Value-at-Risk for Greek Stocks

Timotheos Angelidis
University of Peloponnese, Greece

Alexandros Benos
National Bank of Greece, Greece

This paper analyses the application of several volatility models to forecast daily Value-at-Risk (VaR) both for single assets and portfolios. We calculate the VaR number for 4 Greek stocks, 2 portfolios based on these securities and for the Athens Stock Exchange General Index. We model VaR for long and short trading positions by employing non-parametric methods, such as historical and filtered historical simulation, as well as parametric ones. Especially for the later techniques we use a collection of ARCH models (GARCH, EGARCH and TARCH) based on three distributional assumptions (Normal, Student-*T* and Skewed Student-*T*), while we combine the Extreme Value Theory with a volatility updating technique (via GARCH type-modeling). In order to choose one model among the various forecasting methods, we employ a two-stage backtesting procedure. In the first one, we implement two backtesting criteria (unconditional and conditional coverage) to test the statistical accuracy of the models. In the second stage, we employ standard forecast evaluation methods in order to examine whether any differences between models that have converged are statistica lly significant (JEL: C22; C52; C53; G15).

Keywords: value-at-risk, GARCH, historical simulation, backtesting.

I. Introduction

Value-at-Risk (VaR) has been considered by regulatory authorities and financial institutions as the most important market risk measure. In general, VaR refers to a portfolio's worst outcome that is expected to occur over a predetermined period (one or ten trading days) at a given

* We acknowledge helpful comments from anonymous referees.

(*Multinational Finance Journal*, 2008, vol. 12, no. 1/2, pp. 67–104)
Quarterly publication of the *Multinational Finance Society*, a nonprofit corporation.
© *Global Business Publications*. All rights reserved.
DOI: 10.17578/12-1/2-4

confidence level (e.g., 97.5% or 99%). According to the Basel Committee on Banking Supervision (the Amendment to the Capital Accord to Incorporate Market Risks, January 1996), the VaR methodology can be used by financial institutions to calculate capital charges with respect to their interest rate, equity, foreign exchange and commodity risk.

VaR has, nevertheless, been criticized as a measure of market risk on two grounds. Artzner et al. (1997, 1999) showed that it is not necessarily sub-additive, i.e., the VaR of a portfolio with two instruments maybe greater than the sum of individual VaRs and therefore managing risk by using it may fail to automatically stimulate diversification. Moreover, it does not give any indication about the size of the potential loss given that this loss exceeds the VaR number. In order to remedy the effects of these shortcomings Delbaen (1998) and Artzner et al. (1997) introduced the Expected Shortfall risk measure, which equals the expected value of the losses conditional on a VaR violation. Furthermore, Basak and Shapiro (2001) suggested an alternative risk management procedure that also focuses on the expected loss when (and if) losses occur. They substantiated that the proposed procedure generates losses lower than those of the VaR-based risk management techniques. Last, but not least, the standard VaR measure presumes that asset returns are normally distributed, while it is widely documented that they really exhibit non-zero skewness and excess kurtosis and, hence, the VaR measure either underestimates or overestimates “true risk”.

In order to calculate this infamous metric, a researcher may either use a parametric or a non-parametric method. Under the framework of non-parametric techniques, Historical Simulation methods that are based on the empirical distribution of returns, have been thoroughly examined by several authors without, however, reaching a unanimous conclusion. Hendricks (1996), Vlaar (2000) and Danielsson (2002) argued that sample size affects the precision of VaR estimates, with larger sizes producing the most accurate estimations. On the contrary, Hoppe (1998) proposed the use of smaller sizes, since they can accommodate structural changes of trading behavior, a view also expressed by Frey and Michaud (1997). Lambadiaris et al. (2003) exploited the accuracy of both the historical and the Monte Carlo simulation methods in two different markets. For a stock portfolio, they concluded that the historical simulation method is not appropriate for a risk manager for daily VaR calculation, while, for a bond portfolio, results were mixed, as the best method depends on the backtesting measure and the confidence level chosen. For commodity markets, Cabedo and Moya (2003) developed an ARMA historical simulation method to estimate daily VaR. They compared it to the simple historical one and concluded their technique yielded better estimates.

Other researchers preferred to use parametric methods. We can classify the procedures into two categories. In the mixture case, Venkataraman (1996) and Zangari (1996) suggested a mixture of normal distributions to accommodate the observed skewness and kurtosis of financial time series and hence to describe them better than the standard normal distribution. A more thorough analysis was conducted by Billio and Pelizzon (2000) who estimated a multivariate switching regime model to calculate the VaR for 10 Italian stocks and several portfolios made up from them. Their procedure is different from that of Zangari (1996) in that (a) the regime forecasts are generated by a two state Markov process instead of a Bernoulli one and, (b) volatility clustering is more easily accommodated under their framework. Based on two backtesting measures (proportion of failures and time to first failure), they substantiated that the switching regime specification is more accurate than other known methods (RiskMetrics™ or GARCH with Normal and Student-*t* distributions). In the single state case, authors evaluated the forecasting ability of the most well known volatility techniques (GARCH, APARCH, RiskMetrics™) under several distributional assumptions (Normal, Student-*t*, Skewed Student-*t*). Gurmat and Harris (2002) pointed out that, compared to the GARCH(1,1) specification under both the Normal and the Student-*t* distributions, the proposed exponentially weighted likelihood model improved the estimated daily VaR number at higher confidence levels. Giot and Laurent (2003a, 2003b) estimated daily VaR both for long and short trading positions by employing an APARCH Skewed Student-*t* model to take into account the asymmetry of their dataset. They showed that it performed better than the pure symmetric approach, since it described more accurately the empirical distribution. Brooks and Persand (2003) also consider the issue of asymmetry in the VaR framework and concluded that models, which do not allow for asymmetries either in the unconditional return distribution or in the volatility specification, underestimate the "true" VaR. Finally, Bali and Theodossiou (2004) combined the Skewed Generalized *t*-distribution with 10 GARCH specifications and calculated both VaR and Expected Shortfall numbers. They argued that the *TS*-GARCH, proposed by

Taylor (1986) and Schwert (1989), and EGARCH, introduced by Nelson (1991), had the best overall performance.

So far, all models presented were based on Historical Simulation methods and on variance-covariance techniques. Hull and White (1998) and Barone-Adesi, Giannopoulos and Vosper (1999) introduced the Filtered Historical simulation (*FHS*), which combines both of the above. More specifically, it does not make any distributional assumption about standardized returns, but forecasts variance through a structured volatility model. Hence, it can be considered as a mixture of parametric and non-parametric procedures. Moreover, Barone-Adesi and Giannopoulos (2001) demonstrated the superiority of *FHS* over the historical one, by showing that generated better VaR forecasts. Under the same framework, the Extreme Value Theory (*EVT*) has been recently proposed: it only models the tails of the distribution rather than the entire distribution. Therefore, it focuses on the parts that are essential to VaR.[1]

The purpose of our paper is twofold. First, we want to implement several volatility models (parametric or not) in order to estimate the 97.5% and 99% one-day VaR for both long and short trading positions. We then aim at evaluating the predictive accuracy of various models under a risk management framework. We employ a two-stage procedure to investigate the forecasting power of each volatility forecasting technique. In the first stage, two backtesting criteria (unconditional and conditional coverage) are implemented to test the statistical accuracy of the models. In the second stage, we employ standard forecast evaluation methods to examine whether the differences between models that have sufficiently satisfied the first stage criteria, are statistically significant.

Even if our article looks similar to papers cited above, there are still significant differences. First, we investigate the risk management techniques for both long and short trading positions, while most of the research has focused only on long ones. Given the fact that the financial series exhibit non-zero skewness, it is important for a risk manager to know whether his/her VaR model can be applied to both positions. Second, we employ parametric, non-parametric and semi-parametric techniques in order to investigate their relative performance in a unified environment, contrary to existing literature that, to the best of our

---

1. For more information on *EVT* and VaR see Jondeau and Rockinger (1999), MacNeil and Frey (2000), Jondeau and Rockinger (2003), Ho et al. (2000), Rozario (2002), Seymour and Polakov (2003), Bali (2003), Gençay and Selçuk (2004) and Byström (2004) among others.

knowledge, focuses only on one technique at a time. Third, we implement the two-stage model selection procedure outlined above in an attempt to possibly identify a unique model for each security, trading position and confidence level. Finally, our empirical analysis is carried on a small emerging market, permitting a performance comparison with techniques used in more developed markets.

Our results point out to the need to develop more sophisticated backtesting measures as, in most cases examined, statistical measures cannot identify one model for each case. On the other hand, under an internal loss function framework we developed, we are able to evaluate differences between VaR models and hence choose among alternative risk management practices. Moreover, based on the two backtesting measures used, both *EVT* and *FHS* generate accurate VaR numbers for both trading positions and confidence levels, as they capture more efficiently than parametric methods the characteristics of the empirical distribution. Under the framework of the loss function, the *FHS* should be applied to short trading positions, while for long ones there is no unique specific model that performs better overall. This may imply that even asymmetric models are not sufficiently asymmetric for the returns observed.

The rest of the paper is organized as follows. Section II provides a description of various VaR methods, while section III describes the evaluation framework. Section IV presents preliminary statistics for the dataset, explains the estimation procedure and presents the results of the empirical investigation. Section V concludes.

