

According to this theory, the VaR of a portfolio at a confidence level of $1 - \alpha$ can be calculated as:

$$VaR(\alpha)_t = \mu + \sigma_t G_{k,\xi}^{-1}(\alpha) \qquad (13)$$

where σ_t^2 represents the conditional standard deviation of the financial returns and $G_{k,\xi}^{-1}$ is the α quantile of the GPD.

To estimate the conditional standard deviation of the returns, in this paper, we use the beta-t-EGARCH model and the beta-skewness-t-EGARCH models presented in section 2.

3.2 Backtesting

To check the adequacy of the VaR estimates, we use two alternative approaches: statistical tests that evaluate the accuracy of the estimates and loss functions.

As to the first approach, to test the accuracy of the VaR estimates, we recurre some standard procedures based on statistical hypothesis testing which have been proposed in the literature. In this work, we use to test the accuracy of the VaR estimates five standard tests: (i) unconditional coverage test (Kupiec, 1995), (ii) Backtesting Criterion (BTC), (iii) conditional coverage test (Christoffersen, 1998), (iv) statistics for serial independence (LRind) and (v) dynamic quantile (Engle and Manganelli, 2004).

Before implementing these tests, we must define an exception indicator. This indicator is calculated as follows:

$$I_{t+1} = \begin{cases} 1 & si \quad r_{t+1} < VaR(\alpha) \\ 0 & si \quad r_{t+1} > VaR(\alpha) \end{cases} \qquad (14)$$

Kupiec (1995) shows that if we assume that the probability of getting an exception is constant, then the number of exceptions $x = \sum I_{t+1}$ follows a binomial distribution $B(N, \alpha)$, where N represents the number of observations. An accurate measure VaR should produce an unconditional coverage ($\hat{\alpha} = \frac{\sum I_{t+1}}{N}$) equal to α percent.

Unconditional coverage test has a null hypothesis $\hat{\alpha} = \alpha$, with a likelihood ratio statistic:

$$LR_{uc} = 2[log(\hat{\alpha}^x(1-\hat{\alpha})^{N-x}) - log(\alpha^x(1-\alpha)^{N-x})] \qquad (15)$$

which follows an asymptotic $\chi^2(1)$ distribution.

A similar test for the significance of the departure of $\hat{\alpha}$ from α is the back-testing criterion statistic (BTC):

$$Z = (N\hat{\alpha} - N\alpha)/\sqrt{N\alpha(1-\alpha)} \qquad (16)$$

which follows an asymptotic N(0,1) distribution

The conditional coverage test, developed by Christoffersen (1998), jointly examines whether the percentage of exceptions is statistically equal to the one expected ($\hat{\alpha} = \alpha$) and the serial independence of the exception indicator. The likelihood ratio statistic of this test is given by $LR_{cc} = LR_{uc} + LR_{ind}$ which is asymptotically distributed as $\chi^2(2)$ and the LR_{ind} statistic is the likelihood ratio statistic for the hypothesis of the serial independence against first-order Markov dependence[7].

---

[7] The LR$_{ind}$ statistic is $LR_{ind} = 2[\log L_A - \log L_0]$ and has an asymptotic $\chi^2(1)$ distribution. The likelihood function under the alternative hypothesis is $L_A = (1-\pi_{01})^{N_{00}} \pi_{01}^{N_{01}} (1-\pi_{11})^{N_{10}} \pi_{11}^{N_{11}}$ where N$_{ij}$ denotes the number of observations in state j after having been in state i in the previous period, $\pi_{01} = N_{01}/(N_{00}+N_{01})$ and $\pi_{11} = N_{11}/(N_{10}+N_{11})$. The likelihood function under the null hypothesis is ($\pi_{01} = \pi_{11} = \pi = (N_{11}+N_{01})/N$) is $L_0 = (1-\pi)^{N_{00}+N_{01}} \pi^{N_{01}+N_{11}}$.

Finally, the dynamic quantile test proposed by Engle and Manganelli (2004) examines if the exception indicator is uncorrelated with any variable that belongs to the information set Ω_{t-1} available when the VaR is calculated. This is a Wald test of the hypothesis that all slopes are zero in the regression model:

$$I_t = \beta_0 + \sum_{i=1}^{p} \beta_i I_{t-i} + \sum_{j=1}^{q} \mu_j X_{t-j} \tag{17}$$

where X_{t-j} are explanatory variables contained in Ω_{t-1}.

This statistic introduced as explanatory variables lags of VaR. Under null hypothesis, the exception indicator cannot be explained by the level of VaR, i.e. $VaR(\alpha)$ is usually an explanatory variable to test if the probability of an exception depends on the level of the VaR.

In a second stage, we evaluate the magnitude of the losses experienced. For this purpose, we have considered three loss functions: the regulator's loss function, the firm 's loss function and daily capital requirement.

Lopez (1998, 1999) proposed a loss function that reflects the utility function of a regulator. This function assigns a quadratic specification when the observed portfolio losses exceed the VaR estimate. Thus, we penalise only when an exception occurs according to the following specification:

$$RLF_t = \begin{cases} (VaR_t - r_t)^2 & if \ r_t < VaR(\alpha) \\ 0 & otherwise \end{cases} \tag{18}$$

This loss function gives higher scores when failures take places and considers the magnitude of these failures. In addition, the quadratic term ensures that large failures are penalised more than small failures. Among several models, that provides the lowest loss function value is the best model.

