YOUNG ECONOMISTS' SEMINAR to
14th Dubrovnik Economic Conference

Organised by

CROATIAN NATIONAL BANK

Saša Žiković

Friends and Foes: A Story of Value at Risk and Expected Tail Loss

Hotel "Grand Villa Argentina",
Dubrovnik
June 25, 2008

Draft version

Please do not quote

Friends or foes: A story of Value at Risk and Expected tail loss[1]

Saša Žiković[2]

First Draft: April 2008

Abstract

Since the introduction by Artzner (1997, 1999) of coherent risk measures and the new dawn of measuring extreme losses it seems as if the academic community is willing to sacrifice and erase all the advances made in the field of measuring VaR. Expected tail loss (ETL) as a risk measure is superior to VaR since VaR is not a coherent risk measure, in that it does not encourage diversification and completely ignores extreme events. On the other hand, sophistication of VaR models has reached formidable levels, while ETL estimation is still in its beginnings and usually it is calculated via parametric approach based on Generalized extreme value distribution or nonparametric approaches such as fitting kernels to historical simulation. I propose using advanced VaR calculation techniques to produce superior ETL forecasts that are reactive to changes in volatility levels, do not presume any theoretical distribution of the tails and are not limited to the historical data set. The proposed ETL measure is semi-parametric in its nature and uses volatility transformed tail data. The paper also proposes a new loss function that relates tail losses to ETL forecasts. Testing of VaR and ETL models is performed on US and European stock indexes (DOW JONES, NASDAQ, S&P500 (US), FTSE (Great Britain), DAX (Germany) and CAC (France)).

Key words: Coherent risk measures, Expected tail loss, Value at Risk, Extreme value theory, Bootstrapping

JEL classification: G24, C14, C22, C52, C53

[1] This research is supported by Ministry of Science, Education and Sports of the Republic of Croatia, research project No. 081-0000000-1264.
[2] Department of Finance and Banking, Faculty of Economics, University of Rijeka, Croatia, E-mail: szikovic@efri.hr

1. Introduction

Since its' very introduction VaR as a risk measure is open to criticism from different directions. With the latest market turmoil stemming from US sub-prime mortgage crises it is clear that there is a need for an approach that comes to terms with problems posed by extreme event estimation. Artzner et al. (1997, 1999) used an axiomatic approach to the problem of defining a satisfactory risk measure. They defined attributes that any good risk measure should satisfy, and call risk measures that satisfy these axioms "coherent". It turned out that VaR is not a coherent risk measure because it does not necessarily satisfy the sub-additivity condition set out by Artzner et al. (1999). VaR can only be made sub-additive if a usually implausible assumption is imposed on returns being normally (or more generally, elliptically) distributed. Sub-additivity expresses the fact that a portfolio will risk an amount, which is at most the sum of the separate amounts risked by its subportfolios. Even though VaR's theoretical flaws outweigh its practical advantages, it is a regulatory obligation. Banks have to calculate their VaR figures to construct adequate capital requirements. A serious shortcoming of VaR is that it provides no handle on the extent of the losses that might be suffered beyond a certain threshold confidence level. VaR is incapable of distinguishing between situations where losses in the tail are only a bit worse, and those where they are overwhelming. It provides the lowest bound for losses in the tail of the loss distribution and has a bias toward optimism instead of the conservatism that ought to prevail in risk management. An alternative measure that does quantify the losses that might be encountered in the tail is the expected tail loss (ETL). While VaR represents a maximum loss one expects at a determined confidence level during a given holding period, ETL is the loss one expects to suffer, provided that the loss is equal to or greater than VaR. ETL is a coherent measure of risk, while VaR is not. At present, ETL is not approved by the regulators as a risk measure that can be used to calculate economic capital. The field of ETL estimation and model comparison is just beginning to develop and there is an obvious lack of empirical research. After all, VaR and ETL are inherently connected in the sense that from the VaR surface of the tail ETL figures can be easily calculated. We are only beginning to investigate this new and exciting area in risk management science. Advances that have been made in VaR should not be lost with the probable (and well deserved) adoption of coherent risk measures into regulatory framework. Superior quality of VaR techniques can be employed to yield superior ETL forecasts. Since ETL estimations can be significantly improved by using the knowledge obtained from advances in VaR estimation VaR and ETL should be regarded as partners not rivals.

Since ETL is a relatively new risk measure there is a limited number of papers dealing with empirical testing of ETL models. The goal of this paper is to extend the advances that have been made in the field of VaR to ETL estimation. Contributions of this paper are the following: an empirical investigation into relatively uncharted waters of tail risk assessment based on leading US and European stock indexes, a new hybrid approach to estimating ETL is developed, a new loss function is proposed that is tailored made for evaluating ETL forecasts. The main goal of the paper is to find a ETL-VaR model that at the same time accurately calculates VaR figures and gives the best approximation to tail losses i.e. minimizes the deviation of ETL forecasts from actual extreme losses. The rest of the paper is organized as follows: In section 2 a review of the literature on VaR and ETL estimation and model comparison is presented. In section 3 the logic behind VaR is explained as well as its advantages and disadvantages. Section 4 introduces the concept of coherent risk measures, extreme value theory and a measure of average expected loss in the tail (ETL). The consequences of coherence are discussed and strong points of ETL are presented. Section 4 also presents a new approach to measuring ETL, based on bootstrapping from transformed tail losses. Section 5 outlines the methodology used in the testing of VaR and ETL models, and develops a new loss function that translates ETL estimates to realized tail losses. Findings and backtesting results are presented and discussed in section 6. Section 7 concludes.

2. Literature review

Literature about VaR model comparison is extensive and very broadly based, covering risks inherent in stock, bonds, derivatives, FX, commodities etc. The findings are often contradictory and vary substantially depending on loss function, time period and confidence levels. Wong, Cheng, Wong (2002) conclude that GARCH models, often found superior in forecasting volatility, consistently fail the Basel backtest. Several papers investigate the issue of trade-off in model choice; for example Caporin (2003) finds that the EWMA compared to GARCH-based VaR forecast provides the best efficiency at a lower level of complexity. Bams, Wielhouwer (2000) draw similar conclusions, although sophisticated tail modeling results in better VaR estimates but with more uncertainty. They conclude that if the data generating process is close to be integrated, the use of the more general GARCH model introduces estimation error, which might result in the superiority of EWMA. Guermat, Harris (2002) find that EWMA-based VaR forecasts are excessively volatile and unnecessarily high, when returns do not have conditionally normal distribution but fat tails. According to Brooks, Persand (2003) relative performance of different models depends on the loss function used but GARCH models provide reasonably accurate VaRs. Christoffersen, Hahn, Inoue (2001) show that different models (EWMA, GARCH, Implied Volatility) might be optimal for different probability levels. Berkowitz, O'Brien (2002) examine VaR models used by six leading US banks. Their results indicate that these models are in some cases highly inaccurate. Their results indicate that banks' models have difficulty dealing with changes in volatility. In addition, a comparison of banks' models with a simple univariate parametric GARCH model indicates that the latter gives roughly comparable coverage of high losses, but also tends to produce lower VaR figures and is much better at dealing with volatility changes. Žiković (2007) finds that widespread VaR model consistently underpredict the true levels of risk especially at higher confidence intervals and that semi-parametric models provide superior VaR forecasts in transitional economies.