II. Value-at-Risk

In this section, we present various parametric and non-parametric methods that we apply in order to estimate the daily VaR number. We will differentiate the former by conditional variance structure and by underlying distribution. Specifically, we use three distributional assumptions (Normal, Student-*t* and Skewed Student-*t*) combined with GARCH, EGARCH and TARCH specifications. For non-parametric methods on the other hand, we will take advantage the quantiles of the empirical distribution.

The 1-day VaR is defined as $Pr\ (y_t + 1 < VaR_{t+1/t}) = \alpha$, where y_{t+1} is the future change of the portfolio's value, while α is one minus the VaR confidence level. More formally, VaR is calculated based on the following equation:

$$\text{VaR}_{t+1|t} = F(\alpha)\sigma_{t+1|t} \,, \tag{1}$$

given that F (α) is the corresponding quantile of the assumed distribution and $\sigma_{t+1|t}$ is the predicted conditional standard deviation at time t. Under the assumption that portfolio returns are normally distributed, the calculation of VaR is greatly simplified, as both $\sigma_{t+1|t}$ and F (α) have tractable expressions. This Variance-Covariance (*VC*) method, nevertheless, usually underestimates true VaR, since the normality assumption is usually rejected for financial series. Therefore, we need to make conjectures about (a) the underlying distribution and (b) the conditional variance innovation. We will fully exploit these paths in the following sections.

A. Parametric Volatility Forecasting Models

Let $y_t = \ln (S_t / S_{t-1})$ denote the continuously compounded rate of return from time $t - 1$ to:

$$y_t = \mu + e_t, \tag{2}$$

where S_t is the asset price at time t, μ is the conditional mean, and the unpredictable component, e_t, can be expressed as:

$$e_t = \varepsilon_t \, \sigma_t, \tag{3}$$

where ε_t is *iid*.

Given that most volatility models have been thoroughly examined by several authors and that our main focus is to evaluate their forecasting performance under a VaR framework, we briefly present them in table 1.[2]

We now turn to the several distributional assumptions one can make for ε_t. Engle (1982) introduced the ARCH process under the assumption of normality.

$$D(\varepsilon_t) = (2\pi)^{(-1/2)} e^{\left(-\varepsilon_t^2/2\right)} \tag{4}$$

The excess kurtosis generated from a GARCH process is not often able

2. For more information on volatility forecasting in financial markets see Poon and Granger (2003) among others.

TABLE 1. Volatility Forecasting Models

Model	Equation
GARCH(*p*, *q*)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i e_{t-i}^2 + \sum_{k=1}^{p} b_k \sigma_{t-k}^2$
RiskMetrics™	$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda)\varepsilon_{t-1}^2$
EGARCH(*p*, *q*)	$\ln(\sigma_t^2) = a_0 + \sum_{i=1}^{q}\left(a_i \left\lvert\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right\rvert + \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right) + \sum_{j=1}^{p}\left(b_j \ln(\sigma_{t-j}^2)\right)$
TARCH(*p*, *q*)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 d_{t-1} + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$

Note: This table summarizes the volatility forecasting models. RiskMetrics™ sets λ equal to 0.94 for daily volatility forecasting. d_t is a dummy variable which takes the value 1 if $\varepsilon_t > 0$ and 0 otherwise.

to describe the fat tails of time series. Consequently, Bollerslev (1987) proposed the standardized symmetric Student-*t* distribution with $v > 2$ degrees of freedom:

$$D(\varepsilon_t; v) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(1+\frac{\varepsilon^2}{v-2}\right)^{-\frac{v+1}{2}}, \qquad (5)$$

where $\Gamma(\cdot)$ is the gamma function. The Student-*t* distribution is symmetric around zero and for $v > 4$ the conditional kurtosis equals $3(v-2)(v-4)^{-1}$, which exceeds the normal value of three. For $v \to \infty$, however, the density function of that distribution converges to the standard normal one. Still, many authors prefer to use asymmetric distributions since the Student-*t* distribution cannot accommodate the observed skewness of financial time series.[3] Based on the work of Lambert and Laurent (2001), we will follow the same path and use the

3. See Theodossiou (1998).

standardized Skewed Student-*t* distribution:

$$D(\varepsilon_t;\xi,\nu)=\begin{cases}\frac{2}{\xi+\frac{1}{\xi}}sD\left[\xi\left(s\varepsilon_t+m\right);\nu\right] & \text{if} \quad \varepsilon_t<-\frac{m}{s}\\ \frac{2}{\xi+\frac{1}{\xi}}sD\left[\frac{s\varepsilon_t+m}{\xi};\nu\right] & \text{if} \quad \varepsilon_t\geq-\frac{m}{s}\end{cases} \tag{6}$$

where *D* (.;*ν*) is defined in 5, ξ is the asymmetry coefficient, while $m=\frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(\xi-\frac{1}{\xi}\right)$ and $s^2 = (\xi^2 + 1 / \xi^2 - 1) - m^2$ are the mean and variance of the non-standardized Skewed Student-*t* distribution, respectively. As Lambert and Laurent (2000) noted, the density is skewed to the right (left) if $log(\xi) > 0$ (< 0). Also, they produced the α-quantile function, $st^*_{\alpha,\nu,\xi}$, of the non-standardized Skewed Student-*t* distribution as follows:

$$st^*_{\alpha,\nu,\xi}=\begin{cases}\frac{1}{\xi}t_{a,\nu}\left[\frac{\alpha}{2}\left(1+\xi^2\right)\right] & \text{if} \quad \alpha<\frac{1}{a+\xi^2}\\ -\xi t_{\alpha,\nu}\left[\frac{1-\alpha}{2}\left(1+\xi^2\right)\right] & \text{if} \quad \alpha\geq\frac{1}{a+\xi^2}\end{cases} \tag{7}$$

It is now straightforward to estimate the VaR number.

B. Historical Simulation

Historical simulation (*HS*) has received much attention because of its simplicity and its relative lack of theoretical burden. It uses historical returns and derives the VaR number for a specific confidence interval as the corresponding quantile of the empirical historical distribution:

$$\text{VaR}^p_{t+1/t}=\text{Quantile}\left\{\left\{y_t\right\}^n_{t=1},100p\right\} \tag{8}$$

Specifically, it assumes that the future distribution of y_t is well described by the empirical (historical) one. By relying on actual prices, it accommodates non-normal distributions and, therefore, accounts for “fat tails” and non-zero skewness. This simplicity does not come without a cost, however, as the choice of the sample size, *n*, affects the estimates. If *n* is too large, then the most recent observations, which probably describe the future distribution better, carry the same weight as the earlier ones, which most probably are not as important as the latest ones. In case *n* is too small, the following may occur: either too few or insufficient extreme events will be observed. In both cases, the sample size is a hinder factor and hence "true" VaR is either underestimated or overestimated. This remark was confirmed by Van den Goorbergh and Vlaar (1999) who argued that VaR estimates for Dutch equity were extremely sensitive to sample length.

C. Filtered Historical Simulation

In the case of parametric methods, the distributional choice is crucial, while for non-parametric ones, we see there is no consistent method of estimating the volatility innovation. The Filtered Historical Simulation method (*FHS*), introduced by Hull and White (1998) and Barone-Adesi, Giannopoulos and Vosper (1999), replicates the tails of the distribution in the way proposed by Barone-Adesi and Giannopoulos (2001). Using the quantiles of standardized residuals and the conditional standard deviation forecast from a volatility model, the VaR number is calculated as:

$$\mathrm{VaR}_{t+1|t} = \mathrm{Quantile}\left[\left\{\varepsilon_t\right\}_{t=1}^{n}, 100\,p\right]\sigma_{t+1|t} \qquad (9)$$

For empirical investigation purposes, we assume the volatility estimates and the corresponding quantiles are being generated via a GARCH (1,1) process. The combination of the two methods might alleviate the problems faced by "classica" approaches, since we accommodate volatility clustering, observed "fat" tails and non-zero skewness of the empirical distribution.

D. Extreme Value Theory

The study of extremes in financial series has gradually grown in the last few years. Of course, it is obvious that any statistical procedure attempting to model extremes should benefit from the appropriate choice of the underlying distribution. Therefore, the Generalized Pareto Distribution (*GPD*) may well describe the behavior of extremes, as summarized by the so-called "tail index" τ. We apply the *EVT* method on standardized portfolio returns ($\varepsilon_t = y_t / \sigma_t \sim i.i.d.D\,(0, 1)$) because, for non-*iid* returns, the estimated parameters of the *GDP* density are biased.

Moreover, following McNeil and Frey (2000), we filter the return series via a GARCH (1,1) process, in order to catch the empirical distribution.

The estimation technique implemented attempts to model the breaking of a threshold *u*, also known as the peaks over threshold method. The probability that standardized returns are greater than *u* is given by:

$$F_u(x) \equiv Pr\{z - u \le x \mid z > u\} = \frac{F(x+u) - F(u)}{1 - F(u)} ,$$

where $x > u$. If one lets the threshold *u* get large, then *GPD* is the limiting distribution of the number of excesses.[4] The density function of *GPD*, $G(x;\tau,\psi)$, is described by:

$$G(x;\tau,\gamma) = \begin{cases} 1 - \left(1 + \frac{\tau x}{\psi}\right) & \text{if} \quad \tau \ne 0 \\ 1 - \exp\left(\frac{-x}{\psi}\right) & \text{if} \quad \tau = 0, \end{cases} \tag{10}$$

where $\psi > 0$ is a scale factor and

$$\begin{cases} x \ge u & \text{if} \quad \tau \ge 0 \\ u \le x \le u - \frac{\psi}{\tau} & \text{if} \quad \tau < 0 . \end{cases}$$

The *GPD* covers a wide range of distributions: for example, if $\tau > 0$, it addresses the heavy tailed ones, while, if $\tau < 0$, it includes the short tailed distributions, less frequently used in financial studies. Finally, it converges to the density function of the standard normal distribution when $\tau = 0$.

