UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 08-73
Statistics and Econometrics Series 26
December 2008

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

MEASURING FINANCIAL RISK: COMPARISON OF ALTERNATIVE PROCEDURES TO ESTIMATE VaR AND ES

María Rosa Nieto and Esther Ruiz

Abstract

We review several procedures for estimating and backtesting two of the most important measures of risk, the Value at Risk (VaR) and the Expected Shortfall (ES). The alternative estimators differ in the way the specify and estimate the conditional mean and variance and the conditional distribution of returns. The results are illustrated by estimating the VaR and ES of daily S&P500 returns.

Keywords: Backtesting, extreme value, GARCH models, leverage effect.

Departamento de Estadística. Universidad Carlos III de Madrid, C/ Madrid 126, 28903, Getafe, Madrid, Spain. e-mail addresses: mrnieto@eco.uc3m.es (María Rosa Nieto) and ortega@est-econ.uc3m.es (Esther Ruiz).

MEASURING FINANCIAL RISK: COMPARISON OF ALTERNATIVE PROCEDURES TO ESTIMATE VaR AND ES

Maria Rosa Nieto*and Esther Ruiz

December 2008

Abstract

We review several procedures for estimating and backtesting two of the most important measures of risk, the Value at Risk (VaR) and the Expected Shortfall (ES). The alternative estimators differ in the way the specify and estimate the conditional mean and variance and the conditional distribution of returns. The results are illustrated by estimating the VaR and ES of daily S&P500 returns.

Keywords: Backtesting, extreme value, GARCH models, leverage effect.

1 Introduction

This paper reviews and compares estimators of measures of financial risk. There are many definitions of financial risk because there are different groups of people interested in the money market and each group has its own attitude about risk; see Granger (2002). Here, we focus on market risk. In a simple situation, if we buy an asset at price P_{t-1} at time $t-1$, and sold it at price P_t at time t, we get a return calculated as the first difference of logarithm of prices, $R_t = \log(P_t) - \log(P_{t-1})$. In $t-1$, R_t is unknown and the risk is caused by this uncertainty. The return at time t can be considered unsatisfactory by the investor when it is negative or inferior to the return of some kind of governmental bond.

There are two main issues involved in estimating risk. First, one should consider measures of risk with adequate properties from a theoretical point of view. Second, once we have decided

*Corresponding author, Deparment of Statistics, Universidad Carlos III de Madrid, e-mail:mrnieto@est-econ.uc3m.es. Financial support from Project SEJ2006-03919 by the Spanish Government is gratefully acknowledged

how to measure risk, we should chose estimators of the corresponding measure with appropriate statistical properties.

There are different measures of market risk proposed in the literature. Luce (1981) suggested to measure risk by assigning different weights to the two halves of the distribution of returns. Therefore, if $f(R)$ is the density of returns, the risk is given by

$$R(f) = A_1 \int_0^\infty R_t^\theta f(R)\, dR_t + A_2 \int_{-\infty}^0 |R_t|^\theta f(R)\, dR_t \tag{1}$$

where $A_1, A_2 \geq 0$ and $\theta > 0$. Depending on whether $f(R)$ is the marginal density or the density of R_t conditional on past observations, we obtain marginal or conditional moments. In this paper, we consider conditional distributions of returns because it is a more efficient use of the information contained on the data. When the weights in (1) are equal, we have the class of volatility measures given by

$$V_\theta = \underset{t-1}{E}\left[|R_t - \mu_t|^\theta\right] \tag{2}$$

where $\mu_t = \underset{t-1}{E}[R_t]$ and the $t-1$ under the expectation means that it is taken conditional on the available information up to time $t-1$. V_θ includes the two most popular measures of risk, the variance, when $\theta = 2$, and the mean absolute deviation, when $\theta = 1$. However, only when the utility function is quadratic or the distribution of returns is Normal or log-Normal, the variance is an appropriate measure; see Tobin (1969), Tsiang (1972), Machina and Rothschild (1987) and Levy (1992). The assumption of normal conditional distribution could be adequate in many financial returns. However, the utility cannot be assumed to be quadratic as the investor has different attitudes depending on whether the returns are over or under their means. In that sense, there is uncertainty in the upper part of the distribution, but the risk only exists in the lower part, which means that the investors do not diversify for reducing the possibility of an unexpected positive return, just if it is negative; see Granger (2002). Therefore, measures based on V_θ are not in general adequate to measure risk.

One of the most popular alternative measures of risk is what is known as the Value at Risk (VaR). The VaR appears as a consequence of some adverse results along history which force the agencies that regulate financial activity, to look for a quantitative way to define the risk associated to a position in the market. The VaR is defined as the minimal potential loss that a portfolio can suffer in the $100\alpha\%$ worst cases with $\alpha \in (0,1)$, on some fixed time horizon. In particular, the VaR is given by

$$VaR_t^\alpha = -\sup\left[r \mid \underset{t-1}{P}[R_t \leq r] \leq \alpha\right] \tag{3}$$

Among the main advantages of the VaR are simplicity, wide applicability and universality; see Jorion (1990, 1997) and Embrechts et al. (2000). Consequently, since the 80's, the regulatory agencies have used the VaR to measure the risk of financial institutions. According to the Basel Committee on Banking Supervision, banks have to accomplish some requirements when calculating VaR. They require to compute the VaR for $\alpha = 0.01$ and for returns corresponding to 10 trading days. Furthermore, the VaR should be computed with observations corresponding to at least one year. However, the VaR has fundamental limitations from the point of view of its theoretical properties. The most important inconvenient is that the VaR of a diversified portfolio can be greater than the sum of the $VaR's$ of the individual portfolios; see Acerbi and Tasche (2002). Furthermore, the VaR does not measure losses exceeding itself. Consequently, we can have two distributions with heavy tails and the same VaR, but the losses that exceed VaR could be totally different; see Acerbi et al. (2001). Additionally, from the point of view of optimization problems, the VaR is not useful because it is not convex; see Szegö (2002).

As a result of the limitations of the VaR as a measure of risk, Artzner et al. (1997) defined what is known as Coherent Measures of Risk. A Coherent Measure of Risk must satisfy certain desirable properties[1]. The most distinctive of these properties is the Sub-additivity which implies that a portfolio which is made of sub-portfolios would have at most the same risk as the sum of the risks of sub-portfolios. Note that, as we mentioned above, the VaR is not sub-additive.

One Coherent Measure of Risk proposed by Artzner et al. (1999) is the Tail Conditional Expectation, also called Conditional Value at Risk ($CVaR$). The $CVaR$ measures the *expected loss* in the $100\alpha\%$ worst cases and is given by

$$CVaR_t^{\alpha} = -\underset{t-1}{E}\left\{R_t | R_t \leq -VaR_t^{\alpha})\right\}. \tag{4}$$

The $CVaR$ is a coherent measure of risk when it is restricted to continuous distributions. However, it can violate sub-additivity with non-continuous distributions. Consequently, Acerbi and Tasche (2002) proposed the Expected Shortfall (ES) as a coherent measure of risk. The ES is given by

$$ES_t^{\alpha} = CVaR_t^{\alpha} + (\lambda - 1)\left(CVaR_t^{\alpha} - VaR_t^{\alpha}\right) \tag{5}$$

where $\lambda \equiv \dfrac{\underset{t-1}{P}\left[R_t \leq -VaR_t^{\alpha}\right]}{\alpha} \geq 1$. Note that $CVaR = ES$ when the distribution of returns is continuous. However, the ES is still coherent when the distribution of returns is not continuous. The ES has also several advantages when compared with the more popular VaR. First of all, the ES is free of tail risk in the sense that it takes into account information about the tail of the underlying distribution. The use of a risk measure free of tail risk avoids extreme loss in the tail. Therefore, the ES is an excellent candidate for replacing VaR for financial risk management purposes. However, the effectiveness of ES depends on the stability of its estimation and the choice of efficient backtesting methods; see Fabozzi and Tunaru (2006).

Although we have already mentioned that the VaR has important theoretical limitations as a measure of risk, it is still the measure most extensively implemented by banks and financial

[1]The concept of Coherent Measure of Risk has become very popular between theorist. For example, Acerbi and Simonetti (2002) found a relation among the concept of coherent measure of risk and different economical parameters. Longin (2001), Acerbi and Tasche (2002) y Tasche (2002) have studied examples of coherent measures of risk. However, practitioners extensively use the VaR to measuring risk.

institutions. Therefore, there is a huge interest on its estimation. In this paper, we review several alternative estimators of VaR and its alternative ES. The advantages and disadvantages of the estimators considered are illustrated by implementing them to the estimation of the VaR and ES of a real time series of returns. We also revise and compare alternative methods to test for the adequacy of VaR and ES. The literature on the estimation of the VaR and ES is so large that it is unfeasible trying to cover all the available contributions. Consequently, our objective in this paper is to describe the main contributions updating other surveys previously published; see, for example, Manganelli and Engle (2001), Angelidis et al. (2005), Kuester et al. (2006) and Lima and Néri (2007). In this paper, we extend and update these surveys by providing a more comprehensive comparison of methods. We consider a larger number of: i) models for the conditional variance and ii) error distributions. Finally, we also compare several estimators proposed in the literature to estimate the ES.

This paper has been organized as follows. Section 2 describes several estimation methods for VaR and backtesting procedures to measure its adequacy. Section 3 is devoted to reviewing the estimation and backtesting methods for ES. Section 4 illustrates the estimation methods described in the two previous sections by implementing them to estimate the VaR and ES of a series of S&P500 index returns. Additionally, these procedures are compared through backtesting. Finally, Section 5 concludes the paper with the main conclusions and suggestions for further research.

2 Estimation and testing of Value at Risk (VaR)

This section describes some of the more popular methods to estimating the VaR, focusing on the weakness and strengths of each of them. The estimation of VaR is a difficult computational task due to, among other reasons, the complexity of financial instruments, the dimension of portfolio, the assessment of market probabilities, the approximations introduced to speed up computations and the statistical error on its estimation; see Ju and Pearson (1999), Acerbi et al. (2001), Longin (2001), Krause (2003), and Bao and Ullah (2004) among others. When measuring the risk of a portfolio, this portfolio can be considered as a multivariate system of individual returns or as a univariate return of the whole portfolio. In this paper, we focus on the estimation of the VaR of a univariate series of returns. Additionally, in this section, we describe some backtesting methods used to evaluate the performance of the VaR estimates.

2.1 Estimation Methods for VaR

The oldest and still very popular estimator of the VaR is based on Historical Simulation (HS). The VaR is estimated it as the αth quantile of the empirical distribution of losses, $\widehat{VaR}_t^{\alpha} = -R_{\omega:T}$, where $R_{\omega:T}$ is the ωth-order statistic of the data and $\omega = [T\alpha] = \max\{m \mid m \leq T\alpha, m \in \mathbb{N}\}$; see Acerbi and Tasche (2002). HS is simple and it does not assume any particular distribution of returns. However, it is based on assuming that returns are iid which is an empirically inadequate assumption. Furthermore, it is well known, that the empirical quantile is not an efficient estimator for extreme quantiles. In spite of these several limitations, several authors conclude that, in practice, HS could generate adequate estimates of the VaR depending on the length of the data and the VaR level, α; see, for example, Hendricks (1996) and Vlaar (2000) who obtains satisfactory results when $T = 2,550$ observations and $\alpha = 5\%$[2].

Another popular estimator of the VaR based on the *iid* assumption is based on boostrapping. To compute the VaR, B series of bootstrap returns $R^* = (R_1^*, ..., R_T^*)$, are drawn with replacement from the original series of returns, with each return having the same probability of being chosen. Then, the αth empirical quantile for each of the B replicates is calculated as in HS. Finally, the VaR is estimated as the average of these αth empirical quantiles; see Barone-Adesi and Giannopoulos (2001) for an illustrative example. Note that using this procedure, it is possible to obtain confidence intervals for the estimated VaR.

As we mentioned above, the *iid* assumption is not adequate for real daily returns. Consequently, there are many alternative estimators based on assuming particular specifications for the conditional distribution of returns. Consider the following model of returns

$$R_t = \mu_t + \epsilon_t \sigma_t \tag{6}$$

where μ_t and σ_t are the conditional mean and the conditional standard deviation of returns respectively, and $\{\epsilon_t\}$ are *iid* disturbances with zero mean and variance 1. Thus, the $100\alpha\%$ one-step ahead VaR conditional on information available at time $t-1$ is given by

$$VaR_t^{\alpha} = \mu_t + q_{\alpha} \sigma_t \tag{7}$$

where q_{α} is the $100\alpha\%$ quantile of $f(\epsilon_t)$, the density of standardized returns, ϵ_t.

In order to estimate the VaR in (7) one needs to specify and estimate the conditional mean and the conditional variance of returns and to assume a particular distribution for ϵ_t. Table 1 contains a summary of different assumptions on μ_t, σ_t and the distribution of ϵ_t. The first conclusion from this table is that the most popular assumption for the conditional mean of returns is to specify it as an $ARMA(p, q)$ model given by

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j} \tag{8}$$

where $a_t = R_t - \mu_t = \sigma_t \epsilon_t$; see McNeil and Frey (2000), Bali and Theodossiou (2007) and Kuester et al. (2006) among others. Furthermore, given that the dependence on the conditional mean of returns is usually very simple, most authors have represented it by $AR(1)$ or $MA(1)$ models. On the other hand, looking at the specifications of the conditional variance, Table 1 shows that many authors chose models within the $GARCH$ family. The simplest of these models is the

[2]Remember that the Basel Committee requires the supervision of the VaR for $\alpha = 0.01$

$GARCH\,(1,1)$ model of Bollerslev (1986) that is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{9}$$

where $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$, and $(\alpha_1 + \beta_1) < 1$; see Barone-Adesi et al. (1999), McNeil and Frey (2000), Nystrom and Skoglund (2002), Angelidis et al. (2005), Bali and Theodossiou (2007), Christoffersen and GonÇalves (2005), Giannopoulos and Tunaru (2005) and Kuester et al. (2006) among many others[3].