This loss function could be optimal from the perspective of the regulator, but not from the point of view of the firm since a model that generates a VaR too high leads the firm to incur high capital costs.

Taking this into account, Sarma, Thomas and Shah (2003) defines the firm's loss function as follows:

$$FLF_t = \begin{cases} (VaR_t - r_t)^2 & if \ r_t < VaR(\alpha) \\ -\beta VaR_t & otherwise \end{cases} \tag{19}$$

where β is the opportunity cost of capital.

This function penalises the cases in which there are no exceptions multiplying the VaR by a factor β. As Abad, Benito and López-Martín (2015) indicate, this product does not capture the opportunity cost of capital in an exact way. Therefore, in line with Sarma et al. (2003), they propose a new firm's loss function that is expressed as follows:

$$FABL_t = \begin{cases} (VaR_t - r_t)^2 & if \ r_t < VaR(\alpha) \\ \beta \ (r_t - VaR_t) & otherwise \end{cases} \tag{20}$$

As can be determined in this function, the exceptions are penalized as usual in the literature, following the instructions of the regulator. But when there are no exceptions, the loss function penalizes the difference between the VaR and returns weighted by a factor that represents an interest rate. This product is exactly the opportunity cost of the capital, ie, the excess capital held by the firm.

Finally, we evaluate the VaR estimate on the bases of daily capital requirement (see McAleer, Jimenez & Pérez-Amaral, 2013). These authors adapt to daily term the function used by the financial institutions for calculating market risk capital requirement in a 10 days' horizon (Basel II). The daily capital requirement at time t can be calculated as follow (BCBS, 1996; 2006):

$$DCR_t = sup\{-k \times \overline{VaR}_{60}, -VaR_{t-1}\} \tag{21}$$

where DCR_t represents the daily market capital requirement at time t, which is the higher of $k \times \overline{VaR}_{60}$ and $-VaR_{t-1}$; $\overline{VaR}_{60}$ is the mean VaR over the previous 60 working days; $(3 \leq k \leq 4)$ is the Basel II violation penalty (see Table 1).

Table 1. Basel Accord Penalty Zones

Zone	Number of exceptions	k
Green	0 to 4	3
	5	3.4
	6	3.5
Yellow	7	3.65
	8	3.75
	9	3.85
Red	10 or more	4

Note: The number of exceptions is given for 250 trading days

4. Empirical Application

4.1 Data Analysis

The data set consists of daily returns of six companies from 1/04/2000 to 12/31/2013 (approximately 3520 observations). The considered companies are ADP, Amazon, Apple, Cerner, Microsoft and Telefonica. The computation of the returns (r_t) is based on the formula $r_t = ln(P_t/P_{t-1})$ where P_t is the price at time t. The evolution of daily prices and yields are represented in Figure 1.

Table 2 shows descriptive statistics of the return series. For each company, the unconditional mean is very close to zero. The highest unconditional standard deviation are 0.0368 (Amazon) and the lowest 0.0167 (ADP). For the rest series of return, the standard deviation moves between 0.0208 and 0.0275. The skewness statistic is negative for four of the series considered. This means that in most cases the distribution of those returns is skewed to the left. For all the series, the excess kurtosis statistics is above 3, implying that the distributions of those returns have much thicker tails than the normal distribution. Similarly, the Jarque-Bera statistic is significant rejecting the assumption of normality.

Table 2. Descriptive Statistics

Assets	Mean (%)	Median (%)	Maximum	Minimum	Std. Dev. (%)	Skewness	Kurtosis	Jarque Bera
ADP	0.013	0.023	0.11179	-0.26892	1.665	-1.148* (0.041)	26.678* (0.083)	83004 (0.001)
Amazon	0.042	-0.004	0.29618	-0.28457	3.681	0.450* (0.041)	13.170* (0.083)	15289 (0.001)
Apple	0.066	0.077	0.12239	-0.21999	2.733	-0.446* (0.041)	8.041* (0.083)	3844 (0.001)
Cerner	0.066	0.063	0.22063	-0.19159	2.746	0.441* (0.041)	11.995* (0.083)	11981 (0.001)
Microsoft	-0.035	0.000	0.16455	-0.18623	2.081	-0.497* (0.041)	12.480* (0.083)	13327 (0.001)
Telefónica	0.029	0.000	0.20096	-0.21279	2.474	-0.011* (0.041)	8.483* (0.083)	4410 (0.001)

Note: This table presents the descriptive statistics of the daily percentage returns of ADP, Amazon, Apple, Cerner, Microsoft and Telefónica. The sample period is from January 3rd 2000 to December 31th, 2013. The returns are calculated as Rt=ln(Pt/Pt-1). Standard error of the skewness an kurtosis coefficients are calculated as $\sqrt{(6/n)}$ and $\sqrt{(12/n)}$ respetively. The JB statistic is distributed as the Chi-square with two degrees of freedom; in brackets their p-value. * denotes significance at 1% level.