Under the assumption of $\tau > 0$, a reasonable one for most financial time series, the Hill estimator of the tail index (τ) equals:[5]

$$\tau = \frac{1}{\mathrm{T}_u} \sum_{i=1} T_u \ln\left(\frac{y_i}{u}\right), \tag{11}$$

---

4. See Balkema and de Hann (1974) and Pickands (1975) among others.

5. For more information, see Christoffersen (2003).

where T_u is the number of observations above the threshold u, which is assumed to be equal to 5% of the total sample size (T). Hence, under this framework, the VaR is calculated as:

$$\mathrm{VaR}_{t+1|t} = \sigma_{t+1|t} u \left[\frac{p}{T_u / T} \right]^{-\tau}, \tag{12}$$

where p denotes the VaR confidence level.

III. Evaluation Framework

The objective of this section is to evaluate the adequacy of VaR forecasts as "risk predictors" in a risk management environment. Two backtesting procedures (unconditional and conditional coverage) will serve as the final diagnostic check in order to judge the "quality" of the VaR forecasts. The purpose of backtesting is twofold. First, we would like to test whether the average number of realized VaR violations, in an out-of-sample time period, is statistically equal to the expected one.[6] It is important to note that the estimated Value-at-Risk number must neither overestimate nor underestimate, on average, the "true" but unobservable Value-at-Risk. In the former case, the financial institution may keep costly capital idle and, surely, does not use it efficiently; in the latter, its capital may not be enough to cover possible losses (in the proposed confidence level). Second, given the fact that an adequate model must widen VaR forecasts during high volatility periods and narrow them in low volatility ones, it is necessary to examine whether violations are also randomly distributed.

Unfortunately, in most cases, there are more than one risk models that satisfy both backtesting procedures. Therefore, a risk manager will not be able to select a unique volatility forecasting technique, solely based on them. Hence, in order to achieve such a goal, we should compare the best performing models among all those that have passed the test, via a loss function, in an attempt to select one of them among the various candidates.

---

6. A violation occurs if the predicted VaR is not able to cover realized losses.

TABLE 2. Kupiec's (1995) Unconditional Coverage Test

Confidence level	Evaluation sample size 250	500	750	1000
5.0%	$7 \le N \le 19$	$17 \le N \le 35$	$27 \le N \le 49$	$38 \le N \le 64$
1.0%	$1 \le N \le 6$	$2 \le N \le 9$	$3 \le N \le 13$	$5 \le N \le 16$
0.5%	$0 \le N \le 4$	$1 \le N \le 6$	$1 \le N \le 8$	$2 \le N \le 9$
0.1%	$0 \le N \le 1$	$0 \le N \le 2$	$0 \le N \le 3$	$0 \le N \le 3$

Note: This table presents Kupiec's (1995) unconditional coverage test. No rejection regions for a 5% test size.

A. Unconditional Coverage

Let I_{t+1} be a sequence of VaR violations that can be described as:

$$I_{t+1} = \begin{cases} 1, & \text{if } y_{t+1} < \text{VaR}_{t+1|t} \\ 0, & \text{if } y_{t+1} < \text{VaR}_{t+1|t} \end{cases}$$

and therefore $N = \sum_{t=1}^{T} I_t$ is the number of days over a T period that the portfolio loss was greater than the estimated VaR.

As Kupiec (1995) stated, the failure number follows a binomial distribution and, consequently, the appropriate likelihood ratio statistic, under the null hypothesis that the observed exception frequency equals to the expected one ($N / T = p$), will be given by:

$$LR_{uc} = 2\ln\left[\left(1-\frac{N}{T}\right)^{T-N}\left(\frac{N}{T}\right)^{N}\right] - 2\ln\left[\left(1-p\right)^{T-N} p^{N}\right] \sim x_1^2 \quad (13)$$

Table 2 presents the no rejection regions of N for various sample sizes and confidence levels. Although this test can reject a model that has generated too many or too few VaR violations, its power is generally poor since, especially for high confidence levels, as it cannot indicate the inadequate model, even if the difference between the observed and the expected failure turns out to be significant.

B. Conditional Coverage

For all these reasons, another test was developed by Christoffersen (1998) to jointly examine the hypotheses that (a) the total number of failures is statistically equal to the expected one and, (b) that the VaR violations are independent.[7] If a risk model is well specified and hence incorporates the characteristics of the conditional distribution (time-varying volatility, kurtosis and skewness), the exception indicator (I_{t+1}) must be unpredictable. Under the null hypothesis that the failure process is independent and the expected proportion of violations is equal to p, the appropriate likehood ratio is given by:

$$-2\ln\left[(1-p)^{T-N}\, p^{N}\right]+ \\ 2\ln\left[(1-\pi_{01})^{n_{00}}\, \pi_{02}^{n_{01}}\, (1-\pi_{11})^{n_{10}}\, \pi_{11}^{n_{11}}\right] \sim x_2^2\ , \qquad (14)$$

where n_{ij} is the number of observations with value i followed by j, for i, j = 0, 1,

$$\pi_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$$

are the corresponding probabilities. When ij = 1, it means that a violation has been observed, while ij = 0 indicates the opposite. If the sequence of I_t is independent, the probabilities to observe or not a VaR violation in the next period must be equal, which can be written formally as $\pi_{01} = \pi_{11} = p$. Contrary to Kupiec's (1995) test, the conditional coverage procedure can reject a VaR model that generates too many or too few clustered violations.

C. Loss Functions

In addition to the two backtesting measures outlined previously, a risk manager must also be able to evaluate the proposed models according to a utility function. Lopez (1998) introduced the magnitude loss function, which incorporates both the total number of violations and

7. A similar test has also been developed by Engle and Manganelli (2003).

their magnitudes. It is defined as:

$$\Psi_{t+1}^{\text{Lopez}} = \begin{cases} 1 = \left(y_{t+1} - \text{VaR}_{t+1|t}\right)^2, & \text{if } y_{t+1} < \text{VaR}_{t+1|t} \\ 0, & \text{if } y_{t+1} \geq \text{VaR}_{t+1|t} \end{cases} .$$

The magnitude term, $(y_{t+1} - \text{VaR}_{t+1|t})^2$, ensures that the larger the failure is, the higher the penalty added. At the same time, as in Kupiec's (1995) test, a score of one is added whenever a violation occurs. According to Lopez's (1998) loss function, a model that minimizes the total loss ($\Psi = \sum_{t=1}^{T}\Psi_t^{\text{Lopez}}$) should be preferred over others.

Moreover, an inherent problem of risk models is that the "true" Value-at-Risk is never observed, not even after the realization of the actual return. However, this "true" VaR can be proxied using the empirical distribution of realized returns. For example, if T observations are available for the out-of-sample evaluation, their p-quantile will approximate the "true" VaR.[8] The proposed loss function, named Quantile Loss (QL), has the following form:

$$\Psi_{t+1}^{\text{QL}} = \begin{cases} \left(y_{t+1} - \text{VaR}_{t+1|t}\right)^2, & \text{if } y_{t+1} < \text{VaR}_{t+1|t} \\ \left(\text{Quantile}\{y, 100p\}_1^T - \text{VaR}_{t+1|t}\right)^2, & \text{if } y_{t+1} \geq \text{VaR}_{t+1|t} \end{cases} , \quad (15)$$

Given the QL function, a model is penalized either by the magnitude $(y_{t+1} - \text{VaR}_{t+1})^2$ term, if a violation occurs, or by the distance between the p-quantile of the realized future returns and the calculated VaR $(\text{Quantile } \{y, 100p\}_1^T - \text{VaR}_{t+1|t})^2$. Contrary to Lopez's (1998) loss function, QL does not add a score of one if the predicted VaR cannot cover the future loss, since, under the proposed framework, we evaluate only models that have not been rejected by the two statistical backtesting measures.

Furthermore, following the work of Diebold and Mariano (1995), Sarma, Thomas and Shah (2003) and Angelidis, Benos and Degiannakis (2004), we investigate the forecasting ability of the models according

---

8. This proxy will at least meet the unconditional coverage requirement since, by definition, the total number of violations will be equal to the expected one.

to their loss differential $z_{t+1} \equiv \Psi^{\mathrm{QL}}_{A_{t+1}} - \Psi^{\mathrm{QL}}_{B_{t+1}}$, where $\Psi^{\mathrm{QL}}_{A_{t+1}}$ and $\Psi^{\mathrm{QL}}_{B_{t+1}}$ are the loss function indicators of models A and B respectively. The Diebold-Mariano (1995) statistic is the "*t*-statistic" of a regression of z_{t+1} on a constant with heteroskedastic and autocorrelated consistent standard errors.[9]

IV. Empirical Investigation

A. Data

To evaluate all these volatility models, we generate out-of-sample VaR forecasts for four individual Greek shares (Alpha Bank, Emporiki Bank, National Bank of Greece - NBG and Titan Cement Co.), for two equally weighted share portfolios defined below and, finally, for the General Athens Stock Exchange (ASE) index, obtained from DataStream for the period of January 2, 1991 to December 18, 2003.[10] The first portfolio, P_Small, is based on the returns of these four shares, while the second one, P_All, is calculated from all stocks that currently belong to the FTASE-20 Index.[11]

For all equities and portfolios, we compute daily log returns and plot them. Volatility clustering is clearly visible in figure 1, while figure 2 presents the *QQ*-plot against the normal distribution: it shows that all log returns exhibit non-symmetrical fat tails. Hence, any VaR model must account for volatility clustering, excess kurtosis and skewness at the same time. Table 3 provides summary statistics, the Jarque-Bera statistic as well as the Ljung-Box (1978) statistic for serial correlation of the squared returns for up to 10th order. In all cases, the normality hypothesis is rejected at any level of significance, as there is clear evidence of significant excess kurtosis and positive skewness. Moreover, the Ljung-Box test statistic Q^2 (10) indicates the presence of conditional heteroskedasticity. These preliminary descriptive statistics

---

9. For more details about heteroskedastic and autocorrelated consistent (*HAC*) standard errors, see White (1980) and Newey and West (1987).