Although VaR is useful for financial institutions to see the contours of the risks they face, a growing number of papers clearly show that VaR is not an adequate risk measure. As a result, more general convex measures of risk have been proposed. Among them, ETL became the most popular alternative to VaR. The literature that compares the VaR and ETL is very rare, especially empirical comparison. Besides not being sub-additive VaR does not indicate the size of the potential loss, provided that the loss exceeds VaR. To remedy these shortcomings, Artzner et al. (1997) introduced the Expected Shortfall (ES) risk measure, which equals the expected value of the loss, given that a VaR violation occurred. Basak, Shapiro (2001) suggested an alternative risk management procedure, namely limited expected losses based risk management (LEL-RM), which focuses on the expected loss also when (and if) losses occur. They substantiated that the proposed procedure generates losses lower than what VaR based risk management techniques generate. ETL turned out to be the most attractive coherent risk measure and has been studied by a number of authors (see Acerbi et al. 2001 and Inui, Kijima, 2005). Gilli, Kellezi (2003) advocate the use of Extreme value theory due to its firm theoretical grounds to compute both VaR and ETL. Yamai and Yoshiba (2002a, b) compared the two measures and argued that VaR is not reliable during market turmoil as it can mislead rational investors, whereas ETL can be a better choice overall. They point out that gains by using the ETL measure are substantial whenever its estimation is accurate. In other cases, they advise the market practitioners to combine the two measures for best results. Angelidis, Degiannakis (2007) test the performance of various VaR and ETL model on S&P500 index, Gold Bullion $ per Troy Ounce and US dollar/British pound exchange rate. In their paper they actually tested the impact of different volatility forecasting models within a strictly parametric framework so their results are not comparable with the results of this paper. They find that different volatility models are "optimal" for different assets. One can conclude that although ETL is a superior risk measure to VaR, it lacks the depth of the theoretical and empirical research that VaR measure has. Investigation into the theoretical properties of ETL is still in its early stages, and empirical investigation is only beginnings. Instead of fighting for supremacy VaR and ETL should be used together, combined, giving a better insight into the risks from taking a market position.

3. Value at Risk

VaR is usually defined as:
"VaR is the maximum potential loss that a portfolio can suffer within a fixed confidence level (cl) during a holding period"

Let $(X_t, t \in Z)$ be a strictly stationary time series representing daily observations of the negative log return on a financial asset price. The dynamics of *X* are given by:

$$X_t = \mu_t + \sigma_t Z_t \qquad (1)$$

where the innovations *Z* are IID with zero mean, unit variance and marginal distribution function $F_z(z)$. Assume that μ_t and σ_t are measurable with respect to ψ_{t-1} (information set about the return process available up to time *t-1*). Let $F_x(x)$ denote the marginal distribution of (X_t) and for a horizon *hp* let $F_{X_{t+1}+\ldots+X_{t+hp}|\psi_t}(x)$ denote the predictive distribution of the return over the next *hp* days, given information set up to and including day *t*. Looking from a tail events perspective for $0 < cl < 1$, unconditional $VaR_{cl}(X)$ is a quantile of the marginal distribution denoted by:

$$VaR_{cl}(X) = \inf\{x \in R : F_X(x) \geq cl\} \qquad (2)$$

and conditional $VaR_{cl}(X)$ is a quantile of the predictive distribution for the return over the next *hp* days denoted by:

$$VaR^t_{cl,hp}(X) = \inf\{x \in R : F_{X_{t+1}+\ldots+X_{t+hp}|\psi_t}(x) \geq cl\} \qquad (3)$$

From the perspective of 100cl% best cases, VaR at the 100(1-cl)% confidence level is defined as the upper 100cl percentile of the loss distribution. Following Artzner et al. (1999), VaR is defined at the 100(1-cl)% confidence level ($VaR_{cl}(X)$) as:

$$VaR_{cl}(X) = \sup\{x \mid P[X \geq x] > cl\} \qquad (4)$$

where sup{x | A} is the upper limit of *x* given event *A*, and $\sup\{x \mid P[X \geq x] > cl\}$ indicates the upper 100cl percentile of loss distribution. Unfortunately this definition can sometimes be misleading because VaR does not represent the "maximum" loss since a portfolio can lose much more than suggested by VaR depending on the shape of the tail distribution. A more insightful definition of VaR based on equation (2) and (4) would be:

"*VaR is the minimum potential loss that a portfolio can suffer in the 100(1-cl)% worst cases during a holding period*„ OR
"*VaR is the maximum potential loss that a portfolio can suffer in the 100(1-cl)% best cases during a holding period*".

These definitions cast doubt on the claims made by the advocates of VaR, that VaR is really the most appropriate measure to describe the risks associated with holding a certain position in securities. VaR can be thought of as "the best of worst cases scenario", and it therefore systematically underestimates the potential losses associated with the specified level of probability. Among the statistics of the tails VaR is probably the least appropriate to characterize its risks.

In financial risk management VaR has certainly represented a significant step forward with respect to the more traditional measures, based mostly on sensitivities to market variables. The strength of VaR relies in:

- VaR applies to any financial instrument and can be expressed in any voluntary unit of measure, be it percentage points of monetary units. More traditional measures, such as "greeks", are measures created ad hoc for specific instruments or risk variables and are expressed in different units. The comparison of relative riskiness between an equity portfolio and a fixed income portfolio is not easy with "greeks", while it is a straight comparison knowing their VaRs.
- VaR includes an estimate of future events and allows the risk of the portfolio to be expressed in a single number. "Greeks" amount to a "what if" risk measure that does not make any connection between the probability and severity of future events, wile VaR is excellent at doing this.

Not minimizing the merits of VaR in popularizing quantitative risk management, the truth of the matter is that VaR is badly conceived and pretty useless on its own as a risk measure. An indication of ill-logic behind VaR can be seen when looking at the risk behind holding a portfolio that earns a small amounts with a high level of probability and suffers very large amounts with very small probability. Such a portfolio will have negative VaR whenever the probability of loss is smaller than a specified confidence level at which VaR is measured. VaR for such portfolios is completely unable to reveal any risk. It seems logical to resort to the central estimate of the tail rather than its minimum in order to better measure such risks.

VaR models that are analyzed in this paper are: Normal simple moving average (VCV) VaR, RiskMetrics system, Historical simulation with rolling windows of 250 and 500 days, RiskMetrics system augmented with GARCH type volatility forecasting and Hybrid Historical simulation (HHS) developed in Žiković (2007).

Normal simple moving average (VCV) VaR:

$$VaR_{cl} = \alpha_{cl}\sigma_r \qquad (5)$$

$$\sigma = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(r_t - \bar{r})^2}$$

α_{cl} - corresponding quantile of standard normal variate

RiskMetrics VaR:

$$VaR_{cl} = \alpha_{cl}\sigma_r \tag{6}$$
$$\sigma_t = \sqrt{0.94\sigma_{t-1}^2 + 0.06\varepsilon_t^2}$$

GARCH-RiskMetrics VaR:

$$VaR_{cl} = \alpha_{cl}\sigma_r \tag{7}$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\sigma_{t-i}^2$$

Historical simulation VaR:

Let F_n denote the empirical process of the observed losses $\{X_i\}$, that is:

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le t) \tag{8}$$

where $I(\bullet)$ is the indicator function, $\{X_i\}$ is IID drawn from unknown distribution F. The cl quantile $F^{-1}(cl)$ is estimated by:

$$VaR_{(cl)} = F^{-1}(cl) = X_{(i)} \tag{9}$$

where $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ are order statistics.

Hybrid historical simulation (HHS) VaR:

$$VaR_{(cl)} = F^{-1}(cl) = \hat{Z}_{(i)} \tag{10}$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\sigma_{t-i}^2$$
$$z_t = \frac{\varepsilon_t}{\sigma_t}$$
$$\{\hat{z}_{t+1}\} = \{z_t\} \times \hat{\sigma}_{t+1}^2 \qquad \hat{Z} = \{\hat{z}_t\}$$

where $\hat{Z}_{(1)} \le \hat{Z}_{(1)} \le \ldots \le \hat{Z}_{(n)}$ are order statistics from volatility scaled bootstrapped series $\hat{Z}$.