In Figure 1, we observe that the range of the fluctuation of the returns changes in the time and these variations evolve according to the idea of cluster in volatility (Mandelbrot, 1963). To capture this effect and others characteristics of the return as the leverage effect (Black, 1976) we use the beta-t-EGARCH model, proposed by Creal *et al.* (2008, 2011) and Harvey and Chakravarty (2008), and the beta-skewness-t-EGARCH proposed by and Harvey and Sucarrat (2013). In Table 3, we present the coefficient estimations of both models for each asset in the whole period.

All the parameters estimated are positive and statistically significant. Just in the case of Amazon k_2 is not significant. In addition, the estimations of φ_2 and φ_1 satisfy the identifiability and stationary conditions ($\varphi_2 < \varphi_1 \leq 1$). The parameter k^*, which capture the "leverage effect" is positive in all cases, indicating that volatility tends to be higher after negative returns. The value of this parameter moves between 0.015 (Microsoft) and 0.034 (Apple). To last, in the case of beta-skewness-t-EGARCH model, the parameter γ is inferior to one ($\gamma <1$) in the case of ADP, Cerner, Apple and Telefónica. This means that the distributions of those assets are skewed to the left. For the case of Amazon and Microsoft the distribution is skewed to the right ($\gamma >1$).

Table 3. Estimation of the parameters

		δ (se)	ϕ_1 (se)	ϕ_2 (se)	k_1 (se)	k_2 (se)	κ^* (se)	df (se)	γ (se)	LogL (BIC)
ADP	beta-t-EGARCH[(i)]	0.759 (0.253)	1.000 (0.000)	0.986 (0.006)	0.012 (0.006)	0.018 (0.007)	0.024 (0,004)	5.352 (0.458)	--	-5981.5 (3.416)
	beta-skewness-t-EGARCH[(ii)]	0.357 (0,143)	0.988 (0.009)	0.997 (0.001)	0.024 (0.011)	0.008 (0.011)	0.022 (0.003)	5.386 (0.464)	0.953 (0.020)	-5982.1 (3.418)
Amazon	beta-t-EGARCH	1.587 (0.234)	1.000 (0.000)	0.922 (0.051)	0.022 (0.006)	0.011 (0.009)	0.024 (0.007)	4.179 (0.269)	--	-8522.3 (4.860)
	beta-skewness-t-EGARCH	1.592 (0.236)	1.000 (0.000)	0.899 (0.052)	0.023 (0.005)	0.010 (0.009)	0.023 (0.007)	4.202 (0.270)	1.065 (0.023)	-8518.17 (4.860)
Apple	beta-t-EGARCH	1.435 (0,197)	1.000 (0.000)	0.882 (0.033)	0.015 (0.004)	0.025 (0.008)	0.034 (0.006)	6.136 (0.567)	--	-8013.3 (4.571)
	beta-skewness-t-EGARCH	1.432 (0.196)	1.000 (0.000)	0.885 (0.032)	0.015 (0.004)	0.026 (0.008)	0.034 (0.006)	6.115 (0.0566)	0.978 (0.022)	-8012.9 (4.573)
Cerner	beta-t-EGARCH	1.462 (0.221)	1.000 (0.000)	0.964 (0.012)	0.014 (0.004)	0.027 (0.008)	0.023 (0.004)	4.742 (0.354)	-	-7700.91 (4.393)
	beta-skewness-t-EGARCH	1.465 (0.221)	1.000 (0.000)	0.964 (0.012)	0.014 (0.004)	0.027 (0.008)	0.024 (0.005)	4.726 (0.354)	0.989 (0.021)	-7700.8 (4.395)
Microsoft	beta-t-EGARCH	0.826 (0.304)	1.000 (0.001)	0.971 (0.013)	0.019 (0.005)	0.030 (0.008)	0.015 (0.004)	4.780 (0.350)	--	-6675.8 (3.810)
	beta-skewness-t-EGARCH	0.319 (0.121)	0.943 (0.023)	0.996 (0.002)	0.032 (0.008)	0.020 (0.006)	0.015 (0.004)	4.830 (0.356)	1.004 (0.022)	-6672.8 (3.811)
Telefónica	beta-t-EGARCH	1.185 (0.192)	1.000 (0.000)	0.972 (0.008)	0.007 (0.003)	0.036 (0.006)	0.020 (0.004)	8.034 (0.927)	--	-7716.9 (4.402)
	beta-skewness-t-EGARCH	1.193 (0.191)	1.000 (0.000)	0.973 (0.008)	0.007 (0.003)	0.037 (0.006)	0.021 (0.004)	8.084 (0.945)	0.973 (0.022)	-7716.2 (4.404)

Note: The table reports the parameters estimates of the beta-t-EGARCH model (equation (2)) and beta-skewness-t-EGARCH model (equation (6)). (se) denotes the standard deviation (in parentheses). Log-L is the maximum likelihood value and BIC is the Bayesian Information Criterion. κ^* is the parameter that capture the "leverage effect". γ is the parameter that capture skewness in distribution. $\gamma < 1$, ($\gamma > 1$) denote skewed to the left (right).