10. These four stocks represented 24% of the total market capitalization of ASE in 1991, while this percentage was lower (15%) during 2003.

11. The FTASE-20 Index is the large capitalization index of the ASE, made up from the 20 largest companies listed in the Athens Stock Exchange.

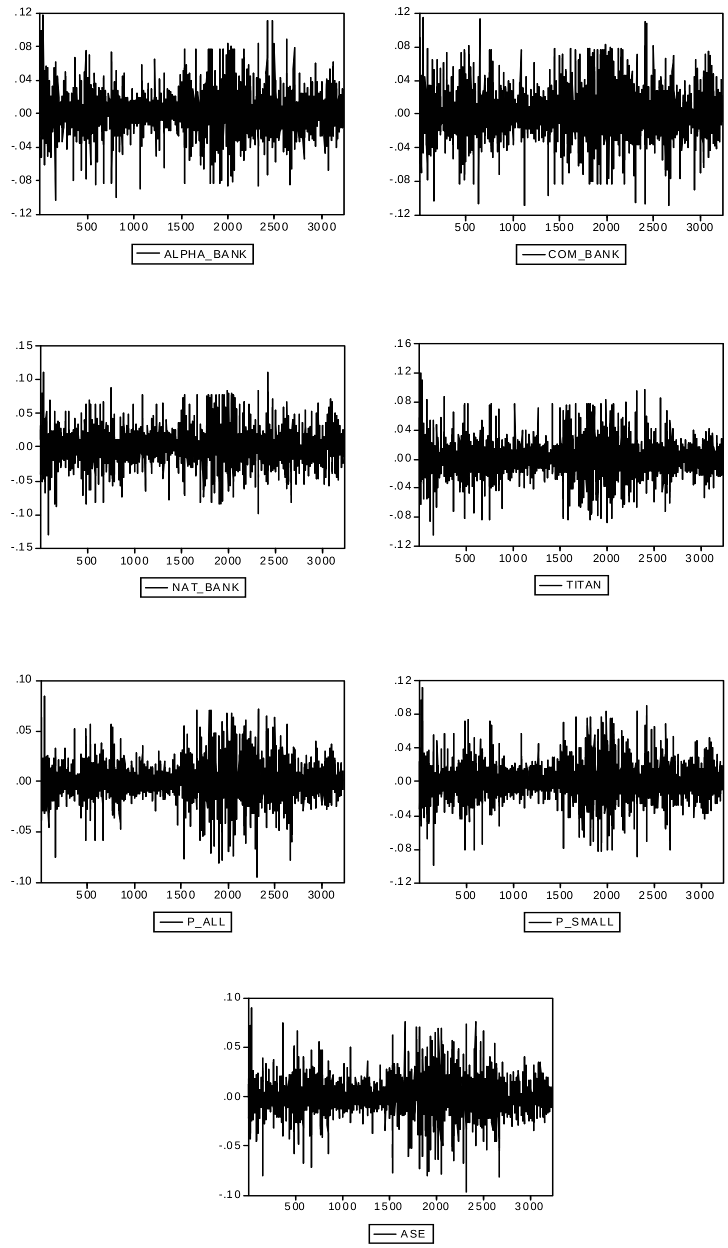

FIGURE 1—Continuously Compounded Daily Returns from January 2, 1991 to December 18, 2003.

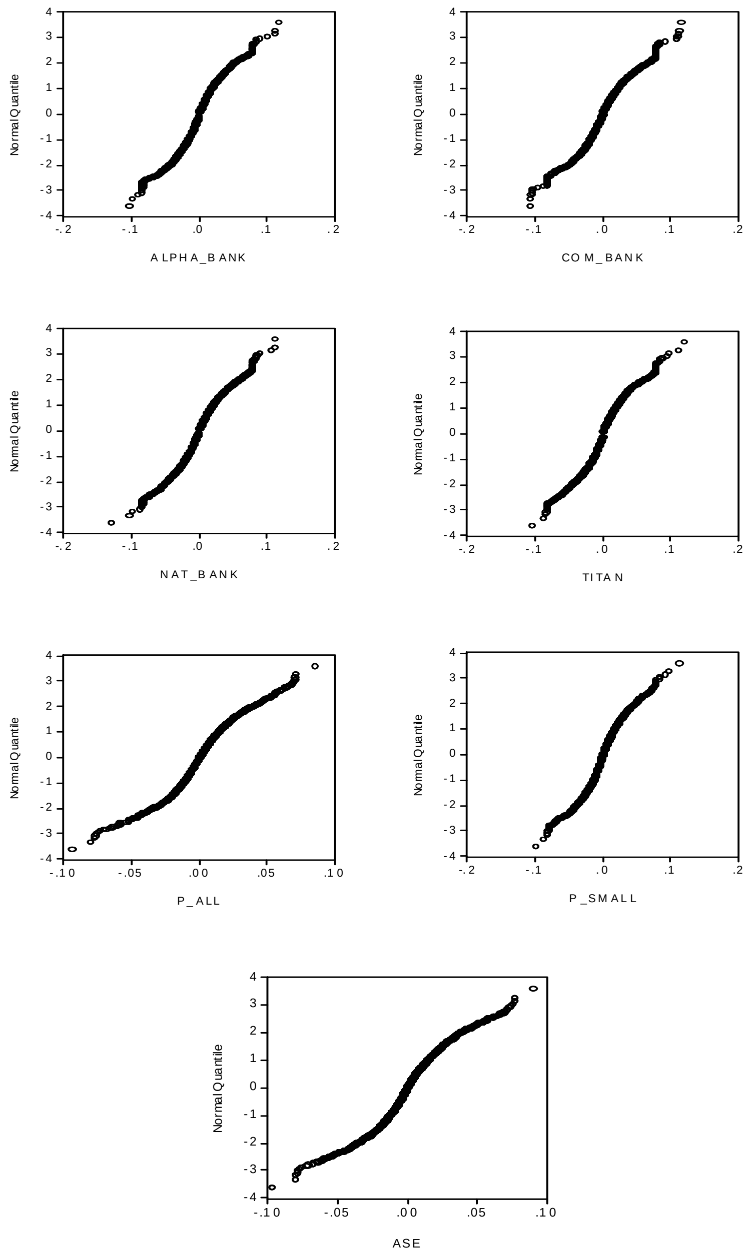

FIGURE 2—*QQ*-plot against the normal distribution. The time period is from January 2, 1991 to December 18, 2003.

TABLE 3. Summary Statistics

	Alpha Bank	Emporiki	NBG	Titan	P_All	P_Small	Ase
Mean	0.05%	0.03%	0.04%	0.06%	0.04%	0.05%	0.03%
Median	0.00%	0.00%	0.00%	0.00%	–0.02%	–0.06%	–0.01%
Maximum	11.67%	11.50%	11.04%	12.00%	8.48%	11.21%	8.98%
Minimum	–10.40%	–10.84%	–13.06%	–10.50%	–9.39%	–9.93%	–9.69%
Std. Dev.	2.18%	2.54%	2.34%	2.15%	1.64%	1.96%	1.70%
Skewness	0.216	0.208	0.210	0.256	0.028	0.212	0.027
Kurtosis	6.170	5.508	5.507	6.308	6.405	6.238	6.516
Jarque-Bera	1380	871	871	1510	1563	1437	1666
Q^2 (10)	642	757	646	633	820	803	733
Observations	3,234	3,234	3,234	3,234	3,234	3,234	3,234

Note: This table presents descriptive statistics of the daily log returns, for the period of January 2, 1991 to December 18, 2003.

demonstrate, on the one hand, that the right and left tails of the empirical distributions are different and therefore, it is interesting to evaluate all risk models both for short and long trading positions, while on the other hand they suggest the use of GARCH modeling, which recognizes temporal dependence in the second moment of daily log returns.

Although there is evidence that stock returns have an asymmetric effect on volatility, one has to perform a formal test to examine the sign and size bias, according to Engle and Ng's (1993) diagnostic procedure:

$$\hat{e}_t^2 = \phi_0 = \phi_1 S_{t-1}^- + \phi_2 S_{t-1}^- \hat{e}_{t-1} + \phi_2 S_{t-1}^+ \hat{e}_{t-1} + \xi_t \,, \tag{16}$$

where S_{t-1}^- is a dummy variable, taking the value 1 if $\hat{e}_{t-1} < 0$ and 0 otherwise. The variable S_{t-1}^+ is simply equal to $1 - S_{t-1}^-$ and, finally, $\hat{e}_{t-1}$ are the residuals of equation 2. If φ_1 is statistically significant, then $\hat{e}_t^2$ depends on the sign of $\hat{e}_{t-1}$, whereas a significant φ_2 or φ_3 indicates that also the size of the shock (e_t) affects the conditional variance. Therefore, a joint test of sign and size effects ($\varphi_1 = \varphi_2 = \varphi_3 = 0$) can be performed based on equation 16. Table 4 shows a significant sign and size effect in the conditional variance and therefore the inclusion of asymmetric components in the volatility specification is supported.