4. Coherent risk measures and Expected tail loss

VaR as a risk measures is open to criticism from many directions. Hoppe (1999) argues that the underlying statistical assumptions are violated because they cannot capture many features of the financial markets such as intelligent agents. Artzner et al. (1997, 1999) have used an axiomatic approach to the problem of defining a satisfactory risk measure. They defined attributes that a good risk measure should satisfy, and call risk measures that satisfy these axioms "coherent". A coherent risk measure ρ assigns to each loss X a risk measure $\rho(X)$ such that the following conditions are satisfied (Artzner et al. 1999):

$\rho(tX) = t\rho(X)$ (homogeneity) (11)

$\rho(X) \geq \rho(Y)$, if $X \leq Y$ (monotonicity) (12)

$\rho(X + n) = \rho(X) - n$ (risk-free condition) (13)

$\rho(X) + \rho(Y) \leq \rho(X + Y)$ (sub-additivity) (14)

for any number n and positive number t. These conditions guarantee that the risk function is convex, which in turn corresponds to risk aversion. That is:

$$\rho(tX + (1 - t)Y) \leq t\rho(X) + (1 - t)\rho(Y) \qquad (15)$$

Homogeneity and monotonicity conditions are reasonable conditions to impose a priori, and together imply that the function $\rho(\cdot)$ is convex. The risk-free condition means that the addition of a riskless asset to a portfolio will decrease its risk because it will increase the value of end-of-period portfolio. According to the last condition a risk measure is sub-additive if the measured risk of the sum of positions X and Y is less than or equal to the sum of the measured risks of the individual positions considered on their own.

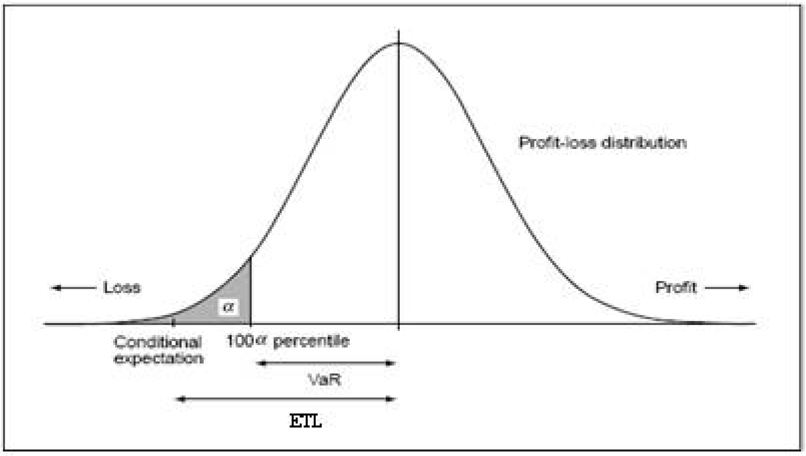
VaR is not a coherent risk measure because it does not necessarily satisfy the sub-additivity condition. VaR can only be made sub-additive if a usually implausible assumption is imposed on returns being normally (or slightly more generally, elliptically) distributed. Subadditivity expresses the fact that a portfolio made of subportfolios will risk an amount, which is at most the sum of the separate amounts risked by its subportfolios. This is maybe the most characterizing feature of a coherent risk measure and represents something that belongs to everybody's concept of risk. The global risk of a portfolio will be the sum of the risks of its parts only in the case when the latter can be triggered by concurrent events, namely if the sources of these risks may conspire to act altogether. In all other cases, the global risk of the portfolio will be strictly less than the sum of its partial risks thanks to risk diversification. This axiom captures the essence of how a good risk measure should behave under the composition/addition of portfolios. It is the key test for checking whether a measurement of a portfolio's risk is consistent with those of its subportfolios. For a sub-additive measure, which ETL is, portfolio diversification always leads to risk reduction, while for measures which violate this axiom, such as VaR, diversification may produce an increase in their value even when partial risks are triggered by mutually exclusive events. Sub-additivity is not some trivial academic invention but a crucial part of any risk measure for a number of reasons:

- If risks are sub-additive, then adding risks together would give an overestimate of combined risk, and this means that a sum of risks can be used as a conservative estimate of combined risk. This facilitates decentralized decision-making within a firm, because a supervisor can always use the sum of the risks of the units reporting to him as a conservative risk measure. But if risks are not sub-additive, adding them together gives an underestimate of combined risks, and this makes the sum of risks effectively useless as a risk measure.
- If regulators use non-sub-additive risk measures to set capital requirements, a bank might be tempted to break itself up to reduce its regulatory capital requirements, because the sum of the capital requirements of the smaller units would be less than the capital requirement of the bank as a whole.
- Non-sub-additive risk measures can also inspire traders to break up their accounts, with separate accounts for separate risks, in order to reduce their margin requirements. This could be a matter of serious concern for the exchange because the margin requirements on the separate accounts would no longer cover the combined risks.

A very serious shortcoming of VaR is that it provides no handle on the extent of the losses that might be suffered beyond the threshold amount indicated by it. VaR is incapable of distinguishing between situations where losses in the tail are only a bit worse, and those where they are overwhelming. Indeed, VaR merely provides a lowest bound for losses in the tail of the loss distribution and has a bias toward optimism instead of the conservatism that ought to prevail in risk management. An alternative measure that does quantify the losses that might be encountered in the tail is ETL. Both VaR and ETL require the user to a priori specify confidence level and holding period. While VaR represents a maximum loss one expects at a determined confidence level for a given holding period, ETL is the loss one expects to suffer, provided that the loss is equal to or greater than VaR, that is:

"ETL is the expected value of the loss of the portfolio in the 100(1-cl)% worst cases during a holding period."

Figure 1 - ETL and VaR



Following the representation from equation (2) the unconditional ETL is defined as:

$$ETL_{cl}(X) = E[X \mid X > VaR_{cl}(X)] = -cl^{-1}\int_{-\infty}^{VaR} xf(x)dx \qquad (16)$$

and the conditional ETL can be expressed as:

$$ETL^{t}_{cl,hp}(X) = E\left[\sum_{j=1}^{hp} X_{t+j} \mid \sum_{j=1}^{hp} X_{t+j} > VaR^{t}_{cl,hp}(X), \psi_t\right] \qquad (17)$$

ETL is very appealing as a risk measure in that it sums all values of x, weighted by $f(x)$, from minus infinity to VaR, thus taking into account of the sizes of losses beyond the VaR level. ETL can be encountered in the literature under many names such as: Expected shortfall (ES), Conditional VaR (CVaR), tail VaR, tail conditional expectation, mean excess loss etc. ETL measure has been used by insurance practitioners, especially casualty insurers for a long time as conditional average claim size. For continuous loss distributions, ETL at a given confidence level is the expected loss given that the loss is greater or equal to VaR at that level. For distributions with possible discontinuities it has a more subtle definition and can differ from either of those quantities, which for convenience in comparison can be designated by $CVaR^{+}$ and $CVaR^{-}$, respectively. $CVaR^{+}$ is also known as "mean shortfall", although the seemingly identical term "expected shortfall" has been interpreted in other ways in Acerbi, Nordio, Sirtori (2001), with the latter paper taking it as a synonym for CVaR itself). "Tail VaR" is a term that has been suggested for $CVaR^-$ (Artzner et al. 1999). Unlike $CVaR^+$ and $CVaR^-$, ETL(CVaR) is a coherent measure of risk in the sense of Artzner et al. (1999). However, ETL is no panacea and has its own flaws; Yamai and Yoshiba (2002b) find that both VaR and ETL are not reliable during market turmoil and can give misleading results, although ETL is a better choice than VaR. Kondor, Varga-Haszonits (2008) find that whenever there is an asset in a portfolio that dominates over others in a given sample the portfolio cannot be optimized under any coherent measure on that sample, including ETL.