(i) beta-t-EGARCH model

$$r_t = exp(\lambda_{t|t-1})(\varepsilon_t), \quad \varepsilon_t \sim t(0, \sigma_\varepsilon, v) \qquad v > 2$$
$$\lambda_{t|t-1} = \delta + \lambda^+_{1,t|t-1} + \lambda^+_{2,t|t-1}$$
$$\lambda^+_{1,t|t-1} = \varphi_1 \lambda^+_{1,t-1|t-2} + k_1 u_{t-1}$$
$$\lambda^+_{2,t|t-1} = \varphi_2 \lambda^+_{2,t-1|t-2} + k_2 u_{t-1} + k^*(-r_{t-1})(u_{t-1} + 1)$$

(ii) beta-skewed-t-EGARCH model

$$r_t = exp(\lambda_{t|t-1})(\varepsilon_t - \mu_\varepsilon), \quad \varepsilon_t \sim st(\mu_\varepsilon, \sigma_\varepsilon, v, \gamma) \qquad v > 2, \ \gamma \ \epsilon (0, \infty)$$

$$\lambda_{t|t-1} = \delta + \lambda^+_{1,t|t-1} + \lambda^+_{2,t|t-1}$$
$$\lambda^+_{1,t|t-1} = \varphi_1 \lambda^+_{1,t-1|t-2} + k_1 u_{t-1}$$
$$\lambda^+_{2,t|t-1} = \varphi_2 \lambda^+_{1,t-1|t-2} + k_2 u_{t-1} + k^*(-(r_{t-1}))(u_{t-1} + 1)$$

4.2 VaR Applications

In this section, the beta-t-EGARCH model and the beta-skewness-t-EGARCH model presented in section (2) are used to calculate the VaR one day ahead at 1% probability. The obtained estimations are compared. The analysis period, marked in dark in Figure 1, ran from January 1st, 2008, to December 31th, 2013 (N=1511).

The comparison of the VaR estimates has been conducted in terms of evaluating the accuracy of the VaR estimates and of the loss function. In Figure 2, we present the returns and the VaR estimates obtained from both volatility models and for all assets considered. As we can observe, the risk assumed by the companies varies along the sample being especially high in 2008-2009. The naked eye, it seems that there are no significant differences between the VaR estimation from both models

To evaluate the accuracy of the VaR estimates, several standard tests are used. The results of these tests are presented in Table 4. For each index, it is presented the number and the percentage of exceptions obtained with

each volatility model considered. The percentages of exceptions are marked in bold. Below the percentages, the statistics used to test the accuracy of the VaR estimates are presented. These statistics are as follows: (i) the unconditional coverage test (LRuc); (ii) backtesting criterion (BTC), (iii) statistics for serial independence (LRind); (iv) the conditional coverage test (LRcc) and (v) the Dynamic Quantile test (DQ). When the null hypothesis that "the VaR estimate is accurate" has not been rejected by any test, we shaded the region of the figure that represents the number and percentage of the exceptions.

The percentage of exceptions goes below 1% in almost all cases indicating that both models overestimate risk. Just in the case of apple the risk is underestimated. However, it is worth to note that the number and the percentage of the exceptions are close to the expected one. The accuracy tests used to test formally the performance of the volatility models in terms of VaR corroborate this hypothesis. These results indicate that both models provide accurate VaR estimates in all cases.

Table 4. Accuracy test

	Beta-t-EGARCH	Beta- skewness-t-EGARCH	Beta-t-EGARCH	Beta- skewness-t-EGARCH
	ADP		**Cerner**	
N ºexceptions	11	11	12	12
% exceptions	**0.73**	**0.73**	**0.79**	**0.79**
LR uc	0.46	0.46	0.58	0.58
BTC	0.23	0.23	0.29	0.29
LR ind	0.79	0.79	0.77	0.77
LR cc	0.74	0.74	0.82	0.82
DQ	0.98	0.08	0.14	0.14
	Amazon		**Microsoft**	
N ºexceptions	10	10	12	12
% exceptions	**0.66**	**0.66**	**0.79**	**0.79**
LR uc	0.35	0.35	0.58	0.58
BTC	0.17	0.17	0.29	0.29
LR ind	0.81	0.81	0.77	0.77
LR cc	0.63	0.63	0.82	0.82
DQ	0.99	0.98	1	1
	Apple		**Telefónica**	
N ºexceptions	18	16	14	17
% exceptions	**1.19**	**1.06**	**0.93**	**1.13**
LR uc	0.63	0.88	0.85	0.75
BTC	0.3	0.39	0.38	0.35
LR ind	0.66	0.7	0.74	0.68
LR cc	0.81	0.92	0.93	0.87
DQ	0.53	0.98	0.84	0.89

Note: VaR violation ratios of the daily returns (%) are boldfaced. The table reports the p-values of the following tests: (i) the unconditional coverage test (LRuc); (ii) the backtesting criterion (BTC); (iii) statistics for serial independence (LRind), (iv) the conditional coverage test (LRcc) and (v) Dynamic Quantile test (DQ). A p-value greater than 5% indicates that the forecasting ability of the VaR model is accurate. The shaded cells indicate that the null hypothesis that the VaR estimate is accurate is not rejected by any test.

Additionally, with the aim of detecting differences between both models (beta-t-EGARCH and beta-skewness-t-EGARCH) we follow Gerlach ,Chen and Chan (2011) and focus on analyzing the ratio VRate/α and some statistics of it. This ratio is calculated as the quotient of percentage exception by the value of α, which is 1%. The beta-t-EGARCH model provides a VRate/α more close to one for three assets (ADP, Cerner and Telefónica) the same the beta-skewness-t-EGARCH (ADP, Cerner and Apple) (Table 5). Table 6 displays summary statistics for VRate/α for each model across the 6 assets. The Std(1) column shows the standard deviation from expected ratio of 1 (not mean sample), while 1st column counts the assets where the model ranked had VRate/α closest to 1.