B. Statistical Evaluation of the VaR Models

For all models, all single equities and all portfolios we use a rolling sample of 1000 observations, in order to forecast the 97.5% and 99% VaR values of both long and short trading positions.[12] At each iteration, we compare the predicted VaR number with the realized return, construct the exception variable (I_{t+1}) and the corresponding loss function (Ψ_{t+1}) and use both of them to assess the statistical accuracy of the various risk management techniques.

Exception rates at both confidence levels and the *p*-values of the backtesting measures are presented in tables 5 and 6. Results can be summarized as follows:

a. We often find that the *VC* and the RiskMetrics (*RM*) methods are not appropriate risk management techniques in practice, since, for all

12. In all cases, we work with 3234 observations and generate 2234 out-of-sample forecasts.

TABLE 4. Volatility Specification Test

	φ_0	φ_1	φ_2	φ_3	$\chi^2(3)$
Alpha Bank	0.00018*	0.00004	0.01553*	–0.01951*	56.36*
	(0.00004)	(0.00006)	(0.00301)	(0.00340)	
Emporiki	0.00022*	0.00013	0.01733*	–0.02145*	57.98*
	(0.00007)	(0.00009)	(0.00349)	(0.00454)	
NBG	0.00025*	0.00002	0.01793*	–0.01548*	53.04*
	(0.00005)	(0.00007)	(0.00326)	(0.00362)	
Titan	0.00022*	0.00007	0.01281*	–0.01507*	77.62*
	(0.00003)	(0.00005)	(0.00270)	(0.00228)	
P_All	0.00007**	0.000081**	0.00989*	–0.01729*	73.98*
	(0.00003)	(0.00003)	(0.00163)	(0.00303)	
P_Small	0.00010*	0.000107***	0.01364*	–0.01830*	78.90*
	(0.00004)	(0.00006)	(0.00262)	(0.00324)	
Ase	0.00008*	0.0000823**	0.01040*	–0.01732*	73.19*
	(0.00003)	(0.00004)	(0.00190)	(0.00306)	

Note: This table presents the Engle and Ng (1993) volatility specification test. The test is performed to the residual of equation 2. Standard errors are in parenthesis; *, **, *** indicate significance at the 1%, 5% and 10% levels, respectively. The period runs from January 2, 1991 until December 18, 2003.

cases, they tend to underestimate the "true" VaR and hence are rejected by the two backtesting measures. For the short trading positions and for both confidence levels, the failure rates are statistically different from their theoretical values, due to excess kurtosis and positive skewness of returns (see table 3).

b. As expected, models based on the Normal distribution, such as the GARCH, EGARCH and TARCH, perform better than the *VC* and the *RM* methods. For long positions and for the 97.5% confidence level specifically, the failure rates are statistically equal to their theoretical values. Generally speaking, for short positions and for both two confidence levels, these models still do not produce acceptable VaR forecasts, as they underestimate the "true" VaR (in percentages ranging from 29% to 82%).

c. GARCH models under the Student-*t* and its corresponding Skewed distribution overestimate VaR numbers at both the 97.5% and the 99% level, a result also documented by several studies (see Guermat and Harris [2002], Billio and Pelizzon [2000] among others). Even at the 99% level, they do not offer a major improvement, as average realized exception rates are significantly lower than expected ones. Moreover, in terms of the two backtesting measures, there are no statistical differences between the two distributions used, since for all portfolios, the asymmetry parameter (*log* (ξ)) is very close to zero and consequently, the Skewed Student-*t* distribution is equivalent to the symmetric one.

d. The Historical Simulation method, although it satisfies the “unconditional coverage” prerequisite, it does not meet that of “conditional coverage,” since, for almost all cases, the *p*-value of the corresponding test is less than 10%. More specifically, if a VaR violation occurs one day, the probability to observe another one the following day is high. Hence, we observe clustered violations, as HS does not update the VaR number quickly enough when market volatility rises.

e. The *FHS* and the *EVT* procedures seem to offer a major improvement over parametric methods. Generally speaking, exception rates are too close to the theoretically expected ones, both for long and short trading positions. For example, the average proportion of failures of the *EVT* method at the higher confidence level is 1% exactly (!), while the corresponding proportion for *FHS* is slightly smaller. The improved performance of these models is due to the empirical quantiles being higher than those of the Normal distribution.

f. Consequently, *FHS* and *EVT* structures seem to describe more efficiently the tails of the empirical distribution than the corresponding parametric or non-parametric models. There is strong evidence that the GARCH model under a Normal distribution underestimates the risk at the 99% confidence level, while under the Student-t distribution overestimates it. The introduction of the asymmetry parameter (ξ) does not appear to improve VaR estimations, as it is close to one in most cases.

C. Model Selection

The two backtesting measures cannot directly compare different VaR models, as a greater *p*-value of one model does not indicate the superiority of that model over its competitors. Therefore, in order to evaluate statistically the reported differences, we compute the *QL*

TABLE 5. Evaluation Results: 97.5% VaR Confidence Level

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
A. Exception Rates - Long Positions (%)														
Alpha Bank	2.86	2.69	2.37	2.33	2.33	0.63	0.45	0.45	0.81	0.63	0.85	2.64	2.42	2.64
Emporiki	3.27	2.95	2.46	2.51	2.51	0.76	0.72	0.67	0.98	0.90	0.85	2.91	2.60	2.78
NBG	2.69	3.09	2.46	2.42	2.46	0.72	0.67	0.67	0.85	0.94	0.85	2.55	2.42	2.73
Titan	2.95	3.40	2.73	2.95	2.69	0.18	0.13	0.18	0.18	0.18	0.22	2.64	2.42	2.86
P_All	3.04	3.13	2.01	2.10	1.88	1.30	1.07	1.16	1.30	1.16	1.30	2.86	2.24	2.33
P_Small	3.09	3.04	2.10	2.42	2.15	1.07	1.07	1.07	1.25	1.34	1.16	2.91	2.06	2.33
Ase	3.04	2.91	2.60	2.51	2.51	0.98	1.07	0.98	0.98	1.16	0.98	2.86	2.42	2.82
Average	2.99	3.03	2.39	2.46	2.36	0.81	0.74	0.74	0.91	0.90	0.89	2.77	2.37	2.64
Exception Rates - Short Positions (%)														
Alpha Bank	4.16	3.98	3.27	3.49	3.18	0.90	0.94	0.90	0.63	0.72	0.76	3.00	2.69	3.09
Emporiki	4.79	3.85	3.76	3.80	3.80	0.67	0.72	0.72	0.63	0.58	0.63	3.22	2.82	3.18
NBG	4.57	4.43	3.67	3.94	3.67	0.94	0.90	0.94	0.81	0.85	0.76	2.60	2.46	2.86
Titan	3.54	3.72	3.40	3.58	3.49	0.72	0.81	0.67	0.63	0.72	0.58	2.55	2.82	3.13
P_All	4.43	3.40	2.86	3.09	2.82	1.30	1.25	1.43	1.30	1.39	1.34	3.22	2.46	2.73
P_Small	4.79	3.80	3.04	3.40	3.13	1.25	1.21	1.21	0.98	1.07	1.03	3.36	2.55	2.78
Ase	3.85	3.67	2.51	2.69	2.42	1.07	1.03	1.12	1.03	1.07	1.16	3.00	2.37	2.42
Average	4.30	3.84	3.22	3.43	3.22	0.98	0.98	1.00	0.86	0.91	0.90	2.99	2.60	2.88

(*Continued*)

TABLE 5. (Continued)

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
B. Unconditional Coverage - Long Positions (%)														
Alpha Bank	28.03	57.83	69.69	59.77	59.77	0.00	0.00	0.00	0.00	0.00	0.00	67.23	80.10	67.23
Emporiki	2.64	18.10	90.81	98.38	98.38	0.00	0.00	0.00	0.00	0.00	0.00	22.66	77.22	41.27
NBG	57.83	8.55	90.81	80.10	90.81	0.00	0.00	0.00	0.00	0.00	0.00	87.66	80.10	49.16
Titan	18.10	0.96	49.16	18.10	57.83	0.00	0.00	0.00	0.00	0.00	0.00	67.23	80.10	28.03
P_All	11.11	6.49	12.83	21.78	4.98	0.01	0.00	0.00	0.01	0.00	0.01	28.03	41.97	59.77
P_Small	8.55	11.11	21.78	80.10	27.59	0.00	0.00	0.00	0.00	0.01	0.00	22.66	16.88	59.77
Ase	11.11	22.66	77.22	98.38	98.38	0.00	0.00	0.00	0.00	0.00	0.00	28.03	80.10	34.22
Unconditonal Coverage - Short Positions (%)														
Alpha Bank	0.00	0.00	2.64	0.46	4.87	0.00	0.00	0.00	0.00	0.00	0.00	14.27	57.83	8.55
Emporiki	0.00	0.02	0.04	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	3.60	34.22	4.87
NBG	0.00	0.00	0.09	0.01	0.09	0.00	0.00	0.00	0.00	0.00	0.00	77.22	90.81	28.03
Titan	0.31	0.06	0.96	0.21	0.46	0.00	0.00	0.00	0.00	0.00	0.00	87.66	34.22	6.49
P_All	0.00	0.96	28.03	8.55	34.22	0.01	0.00	0.04	0.01	0.02	0.01	3.60	90.81	49.16
P_Small	0.00	0.02	11.11	0.96	6.49	0.00	0.00	0.00	0.00	0.00	0.00	1.36	87.66	41.27
Ase	0.02	0.09	98.38	57.83	80.10	0.00	0.00	0.00	0.00	0.00	0.00	14.27	69.69	80.10