The basic $GARCH(1,1)$ model in (9) has been extended in several directions to cope with features of returns observed when analyzing real data. One of the most interesting of these features is the asymmetric response of volatility to positive and negative returns. The volatility is larger when past returns are negative than when they are positive; see Black (1976). This characteristic is known as leverage effect. Hentschel (1995) proposed the following specification of the conditional variance which nest several popular $GARCH$ specifications with leverage effect

$$\frac{\sigma_t^\delta - 1}{\delta} = \alpha_0 + \alpha_1 \sigma_{t-1}^\delta g^\nu\left(\epsilon_{t-1}\right) + \beta_1 \frac{\sigma_t^\delta - 1}{\delta} \tag{10}$$

where $g\left(\epsilon_t\right) = |\epsilon_t - b| - c\left(\epsilon_t - b\right)$. Model (10) encompasses many popular asymmetric models for volatilities. Among the most useful models implemented to estimate the VaR, one can find the Exponential $GARCH$ ($EGARCH$) of Nelson (1991) and the Asymmetric Power ARCH ($APARCH$) of Ding et al. (1993); for example, Angelidis et al. (2005) and Bali and Theodossiou (2007) conclude that the $EGARCH$ model has the best performance when estimating the VaR while Giot and Laurent (2003) fits the $APARCH$ model. The $EGARCH$ model is obtained when $\delta = 0$, $\nu = 1$ and $b = 0$. When $\delta = \nu$, $b = 0$ and $|c| \leq 1$, we obtain the $APARCH$ model. There are another two very popular models that can be obtained as particular cases of the $APARCH$ model. When the parameters are $\delta = \nu = 1$, the The Threshold $GARCH$ ($TGARCH$) of Zakoian (1994) is obtained and when $\delta = \nu = 2$, we obtain the GJR model of Glosten et al. (1993). Taking into account that the estimates of the power parameters are usually very close to 1, the results obtained from the $APARCH$ and $TGARCH$ models should be very similar.

The model has been considered by Angelidis et al. (2005) and Bali and Theodossiou (2007) as having the best performance when estimating VaR. Finally, the has been implemented by Giot and Laurent (2003) to estimate VaR and ES. can be obtained when $\delta = 1$, $\nu = 1$, $b = 0$ and $|c| \leq 1$. When $\delta = 2$, $\nu = 2$ and $b = 0$, is obtained.

The third component needed to compute the VaR in equation (7) is q_α which is obtained from the distribution of the standardized returns ϵ_t. There are two main alternatives to obtain the value of q_α. First, it is possible to assume a particular distribution for ϵ_t and, consequently, q_α will be the αth quantile of this distribution. Alternatively, several authors propose to avoid

[3]The Riskmetrics model for the conditional variance is the $GARCH(1,1)$ model in (9) with the restriction $\alpha + \beta = 1$; see see Longerstaey and More (1995) and Morgan (1995).

assuming a particular distribution for ϵ_t and to estimate directly the quantile. Within the first group of authors that assume a particular distribution for ϵ_t, the most popular one is Normality; see Morgan (1995), Bali and Theodossiou (2007) and Giot and Laurent (2003) among many others. However, it has often been observed that when the conditional variance is specified as a $GARCH$-type model, the distribution of ϵ_t has fat tails. Therefore, when estimating the VaR, several authors have proposed leptokurtic distributions of ϵ_t; see, for example, Pownall and Koedijk (1999), Mittnik and Paolella (2000), Manganelli and Engle (2001) and Angelidis et al. (2005). These authors generally assume that the distribution of ϵ_t is a standardized Student-ν or a GED distribution. Furthermore, to introduce skewness into the marginal distribution of returns several authors have proposed asymmetric conditional distributions of ϵ_t[4]. For example, Giot and Laurent (2003) propose the standardized skewed-Student distribution of Hansen (1994) given by

$$f\left(\epsilon_t|\xi,\nu\right)=\begin{cases}\dfrac{2}{\xi+\dfrac{1}{\xi}}sg\left[\xi\left(s\epsilon+m\right)|\nu\right] & \text{if} \quad \epsilon<-\frac{m}{s}\\ \dfrac{2}{\xi+\dfrac{1}{\xi}}sg\left[\left(s\epsilon+m\right)/\xi|\nu\right] & \text{if} \quad \epsilon\geq-\frac{m}{s}\end{cases} \tag{11}$$

where $g(.|\nu)$ is the standardized Student density with ν degrees of freedom, ξ is the coefficient of asymmetry, and m and s^2 are the mean and the variance of the non-standardized skewed Student given by $m=\dfrac{\Gamma\left(\dfrac{\nu-1}{2}\right)\sqrt{\nu-2}}{\sqrt{\pi}\Gamma\left(\dfrac{\nu}{2}\right)}\left(\xi-\dfrac{1}{\xi}\right)$ and $s^2=\left(\xi^2+\dfrac{1}{\xi^2}-1\right)-m^2$, respectively. When $\xi>0$, the density is skewed to the right while when $\xi<0$, it is skewed to the left. Lambert and Laurent (2000) show that the $100\alpha\%$ quantile of the standardized skewed-Student density is given by $q_\alpha=\dfrac{q_\alpha^*-m}{s}$, where q_α^* is the corresponding quantile of the skewed-Student density given by

$$q_\alpha^*=\begin{cases}\dfrac{1}{\xi}t_\alpha\left[\dfrac{\alpha}{2}\left(1+\xi^2\right)\right] & \text{if } \alpha<\dfrac{1}{1+\xi^2}\\ -\xi t_\alpha\left[\dfrac{1-\alpha}{2}\left(1+\xi^{-2}\right)\right] & \text{if } \alpha\geq\dfrac{1}{1+\xi^2}\end{cases}$$

and t_α is the $100\alpha\%$ quantile of the standardized Student-ν density. As an illustration, Figure 1 plots the skewed-Student distribution for different degrees of freedom and asymmetry parameter $\xi=0.75,-0.75$. For small values of ν the density is more peaked and it becomes flatter as long as it increases.

Another asymmetric distribution is the skewed-generalized-t (SGT) distribution proposed by Theodossiou (1998). The SGT distribution has the attractive of encompassing most of the distributions usually assumed for standardized returns. For example, the Normal, GED, Student-ν

[4]Alternatively, He et al. (2005) propose to introduce skewness in the marginal distribution of returns by assuming an asymmetric conditional mean.

and skewed-Student-ν distributions can be obtained as particular cases. However, in our experience, the maximization of the log-likelihood based on a SGT distribution is very complicated. Consequently, we will not consider further this distribution; see Bali and Theodossiou (2007) for an application of the SGT distribution in the estimation of the VaR.

As we mentioned above, instead of assuming a particular distribution for ϵ_t, several authors propose to estimate directly q_α. For example, Danielsson and de Vries (2000) and McNeil and Frey (2000) among others, use Extreme Value Theory (EVT) for the tails of the distribution of the standardized residuals. This procedure is based on taking into account that when the conditional mean and variance are correctly specified, the standardized residuals, $\widehat{\epsilon}_t = \dfrac{R_t - \widehat{\mu}_t}{\widehat{\sigma}_t}$ are *iid*. Then, they can be used to build the distribution function of the tail. Let F be the distribution of standardized returns. The excess distribution above the threshold u is given by $F_u(y) = P[X - u \leq y \mid X > u] = \dfrac{F(y+u) - F(u)}{1 - F(u)}$. Therefore,

$$1 - F(x) = (1 - F(u))(1 - F_u(x - u)). \tag{12}$$

The function $(1 - F(u))$ can be estimated by the proportion of observations over the threshold, i.e. by N/T, where N is the number of observations in the sample that exceed u. On the other hand, $1 - F_u(x - u)$ can be estimated by ML by assuming that the excess residuals over the threshold have a Generalized Pareto distribution (GPD) given by

$$G_{\xi,\beta}(y) = \begin{cases} 1 - (1 + \xi y/\beta)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-y/\beta) & \text{if } \xi = 0 \end{cases}$$

where β is the scale parameter, ξ is the shape parameter such that if $\xi > 0$ the distribution has heavy tails. The probability density function is

$$g_{\xi,\beta}(y) = \frac{1}{\beta}\left[1 + \frac{\xi y}{\beta}\right]^{-\dfrac{1+\xi}{\xi}}.$$

In practice, we fix the number of observations in the tail to be $N = k$, where $k << T$, obtaining a threshold at the $(k+1)\,th$ order statistic. Consequently, if $\widehat{\epsilon}_{(1)} \geq ... \geq \widehat{\epsilon}_{(T)}$ are the ordered residuals, the threshold is $\widehat{\epsilon}_{(k+1)}$ and the GPD is fitted to $\left(\widehat{\epsilon}_{(1)} - \widehat{\epsilon}_{(k+1)}, ..., \widehat{\epsilon}_{(k)} - \widehat{\epsilon}_{(k+1)}\right)$. Using (12) we get the following tail estimator for $x > u$

$$\widehat{F}(\widehat{\epsilon}_t) = 1 - \frac{k}{T}\left(1 + \widehat{\xi}\,\frac{\widehat{\epsilon}_t - \widehat{\epsilon}_{(k+1)}}{\widehat{\beta}}\right)^{-1/\widehat{\xi}}. \tag{13}$$

Finally, if $\alpha < \dfrac{k}{T}$, the quantile $(1 - \alpha)$ can be obtained from (13) as follows

$$\widehat{q}_{1-\alpha} = \widehat{\epsilon}_{(k+1)} + \frac{\widehat{\beta}}{\widehat{\xi}}\left(\left(\frac{\alpha}{k/T}\right)^{-\widehat{\xi}} - 1\right). \tag{14}$$

The use of the GPD for the excess residuals is just an example of a heavy-tailed distribution. Gnedenko (1943) characterized all such distributions with the following formula for $x > u$

$$1 - F(x) = x^{-1/\xi} L(x).$$

Applying this formula to the ordered residuals beyond the $(k+1)th$ order statistic and choosing $L(\widehat{\epsilon}) = \frac{k}{T}\left(\widehat{\epsilon}_{(k+1)}\right)^{1/\xi}$, the following tail distribution is obtained

$$F(\widehat{\epsilon}_t) = 1 - \frac{k}{T}\left(\frac{\widehat{\epsilon}_t}{\widehat{\epsilon}_{(k+1)}}\right)^{-1/\xi}.$$

In this case, the shape parameter ξ can be estimated using the estimator proposed by Hill (1975) that is given by

$$\widehat{\xi}^{(H)} = \frac{1}{k}\sum_{j=1}^{k} \log\left(\widehat{\epsilon}_{(k)}\right) - \log\left(\widehat{\epsilon}_{(k+1)}\right).$$

The estimation of the quantile is then

$$\widehat{q}_{1-\alpha} = \widehat{\epsilon}_{(k+1)}\left(\frac{\alpha}{k/T}\right)^{-\widehat{\xi}^{(H)}}. \tag{15}$$

One important issue of the Hill estimator is the choice of the number of observations in the tail, k. In this sense, McNeil and Frey (2000) show that the EVT method based on the GPD distribution gives more stable quantile estimates than the Hill estimator. To illustrate this point, Figure 2 plots Hill estimates of the 1% quantile of the S&P500 index observed from 29/08/1995 to 20/10/2005 for different values of k. This figure shows that, as expected, when the number of observations over the threshold is small, the Hill estimator of $q_{0.99}$ is very unstable. However, the estimator of $q_{0.99}$ is an increasing function of k. This figure also plots the estimates based on the GPD distribution. Once more, we observe that the estimator is very unstable for small k. However, when $k > 30$, the estimate of $q_{0.99}$ is approximately 2.6 regardless of k. Note that the same estimate is obtained by the Hill estimator when $30 < k < 250$. Only for very large values of k the Hill estimator generates estimates of $q_{0.99}$ well over 2.6. Therefore, if the number of observations in the tail is moderate, i.e. between 30 and 250, both estimators should give the same answer.

Chan and Xia (2007) derive the asymptotic distribution of the quantile estimator of McNeil and Frey (2000) in (14) without assuming a specific parametric distributional assumption on the distribution of ϵ_t, just that it is a heavy tailed distribution. Then, they propose two alternative methods to construct confidence intervals of the VaR. The first method is the traditional method based on the asymptotic Normality of the VaR estimator. Alternatively, they propose to construct the confidence interval by the tilting method of Hall and Yao (2003) and Peng and Qi (2003). Note, that the confidence intervals for the VaR constructed in this way do not incorporate the uncertainty due to the estimation of the parameters of the conditional mean and standard deviation.

Alternatively, the quantile q_α can be estimated using bootstrap methods that do not assume any particular distribution of the errors and incorporate the uncertainty of the estimated parameters; see Ruiz and Pascual (2002) for a review of the literature on using bootstrap procedures in financial time series and, in particular, for the estimation of the VaR. In particular, Hull and White (1998) and Barone-Adesi et al. (1999) propose a bootstrap method called Filtered Historical Simulation (FHS) based on using random drawns with replacement from the standardized residuals, $\{\epsilon_t^*\}$. The series of bootstrap residual returns are obtained as follows:

$$a_t^* = \epsilon_t^* \widehat{\sigma}_t^* \tag{16}$$

Then, the VaR is calculated as in the HS procedure by averaging the αth empirical quantiles of the B bootstrap replicates.

In this procedure, the estimated parameters are kept fixed in all bootstrap replicates. Thus, this procedure has been extended by Pascual et al. (2006) who propose to estimate the parameters in all bootstrap replicates in order to take into account the uncertainty due to parameter estimation.

Bootstrap procedures have the advantage of allowing to obtain confidence intervals for the estimated VaR and for the conditional volatility. For example, Christoffersen and GonÇalves (2005) implement bootstrap procedures to obtain confidence intervals for the VaR estimates obtained by HS and when the conditional volatility is assumed to follow a $GARCH(1,1)$ model and the distribution of the errors is Normal, Student, EVT, FHS or approximated by a Gram-Charlier or Cornish-Fisher expansion. They show that the confidence intervals for HS are too narrow and do not contain the true VaR value with the desire frequency while the methods that properly account for conditional variance dynamics imply confidence intervals with coverages close to the nominal. It is interesting to note that Christoffersen and GonÇalves (2005) use the bootstrap procedure proposed by Pascual et al. (2006) to obtain confidence intervals for the VaR estimated by other methods. However, the procedure of Pascual et al. (2006) can be directly implemented to obtain the VaR together with its confidence interval. Bootstrap procedures have also been implemented by Hartz and Paolella (2006) who additionally propose a bias-correction method for improving the VaR forecasting ability of the $Normal - GARCH$ model.