Table 5. Ratio Vrate/alfa at alfa=1% for each VaR model

	ADP	Amazon	Apple	Cerner	Microsoft	Telefónica
Beta-t-EGARCH	0.73	0.66	1.19	0.79	0.79	0.93
Beta-skewness-t-EGARCH	0.73	0.66	1.06	0.79	0.79	1.13

Note: Shaded cells indicate closest to 1 in that index.

Table 6. Summary statistics for the ratio Vrate/alfa at alfa=1%

	Mean	Median	Std(1)	1st
Beta-t-EGARCH	0.82	0.79	0.25	3
Beta-skewness-t-EGARCH	0.84	0.79	0.23	3

Note: Shaded cells indicate the most favoured in each column. Std (1) is the standard deviation in ratios from an expected value of 1. 1st indicates the number of markets where that model's VRate/α ratio ranked closest to 1.

According to these statistics the beta-skeweness-t-EGARCH model provides the better results as the mean of the ratio is closer to one and the std(1) is lower than those provide by the beta-t-EGARCH model.

Thus, although the evidence is weak the results indicate that in terms of accurate the beta-skewness-t- EGARCH model may outperform the beta-t-EGARCH which have been estimate below a symmetric distribution.

Another way to compare the VaR estimates that is often used in the VaR literature is through a loss function. The loss function measures the magnitude of the losses experienced. A model that minimises the total loss is preferred to other models. For this purpose, we have considered two loss functions: the regulator's loss function proposed by Lopez (1998, 1999) and the firm's loss function (ABL loss function) (Abad et al., 2015). The results of these loss functions are presented in tables 7[8].

Table 7. Loss functions

Panel (a): Lopez Loss Function (%)						
	ADP	Amazon	Apple	Cerner	Microsoft	Telefónica
Beta-t-EGARCH	0.0721	0.5856	3.1905	0.8017	1.3473	0.7175
Beta-skewness-t-EGARCH	0.0793	0.6757	3.1416	0.8002	1.2492	0.7367
Panel (b): ABL Loss Function (%)						
	ADP	Amazon	Apple	Cerner	Microsoft	Telefónica
Beta-t-EGARCH	0.0708	0.1312	0.10218	0.09079	0.09398	0.10219
Beta-skewness-t-EGARCH	0.07066	0.1301	0.102	0.09037	0.0938	0.09838

Note: The table reports the values of the different loss functions of each VaR model at 99% confidence levels. In both cases, the table shows the average of the losses. The shaded cells denote the minimum value for the different loss functions.

According to the regulator´s loss function there is no model superior to other. The beta-t-EGARCH model provides the lowest losses for ADP, Amazon and Telefónica while the beta-skewness-t-EGARCH model provides the lowest losses for Apple, Cerner and Microsoft. Thus, from the point of view of the regulator both models seem to be equivalent. However, according to the firm´s loss function, which takes the opportunity cost of capital into account, the beta-skewness-t-EGARCH model outperform the beta-t-EGARCH model by providing the lowest losses for all assets considered. Although in daily terms these differences are reduce in annual and monetary terms become high. For instance, for a portfolio value of 10 million of dollars the differences in annual capital opportunity cost provide for both models move around 3500 dollars (ADP) and 92250 dollars (Telefónica). Moreover, for a portfolio value of 100 millions of dollars these differences move around 350000 and 952500 dollars[9]. These data reflect that, although the differences in daily capital opportunity cost are small, in annual and monetary terms become high.

To last, we compare VaR estimates in terms of daily capital requirement (DCR) which have been calculated according to equation (21). The average of the DCR moves around 10% and 20% depending of the asset (Table 8). For almost all assets considered the Beta-t-skewness-EGARCH model provide the lowest daily capital requirement. The difference between these models moves around 0.01%-0.24% depending on the asset.

Table 8. Daily Requirement Capital

	ADP	Amazon	Apple	Cerner	Microsoft	Telefónica
Beta-t-EGARCH	11.70%	20.98%	16.78%	15.92%	15.10%	16.69%
Beta-skewness-t-EGARCH	11.82%	20.72%	16.57%	15.88%	15.11%	16.36%

Note: The table reports the average daily capital requirement (DCR) obtained according to equation 21. For each asset the shaded cells denote the model that provide the lowest average of the DCR

---

[8]In order to calculate the firm's loss function, we proxy the cost of capital with the interest rate of the Eurosystem monetary policy operations of the European Central Bank.

[9]The annual capital opportunity cost is calculated by multiplying the average of the daily opportunity cost by 250 days. And the average of the daily opportunity cost is calculated as the date included en Table 7 (panel(b)) divide by 100 times the portfolio value.

As a resume we can conclude that, in the framework of the conditional extreme value theory considering, a skewness *t*-Student distributions for the returns does not contribute to improve the accuracy VaR estimations respect to the symmetric *t*-Student distribution. However, the results obtained by the loss function indicate that this kind of distributions may be preferred by the financial companies, as they provide opportunity capital cost lowest. In addition, for the banks, using a skewness *t*-Student distribution take them to keep market risk capital requirement lowest.