Note: For all models, this table presents the average exception rate. The models are successively Variance-Covariance (*VC*), RiskMetrics (*RM*), Normal Garch (G-N), Normal Egarch (E-N), Normal Tarch (T-N), Student-T Garch (G-T), Student-T Egarch (E-T), Student-T Tarch (T-T),

(*Continued*)

TABLE 5. (Continued)

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	HS	*FHS*	*EVT*
C. Conditional Coverage - Long Positions (%)														
Alpha Bank	0.00	0.00	0.56	0.06	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.33
Emporiki	0.00	0.00	84.09	4.86	17.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	50.58	2.04
NBG	0.37	1.69	84.09	13.78	84.09	0.00	0.00	0.00	0.00	0.00	0.00	5.51	79.93	74.27
Titan	6.51	0.02	48.43	39.64	50.13	0.00	0.00	0.00	0.00	0.00	0.00	6.62	40.13	38.89
P_All	0.00	4.21	18.39	2.84	7.23	0.01	0.00	0.00	0.01	0.00	0.01	0.00	22.34	2.33
P_Small	0.00	5.43	11.21	13.78	14.46	0.00	0.00	0.00	0.01	0.05	0.00	0.00	8.47	31.35
Ase	0.00	0.11	20.31	1.06	17.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	13.78	2.11
Conditional Coverage - Short Positions (%)														
Alpha Bank	0.00	0.00	2.51	0.13	3.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	50.13	4.83
Emporiki	0.00	0.00	0.06	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.02	20.51	3.61
NBG	0.00	0.00	0.02	0.01	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.70	19.35
Titan	0.00	0.23	3.12	0.73	1.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	61.03	17.46
P_All	0.00	1.28	38.89	10.94	61.03	0.02	0.01	0.13	0.02	0.08	0.04	0.00	91.84	74.27
P_Small	0.00	0.01	12.72	2.35	15.29	0.01	0.01	0.01	0.00	0.00	0.00	0.00	87.40	45.88
Ase	0.00	0.11	23.08	79.64	24.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	87.84	24.79

Note: Skewed Student-T Garch (G-ST), Skewed Student-T Egarch (E-ST), Skewed Student-T Tarch (T-ST), Historical Simulation (*HS*), Filtered Historical Simulation (*FHS*) and Extreme Value Theory (*EVT*). Note that a *P*-value greater than 10% indicates that the forecasting ability of the corresponding VaR model is adequate.

TABLE 6. Evaluation Results: 99% VaR Confidence Level

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
A. Exception Rates - Long Positions (%)														
Alpha Bank	1.92	1.61	1.34	1.43	1.25	0.18	0.13	0.18	0.22	0.22	0.22	1.07	0.90	0.94
Emporiki	2.28	1.57	1.52	1.52	1.52	0.31	0.27	0.31	0.40	0.45	0.40	1.48	1.34	1.16
NBG	1.48	1.52	1.34	1.43	1.39	0.13	0.13	0.13	0.27	0.22	0.27	0.90	0.90	0.94
Titan	1.79	1.79	1.52	1.57	1.57	0.00	0.00	0.00	0.00	0.00	0.00	1.07	0.94	0.98
P_All	2.28	1.79	1.30	1.21	1.25	0.49	0.54	0.54	0.58	0.63	0.58	1.30	0.98	0.98
P_Small	1.92	1.75	1.25	1.30	1.21	0.27	0.27	0.31	0.40	0.36	0.40	1.30	0.90	1.03
Ase	2.06	1.88	1.34	1.52	1.30	0.40	0.49	0.40	0.45	0.45	0.45	1.07	0.90	0.98
Average	1.96	1.70	1.37	1.43	1.36	0.26	0.26	0.27	0.33	0.33	0.33	1.17	0.98	1.00
Exception Rates - Short Positions (%)														
Alpha Bank	2.86	2.42	1.75	1.97	1.84	0.27	0.22	0.27	0.22	0.18	0.22	1.39	1.03	0.98
Emporiki	3.27	2.19	2.01	1.92	1.97	0.31	0.27	0.31	0.22	0.18	0.22	1.43	1.03	1.07
NBG	2.82	2.33	2.15	2.19	2.15	0.22	0.22	0.18	0.13	0.13	0.13	1.39	1.07	1.25
Titan	2.42	2.10	2.15	2.10	2.01	0.27	0.27	0.27	0.27	0.27	0.27	1.12	0.98	0.94
P_All	2.55	1.79	1.34	1.57	1.34	0.45	0.40	0.45	0.36	0.40	0.31	1.30	0.90	0.94
P_Small	2.91	2.19	1.75	1.57	1.61	0.40	0.40	0.54	0.40	0.40	0.40	1.07	0.94	0.90
Ase	2.28	1.84	1.25	1.39	1.34	0.40	0.36	0.45	0.40	0.36	0.40	1.07	0.98	0.94
Average	2.73	2.12	1.77	1.82	1.75	0.33	0.31	0.35	0.29	0.28	0.28	1.25	0.99	1.00

(*Continued*)

TABLE 6. (Continued)

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
B. Unconditional Coverage - Long Positions (%)														
Alpha Bank	0.01	0.76	12.17	5.37	24.69	0.00	0.00	0.00	0.00	0.00	0.00	72.73	61.25	77.35
Emporiki	0.00	1.29	2.13	2.13	2.13	0.01	0.00	0.01	0.13	0.32	0.13	3.43	12.17	44.81
NBG	3.43	2.13	12.17	5.37	8.20	0.00	0.00	0.00	0.00	0.00	0.00	61.25	61.25	77.35
Titan	0.07	0.07	2.13	1.29	1.29	0.00	0.00	0.00	0.00	0.00	0.00	72.73	77.35	94.22
P_All	0.00	0.07	17.57	33.73	24.69	0.75	1.59	1.59	3.12	5.69	3.12	17.57	94.22	94.22
P_Small	0.01	0.14	24.69	17.57	33.73	0.00	0.00	0.01	0.13	0.04	0.13	17.57	61.25	88.89
Ase	0.00	0.02	12.17	2.13	17.57	0.13	0.75	0.13	0.32	0.32	0.32	72.73	61.25	94.22
Unconditional Coverage - Short Positions (%)														
Alpha Bank	0.00	0.00	0.14	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	8.20	88.89	94.22
Emporiki	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	5.37	88.89	72.73
NBG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.20	72.73	24.69
Titan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	57.89	94.22	77.35
P_All	0.00	0.07	12.17	1.29	12.17	0.32	0.13	0.32	0.04	0.13	0.01	17.57	61.25	77.35
P_Small	0.00	0.00	0.14	1.29	0.76	0.13	0.13	1.59	0.13	0.13	0.13	72.73	77.35	61.25
Ase	0.00	0.04	24.69	8.20	12.17	0.13	0.04	0.32	0.13	0.04	0.13	72.73	94.22	77.35

Note: For all models, this table presents the average exception rate. The models are successively Variance-Covariance (*VC*), RiskMetrics (*RM*), Normal Garch (G-N), Normal Egarch (E-N), Normal Tarch (T-N), Student-T Garch (G-T), Student-T Egarch (E-T), Student-T Tarch (T-T),

(*Continued*)

TABLE 6. (Continued)

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
C. Conditional Coverage - Long Positions (%)														
Alpha Bank	0.00	0.00	21.54	3.49	33.41	0.00	0.00	0.00	0.00	0.00	0.00	0.04	34.28	40.58
Emporiki	0.00	3.83	5.79	4.11	4.11	0.07	0.02	0.07	0.53	1.25	0.53	0.00	21.54	54.57
NBG	0.05	5.79	21.54	11.97	16.35	0.00	0.00	0.00	0.02	0.00	0.02	0.20	34.28	77.86
Titan	0.31	0.01	5.79	3.83	3.83	0.00	0.00	0.00	0.00	0.00	0.00	71.72	77.86	79.34
P_All	0.00	0.00	27.35	39.24	33.41	2.64	5.10	5.10	9.04	14.84	9.04	0.00	45.52	45.52
P_Small	0.00	0.24	33.41	27.35	39.24	0.02	0.02	0.07	0.53	0.20	0.53	0.00	34.28	48.52
Ase	0.00	0.01	5.47	0.04	6.46	0.53	2.64	0.53	1.25	1.25	1.25	0.00	34.28	45.52
Conditional Coverage - Short Positions (%)														
Alpha Bank	0.00	0.00	0.54	0.02	0.02	0.02	0.00	0.02	0.00	0.00	0.00	0.07	77.15	79.34
Emporiki	0.00	0.00	0.01	0.05	0.02	0.07	0.02	0.07	0.00	0.00	0.00	0.56	77.15	71.72
NBG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	71.72	35.40
Titan	0.00	0.00	0.00	0.00	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.76	45.52	77.86
P_All	0.00	0.31	19.79	2.56	19.79	1.25	0.53	1.25	0.20	0.53	0.07	0.07	72.76	77.86
P_Small	0.00	0.00	0.54	2.56	2.46	0.53	0.53	5.10	0.53	0.53	0.53	0.04	77.86	72.76
Ase	0.00	0.17	35.40	14.04	19.79	0.53	0.20	1.25	0.53	0.20	0.53	0.66	79.34	77.86

Note: Skewed Student-T Garch (G-ST), Skewed Student-T Egarch (E-ST), Skewed Student-T Tarch (T-ST), Historical Simulation (*HS*), Filtered Historical Simulation (*FHS*) and Extreme Value Theory (*EVT*). Note that a *P*-value greater than 10% indicates that the forecasting ability of the corresponding VaR model is adequate.