With the dawn of a new "coherent" risk measure there exists a need for a new risk management paradigm. Risk management is primarily concerned with the risk of low-probability events that could result in catastrophic losses. Traditional VaR models tend to ignore extreme events and focus on modeling the entire empirical distribution of returns. By wrongly using the Central limit theorem it is often assumed that returns are normally or lognormally distributed, but little attention is paid to the distribution of the tails. Such risk models are prone to fail just when they are needed the most - in large market moves, when large losses occur. Estimation of risks associated with rare events with limited data is inevitably problematic, and these difficulties increase as the events concerned become rarer. Inference about the extreme tails is always uncertain, because of low number of observations and sensitivity to the values of individual extreme observations. The key to estimating the distribution of such events is the extreme value theorem, which governs the distribution of extreme values, and shows how this distribution looks like asymptotically. The application of extreme value theory (EVT) to risk management is just the thing that the new "coherent" risk measure was missing to establish itself as a pretender to the risk management throne currently held by "non-coherent" VaR. EVT models were primarily used in the field of civil engineering: engineers are required to design their structures to withstand the forces that might be reasonable to expect but are rarely experienced. Another standard field of application of EVT is hydrology, where engineers have long struggled with the question of how high dams, sea-walls and dikes should be to contain the probabilities of floods within reasonable limits. They have to do their calculations with even fewer observations than risk practitioners, and their quantile estimates are typically well out of the range of their sample data. EVT provides a framework in which an estimate of anticipated forces could be made using historical data. By definition, extreme events are rare, meaning that their estimates are often required for levels of a process that are greater that those in the available data set. This implies an extrapolation from observed levels to unobserved levels and extreme value theory provides a class of models to enable such extrapolation. In lieu of an empirical basis, asymptotic argument is used to generate EV models. Today EVT is used in telecommunications, ocean wave modeling, thermodynamics of earthquakes, memory cell failure and many other fields. It is important to be aware of the limitations implied by the adoption of the EV paradigm. EV models are developed using asymptotic arguments, which should be kept in mind when applying them to finite samples. EV models are derived under idealized circumstances, which need not be true for a process being modeled.

Presuming n observations of P&L time series, if X is IID drawn from some unknown distribution $F(x) = P(X \leq x)$, estimating extreme value (EV) VaR or ETL posses a significant problem because the distribution $F(x)$ is unknown. Help comes from Fisher-Tippett theorem (1928), which shows that as n gets large the distribution of tail of X converges to Generalized extreme value distribution (GEV):

$$H_{\mu,\sigma,\xi}(x) = \begin{vmatrix} \exp\left(-\left[1+\xi(x-\mu)/\sigma\right]^{-1/\xi}\right) & & \xi \neq 0 \\ \exp\left(-e^{-(x-\mu)/\sigma}\right) & if & \xi = 0 \end{vmatrix} \tag{18}$$

where x satisfies the condition $1 + \xi(x\text{-}\mu)/\sigma > 0$. GEV distribution has three parameters: location parameter (μ), which is a measure of central tendency, scale parameter (σ), which is a measure of dispersion and tail index (ξ), which is a measure of the shape of the tail. GEV distribution has three special cases:

- If $\xi > 0$, GEV distribution becomes a Fréchet distribution, meaning that *F(x)* is leptokurtotic.
- If $\xi = 0$, GEV distribution becomes a Gumbel distribution, meaning that *F(x)* has normal kurtosis.
- If $\xi < 0$, GEV distribution becomes a Weibull distribution, meaning that *F(x)* is platokurtotic, which is usually not the case with financial data.

Mean and variance are related to location and scale parameters of GEV distribution:

$$Mean = \mu + \left[\frac{\Gamma(1-\xi)-1}{\xi}\right]\sigma \text{ which converges to } \mu + 0{,}577216\sigma \text{ as } \xi \to 0 \tag{19}$$

$$Variance = \left[\frac{\Gamma(1-2\xi)-\Gamma^2(1-\xi)}{\xi^2}\right]\sigma^2 \text{ which converges to } \frac{\pi^2}{6}\sigma^2 \text{ as } \xi \to 0 \tag{20}$$

It is easy to obtain mean and variance from μ and σ, but one must be careful not to confuse the two since they differ significantly. Quantiles of GEV distribution can be obtained by taking the log of equation (18):

$$\log(cl) = \begin{cases} -\left(1+\xi(x-\mu)/\sigma\right)^{-1/\xi} & \xi \neq 0 \\ -\exp(-(x-\mu)/\sigma) \quad if & \xi = 0 \end{cases} \tag{21}$$

Value of x is then calculated to get the quantiles or VaRs associated with a desired confidence level. EV VaR is calculated as (Dowd, 2002):

$$VaR_{cl} = \mu - \frac{\sigma}{\xi}\left[1-(-\log(cl))^{-\xi}\right] \qquad \text{(Fréchet VaR, } \xi > 0\text{)} \tag{22}$$

$$VaR_{cl} = \mu - \sigma \log\left[\log(1/cl)\right] \qquad \text{(Gumbel VaR, } \xi = 0\text{)} \tag{23}$$

There are no closed form ETL formulas for Fréchet and Gumbel distributions but EV ETL can be derived from EV VaR estimates using "average-tail VaR" algorithm set out in Dowd (2002). The fact that ETL is a probability weighted average of tail losses suggests that ETL can be estimated as an average of tail VaRs. The approach suggested by Dowd (2002) to calculating ETL is to divide the tail of the P&L distribution into a large number (n > 500) of equally distant slices, each of which has the same probability mass and then calculate the VaR for each slice. The mean of calculated tail VaRs gives ETL. It is easily shown that ETL is indeed estimable in a consistent way as the "average of 100cl% worst cases:

$$ETL_{cl}^{(n)}(X) = -\frac{1}{[ncl]}\sum_{i=1}^{[ncl]} X_{i:n} \qquad \text{where } X_{i:n} \text{ are order statistics} \tag{24}$$

$$ETL_{cl}^{(n)}(X) \xrightarrow[n\to\infty]{} ETL_{cl}(X) \tag{25}$$

To estimate EV risk measures it is necessary to estimate EV parameters – μ, σ, and in the case of Fréchet distribution the tail index (ξ). Unlike the estimation of mean and variance parameters estimation of tail index is a bit more demanding. Embrechts et al. (1997) suggests determining the tail index of the distribution via Hill estimator:

$$\hat{\xi}^{(H)}_{n,k} = k^{-1}\sum_{j=1}^{k}\ln X_{j,n} - \ln X_{k+1,n} \tag{26}$$

where k, the tail threshold used in the Hill estimation has to be chosen arbitrarily, which presents a major source of problems in practice. The Hill estimator is the average of the k most extreme observations, minus ($k+1$)th observation, which is next to the tail. There are two approaches to handling the trade off between bias and variance. The first approach, recommended by Embrechts et al. (1997), is based on estimating the Hill estimator for a range of k values and selecting the k values where the plot of the Hill estimator against k flattens out. Danielson, de Vries (1997) suggest finding an optimal value of k that minimizes MSE loss function and, in regards to MSE, reflects an optimal trade off between bias and variance. Their procedure takes a second-order approximation to the tail of the distribution and uses the fact that k is optimal (in the MSE sense) at the point where bias and variance reduce at the same rate. They suggest using a sub-sample bootstrapping procedure for finding the optimal value of k. Unfortunately this approach is impractical since it requires a very large sample size which is very difficult to obtain in practice. For this reason in this paper the approach recommended by Embrechts et al. (1997) is adopted.

When forecasting ETL researchers mostly use either simple moving average models with GEV distribution or plain empirical equally weighted historical returns. In this paper a new approach is used based on bootstrapping transformed tail losses. A similar approach used for VaR estimation on volatile markets yielded significant improvements over parametric and nonparametric approaches (see Žiković, 2007). ETL models analyzed in this paper are: simple moving average model (VCV) with Fréchet distribution, simple moving average model (VCV) with Gumbel distribution, bootstrapped historical simulation with 500 days observation period, GARCH RiskMetrics model with Fréchet distribution, GARCH RiskMetrics model with Gumbel distribution and bootstrapped HHS ETL model. ETL for VCV model with Fréchet and Gumbel distribution and GARCH RiskMetrics model with Fréchet and Gumbel distribution, due to the lack of closed form solutions, are derived from their respective EV VaR estimates using "average-tail VaR" algorithm.

Historical simulation ETL can be expressed as:

$$ETL = E(X \mid X > VaR) = \left(\sum_{i=[ncl]}^{n} X_{n(i)}\right) / (n - [ncl]) \tag{27}$$

where $X_{n(1)} \leq X_{n(2)} \leq \ldots \leq X_{n(n)}$ are order statistics.