5. Conclusion

It is well documented in the literature that the financial return distribution is characterized for being skewed and exhibit an important excess of kurtosis. So that to assume a normal distribution for VaR estimation may take us to underestimate risk.

Taking this into account, the research in the framework of parametric method for VaR estimation has focus in investigating other density functions that capture the skew and kurtosis of financial returns. In this line, recently papers show that in the context of this method assuming a fat tail and skewness distributions improve the results.

In this same line, we evaluate the role of the heavy tail and skewed distribution in VaR estimation in the framework of the conditional extreme value theory. Below the conditional extreme value theory, the value at risk of a portfolio at $(1 - \alpha)\%$ confidence level is calculated as the product of the conditional standard deviation of the return portfolio by the α (k_α) quantile of Pareto generalize distribution. Traditionally, the conditional standard deviation of the return portfolio is estimated by assuming a symmetric distribution for the financial return, normal or student-t distribution. Thus in this paper we analyze if in the framework of this method the estimation of the volatility model below a fat tail and skewness distribution contribute to improve the results in VaR estimation.

The study has been done for six individual assets bellowing to the telecommunication sector: ADP, Amazon, Cerner, Apple, Microsoft and Telefónica. The analysis period runs from January 1st 2008 to the end of December 2013. Although the evidence found is a little bit weak the results obtained seems to indicate that the heavy tail and skewed distribution outperform the symmetric distribution both in terms of accuracy VaR estimations as in terms of firm's loss function and requirement capital.
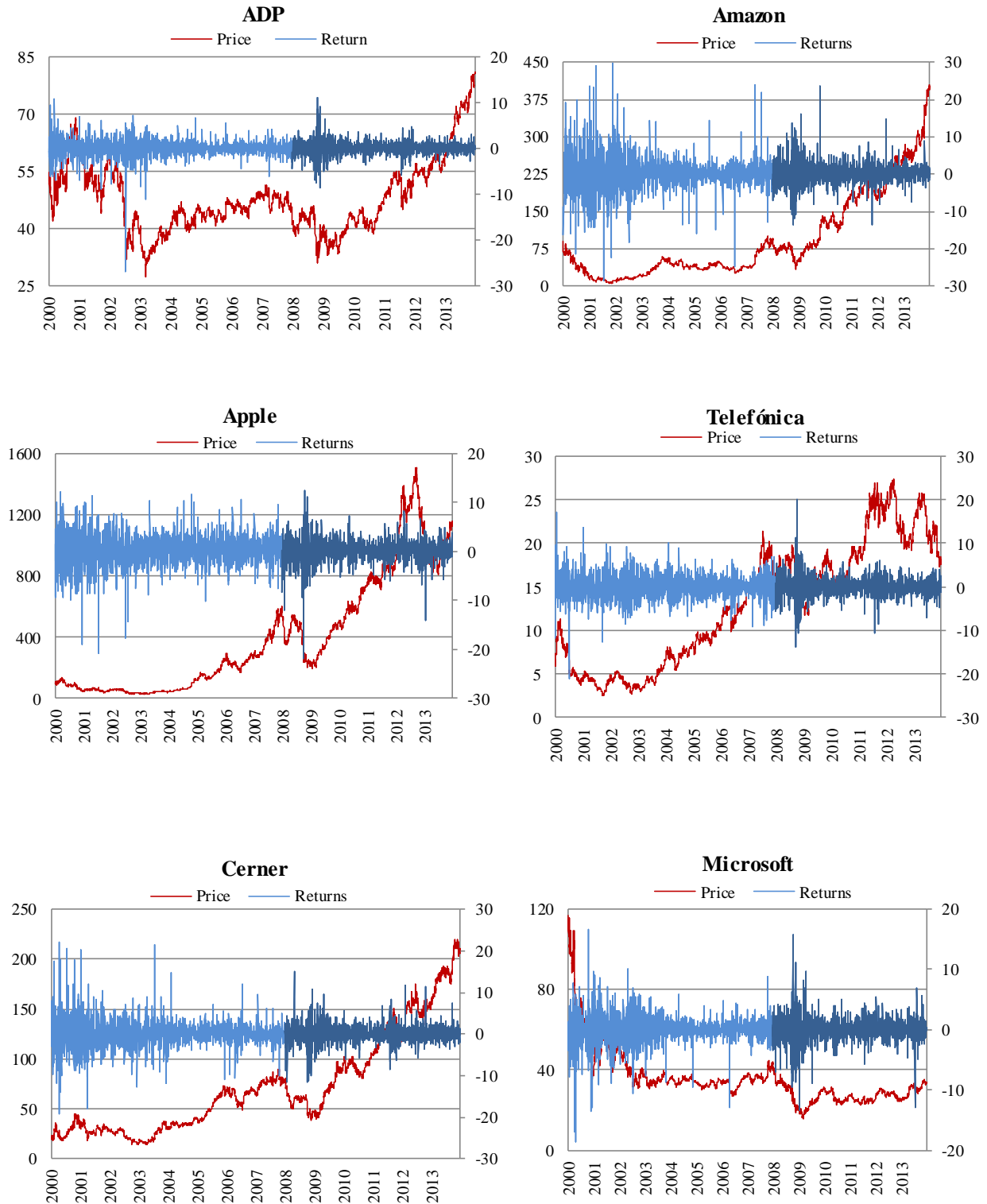
Figure 1. Evolution of daily prices and yields