TABLE 7. Loss Function Results: 97.5% VaR Confidence Level

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
A. Loss Value - Long Positions														
Alpha Bank														
Emporiki			0.66		**0.65**								0.69	
NBG			0.46	**0.40**	0.47								0.46	0.45
Titan			0.71	**0.61**	0.71								0.71	0.68
P_All			**0.48**										0.55	
P_Small			0.53	**0.47**	0.53									0.54
Ase			**0.44**		0.45								0.46	
B. T-Statistic - Long Positions														
Alpha Bank														
Emporiki			–0.35										–2.99*	
NBG			–4.10*		–4.37*								–4.00*	–4.68*
Titan			–4.48*		–4.60*								–3.01*	–4.08*
P_All													–6.74*	
P_Small			–4.44*		–4.42*									–6.18*
Ase					–1.33								–2.31**	

Note: *, **, *** indicate significance at the 1%, 5% and 10% levels, respectively. The models with the lowest loss values are boldfaced. The models are successively Variance-Covariance (*VC*), RiskMetrics (*RM*), Normal Garch (G-N), Normal Egarch (E-N), Normal Tarch (T-N), Student-T

(*Continued*)

TABLE 7. (Continued)

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
C. Loss Value - Short Positions														
Alpha Bank													**0.69**	
Emporiki													**1.03**	
NBG													**0.72**	0.76
Titan													**0.89**	
P_All			0.54		0.55								**0.48**	0.50
P_Small			0.67										**0.62**	0.63
Ase			0.45	**0.37**	0.46								0.40	0.42
D. T-Statistics - Short Positions														
Alpha Bank														
Emporiki														
NBG														–4.18*
Titan														
P_All			–10.95*		–10.46*									–6.56*
P_Small			–3.34*											–1.30
Ase			–4.59*		–4.88*								–1.72***	–2.86*

Note: Garch (G-T), Student-T Egarch (E-T), Student-T Tarch (T-T), Skewed Student-T Garch (G-ST), Skewed Student-T Egarch (E-ST), Skewed Student-T Tarch (T-ST), Historical Simulation (*HS*), Filtered Historical Simulation (*FHS*) and Extreme Value Theory (*EVT*).

TABLE 8. Loss Function Results: 99% VaR Confidence Level

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
A. Loss Value - Long Positions														
Alpha Bank			0.98		**0.97**								1.07	1.05
Emporiki													**1.72**	1.74
NBG			**0.72**										0.74	0.79
Titan												**0.26**	1.10	1.12
P_All			0.86	**0.82**	0.88								0.93	0.93
P_Small			1.04	**0.97**	1.03								1.06	0.99
Ase													1.01	**0.97**
B. T-Statistic - Long Positions														
Alpha Bank			–1.27	–1.48	–1.27									
Emporiki														–1.01
NBG													–0.58	–1.98**
Titan													–11.69*	–11.56*
P_All			–3.19*		–3.84*								–2.60*	–3.27*
P_Small			–4.02*		–3.97*								–1.47	–0.57
Ase														–3.78*

Note: *, **, *** indicate significance at the 1%, 5% and 10% levels, respectively. The models with the lowest loss values are boldfaced. The models are successively Variance-Covariance (*VC*), RiskMetrics (*RM*), Normal Garch (G-N), Normal Egarch (E-N), Normal Tarch (T-N), Student-T

(*Continued*)

TABLE 8. (Continued)

	VC	*RM*	G-N	E-N	T-N	G-T	E-T	T-T	G-ST	E-ST	T-ST	*HS*	*FHS*	*EVT*
C. Loss Value - Short Positions														
Alpha Bank													**1.65**	1.68
Emporiki													**1.39**	1.46
NBG													1.41	**1.40**
Titan													**1.74**	1.88
P_All			0.89		0.90								**0.81**	0.81
P_Small													**1.12**	1.13
Ase			0.91		0.92								**0.79**	0.83
D. T-Statistics - Short Positions														
Alpha Bank														–2.35**
Emporiki														–3.60*
NBG													–0.34	
Titan														–6.23*
P_All			–3.49*	–4.01*										–0.51
P_Small														–1.28
Ase			–7.48*		–6.79*									–4.62*

Note: Garch (G-T), Student-T Egarch (E-T), Student-T Tarch (T-T), Skewed Student-T Garch (G-ST), Skewed Student-T Egarch (E-ST), Skewed Student-T Tarch (T-ST), Historical Simulation (*HS*), Filtered Historical Simulation (*FHS*) and Extreme Value Theory (*EVT*).

TABLE 9. Model Selection Based on the Loss Function Procedure

	Unconditional Coverage	Conditional Coverage	Selected based on the loss function
97.5% VaR Confidence Level - Long Positions			
Alpha Bank	VC, RM, G-N, E-N, T-N, HS, FHS, EVT		
Emporiki	RM, G-N, E-N, T-N, HS, FHS, EVT	G-N, T-N, FHS	G-N, **T-N**
NBG	VC, G-N, E-N, T-N, HS, FHS, EVT	G-N, E-N, T-N, FHS, EVT	**E-N**
Titan	VC, G-N, E-N, T-N, HS, FHS, EVT	G-N, E-N, T-N, FHS, EVT	**E-N**
P_All	VC, RM, G-N, E-N, HS, FHS, EVT	G-N, FHS	**G-N**
P_Small	RM, G-N, E-N, T-N, HS, FHS, EVT	G-N, E-N, T-N, EVT	**E-N**
Ase	VC, RM, G-N, E-N, T-N, HS, FHS, EVT	G-N, T-N, FHS	**G-N**, T-N
97.5% VaR Confidence Level - Short Positions			
Alpha Bank	HS, FHS	FHS	**FHS**
Emporiki	FHS	FHS	**FHS**
NBG	HS, FHS, EVT	FHS, EVT	**FHS**
Titan	HS, FHS	FHS	**FHS**
P_All	G-N, T-N, FHS, EVT	G-N, T-N, FHS, EVT	**FHS**
P_Small	G-N, FHS, EVT	G-N, FHS, EVT	**FHS**, EVT
Ase	G-N, E-N, T-N, HS, FHS, EVT	G-N, E-N, T-N, FHS, EVT	**E-N**

Note: This table summarizes the model selection procedure. The models are successively Variance-Covariance (*VC*), RiskMetrics (*RM*), Normal Garch (G-N), Normal Egarch (E-N), Normal Tarch (T-N), Student-T Garch (G-T), Student-T Egarch (E-T), Student-T Tarch (T-T),

(*Continued*)

TABLE 9. (Continued)

	Unconditional Coverage	Conditional Coverage	Selected based on the loss function
99% VaR Confidence Level - Long Positions			
Alpha Bank	G-N, T-N, HS, FHS, EVT	G-N, T-N, FHS, EVT	G-N, **T-N**, FHS, EVT
Emporiki	FHS, EVT	FHS, EVT	**FHS**, EVT
NBG	G-N, HS, FHS, EVT	G-N, FHS, EVT	**G-N**, FHS
Titan	HS, FHS, EVT	HS, FHS, EVT	**HS**
P_All	G-N, E-N, T-N, HS, FHS, EVT	G-N, E-N, T-N, FHS, EVT	**E-N**
P_Small	G-N, E-N, T-N, HS, FHS, EVT	G-N, E-N, T-N, FHS, EVT	**E-N**, FHS, EVT
Ase	G-N, T-N, HS, FHS, EVT	FHS, EVT	**EVT**
99% VaR Confidence Level - Short Positions			
Alpha Bank	FHS, EVT	FHS, EVT	**FHS**
Emporiki	FHS, EVT	FHS, EVT	**FHS**
NBG	FHS, EVT	FHS, EVT	FHS, **EVT**
Titan	HS, FHS, EVT	FHS, EVT	**FHS**
P_All	G-N, T-N, HS, FHS, EVT	G-N, T-N, FHS, EVT	**FHS**
P_Small	HS, FHS, EVT	FHS, EVT	**FHS**, EVT
Ase	G-N, T-N, HS, FHS, EVT	G-N, T-N, FHS, EVT	**FHS**

Note: Skewed Student-T Garch (G-ST), Skewed Student-T Egarch (E-ST), Skewed Student-T Tarch (T-ST), Historical Simulation (HS), Filtered Historical Simulation (*FHS*) and Extreme Value Theory (*EVT*). The best performing models are boldfaced.

function and carry out the equality test that was described in section III.C for each model that produced a *p*-value for both tests greater than 10%. We preferred such a high cutoff point for *p*-values to ensure that (a) "successful" models will not statistically over / under estimate "true" VaR, as a high (low) VaR will imply in practice that the financial institution allocates more (less) capital than actually necessary and, (b) VaR violations will not be clustered. Furthermore, by increasing the level of significance to 10%, we can more easily reject an incorrect model, which could otherwise be costly for a risk manager.[13]

Tables 7 and 8 summarize results of the loss function approach. For example, in panel A of table 7, the model with the smallest loss value for Emporiki Bank is the TARCH under the Normal distribution (T-N), while the other two models that have not been rejected by the two backtesting criteria are the GARCH with a Normal distribution (G-N) and the *FHS*. Based on panel B of table 7, we conclude that the differences between the T-N and the G-N models are not statistically significant, yet the former and the *FHS* are not equivalent approaches.