Parametric ETL forecasts, even those based on GEV distribution are sensitive to misspecification of the functional form and parameter estimates, especially the tail index. Simpler parametric models can not adequately or timely adapt to sudden changes in levels of volatility. Nonparametric ETL models i.e. calculating ETL from the untransformed historical data set of tail losses are certain to be unreactive to sudden shifts in market regimes and occurrence of extreme events. These are exactly the same critiques that apply to these approaches when using them for VaR calculation. The weak points of risk measurement models cannot be ignored and they continually come back to haunt us even when we switch from one risk measure to another. The problems remain the same regardless whether we are estimating VaR or ETL. Hybrid historical simulation (HHS) model from Žiković (2007) yielded superior results for stock indexes from 12 EU new member states can also be used as a basis for developing a semi-parametric approach to estimation of ETL. From the VaR version of the HHS model a Bootstrapped HHS ETL model can be developed. The model standardizes the tail losses in excess of HHS VaR by the latest GARCH volatility update for that point in time to form a series of standardized tail losses:

$$z_t = \frac{Tail \quad loss_t}{\sigma_t} \tag{28}$$

Since these standardized tail losses are now IID they are suitable for bootstrapping. New discrete PDFs of tail losses are derived through bootstrapping, which are than updated by the latest GARCH volatility forecasts:

$$\hat{F}_{(n)}(t) = F_{(n)}(t) \times \sigma_{t+1} \tag{29}$$

By taking the averages over a great number of volatility updated tail PDFs $\left(\hat{F}_{(n)}(t)\right)$ ETL forecasts are obtained. Besides being reactive to the latest market developments through the use of GARCH volatility updating the HHS ETL approach also provides for an elegant way of calculating confidence intervals for ETL estimates, based on bootstrapping that is free of any distributional assumptions. The only assumption that is made in the model is that the underlying data generating process can be described by a GARCH process. HHS ETL model does not impose any distributional assumptions about the behavior of the tail losses, unlike EV ETL models, and allows for the empirical distribution of tails to evolve over time.

Hybrid historical simulation (HHS) ETL can be expressed as:

$$ETL = E(X \mid X > VaR) = \left(\sum_{i=[ncl]}^{n} \hat{Z}_{n(i)} \right) / \left(n - [ncl]\right) \tag{30}$$

where $\hat{Z}_{n(1)} \leq \hat{Z}_{n(1)} \leq \ldots \leq \hat{Z}_{n(n)}$ are order statistics from volatility scaled bootstrapped series $\hat{Z}$.

The strong points and weaknesses of every model remain with them and that is why knowledge obtained in developing VaR models should not be wasted. VaR estimation techniques can easily be adopted to serve a new “superior” risk measure – ETL. Research in VaR estimation should by no means be discouraged, but instead intensified, because it could now serve a dual purpose – improving VaR estimates but also improving ETL estimates.

5. Data and backtesting methodology

Data used in the analyses of VaR and ETL models are the daily log returns from US and European stock indexes (DOW JONES, NASDAQ, S&P500, CAC, DAX and FTSE). The returns are collected from Bloomberg web site for the period 01.01.2000 - 12.3.2008, which includes the latest US sub prime mortgage crisis and its effects on global stock markets. The calculated VaR and ETL figures are for a one-day ahead horizon and 95 and 99 percent confidence levels. To secure the same out-of-the-sample VaR backtesting period for all of the tested stock indexes, the out-of-the-sample data sets are formed by taking out 1,000 of the latest observations from each stock index. The rest of the observations are used as presample observations needed for VaR starting values and volatility model calibration. Length of the tail losses data set used for backtesting depends on the number of VaR errors generated by each VaR model. The quality of ETL forecasts does not only depend on ETL estimation model but also on the quality of VaR forecast. This can be easily demonstrated by the simple fact that a loss that might be extreme under one VaR model and as such is compared to the ETL forecast might not exceed another, more conservative VaR measure.

Data from analyzed stock indexes shows significant heteroskedasticity for all of the indexes and autoregression in the returns on CAC, DAX and FTSE. GARCH and EGARCH representation of volatility with Gaussian and Student's T distribution is used to capture the dynamics of data generating processes of analyzed indexes. All of the analyzed VaR models are tested in several ways to determine their statistical characteristics and ability to adequately measure market risk in the countries analyzed in this paper.

It often happens that more than one VaR model is deemed adequate and the problem of ranking the models arises. To overcome this shortcoming of the backtesting measures forecast evaluation can be used. Blanco-Ihle (1998) suggests a loss function:

$$C_t = \begin{cases} \dfrac{L_t - VaR_t}{VaR_t} & if \quad L_t > VaR_t \\ 0 & if \quad L_t \le VaR_t \end{cases} \tag{31}$$

This loss function allows for the sizes of tail losses to influence the final rating of VaR model. VaR model that generates higher tail losses would generate higher values under this size adjusted loss function than a VaR model that generates lower tail losses, ceteris paribus. The problem with Blanco-Ihle loss function is that it compares VaR with tail losses, which makes no practical sense because VaR forecasts only the „best“ scenario for the tail losses. Since VaR does not hold any information about the size of the expected tail loss, Blanco-Ihle loss function actually measures the discrepancy between the lowest possible and actual tail losses, which is not especially useful. Blanco-Ihle loss function can be modified to compare ETL with the actual value of the tail loss, which is exactly what the loss function should be measuring. The suggested modification of the Blanco-Ihle loss function can be written as:

$$C_t = \begin{cases} \dfrac{L_t - ETL_t}{ETL_t} & if \quad L_t > VaR_t \\ 0 & if \quad L_t \le VaR_t \end{cases} \tag{32}$$

In order to select superior ETL models, each model will be graded by four symmetrical error statistics: the mean absolute error (MAE), two versions of the root mean squared error (RMSE), and the mean absolute percentage error (MAPE), which is actually modified to represent the proposed ETL modification of Blanco-Ihle (1998) loss function.

$$RMSE(1) = \sqrt{\frac{\sum_{i=1}^{T} | L_i^2 - ETL_i^2 |}{T}} \tag{33}$$

$$RMSE(2) = \sqrt{\frac{\sum_{i=1}^{T}\left(\left|L_i\right| - \left|ETL_i\right|\right)^2}{T}} \tag{34}$$

$$MAE = \frac{1}{T}\sum_{i=1}^{T}\left\|L_i\right| - \left|ETL_i\right\| \tag{35}$$

$$MAPE = \frac{1}{T}\sum_{i=1}^{T}\left|\frac{\left|L_i\right| - \left|ETL_i\right|}{L_i}\right| \tag{36}$$

Out of these error statistics, ETL modification of Blanco-Ihle (1998) loss function (MAPE) measure could be considered as the most indicative since it compares the tail loss to ETL while taking into account the relative size of the tail loss compared to the difference between the two. According to the employed two-stage backtesting procedure, the best performing VaR/ETL model will primarily satisfy both Kupiec and Christoffersen IND test and than provide superior, in the sense of minimizing used error statistics, tail loss forecasts. The optimal VaR/ETL should minimize the modified Blanco-Ihle loss function (C_t).

6. Results

In this section the backtesting results for tested VaR and ETL models are presented, their performance according to different criteria is analyzed and the implications are discussed.