Note: This figure illustrates the daily evolution of price (red) and returns (blue) of six assets (ADP, Amazon, Apple, Cerner, Microsoft and Telefónica) from January 3rd 2000 to December 31th, 2013.
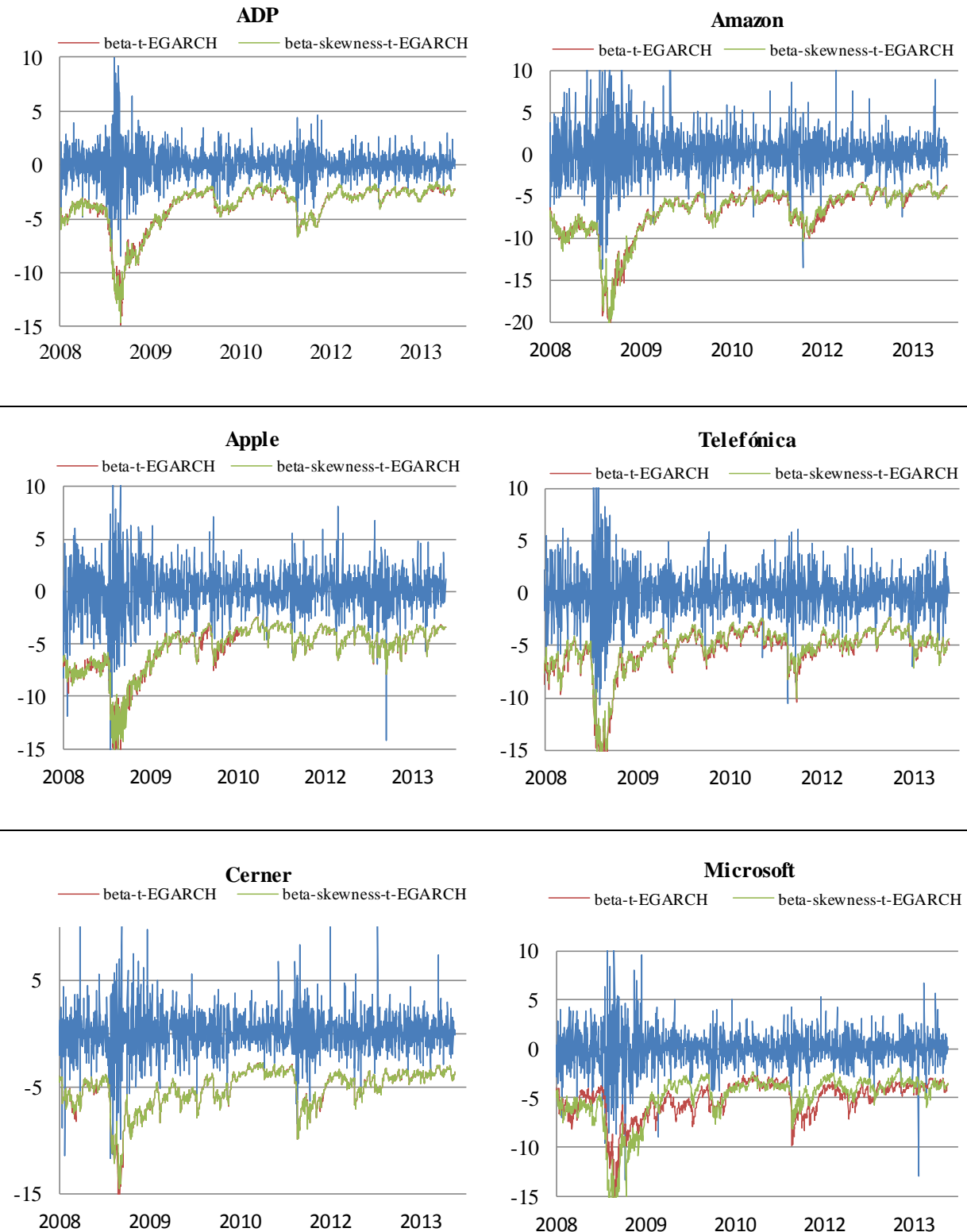
.

Figure 2. Returns and VaR estimations

Note: This figure illustrates the return (blue) and VaR estimations obtained from the beta-t-EGARCH model and the beta skewness-t- EGARCH model. The analysis period goes from January 1rd 2008 to December 31th, 2013.