Semiparametric and nonparametric specifications of the conditional mean and variances have also been considered in the literature. For example, Fan and Gu (2003) introduce a semiparametric model to estimate the volatility using the geometric Brownian motion, a time-dependent diffusion model, as a discretization of the $IGARCH(1,1)$ model of Riskmetrics. In order to estimate the decay factor needed for the Riskmetrics methodology they propose two alternatives, one resulting in a data dependent decay factor which remains constant in the forecasting period, and the other adapts automatically to changes in stock price dynamics, adding flexibility to the first decay factor. Additionally, Fan and Gu (2003) propose a symmetric nonparametric estimation approach to estimate the quantiles of the standardized residuals. On the other hand, Martins-Filho and Yao (2006), based on the two stage approach of McNeil and Frey (2000), propose a nonparametric estimation procedure for the conditional mean and variance using the local linear estimator of Fan (1992). Furthermore, they propose a method based on L-Moment theory instead of the GPD used by McNeil and Frey (2000). These nonparametric methods are more difficult to estimate than the parametric procedures. However, there can be inferential gains when the assumptions of the parametric models are wrong. Another nonparametric procedure is the one developed by Chen and Tang (2005). They propose to calculate the VaR by implementing kernel smoothing on the empirical distribution of returns in such a way that the estimator of the VaR is a weighted average of the order statistics around $R_{\omega:T}$. They also emphasize the importance of the standard error of the VaR estimates and develop a procedure for its estimation based on a kernel estimation of the spectral density function of a series built using the smoother function. More recently, Cai and Wang (2008) developed an nonparametric estimator of the VaR and the ES by obtaining a weighted double kernel local linear estimator of the conditional distribution function. The proposed estimator is a combination of the weighted Nadaraya-Watson method of Cai (2002) and the double kernel local linear method of Yu and Jones (1998).

It is worth to notice that all the results obtained from the procedures described before depend on the specification of the mean and conditional variance and on the uncertainty due to parameter estimation.

Finally, there is another way to calculate the VaR by modelling directly the dynamic of the quantile over time. The Conditional Autoregressive Value at Risk ($CAViaR$) was introduced by Engle and Manganelli (2004) who propose the following equation for the VaR

$$VaR_t^{\alpha} = \beta_0 + \beta_1 VaR_{t-1}^{\alpha} + l\left(\beta_2, R_{t-1}, VaR_{t-1}^{\alpha}\right) \tag{17}$$

where different forms of the function l can be proposed. Some examples can be the asymmetric slope, $l(\cdot) = \beta_2 (R_{t-1})^{+} + \beta_3 (R_{t-1})^{-}$, where $(x)^{+} = \max(x, 0)$, $(x)^{-} = -\min(x, 0)$, and the adaptive, $l(\cdot) = \beta_2 \left\{ \left[1 + \exp\left(G\left[R_{t-1} + VaR_{t-1}^{\alpha}\right]\right)\right]^{-1} - \alpha \right\}$, where G is some positive finite number. The parameters of this model are estimated by the method of regression quantiles developed by Koenker and Bassett (1978). Manganelli and Engle (2001) also incorporate EVT to $CAViaR$. The procedure is the following: first, fit a $CAViaR$ model to get an estimation of the VaR for a large α, for example 10%, then construct the series of standardized quantile residuals as follows: $\dfrac{\widehat{\epsilon}_{t,\alpha}}{\widehat{VaR}_t^{\alpha}} = \left(\dfrac{a_t}{\widehat{VaR}_t^{\alpha}}\right) - 1$ and apply EVT to this series to get an estimation of the tail $\widehat{q}_p$ for $p < \alpha$. Then the VaR is calculated as

$$\widehat{VaR}_t^{p} = \widehat{VaR}_t^{\alpha}\left(1 + \widehat{q}_p\right).$$

Alternatively, DeRossi and Harvey (2006) propose to combine the approach of Engle and

Manganelli (2004) with signal extraction. The idea is to use some of the forms of the function l and approximate them to the filtered estimators of time-varying quantiles.

Other references where the dynamic dependence of the VaR is modelled are Chen and Chen (2005) and Gourieroux and Jasiak (2006). Gourieroux and Jasiak (2006) propose a dynamic adaptive quantile which improves the approach of Engle and Manganelli (2004) by taking into account a property inherent in the behavior of the quantiles, that is, the monotonicity of quantile estimators. This property ensures that the quantile is an increasing function of α. On the other hand, Chen and Chen (2005) make a comparison of the performance of the Riskmetrics approach, the $GARCH(1,1)$ model with the Normal and the Student-ν distributions for estimating VaR and the combination of them with quantile regression. The conclusions are that the quantile regression combined with the $GARCH(1,1)$ and the Student-ν distribution provides the best estimates.

2.2 Backtesting VaR estimates

In order to asses the accuracy of VaR estimates, the Basel Committee on Banking Supervision (1996a) and the amendments of Basel Committee on Banking Supervision (1996b) develop a statistical testing device denominated backtesting. According to their requirements, the backtesting should be based on 250 one step-ahead estimates of the VaR, i.e. estimates over one year. In this section, we review the most important backtesting procedures proposed in the literature. Backtesting is based on testing whether the VaR estimates are statistically accurate. When there are several alternative estimators of the VaR, one may want also to test which is the best among the ones that generate accurate estimates; see, for example, Sarma and Shah (2003) and Angelidis and Degiannakis (2007).

Backtesting procedures are based on the following failure process $\{I_t^\alpha\}$

$$I_t^\alpha = 1\left(R_t < -VaR_t^\alpha\right), \qquad\qquad t = T+1, ...T+n$$

where $1(.)$ is the indicator function, T is the size of the sample used to estimate the parameters of the model and n is the number of one step-ahead VaR's computed. A VaR estimator is accurate if and only if

$$\underset{t-1}{E}\left[I_t^\alpha\right] = \alpha. \tag{18}$$

Most backtesting procedures are based on testing some of the implications of this condition. The most popular backtesting procedure, proposed by Kupiec (1995), is based on the number of failures defined as $x = \sum_{T+1}^{T+n} I_t^\alpha$. which has a binomial distribution with parameters n and α. Kupiec (1995) proposes to test the null hypothesis $H_0 : E\left[I_t^\alpha\right] = \alpha$, using the following likelihood ratio statistic

$$LR_{uc} = 2\log\left[\left(1-\frac{x}{n}\right)^{n-x}\left(\frac{x}{n}\right)^{x}\right] - 2\log\left[(1-\alpha)^{n-x}\alpha^x\right]. \tag{19}$$

Under the null, the LR_{uc} test has asymptotically a $\chi^2_{(1)}$ distribution. It has low power when implemented with small samples. However, note that the null hypothesis is testing whether the unconditional expectation is α, which is not the hypothesis of interest in (18). Consequently, Christoffersen (1998) proposes a test of conditional coverage, where the null hypothesis is given by $H_0 : E\left[I_t^\alpha | I_{t-1}^\alpha\right] = \alpha$. This is equivalent to testing whether I_t^α are *iid* $Ber(\alpha)$ random variables against the alternative of first order Markov dependence. Note that this condition is necessary but not sufficient for the hypothesis in (18). This test considers whether the unconditional coverage is correct and adds a term to consider the serial independence of the failure process $\{I_t\}$. The serial independence term, LR_{ind}, is defined as follows

$$LR_{ind} = 2\log\left[(1-\pi_{01})^{n_{00}}\,\pi_{01}^{n_{01}}\,(1-\pi_{11})^{n_{10}}\,\pi_{11}^{n_{11}}\right] - 2\log\left[(1-\pi)^{n_{00}+n_{10}}\,\pi^{n_{01}n_{11}}\right]$$

where n_{ij} is the number of I_t^α observations with value i followed by an observation with value j, for $i, j = 0, 1$ and $\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}}$, $\pi_{11} = \frac{n_{11}}{n_{10}+n_{11}}$. Under the null hypothesis $\pi_{01} = \pi_{11} = \pi = \frac{n_{01}+n_{11}}{n}$ and the LR_{ind} statistic has an asymptotic $\chi^2_{(1)}$ distribution. Finally, the likelihood ratio for conditional coverage, LR_{cc} is defined as $LR_{cc} = LR_{uc} + LR_{ind}$ which has asymptotically a $\chi^2_{(2)}$ distribution under the null.

Recently, other tests for independence based on the autocovariances $Cov\left(I_t^\alpha, I_{t-j}^\alpha\right)$ have been proposed. For example, Berkowitz et al. (2006) discuss the following Portmanteau test

$$LB(m) = (n)(n+2)\sum_{j=1}^{m}(n-j)^{-1}r_j^2 \tag{20}$$

where r_j is the order j sample autocorrelation of $I_t^\alpha - \alpha$. Under the null $LB(m)$ is asymptotically $\chi^2_{(m)}$. On the other hand, Engle and Manganelli (2004) suggest a dynamic quantile (DQ) test obtained by regressing $I_t^\alpha - \alpha$ against its lagged variables and other values included in the conditioning set and testing whether these variables are significant.

All the methods described above are based on the assumption that the parameters of the models fitted to estimate the VaR are known. However, in practice, these parameters have to be estimated. Escanciano and Olmo (2008) show that the use of standard unconditional and independence backtesting procedures to asses VaR models in out-of-sample environments can be misleading. They quantify the risk associated with the estimation of the parameters in a very general class of dynamic parametric VaR models and propose a correction of the standard backtesting procedures that takes into account such a risk. They show that one of the main determinants of the corrected asymptotic variance is the forecasting scheme used to generate the forecasts of the VaR, i.e. whether one uses recursive, rolling or fix parameter estimates.

As we commented above, the backtesting procedures help to decide whether a particular procedure gives accurate estimates of the VaR. However, when several estimators are available, one wants to decide which estimator is best among those which are accurate. With this goal, Lopez (1999) proposes to choose the procedure that minimizes $C_m = \sum_{t=T+1}^{T+n} C_{m,t}$ where

$$C_{m,t} = \begin{cases} f\left(R_t, VaR^{\alpha}_{m,t}\right) & \text{if } R_t < VaR^{\alpha}_{m,t}, \\ g\left(R_t, VaR^{\alpha}_{m,t}\right) & \text{if } R_t \geq VaR^{\alpha}_{m,t}. \end{cases}$$

where the index m is used to represent the procedure m to estimate the VaR and $f(x, y)$ and $g(x, y)$ are functions such that $f(x, y) \geq g(x, y)$.

Different loss functions has been proposed in the literature; see Lopez (1999). Sarma and Shah (2003) and Angelidis and Degiannakis (2007) use the following Regulatory Loss function that is similar to the Quadratic Loss Function proposed by Lopez (1999)

$$C_{m,t} = \begin{cases} \left(R_t - VaR^{\alpha}_{m,t}\right)^2 & \text{if } R_t < VaR^{\alpha}_{m,t}, \\ 0 & \text{if } R_t \geq VaR^{\alpha}_{m,t}. \end{cases} \tag{21}$$

Alternatively, Angelidis et al. (2005) proposed the Quantile Loss function that additionally penalizes for higher than needed amount of capital and it is defined by

$$C_{m,t} = \begin{cases} \left(R_t - VaR^{\alpha}_{m,t}\right)^2 & \text{if } R_t < VaR^{\alpha}_{m,t}, \\ \left(R_{\omega:n} - VaR^{\alpha}_{m,t}\right)^2 & \text{if } R_t \geq VaR^{\alpha}_{m,t}. \end{cases} \tag{22}$$

Alternatively, Sarma and Shah (2003) propose a testing procedure that allows to measure the superiority between two models with respect to a certain loss function. The hypothesis are

$$H_0 : \{\theta = 0\} \; vs \; H_1 : \{\theta < 0\}$$

where θ is the median of the distribution of the loss differential between procedure i and procedure j, $z_t = C_{i,t} - C_{j,t}$.

The number of non-negative $z's$, S, is defined as $S_{ij} = \sum_{t=T+1}^{T+n} \psi_t$, where $\psi_t = 1\left(z_t \geq 0\right)$. If z_t is iid, under the null hypothesis the exact distribution of S_{ij} is binomial with parameters $(n, 0.5)$ and the asymptotic distribution of the standardized S_{ij} is a $N\left(0, 1\right)$

$$\frac{S_{ij} - 0.5n}{\sqrt{0.25n}} \overset{a}{\sim} N\left(0, 1\right);$$

see Diebold and Mariano (1995).

If H_0 is rejected, the model i is significantly better than model j for the chosen loss function. On the empirical application, the Diebold-Mariano statistic is obtained as the t-statistic of the regression of z_t on a constant using the Newey and West (1987) heteroskedasticity autocorrelation consistent (HAC) standard errors.

On the other hand, Angelidis and Degiannakis (2007) propose to compare alternative models using the test of superior predictive ability of Hansen (2005). The null hypothesis that the benchmark model ($m = 0$) is not inferior than the others is tested with the statistic

$$T_n^{SPA} = \max\left[\max_{m=1,\ldots,M} \frac{n^{1/2}\overline{z}_m}{\widehat{\omega}_m}, 0\right]$$

where $\widehat{\omega}_m^2$ is a consistent estimator of $\omega_m^2 = var\left(n^{1/2}\overline{z}_m\right), \overline{z}_m = n^{-1}\sum_{t=1}^{n} z_{m,t}$ and $z_{m,t} = C_{0,t} - C_{m,t}$. The estimation of ω_m^2 and the p-values of the T_n^{SPA} can be obtained using the stationary bootstrap of Politis and Romano (1994). The optimal block-size can be chosen by the block selection algorithm proposed by Politis and White (2004).

On the other hand, Giacomini and Komunjer (2005) and Bao et al. (2006) compare competing VaR forecasts using the predictive quantile loss function (PQL) based on the methodology of Koenker and Bassett (1978). The PQL function is given by

$$Q_\alpha = \frac{1}{n}\sum_{t=T+1}^{n}\left[\alpha - 1\left(R_t < VaR_t^\alpha\right)\right]\left[R_t - VaR_t^\alpha\right].$$

The selected model is the one that provides the VaR forecast with the minimun Q_α.

3 Estimation and testing of Expected Shortfall (ES)

This section describes different methods proposed in the literature for estimating ES. As we mentioned above, ES is a relatively new measure of risk, and consequently there are relatively few papers dealing with its estimation. Most of the papers actually estimate $CVaR$ instead of ES. Remember that the $CVaR$ only is coherent if the returns have a continuous probability distribution. However, in practice, the distribution of returns is often assumed to be continuous and, in this case, the $CVaR$ and the ES are equivalent; see Giannopoulos and Tunaru (2005). We also describe methods to evaluate the accuracy of the estimated ES.