VaR backtesting results

Backtesting results for VaR at 95% confidence level are presented in tables 1-6 in the appendix. According to the backtesting results, in the case of DOW JONES index VCV model was the worst performer, but RiskMetrics model was the only one to generate dependent VaR failures. VCV, HS250 and RiskMetrics models failed the unconditional coverage condition. According to Lopez test the best performer was the parametric model with EGARCH(1,1) and T distribution. According to Blanco-Ihle loss function HHS model with the same volatility model was the best performer. For NASDAQ index RiskMetrics model was the worst performer. VCV and RiskMetrics models failed the unconditional coverage condition. According to Lopez test the best performer was the HS 500 model. According to Blanco-Ihle loss function HHS model with T distributed EGARCH(1,1) volatility was the best performer. For S&P500 index according to the backtesting results RiskMetrics model was the worst performer. VCV, HS250 and RiskMetrics models failed the unconditional coverage condition. According to Lopez and Blanco-Ihle test the best performer was the HHS model with T distributed EGARCH(1,1) volatility. For CAC index HS250 model was the worst performer. VCV, HS250 and RiskMetrics models failed the unconditional coverage condition According to Lopez test the best performer was the parametric model with AR(1) GARCH(1,1) and T distribution. According to Blanco-Ihle loss function HHS model with the same volatility model was the best performer. For DAX index VCV model was the worst performer. VCV and HS250 models failed the unconditional coverage condition. According to Lopez test the best performer was the HS500 model. According to Blanco-Ihle loss function HHS model with T distributed AR(1) GARCH(1,1) volatility was the best performer. For FTSE index all the models satisfied both the unconditional coverage as well as independence tests. According to Lopez test the best performer was the parametric model with AR(1) EGARCH(1,1) and T distribution. According to Blanco-Ihle loss function HHS model with the same volatility model was the best performer.

Backtesting results of selected VaR models for US and selected European stock indexes at 95% confidence level are consistent, with VCV, HS250 and RiskMetrics models failing the unconditional coverage test and being the worst performers for all of the tested indexes. The RiskMetrics model was the only one to generate dependent VaR failures. HHS and GARCH RM model are the best out of the tested models, with HHS model being the clear winner.

Backtesting results for VaR at 99% confidence level are presented in tables 7-12 in the appendix. For DOW JONES index VCV model was again the worst performer. All of the tested models except the HHS model failed the unconditional coverage condition. According to Lopez and Blanco-Ihle test the best performer was the HHS model with T distributed EGARCH(1,1) volatility. For NASDAQ index VCV model was the worst performer. All of the tested models except the GARCH RM and HHS model failed the unconditional coverage condition. According to Lopez test the best performer was the parametric model with EGARCH(1,1) and T distribution. According to Blanco-Ihle loss function HHS model with the same volatility model was the best performer. For S&P500 index VCV model was the worst performer. All of the tested models except the HHS model failed the unconditional coverage condition. According to Lopez and Blanco-Ihle test the best performer was the HHS model with T distributed EGARCH(1,1) volatility. For CAC index HS250 model was the worst performer. All of the tested models except the HHS model failed the unconditional coverage condition. According to Lopez and Blanco-Ihle test the best performer was the HHS model with T distributed AR(1) GARCH(1,1) volatility. For DAX index VCV model was the worst performer. All of the tested models except the HS500 and HHS model failed the unconditional coverage condition. According to Lopez and Blanco-Ihle test the best performer was the HHS model with T distributed AR(1) GARCH(1,1) volatility. For FTSE index VCV model was the worst performer. All of the tested models except the HHS model failed the unconditional coverage condition. According to Lopez and Blanco-Ihle test the best performer was the HHS model with T distributed AR(1) EGARCH(1,1) volatility.

As expected the backtesting results for VaR models on US and selected European stock indexes at 99% confidence level are even more indicative since five out of six tested VaR models continually failed the unconditional coverage condition. HHS was the only model, out of the tested models, that passed all of the tests, for both the unconditional coverage and independence of VaR failures. The results are again consistent with the results for transitional markets reported in Žiković (2007). Overall the results obtained for the tested US and selected European stock indexes are in line with the results reported by Žiković (2007) for stock indexes from transitional markets. HHS model provided adequate risk coverage at both 95% and 99% confidence levels while other tested models failed. For both developing and developed stock markets simpler VaR models consistently fail their task, and provide the risk managers with falsely optimistic data about the levels of risk that the financial institutions are exposed to. Furthermore, GARCH based volatility models, even at lower confidence level, continually outperform VaR models based on the assumption of simpler models of volatility such as SMA and EWMA which is in line with the conclusions from Wong, Cheng, Wong (2002), Guermat, Harris (2002) and Brooks, Persand (2003), as opposed to conclusion by Bams, Wielhouwer (2000) and Caporin (2003).

ETL backtesting results

Backtesting results for ETL at 95 and 99% cut-off level are presented in tables 13-18 in the appendix. For DOW JONES index at both 95% and 99% cut-off level according to all of the employed error statistics the Bootstrapped HHS ETL approach was the best performer resulting in smallest deviations from realized tail losses, and minimizing the proposed loss function. The worst performer at both cut-off levels was the simple moving average (VCV) approach based on Fréchet distribution with tail index of 0.2, resulting in highest deviations from realized tail losses. For NASDAQ index at 95% cut-off level according to MAE measure and modification of Blanco-Ihle loss function the Bootstrapped HHS ETL approach was the best performer. According to RMSE measures Bootstrapped HS500 approach was the best performer at 95% cut-off level. The worst performer at this cut-off level was the parametric GARCH volatility based (GARCH RM) approach based on Fréchet distribution with tail index of 0.31. At 95% cut-off level according to all of the employed error statistics the Bootstrapped HS500 approach was the best performer. The worst performer at both cut-off levels was the VCV approach based on Fréchet distribution. For S&P500 index at both cut-off levels according to all of the employed error statistics the Bootstrapped HHS ETL approach was the best performer. The worst performer at 95% cut-off level was the VCV approach based on Fréchet distribution with tail index of 0.24. The worst performer at 99% cut-off level was the GARCH RM approach based on Fréchet distribution. For CAC index at both cut-off levels according to all of the employed error statistics the Bootstrapped HHS ETL approach was the best performer. The worst performer at both cut-off levels was the GARCH RM approach based on Fréchet distribution with tail index of 0.23. For DAX index at the 95% according to modification of Blanco-Ihle loss function the Bootstrapped HHS ETL approach was the best performer, while other error statistics chose Bootstrapped HS500 approach as their favorite. At 99% cut-off level according to all of the employed error statistics the Bootstrapped HHS ETL approach was the best performer. The worst performer at 95% cut-off level was the GARCH RM approach based on Fréchet distribution with tail index of 0.3. The worst performer at 99% cut-off level was the VCV approach based on Fréchet distribution For FTSE index at both cut-off levels according to MAE measure and modification of Blanco-Ihle loss function the Bootstrapped HHS ETL approach was the best performer. According to RMSE measures GARCH RM approach with Gumbel distribution was the best performing ETL model at 95% cut-off level. The worst performer at both cut-off levels was the GARCH RM approach based on Fréchet distribution with tail index of 0.22.

Overall the Bootstrapped HHS ETL approach was the best performing ETL measure across all of the tested indexes with the exception of NASDAQ index at 95% cut-off level where Bootstrapped HS500 ETL model was the best performer. Same as with the best performers, the worst performers also stand out with the VCV approach based on Fréchet distribution and GARCH RM approach with Fréchet distribution being the worst ranked ETL models for all of the tested indexes. These models greatly overestimated the expected averages of tail loss. Comparing these results with the performance of the same models that used Gumbel instead of Fréchet distribution provides a clearer picture into the anatomy of the problem. Models that used Gumbel distribution performed far better compared to those with Fréchet distribution so this points to two possible reasons for such overpredictions:

a) Tail indexes have been incorrectly calculated (they are too high) and
b) The use of GEV distributions in ETL estimation provides overly conservative estimates of average tail losses.

Since ETL models using Gumbel distribution also produced visible overpredictions of tail losses we can disregard the problem with estimation of tail indexes for now and conclude that for the tested stock indexes the use of GEV distributions in ETL estimation provides overly conservative estimates of average tail losses. The backtesting performance of tested model shows remarkable consistency in that similar models are ranked as best and worst performers both in VaR and ETL estimation. When analyzing the obtained results it is hard not to draw a parallel between the VaR and ETL estimation. Advanced models and models with weak assumptions consistently outperform simpler and more restrictive models. These findings are exactly the reason why should not The focus of future research should be on improving both VaR and ETL estimation techniques as well as finding optimal combinations of VaR-ETL models, because only such complete information can serve as a solid basis for decision making in financial institutions and reveal the actual risk exposure to investors and regulators.