References

Abad, P., Benito S., López-Martń, C., & Sánchez, M.A. (2016). Evaluating the performance of the skewed distributions to forecast Value at Risk in the Global Financial Crisis. *Journal of Risk*, *18*(5), 1-28. https://doi.org/10.21314/J0R.2016.332

Abad, P., Benito, S., & López-Martń, C. (2014). A comprehensive review of value at risk methodologies. *The Spanish Review of Financial Economic*, *12,* 15-32. https://doi.org/10.1016/j.srfe.2013.06.001

Abad, P., Benito, S., & López-Martń, C. (2015). Role of the loss function in the VaR comparison. *Journal of Risk Model Validation*, *9*(1), 1-19. https://doi.org/10.21314/JRMV.2015.132

Ausń, M., & Galeano, P. (2007). Bayesian estimation of the gaussian mixture GARCH Model. *Computational Statistics & Data Analysis*, *51*(5), 2636-2652. https://doi.org/10.1016/j.csda.2006.01.006

Bali, T., & Theodossiou, P. (2008). Risk measurement performance of alternative Distribution functions. *The Journal of Risk and Insurance*, *75*(2), 411-437. https://doi.org/10.1111/j.1539-6975.2008.00266.x

Basel Committee on Banking Supervision. (1996). Supervisory Framework for the use of backtesting in conjunction with the internal models approach to market risk capital requirements, Basel, Swizerland: BIS. [Online] Available at http://www.bis.org/publ/bcbs22.pdf

Basel Committee on Banking Supervision. (2006). International convergence of capital measurement andcapital standards: a revised framework - comprehensive version, Basel, Swizerland: BIS. [Online] Available at http://www.bis.org/publ/bcbs128.pdf

Black, F. (1976). Studies in stock price volatility changes. Proceedings of the 1976 business meeting of the business and economics statistics section. *American Association,* 177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *21,* 307-327. https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, *69,* 542-547. https://doi.org/10.2307/1925546

Chen, C., Gerlach, R., Lin, E. & Lee, W. (2012). Bayesian forecasting for financial risk management, pre and post the global financial crisis. *Journal of Forecasting, 31*(8), 661-687. https://doi.org/10.1002/for.1237

Christoffersen, P. (1998). Evaluating interval forecasting. *International Economic Review*, *39,* 841-862. https://doi.org/10.2307/2527341

Creal, D., Koopman, S. J., & Lucas, A. (2008). A general framework for observation driven time-varying parameter models, Tinbergen Institute Discussion Paper, TI 2008–108/4, Amsterdam.

Creal, D., Koopman, S. J., & Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics, 29,* 552-563. https://doi.org/10.1198/jbes.2011.10070

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1,* 83-106. https://doi.org/10.1016/0927-5398(93)90006-D

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, *50,* 987-1008. https://doi.org/10.2307/1912773

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

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, *22*(4), 367-381. https://doi.org/10.1198/073500104000000370

Fernandez, C., & Steel, M. F. J. (1996). On bayesian modelling of fat tails and skewness. https://doi.org/10.2139/ssrn.821

Gerlach R. H., Chen, C. W. S., & Chan, N. Y. C. (2011). Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business and Economics Statistics*, *29*(4), 481-492. https://doi.org/10.1198/jbes.2010.08203

Gerlach, R. H., Chen, C. W. S, & Chan, N. Y. C. (2011). Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business and Economics Statistics*, *29*(4), 481-492. https://doi.org/10.1198/jbes.2010.08203

Harvey, A., & Chakravarty, T. (2008). Beta-t-(E)GARCH, Cambridge Working Papers in Economics, Faculty of Economics, University of Cambridge.

Harvey, A., & Sucarrat, G. (2014). EGARCH models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis*, 320-338. https://doi.org/10.1016/j.csda.2013.09.022

Jorion, P. (2001). Value at risk: The new benchmark for managing financial risk. McGraw-Hill.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, *2*, 73-84. https://doi.org/10.3905/jod.1995.407942

Lee, G. G. J., & Engle, R. F. (1993). A permanent and transitory component model of stock return volatility. [Online] Available at SSRN: http://ssrn.com/abstract=5848

Lopez, J. A. (1998). Testing your risk tests. *Financial Survey*, 18-20.

Lopez, J. A. (1999). Methods for evaluating value-at-Risk estimates, *Federal Reserve Bank of San Francisco Economic Review*, *2*, 3-17.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, *36*, 394-419. https://doi.org/10.1086/294632

McAleer, M., Jimenez J. A., & Pérez-Amaral, T. (2013). International evidence on GFC-robust forecast for risk management under the Basel Accord. Econometric Institute Report EI 2011-04, Erasmus University Rotterdam, Econometric Institute. https://doi.org/10.1002/for.1269

McNeil, A. (1998). Calculating quantile risk measures for financial time series using extreme value theory. Department of Mathematics, ETS. Swiss Federal Technical University E-Collection. http://e-collection.ethbib.etchz.ch/

McNeil, A., & Frey, R. (2000). Estimation of tail-related risk measures for Heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, *7*, 271-300. https://doi.org/10.1016/S0927-5398(00)00012-8

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, *59*(2), 347-370. https://doi.org/10.2307/2938260

Polanski, A., & Stoja, E. (2010). Incorporating higher moments into Value-at-Risk forecasting. *Journal of Forecasting*, *29*(6), 523-535. https://doi.org/10.1002/for.1155

Sarma, M., Thomas, S., & Shah, A. (2003). Selection of value-at-risk models. *Journal of Forecasting, 22*(4), 337-358. https://doi.org/10.1002/for.868

Sucarrat, G. (2013). Betategarch: simulation, estimation and forecasting of Beta-Skew-t-EGARCH models. *The R Journal*, *5*(2), 137-147.

Zhang, M. H., & Cheng, Q. S. (2005). An approach to VaR for capital markets with gaussian mixture. *Applied Mathematics and Computations*, *168*, 1079-1085. https://doi.org/10.1016/j.amc.2004.10.004