3.1 Estimation

Acerbi and Tasche (2002) propose to estimate the ES using the VaR estimator based on Historical Simulation. In this case, the estimator is given by

$$\widehat{ES}_t^\alpha = -\overline{R}_{(\omega)}, \tag{23}$$

where $\overline{R}_{(\omega)} = \frac{\sum_{i=1}^{\omega} R_{i:T}}{\omega}$ is the average of the smallest $100\alpha\%$ returns. This estimator has a positive bias attributable to the negative biases of the order statistics. Consequently, Inui and Kijima (2005) has proposed an extrapolation method to adjust the bias and stabilize the estimator.

Later, several authors propose to estimate the ES as the average of observed returns beyond the VaR when the VaR has been estimated by one of the methods described in the previous section; see, for example, Giot and Laurent (2003) and Bali and Theodossiou (2007).

Alternatively, note that if returns are given by equation (6) then the ES is given by

$$ES_t^\alpha = \mu_t + \sigma_t \underset{t-1}{E} \left[\epsilon_t | \epsilon_t < q_\alpha\right]. \tag{24}$$

There are different alternative methods to calculate $E\left[\epsilon_t | \epsilon_t > q_\alpha\right]$. First, one can assume a particular distribution for the innovations and calculate analytically the corresponding expectation. If, for example, they are Normal, then $E\left[\epsilon_t | \epsilon_t > q_\alpha\right] = -\frac{\phi\left(\Phi_\alpha^{-1}\right)}{\alpha}$, where Φ_α^{-1} is the αth quantile of the standard Normal distribution. On the other hand, if the innovations are Student-ν, then

$$E\left[\epsilon_t | \epsilon_t > q_\alpha\right] = \frac{\nu - 2}{\alpha(1-\nu)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)\Gamma\left(\frac{\nu}{2}\right)}} \left(1 + \frac{q_\alpha^2}{\nu - 2}\right)^{\frac{1-\nu}{2}}$$

where q_α is the αth quantile of the Student-ν; see Christoffersen and GonÇalves (2005) for an empirical application. In the case of the GED and the Skewed-t distributions, the corresponding conditional expectations are estimated by Monte Carlo simulations.

Another procedure to estimate the ES is by using EVT in order to estimate the tail of the distribution of the standardized residuals and then, calculate the conditional expectation of the values beyond the quantile q_α; see McNeil and Frey (2000). In this case, if the excess residuals over the threshold u are assumed to follow a GPD distribution with parameters $\xi < 1$ and β, then the expected shortfall is estimated as follows

$$\widehat{ES}_t^\alpha = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1} \widehat{q}_{1-\alpha} \left(\frac{1}{1-\widehat{\xi}} + \frac{\widehat{\beta} - \widehat{\xi}\epsilon_{(k+1)}}{\left(1-\widehat{\xi}\right)\widehat{q}_{1-\alpha}} \right). \tag{25}$$

Alternatively, using the Hill estimator we can obtain the next estimation of the ES

$$\widehat{ES}_t^\alpha = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1} \frac{\widehat{q}_{1-\alpha}}{1-\widehat{\xi}^H}. \tag{26}$$

Another method for estimating ES is FHS as in Giannopoulos and Tunaru (2005). Once the bootstrap distribution of returns is obtained, the ES estimator is calculated as the sample average of the returns that exceed VaR. We can also obtain the bootstrap distribution of returns using the procedure develop by Pascual et al. (2006) and then estimate the ES as the average of the returns that exceed VaR. Christoffersen and GonÇalves (2005) implement the bootstrap procedure of Pascual et al. (2006) to obtain confidence intervals for alternative estimators of the ES. They show that ES measures are generally less accurate than VaR measures and that the confidence bands around ES are also less reliable. However, note that, as we mentioned above for the confidence intervals for the VaR, the bootstrap procedure of Pascual et al. (2006) can be directly implemented to estimate the ES and its confidence interval. Table 2 contains a summary of different assumptions on μ_t, σ_t and the distribution of ϵ_t for estimating the ES.

3.2 Backtesting ES estimates

In order to evaluate the adequacy of the estimated ES, Angelidis and Degiannakis (2007) propose a two stage evaluation framework that extends the evaluation approach of Lopez (1999). In the first stage, the correct conditional coverage of the VaR forecast is tested using the LR_{cc} statistic defined above. In the second stage, the loss function is calculated with respect of the ES instead of the VaR, because the VaR does not give information about the size of the expected loss. The loss function is then

$$C_{m,t} = \begin{cases} \left(R_t - ES^{\alpha}_{m,t}\right)^2 & \text{if } R_t < VaR^{\alpha}_t \\ 0 & \text{if } R_t \geq VaR^{\alpha}_t \end{cases}, \tag{27}$$

where the subindex m refers to the model m. For each model the mean squared error is calculated by $MSE = \dfrac{1}{n}\sum_{t=T+1}^{T+n} C_{m,t}$. We choose the model that minimizes the MSE.

However, in this case we also have the ambiguity about the interpretation of the MSE. In order to overcome this problem, alternative models can be tested by using the test of superior predictive ability of Hansen (2005).

Note that this procedure uses a loss function that involves the ES for testing superiority among the models that provide accurate VaR forecasts, but it is not a method to test the accuracy of the ES forecasts.

4 Empirical Application

In this section, we implement the methods described above to estimate the VaR and ES of a series of daily returns of the S&P500 index observed from 29/08/1995 to 20/10/2005. The series of returns, plotted in the first row of Figure 3, have volatility clustering. Table 3, that reports some descriptive statistics, shows that the returns have excess of kurtosis and skewness. This table also reports the ratio between the Box-Pierce statistic and its corresponding 5% critical value for testing whether the first 20 autocorrelations are jointly equal to zero. We can observe that the ratio is smaller than one and consequently, the null is not rejected. The second row of Figure 3 plots the correlogram of the series of returns together with their 95% confidence bands computed as suggested by Diebold(1988) to account for the presence of conditional heterocedasticity. The dependence in S&P500 returns seems to be well represented by a white noise. Table 3 also reports the ratio of the Rodriguez and Ruiz (2005) statistic for testing whether the first 20 autocorrelations of absolute returns are jointly equal to zero[5]. In this case, the null is

[5]Rodriguez and Ruiz (2005) propose to test for conditional homoscedasticity by using the following statistic $Q^*_i(M) = T\sum_{k=1}^{M-i}\left[\sum_{l=0}^{i}\widetilde{r}(k+l)\right]^2$ where $\widetilde{r}(k+l)$ is the standardized sample autocorrelation of order $k+l$ of absolute returns, $M = [M/3]-1$ is the number of autocorrelations and T is the sample size. This statistics is more powerful than the more popular McLeod and Li (1983) test because it takes into account that under the null the sample autocorrelations have to be equal to zero and mutually uncorrelated.

clearly rejected. This result is in concordance with the correlogram of absolute returns plotted in the third row of Figure 3. The sample correlations of absolute returns are positive and highly persistent, being significantly different from zero even for very long lags. Therefore, S&P500 returns could be conditionally heteroscedastic, possibly with long-memory. Figure 3 also plots the cross-correlogram between returns and squared returns, $Corr\left(y_t, y_{t+h}^2\right)$, $h = ..., -2, -1, 0, 1, 2...$ These cross-correlations suggest that the volatility of S&P500 seems to have a leverage effect. The evidence of an effect of the volatility in the conditional mean of returns is much weaker given that the previous correlations are not significant when $h < 0$.

Therefore, the S&P500 returns seems to be conditionally heteroscedastic with the volatility being larger when the returns are negative than when they are positive. Consequently, we fit several $GARCH$ type models with leverage effect, each of them with alternative assumptions on the error distribution. In particular, we consider the $GARCH$, $TGARCH$, GJR, $EGARCH$ and $APARCH$ models. The distribution of the errors has been assumed Normal, Student-ν, GED and Skewed-t. Table 4 reports the maximum likelihood estimates of the parameters[6]. The first conclusion from this table is that for all combinations between models and distributions considered, the $ARCH$ and the asymmetry parameters are significant. The asymmetry parameter of the Skewed-t distribution is significant in all the models considered and its estimated value is always around -0.11. However, the estimated degrees of freedom of the Student-ν and of the Asymmetric Student-ν distributions are always estimated larger than 10. Therefore, it seems that it is more important the asymmetry than the excess kurtosis of the errors. Figure 4 plots the estimated kernel densities of the standardized residuals when the $GARCH$ and $APARCH$ models are fitted together with each of the densities estimated. The results for all other models considered are very similar and not plotted to save space.

On the other hand, the estimates of the leverage effect parameter are significant in all the models considered. Finally, the estimates of the power parameter in the $APARCH$ model is rather close to one. Therefore, looking at these results, it seems that the $APARCH$ model with Skewed-t errors should provide the best fit.

Table 5 reports diagnostics on all the estimated models. In particular we report the skewness, kurtosis, ratio of the $Q^*_{[20/3]-1}(20)$ and $Q(20)$ statistics and the correlation of order one between standardized residuals and future squared standardized residuals. All the models are successful in explaining the autocorrelations of absolute values which are not any longer significant. On the other hand, the cross-correlations still significant. Additionally, when the distribution of the errors is assumed to be symmetric, the skewness of the standardized residuals is still different from zero.

[6]These estimations were obtained by Matlab codes written by the first author.

4.1 Estimates of the VaR

For each of the models estimated, the $VaR_{T+1}^{0.01}$ has been computed by assuming that the conditional mean is zero and the conditional variance and error distribution are those reported in Table 4[7]. The VaR has also been estimated by assuming that the conditional variance is the one estimated by assuming Normal errors and then, the distribution of the errors estimated by bootstraping and by the Hill and GPD procedures. Finally, we also estimate the VaR using HS and the asymmetric and adaptive versions of the $CAViaR$ model. For the EVT method of McNeil we compute the maximum likelihood estimates of the GPD distribution which are given by $\beta = 0.51$ and $\xi = 0.17$.

Finally, we backtest the models using the procedures described before. Table 6 reports the results of the backtesting tests. The expected number of failures, x, if we have $1,000$ VaR forecasts at the 1% confidence level should be 10 and the nonrejection region for the number of failures is $4 < x < 16$. If x belongs to this interval, then the model correctly measures the VaR. According to the test of correct conditional coverage the models that give accurate forecasts are the asymmetric specification of the $CAViaR$ model, the $GARCH$ model with the $Normal$ distribution and *Bootstrap*, the $TGARCH$, GJR and $EGARCH$ for all the distributions, and the $APARCH$ model for all the distributions except the $Student - \nu$. Results of the DQ test of Engle and Manganelli (2004) and the $LB(5)$ test of Berkowitz et al. (2006) are also presented on Table 6, note that under these tests all the models are accurate except HS. It is important to mention that the conclusions using the $LB(m)$ test can change depending on the choice of the number of lags included in the test. In order to exemplify the previous statement we calculate also the $LB(20)$ test. We observe that, in this case, the asymmetric $CAViaR$ is rejected as a model that provides accurate VaR forecast.

Figure 5 represents scatter-plots of the VaR estimated assuming a $Normal$ distribution with the five models considered. We can observe that the estimated VaR obtained with the $TGARCH$ and $APARCH$ models are identical. The estimates obtained with the $EGARCH$ and GJR models are also similar. The only model that generates estimates clearly different from all others is the $GARCH$ that, as we have seen in Table 6 is rejected by the Christoffersen (1998) test. The results for all the other distributions considered in this paper are similar and are not reported to save space.

In Figure 6 we also represent the VaR's estimated for a given model with different assumptions on the error distribution. In particular, we chose the $EGARCH$ model. First of all, comparing Figures 5 and 6, we can observe that the differences among estimated VaR's are larger when the distribution is fixed and the model changes than when the model is fixed and the distribution changes. Therefore, it seems that it is more important to choose correctly the model and that

[7]Results for other values of α are available from the authors upon request. The conclusions may change depending on α. We report in this paper the results for $\alpha = 1\%$ because this is the value required by the Basel Committee on Banking Supervision.

as far as it is not $Normal$, the particular distribution chosen only has a marginal effect on the estimates of the VaR. Figure 6 also shows that computing the VaR using the $Hill$ or the GPD extreme value estimators provide nearly identical results.

Then, we calculate the regulatory loss function and the quantile loss function in (21) and (22) only for the models that present correct conditional coverage. Table 7 reports these results. If we only based our conclusions on the model that minimizes these loss functions, then we choose, using the regulatory loss function the $GJRGARCH - Skewed\, t$. On the other hand, the quantile loss function provides the best results with the $EGARCH - Normal$ and the $EGARCH - GED$ model.

We also implement the two procedures explained above, the test of Diebold and Mariano (1995) and the test of Hansen (2005). Table 7 shows the results obtained when applying the test of Hansen (2005) using the regulatory loss function and the quantile loss function. According to these results, we can choose the models that are not rejected to overcome its competitors. With the regulatory loss function the selected models are the $TGARCH$, GJR and the $APARCH$ model with the $Skewed - t$ distribution and the $GARCH$ model with the $Normal$ distribution. On the other hand, with the quantile loss function the $EGARCH$ model with the GED, $Normal$ and $Skewed - t$ distributions and the $GARCH - Normal$ are selected.

On Table 9 are presented the results of the test of Diebold and Mariano (1995) with the regulatory loss function. Here, the models that produce accurate VaR forecasts are compared by pairs in such a way that the superiority between the models selected above can be tested. In each case, the null hypothesis of equal predictive accuracy is not rejected, concluding that all these models can be used in order to obtain optimal VaR forecasts.

Finally, with respect to the quantile loss function, Table 9 shows the results for the test of Diebold and Mariano (1995) concluding that in the case of the $GARCH$ model with the $Normal$ distribution, the $EGARCH$ with the $Skewed - t$ distribution and the $EGARCH - Normal$ model, the null hypothesis is not rejected, which implies that comparing with each other, there are not superiority among them when predicting the VaR. On the other hand, the $EGARCH - GED$ is not superior compared with the $GARCH - Normal$ and the $EGARCH - GED$ model, but comparing with the $EGARCH$ with the $Skewed - t$ distribution, the null hypothesis is rejected and the value of the t-statistic is negative. Therefore, the $EGARCH - GED$ model is better than the $EGARCH - Skewed - t$ for predicting the VaR.