7. Conclusion

VaR is not a coherent risk measure because it does not necessarily satisfy the subadditivity condition. Subadditivity expresses the fact that a portfolio will risk an amount which is at most the sum of the separate amounts risked by its subportfolios. This is maybe the most characterizing feature of a coherent risk measure, something which belongs to everybody's concept of risk. A second serious shortcoming of VaR is that it provides no handle on the extent of the losses that might be suffered beyond the threshold. VaR is incapable of distinguishing between situations where losses in the tail are only a bit worse, and those where they are overwhelming. ETL is still not required by the regulators as a risk measure that can be used to calculate economic capital. Advances that have been made in VaR should not be lost with the adoption of coherent risk measures into regulatory framework. Superior quality of VaR techniques should yield superior ETL forecasts showing that VaR and ETL should be regarded as partners not rivals. After all, VaR and ETL are connected in the sense that from the VaR surface of the tail ETL figures can easily be calculated. When forecasting ETL researchers mostly use either simple moving average models with GEV distribution or plain empirical equally weighted historical returns. Parametric ETL forecasts, even those based on GEV distribution are sensitive to the misspecification of the functional form of the losses and parameter estimates. Furthermore, these models can not adequately or timely adapt to sudden changes in levels of volatility. Purely nonparametric ETL estimation approaches, such as calculating ETL from the untransformed historical data set of tail losses, are certain to be unreactive to sudden shifts in market regimes and occurrence of extreme events. These are exactly the same critiques that apply to these approaches when using them for VaR calculation. The weak points of risk measurement models cannot be ignored and they will continually come back to haunt us even when we switch from one risk measure to another. The problems remain the same regardless whether we are estimating VaR or ETL. The proposed, new ETL model based on bootstrapping volatility transformed tail losses, which has been shown to yield significant improvements over parametric and nonparametric approaches when used as a VaR model in volatile and illiquid markets.

Overall the results of VaR model comparison obtained for the tested US and selected European stock indexes are in line with the results reported by Žiković (2007) for stock indexes from transitional markets. HHS model provided adequate risk coverage at both 95% and 99% confidence levels while other tested models failed. For both developing and developed stock markets simpler VaR models consistently fail their task, and provide the risk managers with falsely optimistic data about the levels of risk that the financial institutions are exposed to. GARCH based volatility models, even at lower confidence level, continually outperform VaR models based on the assumption of simpler models of volatility such as SMA and EWMA which is in line with the conclusions from Wong, Cheng, Wong (2002), Guermat, Harris (2002) and Brooks, Persand (2003), as opposed to conclusion by Bams, Wielhouwer (2000) and Caporin (2003). When it comes to ETL comparison the suggested Bootstrapped HHS ETL approach was the best performing ETL measure across all of the tested indexes with the exception of NASDAQ index at 95% cut-off level. The worst performers are VCV approach based on Fréchet distribution and GARCH RM approach with Fréchet distribution, which greatly overestimated the expected averages of tail loss. For the tested stock indexes the use of GEV distributions in ETL estimation provides overly conservative estimates of ETL. The backtesting performance of tested model shows remarkable consistency in that similar models are ranked as best and worst performers both in VaR and ETL estimation. When analyzing the obtained results it is hard not to draw a parallel between the VaR and ETL estimation. Advanced models consistently outperform simpler and more restrictive models. The strong points and weaknesses of every model remain with them and that is way knowledge obtained in developing VaR models must not be wasted. Cutting-edge VaR estimation techniques can easily be adopted to serve a new "superior" risk measure – ETL. Research in VaR estimation should by no means be discouraged, but instead intensified, because it could now serve a dual purpose – improving VaR estimates but also improving ETL estimates. The focus of future research should be on improving both VaR and ETL estimation techniques as well as finding optimal combinations of VaR-ETL models, because only such complete information can serve as a solid basis for decision making in financial institutions and reveal actual risk exposure both to investors and regulators.

References:

1. Acerbi C., Nordio C., Sirtori C.: Expected Shortfall as a Tool for Financial Risk Management. Working Paper, 2001. http://www.gloriamundi.org/var/wps.html
2. Angelidis T., Degiannakis S.: Backtesting VaR Models: An Expected Shortfall Approach. University of Crete, Department of Economics, Working Papers No. 0701, Jan 2007. http://ideas.repec.org/p/crt/wpaper/0701.html
3. Artzner P., Delbaen F., Eber J.M., Heath D.: Thinking coherently. Risk, Vol. 10, No. 11., 1997. p. 68–71.
4. Artzner P., Delbaen F., Eber J.M., Heath D.: Coherent measures of risk. Mathematical Finance, Vol. 9, No. 3., 1999. p. 203–228.
5. Bams D., Wielhouwer L. J.: Empirical Issues in Value-at-Risk Time Varying Volatility, Fat Tails and Parameter Uncertainty. Dec 2000. http://www.gloriamundi.org/picsresources/dbjw.pdf
6. Basak S., Shapiro A.: Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices. The Review of Financial Studies. Vol. 14., No. 2., 2001. p. 371-405.
7. Berkowitz J., O'Brien J.: How accurate are VaR models at commercial banks?. Journal of Finance, 2002. http://www.gloriamundi.org/picsresources/jbjo.pdf
8. Blanco C., Ihle G.: How Good is Your VaR Using Backtesting to Assess System Performance. Financial Engineering News, Aug 1998. p. 1-2.
9. Brooks C., Persand G.: Volatility forecasting for Risk Management. Journal of forecasting, 22, 2003. p. 1-22.
10. Caporin M.: The Trade Off Between Complexity and Efficiency of VaR Measures: A Comparison of RiskMetric and GARCH-Type Models. 2003. http://www.gloriamundi.org/picsresources/mcga.pdf

11. Christoffersen P., Hahn J., Inoue A.: Testing and Comparing Value-at-Risk Measures. Montreal: CIRANO, Paper 2001s-03, 2001.
12. Danielsson, J., de Vries C.: Tail Index and Quantile Estimation with Very High Frequency Data. Journal of Empirical Finance, 4, 1997. p. 241-257.
13. Danielsson J., Jorgensen N. B., Sarma M., de Vries G. C.: Sub-additivity re-examined: the case for Value-at-Risk. Financial Markets Group, FMG Discussion Papers, No. dp549, Feb 2005.
14. Dowd K.: Measuring Market Risk. New York: John Wiley & Sons Ltd., 2002.
15. Embrechts Paul, Resnick I. Sidney, Samorodnitsky Gennady: Extreme value theory as a risk management tool. „Extremes and Insurance“ XXVIII-th International ASTIN Colloquium, Cairns, 1997.
16. Gilli M., Kellezi E.: An Application of Extreme Value Theory for Measuring Risk. Department of Econometrics, University of Geneva and FAME, Geneva, 2003.
17. Guermat C., Harris D.F. R.: Forecasting value at risk allowing for time variation in the variance and kurtosis of portfolio returns. International Journal of Forecasting, 18, Elsevier, 2002. p. 409-419.
18. Hoppe R.: Finance is not Physics. Risk Professional, Vol. 1, No. 7., 1999. http://www.itrac.com/paper/FinancePhysics.doc
19. Inui K., Kijima M.: On the Significance of Expected Shortfall as a Coherent Risk Measure. Journal of Banking and Finance, 29, 2005. p. 853–864.
20. Kondor I., Varga-Haszonits I.: Feasibility of Portfolio Optimization under Coherent Risk Measures. 2008. http://xxx.tau.ac.il/abs/0803.2283
21. Yamai Y., Yoshiba T.: On the Validity of Value-at-Risk: Comparative Analyses with Expected-Shortfall. Bank of Japan, Institute of Monetary and Economic studies, Jan 2002a. p. 57-86.
22. Yamai Y., Yoshiba T.: Comparative Analyses of Expected-Shortfall and Value-at Risk (3): Their Validity under Market Stress. Bank of Japan, Institute of Monetary and Economic studies, Oct 2002b. p. 181-238.
23. Wong C.S. M., Cheng Y. W., Wong Y.P. C.: Market risk management of banks: Implications from the accuracy of VaR forecasts. Journal of Forecasting, 22, 2002. http://www.gloriamundi.org/picsresources/mwwccw.pdf
24. Žiković S.: Measuring Market Risk in EU New Member States. 13th Dubrovnik Economic Conference, Croatian National Bank, Dubrovnik, Croatia, June 26-27, 2007.