For short positions, in most cases, the *FHS* method is preferred over the others, since it is either the only method producing accurate forecasts, or it minimizes the value of the loss function. For long positions, results are mixed: no model seems to systematically produce globally acceptable VaR estimates, as almost for each equity and each portfolio there is a different model that is characterized as a preferred one. Nevertheless, based on the proposed model selection procedure, we manage to reduce the chosen models to a smaller set.

Finally, in order to summarize the model selection procedure, we present in table 9, the two stages we followed. In the first two columns (2 and 3), we list the models that have not been rejected by the statistical backtesting measures, while in the last one (column 4), we list the volatility methods that were preferred over the others based on the loss function method.

V. Conclusion

In this paper we examine different Value-at-Risk estimation methods. Using an out-of-sample testing framework, we compare:

1. Parametric methods, i.e., the Variance Covariance and the

---

13. The Type II error here is set equal to 10%.

RiskMetricsTM approaches with GARCH, EGARCH and TARCH volatility modeling, under the Normal, the Student-*t* and the Skewed Student-*t* distributions;

2. Non Parametric methods, i.e., the Historical Simulation approach, and;

3. Semi-Parametric methods, i.e., the Filtered Historical Simulation and the *EVT* procedures both for long and short trading positions. At the 99% confidence level, the *FHS* method performs better than the other ones, as forecasts accurately the corresponding VaR values. For the same confidence level, the *EVT* method also produces acceptable results. On the hand, at the lower confidence level, most models give similar and good results.

As backtesting procedures are not powerful enough to identify a unique model for each individual equity or portfolio, at each different confidence level and for each trading position, we go on to explicitly develop a loss function to evaluate those models that have met two standard prerequisites: that of a correct "unconditional" and of "conditional" coverage. Under this new framework, a model that minimizes total loss is preferred over the others. By subsequently implementing a test for the differences of forecast errors, we provide statistical inference for the forecasting ability of all acceptable models. In most cases, there are significant differences between them. Given that we started initially with 14 possible combinations, we managed to reduce them to a smaller set using this procedure and, in some cases, even to identify a unique optimal risk management technique.

References

Angelidis, T., Benos, A. and Degiannakis, S. 2004. The use of GARCH models in VaR estimation. *Statistical Methodology* Vol. 1, No. 2: 105–128.

Artzner, P., F; Delbaen, J.-M. Eber; and Heath, D. 1997. Thinking coherently. *Risk* 10: 68–71.

Artzner, P., F; Delbaen, J.-M. Eber; and Heath, D. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203-228.

Bali, T. G., and Theodossiou, P. 2004. A Conditional-SGT-VaR approach with alternative GARCH models. Forthcoming in Annals of Operations Research.

Bali, T. G. 2003. An extreme value approach to estimating volatility and value at risk. *Journal of Business* 76: 83–108.

Balkema, A., and de Hann, L. 1974. Residual lifetime at great age. *Annals of Probability* 2: 792–804.

Barone-Adesi, G.; Giannopoulos, K.; and Vosper, L. 1999. VaR without correlations for nonlinear portfolios. *Journal of Futures Markets* 19: 583–602.

Barone-Adesi, G., and Giannopoulos, K. 2001. Non-parametric VaR techniques. Myths and realities. Economic Notes by Banca Monte dei Paschi di Siena SpA 30: 167–181.

Basak, S., and Shapiro, A. 2001. Value-at-Risk-based risk management: Optimal policies and asset prices. *Review of Financial Studies* 14: 371–405.

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

Bollerslev, T. 1987. A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics* 69: 542–547.

Brooks, C., and Persand, G. 2003. The effect of asymmetries on stock index return Value-at-Risk estimates. *The Journal of Risk Finance* (Winter): 29–42.

Byström, H.N.E. 2004. Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis* Volume 13, Issue 2, Summer: 133–152

Cabedo, D. J. and Moya, I. 2003. Estimating oil price ‘Value at Risk’ using the historical simulation approach. *Energy Economics* 25: 239–253.

Christoffersen, P. 1998. Evaluating interval forecasts. *International Economic Review* 39: 841–862.

Christoffersen, P. 2003. *Elements of Financial Risk Management*. Academic Press.

Daníelsson, J. 2002. The emperor has no clothes: Limits to risk modelling. *Journal of Banking & Finance* 26: 1273–1296.

Delbaen, F. 1998. Coherent risk measures on general probability spaces. Working Paper, ETH Zurich.

Diebold, F.X. and Mariano, R. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13(3): 253–263.

Engle, R.F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50: 987–1008.

Engle, R.F. and S. Manganelli. (2003). CAViaR: conditional autoregressive Value at Risk by regression quantiles. Forthcoming in *Journal of Business and Economic Statistics*.

Engle, R.F. and Ng, V. 1993. Measuring and testing the impact of news on volatility. *Journal of Finance* 48: 1749–1778.

Frey, R. and Michaud, P. 1997. The effect of GARCH-type volatilities on prices and payoff-distributions of derivative assets - A simulation study. Working Paper, ETH Zurich.

Gençay, R., and Selçuk, F. 2004. Extreme value theory and Value-at-Risk:

Relative performance in emerging markets. *International Journal of Forecasting* Volume 20, Issue 2, April-June: 287–303.

Giot, P. and Laurent, S. 2003a. Value-at-Risk for long and short trading positions. *Journal of Applied Econometrics* 18: 641–664.

Giot, P. and Laurent, S. 2003b. Market risk in commodity markets: A VaR approach. *Energy Economics* 25: 435–457.

Guermat, C. and Harris, R.D.F. 2002. Forecasting Value-at-Risk allowing for time variation in the variance and kurtosis of portfolio returns. *International Journal of Forecasting* 18: 409–419.

Hendricks, D. 1996. Evaluation of Value-at-Risk models using historical data. *Economic Police Review* 2: 39–70.

Ho, L.-C.; Burridge, P.; Cadle, J.; and Theobald, M. 2000. Value-at-Risk: Applying the extreme value approach to Asian markets in the recent financial turmoil. *Pacific-Basin Finance Journal* 88: 249–275.

Hoppe, R. 1998. VAR and the unreal world. *Risk* 11: 45–50.

Hull, J., and White, A. 1998. Incorporating volatility updating into the historical simulation method for VaR. *Journal of Risk* 1: 5–19.

Jondeau, E. and Rockinger, M. 1999. The tail behavior of stock returns: Emerging versus mature markets. Working Paper, HEC-School of Management, Department of Finance.

Jondeau, E. and Rockinger, M. 2003. Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance* 10: 559–581.

Kupiec, P.H. 1995. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives* 3: 73–84.

Lambadiaris, G.; Papadopoulou, L.; Skiadopoulos, G.; and Zoulis, Y. 2003. VAR: history or simulation? *Risk* 16 (September): 122–127.

Lambert, P. and Laurent, S. 2000. Modeling skewness dynamics in series of financial data. Discussion Paper, Institut de Statistique, Louvain-la-Neuve.

Lambert, P. and Laurent, S. 2001. Modeling financial time series using Garch-type models and a skewed student Density. Mimeo, Université de Liège.

Ljung, C. M. and Box, G. E. P. 1978. On measuring of lag in time series models. *Biometrics* 67: 297–303.

Lopez, J.A. 1998. Methods for evaluating Value-at-Risk estimates. Federal Reserve Bank of New York, Economic Policy Review.

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

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

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

Pickands, J. 1975. Statistical inference using extreme order statistics. *Annals of*

Statistic 3: 119–131.
Poon, S.-H., and Granger, W.J. 2003. Forecasting volatility in financial markets: A review. *Journal of Economic Literature* XLI: 478–539.
Rozario, D.,R. 2002. Estimating Value at Risk for the electricity market using a technique from extreme value theory. Working Paper. School of Banking and Finance, University of NSW.
Sarma, M.; Thomas S.; and Shah., A. 2003. Selection of VaR models. *Journal of Forecasting* 22,4: 337–358.
Schwert, G. W. 1989. Why does stock market volatility change over time? *Journal of Finance* 44: 1115–1153.
Seymour, A. J., and Polakow, D. A. 2003. A coupling of extreme-value theory and volatility updating with Value-at-Risk Estimation in emerging markets: A South African test. *Multinational Finance Journal* 7 (1&2): 3–23.
Taylor, S. 1986. Modeling Financial Time Series. New York: Wiley.
Theodossiou, P. 1998. Financial data andd the skewed generalized t distribution. *Management Science* 44 (12-1) (December): 1650–1661.
Venkataraman, S. 1996. Value at Risk for a mixture of normal distributions: The use of quasi-bayesian estimation techniques. Economic Perspectives, Federal Reserve Bank of Chicago (March/April): 2–13.
Van den Goorbergh, R.W.J. and Vlaar, P. 1999. Value-at-Risk analysis of stock returns. Historical simulation, variance techniques or tail index estimation? DNB Staff Reports 40, Netherlands Central Bank.
Vlaar, P. 2000. Value at Risk models for Dutch bond portfolios. *Journal of Banking and Finance* 24: 131–154.
White, H. 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica* 48 :817–838.
Zangari, P. 1996. An improved methodology for measuring VAR. RiskMetrics Monitor, Reuters/JP Morgan.