4.2 Estimates of the ES

The parametric $ES_{T+1}^{0.01}$ is calculated in two different ways, one is making the average of the returns beyond the parametric VaR and the other is calculating the expected value of the returns beyond the quantile using the assumed distribution. The Bootstrap ES is calculated using the predictive distribution of returns, and the $EVT - ES$ is calculated using equations (24) and (25).

Figure 7 represent scatter-plots of the estimated ES when the error distribution is $Normal$ and the expected value of the returns under the VaR is estimated assuming this distribution. This figure ahows that only when the specification of the volatility is assumed to be $GARCH$, the estimates of the ES are clearly different fron the others. We expect this result to be also satisfied when other error distribution are considered. It seems that, as in the VaR, estimate the adequate election of the specification of the variance is more important than the error distribution.

For the backtesting, the 1,000 parametric ES forecasts were calculated as the average of the returns beyond the parametric VaR. The two stage procedure of Angelidis and Degiannakis (2007) is used for backtest the ES. In the first stage, the models that produce accurate VaR forecast are selected. In the second stage, the loss function is calculated using equation (27) and then the test of Hansen (2005) is used for evaluating the models. These results are presented in Table 8. Then, we conclude that the selected models are the GJR model using the Normal and the $EVT-GPD$ distribution, the $EGARCH$ with the $Normal$ distribution and the $APARCH$ model using Bootstrap.

Analogous to the case of the VaR, Table 9 shows the results of the test of Diebold and Mariano (1995). The null hypothesis is not rejected when comparing the four models previously selected. Therefore, any of them can be used in order to estimate accurate VaR and ES.

5 Conclusions

In this paper, we describe and implement several procedures developed in the literature for estimating VaR and ES and also, for backtesting them. According to the requirements of the Basel Committee, the risk measures were calculated at the 1% confidence level instead of, as it is usual in the literature, at the 5% level.

The first conclusion is that the HS model is clearly rejected using all the test for measuring the accuracy of the VaR forecast. On the other hand, the rest of the models are not rejected except for the $GARCH$ model using the Christoffersen (1998) test. Although, using the $LB(m)$ test some other models are rejected because of its sensibility to the choice of the number of lags m.

Our results show that different backtesting procedures may lead to different conclusions on the accuracy of VaR and ES estimates. In general, ir seems that, for the data set considered in this paper, the Christoffersen test is the less conservative.

On the other hand, when looking at the most appropiate parametrization of the model used to estimate the VaR, our results suggest that choosing adequately the specification of the conditional variance is more important than choosing the error distribution as far as this distribution is not Normal.

Therefore, we can save computational time because, in some cases, it is hard to compute the likelihood of certain distributions.

Alternatively to the procedures presented in this paper, the volatility in (7) can also be represented by Stochastic Volatility (SV) models. One example is the log-normal SV model given by

$$R_t = \epsilon_t \exp\left(h_t/2\right)$$
$$h_t = \alpha + \beta h_{t-1} + \eta_t$$

where $h_t = \log\left(\sigma_t^2\right)$, ϵ_t and η_t are mutually independent and η_t are iid disturbances with zero mean and variance σ_η^2. Another example of a SV model is the following Markov switching model proposed by Hamilton (1989) given by

$$R_t = \epsilon_t \exp\left(h_t/2\right)$$
$$h_t = \alpha + \beta s_t$$

where s_t is a two-state first-order Markov chain which can take values $0, 1$. Some applications of these models to the estimation of the VaR can be found in Billio and Pelizzon (2000), Billio and Sartore (2003) and Sadorsky (2005). Another application is the one proposed by Eberlein et al. (2003). They combine the SV models with the hyperbolic distribution introduced by Eberlein and Keller (1995). One line of future research is to combine and compare SV models with the alternatives used in this work.

Figure 1. Skewed Student density function for different degrees of freedom and asymmetry parameter (a) $\xi = 0.75$ and (b) $\xi = -0.75$.

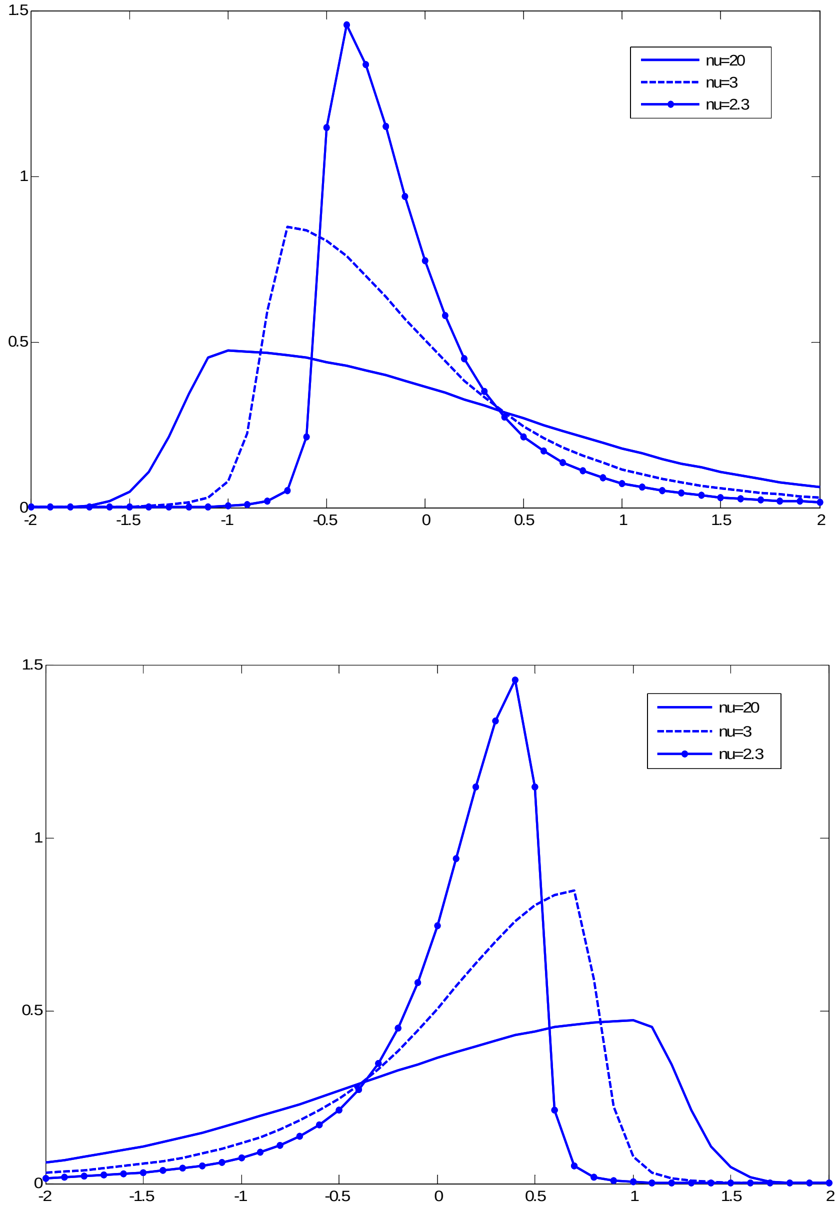
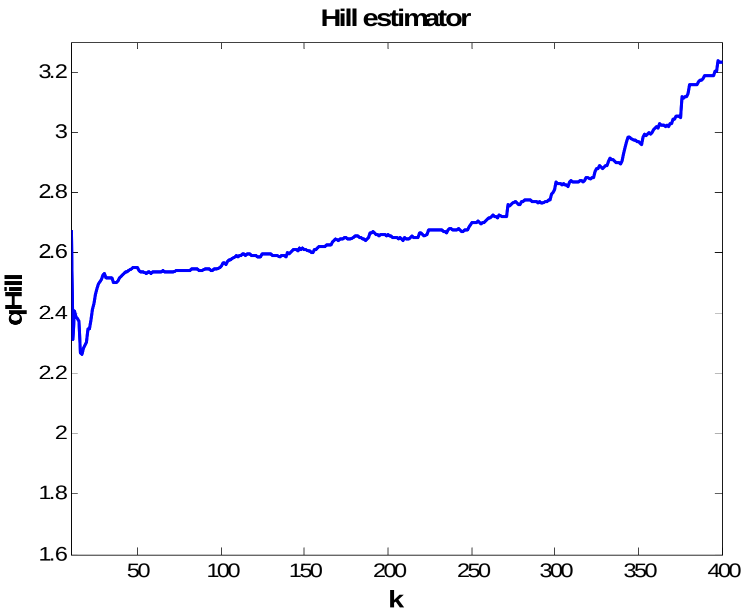


Figure 2. Hill and GPD quantile estimators for SP&500 returns and $\alpha = 1\%$



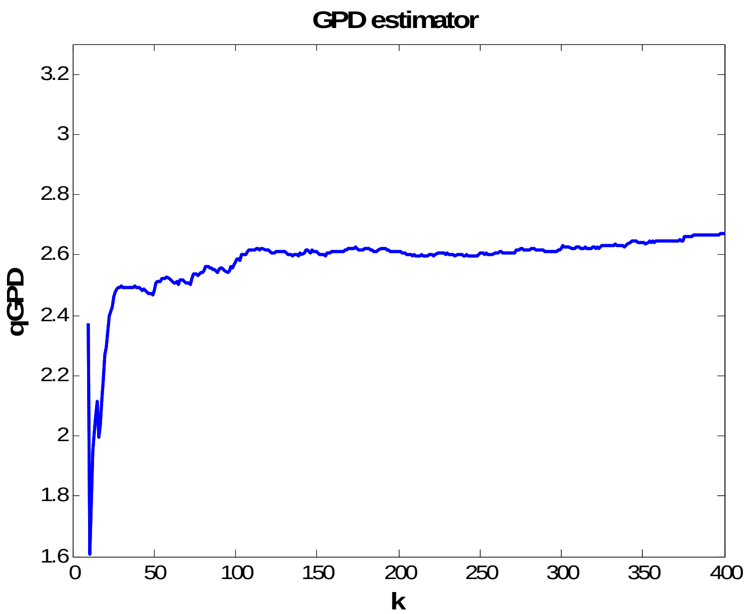


Table 1. Summary of the models proposed to estimate the VaR

References	Conditional Mean	Conditional Variance	Distribution of ϵ_t
Morgan (1995)		$IGARCH\,(1,1)$	Normal
Barone-Adesi et al. (1999)	$ARMA\,(1,1)$	$AGARCH(1,1)$	Bootstrap
Billio and Pelizzon (2000)		SV	Switching Regime
Danielsson and de Vries (2000)		$GARCH(1,1)$	EVT-Hill
McNeil and Frey (2000)	$AR\,(1)$	$GARCH\,(1,1)$	EVT-GDP
Mittnik and Paolella (2000)	$AR\,(1)$	$GARCH(1,1)$, $APARCH\,(1,1)$	Normal, Student
Barone-Adesi and Giannopoulos (2001)	$MA\,(1)$	$AGARCH(1,1)$	Bootstrap
Nystrom and Skoglund (2002)	$ARMA(p,q)$	$GJRGARCH(1,1)$	EVT-GDP, EVT-Hill
Billio and Sartore (2003)		SV	Bootstrap
Eberlein et al. (2003)		SV	Hyperbolic
Giot and Laurent (2003)	$AR\,(p)$	$IGARCH\,(1,1)$, $APARCH\,(1,1)$	Normal, Student, Skewed-t
Angelidis et al. (2005)	$AR\,(1)$	$GARCH(1,1)$, $GJRGARCH(1,1)$, $EGARCH(1,1)$	Normal, Student, GED
Christoffersen and GonÇalves (2005)		$GARCH\,(1,1)$	Normal, Student, EVT-Hill, Bootstrap
Giannopoulos and Tunaru (2005)	$MA\,(1)$	$AGARCH(1,1)$	Bootstrap
Sadorsky (2005)		SV	Normal
Hartz and Paolella (2006)	$ARMA(p,q)$	$GARCH\,(1,1)$	Bootstrap
Bali and Theodossiou (2007)	$AR\,(1)$	All $GARCH$ models included in (12)	Normal, SGT, GED
Engle and Manganelli (2004)	Conditional Quantile		
Chen and Chen (2005)	Dynamic Adaptive Quantile		
Gourieroux and Jasiak (2006)	Quantile Regression		

Table 2. Summary of the models proposed to estimate the ES

References	Conditional Mean	Conditional Variance	Distribution of ϵ_t
McNeil and Frey (2000)	$AR\,(1)$	$GARCH\,(1,1)$	EVT-GDP
Giot and Laurent (2003)	$AR\,(p)$	$IGARCH\,(1,1)$, $APARCH\,(1,1)$	Historical Simulation
Christoffersen and GonÇalves (2005)		$GARCH\,(1,1)$	Normal, Student, EVT-Hill
Giannopoulos and Tunaru (2005)	$MA\,(1)$	$AGARCH(1,1)$	Bootstrap
Bali and Theodossiou (2007)	$AR\,(1)$	All $GARCH$ models included in (12)	Historical Simulation

Figure 3. S&P500 returns, correlograms of returns, absolute returns and cross-correlogram of returns and squared returns

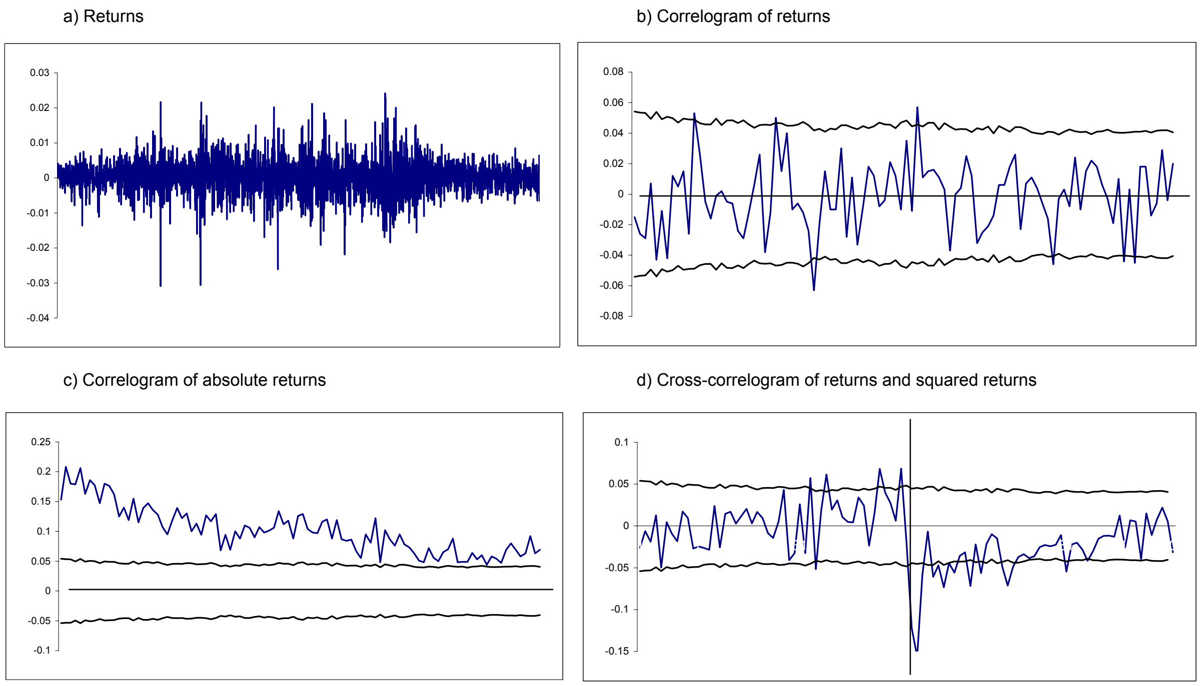


Table 3. Descriptive statistics of daily returns of S&P500 observed from 29th August 1995 until 20th October 2005. The quantities reported for $Q(20)$ and $Q^*_{[M/3]-1}(20)$ are the ratios between the value of the statistic and its corresponding 5% critical value.

Sample size	2555
Mean	0.0001
Median	0.0002
Maximum	0.0242
Minimum	-0.0309
Std. Dev.	0.0050
Skewness	-0.0971*
Kurtosis	6.0215*
$Q(20)$	0.90
$Q^*_{[M/3]-1}(20)$	389.06*
$Corr\left(y_t, y_{t+1}^2\right)$	-0.1225

* Significant values at 5% level.

Figure 4. GARCH models with alternative conditional distributions for S&P500

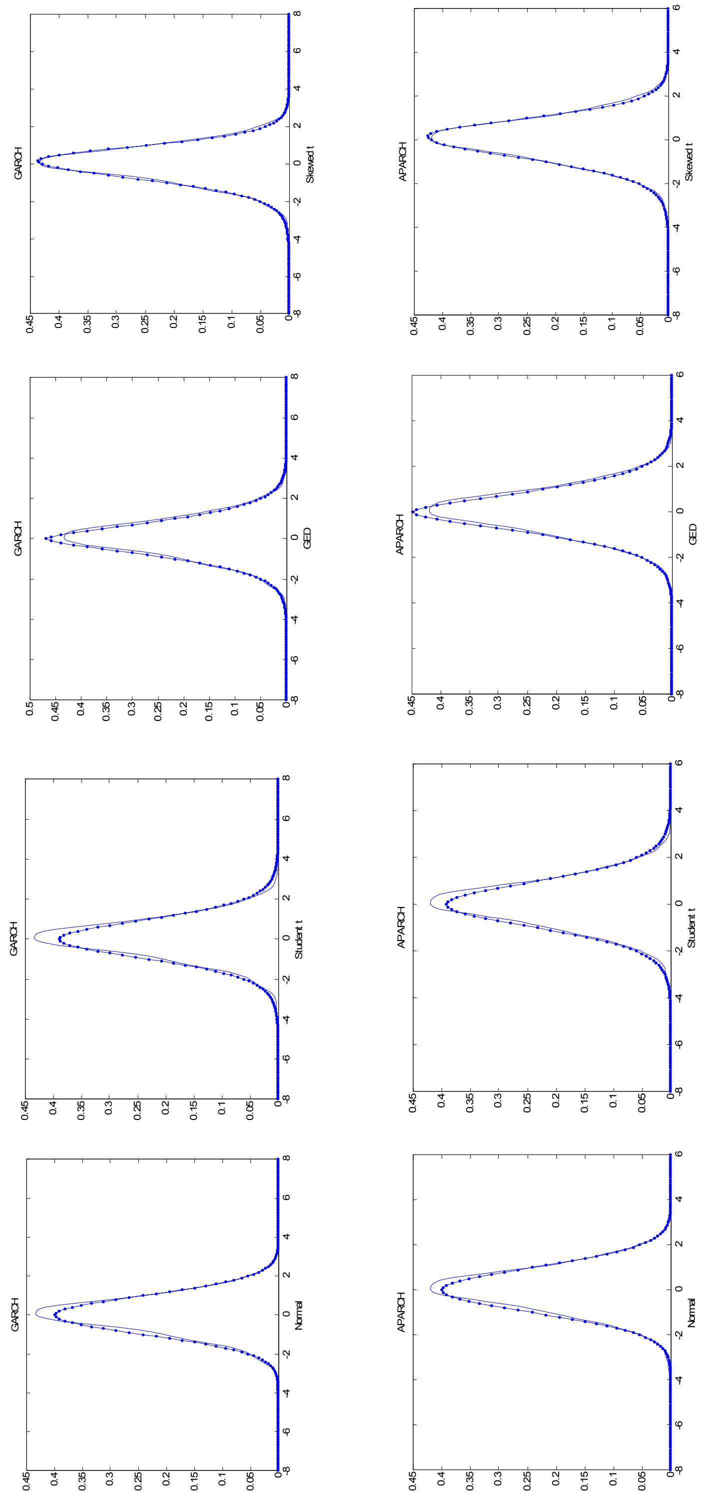


Table 4. Maximum Likelihood Estimates of the alternative GARCH-type models and conditional distributions implemented to daily S&P500.

Model		Normal	Student-ν	GED	Skewed-t
GARCH	α_0	$2.49E-07$ $(8.30E-08)$	$2.10E-07$ $(5.78E-08)$	$1.96E-07$ $(5.78E-08)$	$2.15E-07$ $(1.17E-07)$
	α_1	0.080 (0.014)	0.066 (0.011)	0.071 (0.013)	0.072 (0.021)
	β_1	0.912 (0.013)	0.925 (0.011)	0.922 (0.012)	0.921 (0.0216)
	ν		9.881 (1.993)	1.535 (0.080)	10.236 (2.080)
	ξ				-0.117 (0.025)
TGARCH	α_0	$1.10E-04$ $(3.22E-05)$	$9.99E-05$ $(2.43E-05)$	$1.05E-04$ $(2.93E-05)$	$9.77E-05$ $(4.63E-05)$
	α_1	0.011 (0.011)	0.007 (0.028)	0.009 (0.005)	0.010 (0.020)
	β_1	0.924 (0.015)	0.931 (0.018)	0.927 (0.011)	0.930 (0.019)
	γ	-0.115 (0.021)	-0.110 (0.021)	-0.113 (0.020)	-0.108 (0.032)
	ν		12.705 (3.289)	1.635 (0.085)	13.340 (3.598)
	ξ				-0.110 (0.026)
GJR	α_0	$4.07E-07$ $(1.02E-07)$	$3.31E-07$ $(6.23E-08)$	$3.57E-07$ $(7.44E-08)$	$3.32E-07$ $(1.13E-07)$
	α_1	0.002 (0.019)	0.000 (0.016)	0.000 (0.005)	0.000 (0.008)
	β_1	0.913 (0.023)	0.924 (0.014)	0.921 (0.011)	0.924 (0.016)
	γ	0.140 (0.028)	0.128 (0.022)	0.133 (0.023)	0.128 (0.034)
	ν		12.160 (3.041)	1.620 (0.088)	13.237 (3.730)
	ξ				-0.120 (0.026)
EGARCH	α_0	-0.351 (0.084)	-0.323 (0.068)	-0.337 (0.065)	-0.320 (0.115)
	α_1	0.128 (0.021)	0.119 (0.0168)	0.123 (0.0179)	0.122 (0.029)
	β_1	0.976 (0.006)	0.978 (0.005)	0.977 (0.0052)	0.978 (0.009)
	γ	-0.102 (0.018)	-0.099 (0.015)	-0.101 (0.0151)	-0.097 (0.026)
	ν		12.909 (3.425)	1.638 (0.086)	13.608 (3.660)
	ξ				-0.112 (0.026)
APARCH	α_0	$6.55E-05$ $(7.788E-05)$	$4.12E-05$ $(6.68E-05)$	$1.10E-04$ $(3.50E-05)$	$1.95E-05$ $(4.02E-05)$
	α_1	0.067 (0.013)	0.059 (0.011)	0.066 (0.010)	0.058 (0.019)
	β_1	0.923 (0.014)	0.929 (0.011)	0.925 (0.014)	0.928 (0.020)
	γ	0.834 (0.118)	0.893 (0.151)	0.872 (0.114)	0.847 (0.287)
	ν		12.705 (3.287)	1.635 (0.084)	13.410 (3.658)
	ξ				-0.114 (0.027)
	δ	1.093 (0.202)	1.159 (0.282)	0.999 (0.084)	1.287 (0.363)

Table 5. Residual analysis of GARCH models.

		Skewness	*Kurtosis*	$Q(20)$	$Q^*_{[20/3]-1}(20)$	$Corr\left(y_t, y_{t+1}^2\right)$
GARCH	Normal	-0.4253 (0)	4.5516* (3)	0.6959	0.4613	-0.0953
	Student-ν	-0.4284* (0)	4.6206* (4.02)	0.6424	0.8324	-0.098
	GED	-0.4293* (0)	4.5932* (3.68)	0.6440	0.6494	-0.097
	Skewed-t	-0.4286* (-0.2803)	4.5910* (4.03)	0.6456	0.6301	-0.0969
TGARCH	Normal	-0.3460* (0)	4.0828* (3)	0.7229	0.5275	-0.0583
	Student-ν	-0.3500* (0)	4.1137* (3.68)	0.7193	0.6897	-0.0612
	GED	-0.3478* (0)	4.0969* (3.48)	0.7213	0.6098	-0.0594
	Skewed-t	-0.3519* (-0.2365)	4.1118* (3.69)	0.7185	0.5267	-0.0622
GJR	Normal	-0.3883* (0)	4.2497* (3)	0.7480	0.3278	-0.0592
	Student-ν	-0.3993* (0)	4.3196* (3.73)	0.7421	0.3817	-0.0637
	GED	-0.3957* (0)	4.2970* (3.51)	0.7446	0.3597	-0.062
	Skewed-t	-0.3994* (-0.2589)	4.3204* (3.70)	0.7420	0.3817	-0.0638
EGARCH	Normal	-0.3476* (0)	4.0723* (3)	0.7284	0.4406	-0.0595
	Student-ν	-0.3509* (0)	4.0960* (3.67)	0.7245	0.5164	-0.0619
	GED	-0.3491* (0)	4.0838* (3.48)	0.7266	0.4638	-0.0604
	Skewed-t	-0.3526* (-0.2389)	4.0945* (3.67)	0.7240	0.3984	-0.0627
APARCH	Normal	-0.3504* (0)	4.0951* (3)	0.7298	0.4923	-0.0581
	Student-ν	-0.3583* (0)	4.1409* (3.68)	0.7302	0.6630	-0.0609
	GED	-0.3451* (0)	4.0881* (3.48)	0.7229	0.6695	-0.0579
	Skewed-t	-0.3666* (-0.2445)	4.1646* (3.69)	0.7352	0.4652	-0.0619

* Significant values at 5% level

The figures in parenthesis in the column of skewness and kurtosis, represents the corresponding population moments implied by the estimated distribution.

Table 6. Backtesting VaR methods for S&P500. T=1000 days. VaR Confidence level 1%.

			Christoffersen Likelihood		DQ Test		LB(5)		LB(20)	
		x	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Historical Simulation		6	6.94	0.031	18.983	0.004	26.441	7.33E-05	186.366	5.48E-29
CAViaR Adaptive		2	9.62	0.008	6.709	0.349	0.020	0.999	0.08093	0.999
CAViaR Asymmetric		6	1.96	0.375	2.321	0.888	0.181	0.999	53.7823	6.23E-05
GARCH	Normal	7	1.12	0.57	1.029	0.984	0.246	0.999	0.859	0.999
	Student- v	2	9.62	0.008	6.472	0.372	0.020	0.999	0.081	0.999
	GED	3	6.84	0.032	4.968	0.548	0.045	0.999	0.182	0.999
	Skewed-t	1	13.46	0.001	8.144	0.228	0.005	0.999	0.020	0.999
	Bootstrap-Barone	6	1.96	0.374	1.861	0.932	0.181	0.999	0.617	0.999
	EVT-Hill	2	9.62	0.008	6.508	0.369	0.020	0.999	0.081	0.999
	EVT-GPD	2	9.62	0.008	6.508	0.369	0.020	0.999	0.081	0.999
TGARCH	Normal	5	3.14	0.208	2.561	0.862	0.126	0.999	0.506	0.999
	Student- v	4	4.73	0.093	3.694	0.718	0.080	0.999	0.324	0.999
	GED	4	4.73	0.093	3.671	0.721	0.080	0.999	0.324	0.999
	Skewed-t	4	4.73	0.093	3.710	0.716	0.080	0.999	0.324	0.999
	Bootstrap-Barone	4	4.73	0.094	3.624	0.727	0.080	0.999	0.324	0.999
	EVT-Hill	4	4.73	0.094	3.674	0.721	0.080	0.999	0.324	0.999
	EVT-GPD	4	4.73	0.094	3.672	0.721	0.080	0.999	0.324	0.999
GJR	Normal	5	3.14	0.207	2.561	0.862	0.126	0.999	0.506	0.999
	Student- v	5	3.14	0.207	2.619	0.855	0.126	0.999	0.506	0.999
	GED	5	3.14	0.207	2.587	0.859	0.126	0.999	0.506	0.999
	Skewed-t	4	4.73	0.094	3.760	0.709	0.080	0.999	0.324	0.999
	Bootstrap-Barone	5	3.14	0.208	2.540	0.864	0.126	0.999	0.506	0.999
	EVT-Hill	4	4.73	0.094	3.691	0.718	0.080	0.999	0.324	0.999
	EVT-GPD	4	4.73	0.094	3.691	0.718	0.080	0.999	0.324	0.999
EGARCH	Normal	4	4.73	0.094	3.661	0.722	0.080	0.999	0.324	0.999
	Student- v	4	4.73	0.094	3.691	0.718	0.080	0.999	0.324	0.999
	GED	4	4.73	0.094	3.677	0.720	0.080	0.999	0.324	0.999
	Skewed-t	6	1.96	0.375	3.704	0.717	0.080	0.999	0.324	0.999
	Bootstrap-Barone	4	4.73	0.094	3.650	0.724	0.080	0.999	0.324	0.999
	EVT-Hill	4	4.73	0.094	3.671	0.721	0.080	0.999	0.324	0.999
	EVT-GPD	4	4.73	0.094	3.671	0.721	0.080	0.999	0.324	0.999
APARCH	Normal	5	3.14	0.208	2.567	0.861	0.126	0.999	0.506	0.999
	Student- v	3	6.84	0.033	5.004	0.543	0.045	0.999	0.1821	0.999
	GED	4	4.73	0.094	3.676	0.720	0.080	0.999	0.324	0.999
	Skewed-t	5	3.14	0.208	3.735	0.712	0.080	0.999	0.324	0.999
	Bootstrap-Barone	4	4.73	0.094	3.626	0.727	0.080	0.999	0.324	0.999
	EVT-Hill	4	4.73	0.094	3.677	0.720	0.080	0.999	0.324	0.999
	EVT-GPD	4	4.73	0.094	3.676	0.720	0.080	0.999	0.324	0.999

Figure 5. Scatter-plots of the VaR with the $Normal$ distribution.

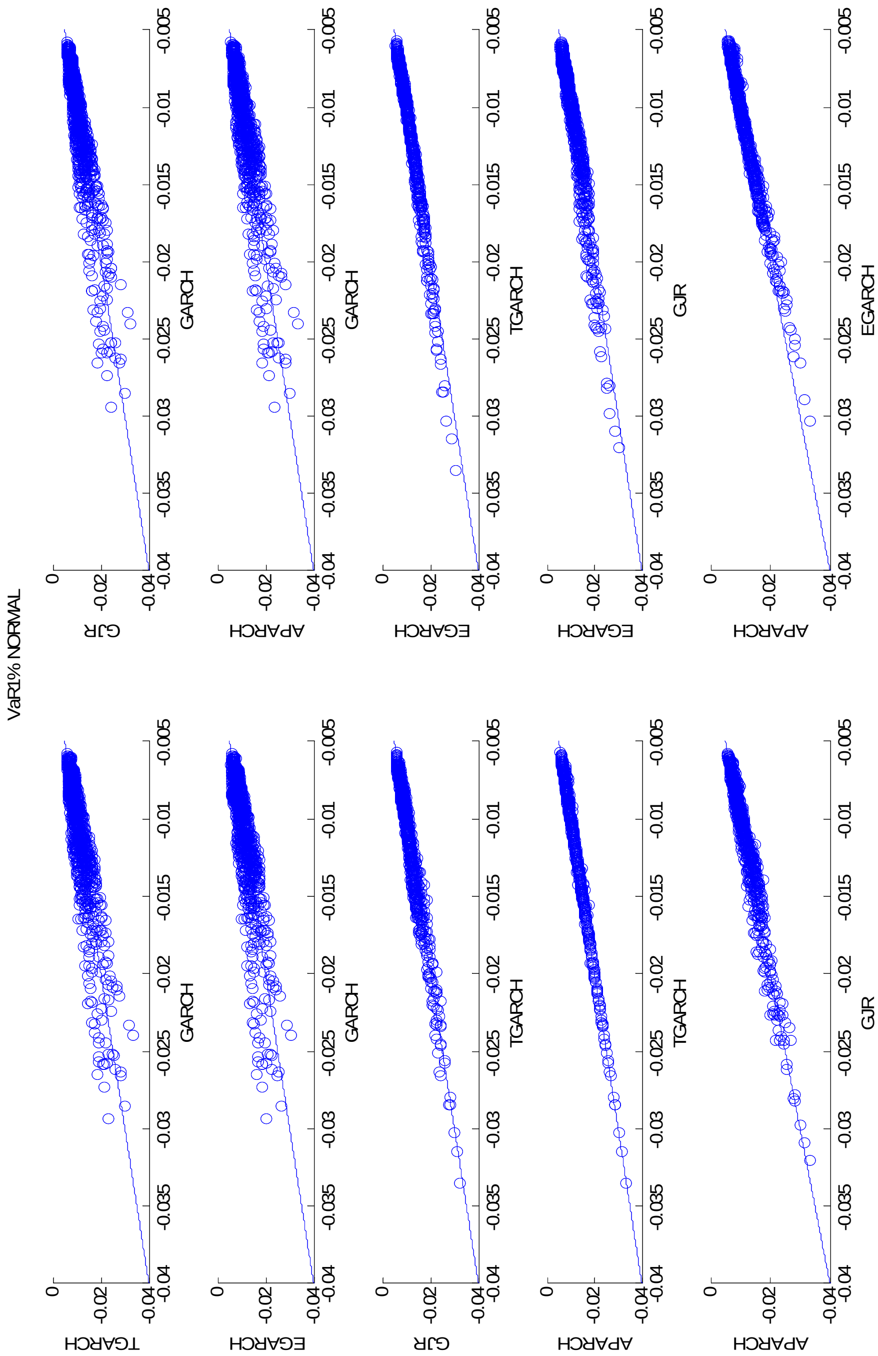


Figure 6. Scatter-plots of the VaR with the $EGARCH$ model.

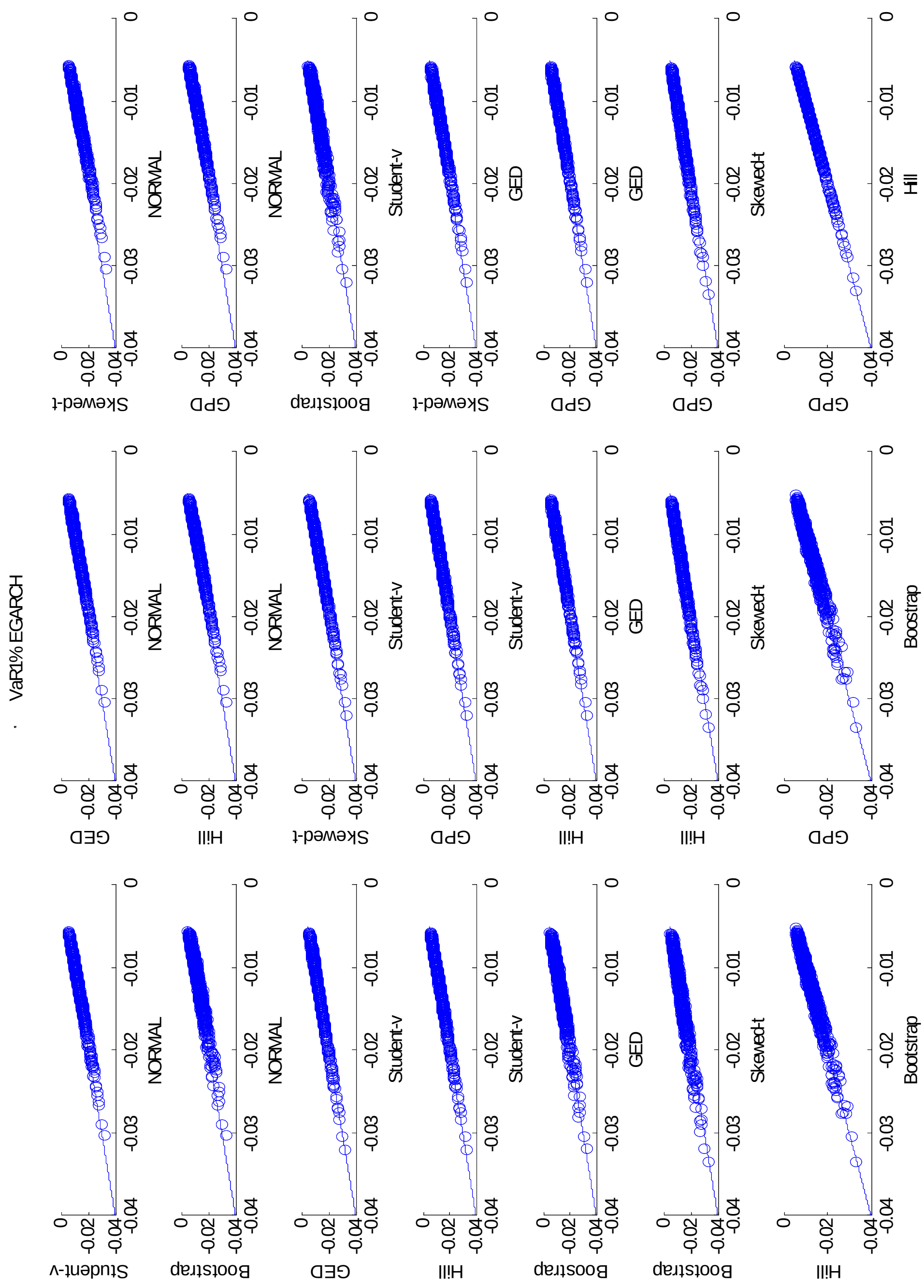


Figure 7. Scatter-plots of the ES with the $Normal$ distribution.

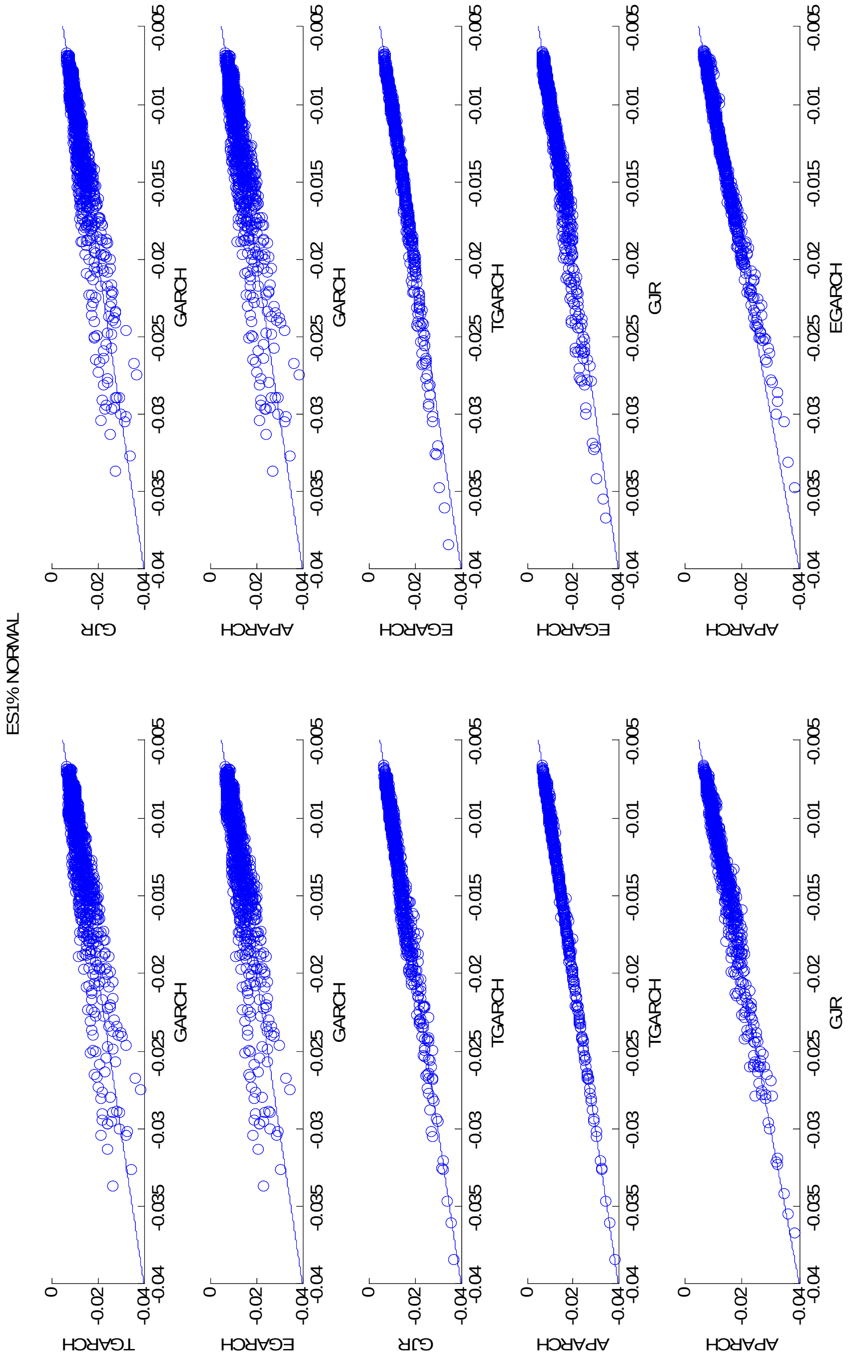


Table 7 Loss Function Value for $VaR_{1\%}$

	Loss Function		SPA test	
Model	RLF	QLF	RLF	QLF
			p-value	p-value
$CAViaR\ Asymmetric$	0.6567	0.0261	0.9268	0.3079
$GARCH-N$	0.2494	0.0214	**0.9876**	**0.9864**
$GARCH-Bootstrap$	0.3041	0.0235	0.9362	0.0002*
$TGARCH-N$	0.3671	0.0214	0.7011	0.0037*
$TGARCH-t$	0.1945	0.0224	0.4795	0.0058*
$TGARCH-GED$	0.2331	0.0219	0.4986	0.2311
$TGARCH-Skewed-t$	0.1116	0.0235	**0.9999**	0.2793
$TGARCH-Bootstrap$	0.2814	0.0236	0.7035	0.0550**
$TGARCH-EVT-GPD$	0.1796	0.0230	0.4846	0.2492
$TGARCH-EVT-Hill$	0.1721	0.0232	0.9254	0.1585
$GJR-N$	0.3838	0.0221	0.5621	0.0016*
$GJR-t$	0.1871	0.0227	0.3610	0.0620**
$GJR-GED$	0.2242	0.0223	0.4298	0.1187
$GJR-Skewed-t$	**0.0836**	0.0238	**0.9999**	0.2967
$GJR-Bootstrap$	0.3001	0.0244	0.6344	0.0090*
$GJR-EVT-GPD$	0.1596	0.0233	0.2329	0.2673
$GJR-Hill$	0.1547	0.0233	0.4833	0.2982
$EGARCH-N$	0.4230	**0.0199**	0.7525	**0.9999**
$EGARCH-t$	0.2523	0.0202	0.6630	0.0025
$EGARCH-GED$	0.2866	**0.0199**	0.7019	**0.9742**
$EGARCH-Skewed-t$	0.1584	0.0211	0.9600	**0.3386**
$EGARCH-Bootstrap$	0.2858	0.0204	0.6620	0.2442
$EGARCH-EVT-GPD$	0.2588	0.0203	0.7353	0.2963
$EGARCH-EVT-Hill$	0.2554	0.0204	0.9055	0.3033
$APARCH-N$	0.3600	0.0215	0.7098	0.0041*
$APARCH-GED$	0.2333	0.0223	0.4768	0.0016*
$APARCH-Skewed-t$	0.3267	0.0236	**0.9858**	0.2004
$APARCH-Bootstrap$	0.1522	0.0235	0.7901	0.0621
$APARCH-EVT-GPD$	0.1790	0.0229	0.5176	0.2839
$APARCH-EVT-Hill$	0.1722	0.0231	0.9168	0.2510

** Indicates that the hypothesis is rejected at 10%

* Indicates that the hypothesis is rejected at 5%

Table 8 Backtesting ES models using the SPA test

1% Confidence Level	
Model	RLF
	p-value
$CAViaR\ Asymmetric$	-
$GARCH-N$	0.4947
$GARCH-Bootstrap$	0.9896
$TGARCH-N$	0.7564
$TGARCH-t$	0.9294
$TGARCH-GED$	0.6662
$TGARCH-Skewed-t$	0.9063
$TGARCH-Bootstrap$	0.9999
$TGARCH-EVT-GPD$	0.2362
$TGARCH-EVT-Hill$	0.8293
$GJR-N$	**0.9999**
$GJR-t$	0.7841
$GJR-GED$	0.7829
$GJR-Skewed-t$	0.9207
$GJR-Bootstrap$	0.8668
$GJR-EVT-GPD$	**0.9999**
$GJR-EVT-Hill$	0.4648
$EGARCH-N$	**0.9999**
$EGARCH-t$	0.7005
$EGARCH-GED$	0.5616
$EGARCH-Skewed-t$	0.8698
$EGARCH-Bootstrap$	0.4162
$EGARCH-EVT-GPD$	0.7618
$EGARCH-EVT-Hill$	0.6522
$APARCH-N$	0.7552
$APARCH-GED$	0.6670
$APARCH-Skewed-t$	0.8867
$APARCH-Bootstrap$	**0.9999**
$APARCH-EVT-GPD$	0.3327
$APARCH-EVT-Hill$	0.7091

** Indicates that the hypothesis is rejected at 10%

* Indicates that the hypothesis is rejected at 5

Table 9. Backtesting VaR and ES methods using the Diebold-Mariano t statistic test

VaR Regulatory Loss Function			
Model	$TGARCH-Skewed-t$	$GJR-Skewed-t$	$APARCH-Skewed-t$
$GARCH-N$	0.732 (0.464)	1.019 (0.308)	0.506 (0.613)
$TGARCH-Skewed-t$		0.654 (0.513)	−0.669 (0.504)
$GJR-Skewed-t$			−1.059 (0.290)

VaR Quantile Loss Function			
Model	$EGARCH-N$	$EGARCH-GED$	$EGARCH-Skewed-t$
$GARCH-N$	0.996 (0.319)	0.954 (0.340)	0.185 (0.853)
$EGARCH-N$		−0.116 (0.907)	−1.122 (0.262)
$EGARCH-GED$			−1.906 (0.056)
$EGARCH-GED$			−1.906 (0.056)

ES Regulatory Loss Function			
Model	$GJR-EVT-GPD$	$EGARCH-N$	$APARCH-Bootstrap$
$GJR-N$	0.122 (0.903)	0.280 (0.780)	0.192 (0.848)
$GJR-EVT-GPD$		0.199 (0.843)	0.126 (0.900)
$EGARCH-N$			−0.185 (0.853)

*p-values in parentheses

References

Acerbi, C., C. Nordio, and C. Sirtori (2001). Expected shortfall as a tool for financial risk management. *Working Paper Milan, AbaxBank*.

Acerbi, C. and P. Simonetti (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance 26*, 1505–1518.

Acerbi, C. and D. Tasche (2002). Expected shortfall: a natural coherent alternative to value at risk. *Economic Notes 31*, 379–388.

Angelidis, T., A. Benos, and S. Degiannakis (2005). The use of garch models in var estimation. *Statistical Methodology 2*, 49–57.

Angelidis, T. and S. Degiannakis (2007). Backtesting var models: An expected shortfall approach. *Working Paper, University of Crete*.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1997). Thinking coherently. *Risk 10*, 68–71.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 2*, 203–228.

Bali, T. and P. Theodossiou (2007). A conditional-sgt-var approach with alternative garch models. *Annals of Operations Research 151*, 241–267.

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

Bao, Y. and A. Ullah (2004). Bias of value-at-risk model. *Finance Research Letters 1*, 241–249.

Barone-Adesi, G. and K. Giannopoulos (2001). Non-parametric var techniques. myths and realities. *Economic Notes 30*, 167–181.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper (1999). Var without correlations for portfolios of derivative securities. *The Journal of Future Markets 19*, 583–602.

Basel Committee on Banking Supervision (1996a). Supervisory framework for the use of backtesting in conjunction with the internal models approach to market risk capital requirements, bank for international settlements. Technical report.

Basel Committee on Banking Supervision (1996b). Supervisory framework for the use of backtesting in conjunction with the internal models approach to market risk capital requirements, bank for international settlements. Technical report.

Berkowitz, J., P. Christoffersen, and D. Pelletier (2006). Evaluating value-at-risk models with desk-level data. *Working Paper Series 010 Department of Economics, North Carolina State University*.

Billio, M. and L. Pelizzon (2000). Value-at-risk: a multivariate switching regime approach. *Journal of Empirical Finance 7*, 531–554.

Billio, M. and D. Sartore (2003). *Stochastic Volatility Models: A Survey with Applications to Option Pricing and Value at Risk*. West Sussex, UK: John Wiley.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the Business and Economic Statistics Section*, 177–181.

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

Cai, Z. (2002). Regression quantiles for time series data. *Econometric Theory 18*, 169–192.

Cai, Z. and X. Wang (2008). Nonparametric estimation of conditional var and expected shortfall. *Journal of Econometrics 147*, 120–130.

Chan, N.H, D. S.-J. P. L. and Z. Xia (2007). Interval estimation for the conditional value at risk based on garch models with heavy tailed innovations. *Journal of Econometrics 137*, 556–576.

Chen, M. Y. and J. Chen (2005). Application of quantile regression to estimation of value at risk. *Review of Financial Risk Management 1*, 1–15.

Chen, S. and C. Tang (2005). Nonparametric inference of value at risk for dependent financial returns. *Journal of Financial Econometrics 3*, 227–255.

Christoffersen, P. (1998). Evaluating interval forecast. *Journal of Business and Economic Statistics 39*, 841–862.

Christoffersen, P. and GonÇalves (2005). Estimation risk in financial risk management. *Journal of Risk 7*, 1–28.

Danielsson, J. and C. de Vries (2000). Value at risk and extreme returns. *Annales d'Economie et Statistique 60*, 239–270.

DeRossi, G. and A. Harvey (2006). Time-varying quantiles. *University of Cambridge, Faculty of Economics, CWPE 0649*.

Diebold, F. X. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

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

Eberlein, E., J. Kallsen, and J. Kristen (2003). Risk management based on stochastic volatility. *Journal of Risk 5*, 19–44.

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*, 281299.

Embrechts, P., C. Klüppelberg, and T. Mikosch (2000). *Modelling Extremal Events for Insurance and Finance*. Londres: Springer.

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

Escanciano, J. and J. Olmo (2008). Backtesting var with estimation risk. *Revise and Resubmit in Journal of Business and Economics Statistics*.

Fabozzi, F. and R. Tunaru (2006). On risk management problems related to a coherence property. *Quantitative Finance 6*, 75–81.

Fan, J. (1992). Design adaptive nonparametric regression. *Journal of the American Statististical Association 87*, 998–1004.

Fan, J. and J. Gu (2003). Semiparametric estimation of value at risk. *The Econometrics Journal 6*, 261–290.

Giacomini, R. and I. Komunjer (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics 23*, 416–431.

Giannopoulos, K. and R. Tunaru (2005). Coherent risk measures under filtered historical simulation. *Journal of Bnking and Finance 29*, 979–996.

Giot, P. and S. Laurent (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics 18*, 641–663.

Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*, 1779–1801.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics 44*, 423–453.

Gourieroux, C. and J. Jasiak (2006). Dynamic quantile models. *York University, Department of Economics Working Paper*.

Granger, C. W. (2002). Some comments on risk. *Journal of Applied Econometrics 17*, 447–456.

Hall, P. and Q. Yao (2003). Data tilting for time series. *Journal of the Royal Statistical Society, Series B 65*, 425–442.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357–384.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review 35*, 705–730.

Hansen, P. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics 23*, 365–380.

Hartz, C., M. S. and M. Paolella (2006). Accurate value-at-risk forecasting based on the normal-garch model. *Computational Statistics & Data Analysis 51*, 2295–2312.

He, C., A. Silvennoinen, and T. Teräsvirta (2005). Parameterizing unconditional skewness in models for financial time series. *Research Paper Series 169, Quantitative Finance Research Centre, University of Technology, Sydney*.

Hendricks, D. (1996). Evaluation of value at risk models using historical data. *Federal Reserve Bank of New York Economic Policy Review, April*, 39–69.

Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric garch models. *Journal of Financial Economics 39*, 71–104.

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics 3*, 1163–1174.

Hull, J. and A. White (1998). Incorporating volatility up-dating into the historical simulation method for var. *Journal of Risk 1*, 5–19.

Inui, K. and M. Kijima (2005). On the significance of expected shortfall as a coherent risk measure. *Journal of Banking and Finance 29*, 853–864.

Ju, X. and N. D. Pearson (1999). Using value-at-risk to control risk taking: How wrong can you be? *Journal of Risk 1*, 5–36.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50.

Krause, A. (2003). Exploring the limitations of value at risk: How good is it in practice? *Journal of Risk Finance 4*, 19–28.

Kuester, K., S. Mittnik, and M. Paolella (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics 4*, 53–89.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives 3*, 73–84.

Lambert, P. and S. Laurent (2000). Modelling skewness dynamics in series of financial data. *Discussion Paper, Institute de Statistique, Louvain-la-Neuve*.

Levy, H. (1992). Stochastic dominance expected utility: survey and analysis. *Management Science 38*, 555–593.

Lima, R. and B. Néri (2007). Comparing value-at-risk methodologies. *Brazilian Review of Econometrics 27*, 1–25.

Longerstaey, J. and L. More (1995). *Introduction to RiskMetrics* (4a edición ed.). New York: Morgan Guaranty Trust Company.

Longin, F. (2001). Beyond the var. *Journal of Derivatives 8*, 36–48.

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

Luce, R. (1981). Corrections to 'some possible measures of risk'. *Theory and Decision 13*, 381.

Machina, M. and M. Rothschild (1987). *Risk*. London: Macmillan.

Manganelli, S. and R. Engle (2001). Value at risk models in finance. *Working Paper, European Central bank*.

Martins-Filho, C. and F. Yao (2006). Estimation of value-at-risk and expected shortfall based on nonlinear models of return dynamics and extreme value theory. *Studies in Nonlinear Dynamics & Econometrics 10*, 107–149.

McLeod, A. and W. Li (1983). Diagnostic checking arma time series models using squared-residual autocorrelations. *Journal of the Time Series Analysis 4*, 269–273.

McNeil, A. J. and R. Frey (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance 7*, 271–300.

Mittnik, S. and M. Paolella (2000). Conditional density and value-at-risk prediction of asian currency exchange rates. *Journal of Forecasting 19*, 313–333.

Morgan, J. (1995). Riskmetrics. *RiskMetrics Technical Document. 3a edición.*

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*, 347–370.

Newey, W. and K. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Nystrom, K. and J. Skoglund (2002). Univariate extreme value theory, garch and measures of risk. *Preprint, Swedbank*.

Pascual, L., J. Romo, and E. Ruiz (2006). Bootstrap prediction for returns and volatilities in garch models. *Computational Statistics and Data Analysis 50*, 2293–2312.

Peng, L. and Y. Qi (2003). Confidence intervals for high quantiles of a heavy tailed distribution. *Technical Report*, Georgia Institute of Technology, United States.

Politis, D. and J. Romano (1994). The stationary bootstrap. *Journal of American Statistical Association 89*, 1303–1313.

Politis, D. and H. White (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Review 23*, 53–70.

Pownall, R. and K. Koedijk (1999). Capturing downside risk in financial markets: the case of the asian crisis. *Journal of International Money and Finance 18*, 853–870.

Rodriguez, J. and E. Ruiz (2005). A powerful test for conditional heteroscedasticity for financial time series with highly persistent volatilities. *Statistica Sinica 15*, 505–525.

Ruiz, E. and L. Pascual (2002). Bootstrapping financial time series. *Journal of Economic Surveys 16*, 271–300.

Sadorsky, P. (2005). Stochastic volatility forecasting and risk management. *Applied Financial Economics 15*, 121–135.

Sarma, M., T. S. and A. Shah (2003). Selection of value at risk models. *Journal of Forecasting 22*, 337–358.

Szegö, G. (2002). Measures of risk. *Jornal of Banking and Finance 26*, 1253–1272.

Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking and Finance 26*, 1519–1533.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science 44*, 1650–1661.

Tobin, J. (1969). Comment on borch and feldstein. *The Review of Economic Studies 36*, 13–14.

Tsiang, S. (1972). The rationale of the mean-stardard deviation analysis, skewness preference, and the demand for money. *The American Economic Review 62*, 354–371.

Vlaar, P. J. G. (2000). Value at risk models for dutch bond portfolios. *Journal of Banking and Finance 24*, 1131–1154.

Yu, K. and M. Jones (1998). Local linear quantile regression. *Journal of American Statistical Association 93*, 228–237.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control 18*, 931–995.