APPENDIX

Table 1 - Backtesting results for VaR forecasts (DOW JONES index, cl = 0.95, period 22.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	67	57	70	66	48	37
Frequency of failures	0.067	0.057	0.07	0.066	0.048	0.037
Kupiec test	0.0074077	0.13892**	0.0023303	0.010591	0.57799**	0.96934**
Christoffersen UC test	0.01876	0.32	0.0060675	0.026572	0.77024	0.048421
Christoffersen IND test	0.10479**	0.14277**	0.07007*	0.033774	0.64512**	0.59946**
Christoffersen CC test	0.016951	0.20839	0.0044906	0.0089846	0.86183	0.12424
Lopez test	17.302	7.3261	20.348	16.27	-1.7617	-12.808
Blanco-Ihle test	27.282	28.857	31.789	26.954	21.148	15.908
RMSE	0.01054	0.011884	0.010161	0.010728	0.011267	0.012469
MAPE	3.5638	4.162	3.4517	2.2719	2.5927	2.8746
Average VaR	-0.011478	-0.01248	-0.01107	-0.011475	-0.012222	-0.013404

Table 2 - Backtesting results for VaR forecasts (NASDAQ index, cl = 0.95, period 22.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	58	52	66	70	59	29
Frequency of failures	0.058	0.052	0.066	0.07	0.059	0.029
Kupiec test	0.11056**	0.3514**	0.010591	0.0023303	0.086732*	0.99928**
Christoffersen UC test	0.2571	0.77305	0.026572	0.0060675	0.20362	0.0009791
Christoffersen IND test	0.16454**	0.183**	0.20803**	0.33469**	0.36243**	0.18787**
Christoffersen CC test	0.20026	0.39529	0.038707	0.014546	0.29439	0.0018355
Lopez test	8.3004	2.2753	16.347	20.318	9.261	-20.888
Blanco-Ihle test	20.466	17.798	23.804	24.23	17.528	6.5933
RMSE	0.014471	0.016391	0.013872	0.01363	0.014258	0.018124
MAPE	3.2275	3.9889	3.6526	2.2342	1.9656	2.8646
Average VaR	-0.016049	-0.01765	-0.015402	-0.015026	-0.015971	-0.019807

Table 3 - Backtesting results for VaR forecasts (S&P500 index, cl = 0.95, period 22.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	62	57	68	69	51	45
Frequency of failures	0.062	0.057	0.068	0.069	0.051	0.045
Kupiec test	0.038393	0.13892**	0.0051087	0.0034744	0.40582**	0.73904**
Christoffersen UC test	0.092747	0.32	0.013059	0.0089635	0.88499	0.46085
Christoffersen IND test	0.54717**	0.66999**	0.51042**	0.2994**	0.8**	0.98411**
Christoffersen CC test	0.20307	0.55696	0.036988	0.019188	0.95834	0.76177
Lopez test	12.323	7.3604	18.371	19.308	1.2617	-4.7697
Blanco-Ihle test	27.087	30.222	31.925	28.966	20.214	17.202
RMSE	0.0112	0.012356	0.010581	0.011294	0.011755	0.012324
MAPE	3.2686	4.444	3.2653	2.2686	1.9112	2.2297
Average VaR	-0.012035	-0.01286	-0.011441	-0.011879	-0.012717	-0.013325

Table 4 - Backtesting results for VaR forecasts (CAC index, cl = 0.95, period 21.4.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	64	61	62	62	55	43
Frequency of failures	0.064	0.061	0.062	0.062	0.055	0.043
Kupiec test	0.02075	0.05111*	0.038393	0.038393	0.21008**	0.82671**
Christoffersen UC test	0.051087	0.1223	0.092747	0.092747	0.47493	0.29854
Christoffersen IND test	0.34409**	0.10185**	0.27468**	0.27468**	0.5078**	0.47681**
Christoffersen CC test	0.095343	0.079497	0.13404	0.13404	0.62217	0.4523
Lopez test	14.497	11.445	12.481	12.385	5.3259	-6.755
Blanco-Ihle test	36.306	32.096	34.104	28.775	20.981	14.419
RMSE	0.012943	0.01603	0.013164	0.013943	0.014194	0.015711
MAPE	3.6726	3.9778	3.5205	2.3984	2.5272	2.2342
Average VaR	-0.013719	-0.016042	-0.014006	-0.014499	-0.015215	-0.016789

Table 5 - Backtesting results for VaR forecasts (DAX index, cl = 0.95, period 14.4.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	65	50	68	60	49	30
Frequency of failures	0.065	0.05	0.068	0.06	0.049	0.03
Kupiec test	0.01493	0.46247**	0.0051087	0.067063*	0.52026**	0.99872**
Christoffersen UC test	0.037108	1	0.013059	0.15895	0.88427	0.0017752
Christoffersen IND test	0.18236**	0.74748**	0.12208**	0.21488**	0.28588**	0.91569**
Christoffersen CC test	0.0468	0.94949	0.013901	0.17184	0.55987	0.0075222
Lopez test	15.41	0.34313	18.44	10.378	-0.70065	-19.786
Blanco-Ihle test	27.402	22.342	30.779	29.687	19.375	12.214
RMSE	0.014434	0.019624	0.013996	0.014141	0.014661	0.016861
MAPE	3.7148	3.232	3.8824	1.8779	1.5927	2.2886
Average VaR	-0.015518	-0.019274	-0.014969	-0.014844	-0.015841	-0.018098

Table 6 - Backtesting results for VaR forecasts (FTSE index, cl = 0.95, period 29.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	61	59	60	60	58	41
Frequency of failures	0.061	0.059	0.06	0.06	0.058	0.041
Kupiec test	0.05111*	0.086732*	0.067063*	0.067063*	0.11056**	0.8937**
Christoffersen UC test	0.1223	0.20362	0.15895	0.15895	0.2571	0.17827
Christoffersen IND test	0.50093**	0.41436**	0.45661**	0.82685**	0.39588**	0.55487**
Christoffersen CC test	0.24164	0.31942	0.28106	0.36203	0.36695	0.33948
Lopez test	11.447	9.4454	10.434	10.335	8.2723	-8.8011
Blanco-Ihle test	38.476	38.207	36.347	27.992	19.597	12.667
RMSE	0.011353	0.013063	0.011301	0.012538	0.012258	0.013725
MAPE	3.758	3.6093	3.6471	1.6082	1.3474	1.9745
Average VaR	-0.011729	-0.013081	-0.011737	-0.012601	-0.013003	-0.014442

Table 7 - Backtesting results for VaR forecasts (DOW JONES index, cl = 0.99, period 22.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	HS 500	VCV	Risk Metrics	GARCH RM	HHS
Number of failures	19	17	28	21	19	14
Frequency of failures	0.019	0.017	0.028	0.021	0.019	0.014
Kupiec test	0.0032884	0.013833	6.397E-07	0.0006518	0.0032884	0.082412*
Christoffersen UC test	0.010956	0.043113	2.744E-06	0.0023116	0.010956	0.23056
Christoffersen IND test	0.39069**	0.44295**	0.81013**	0.07494*	0.39069**	0.52813**
Christoffersen CC test	0.027197	0.096347	1.633E-05	0.001974	0.027197	0.39944
Lopez test	9.0735	7.0799	18.137	11.093	9.0909	4.0476
Blanco-Ihle test	4.4985	4.4843	9.1215	7.5062	6.0212	2.9546
RMSE	0.016036	0.018566	0.014562	0.015708	0.016607	0.019198
MAPE	1.3607	1.3674	2.0888	1.4218	1.3252	0.99445
Average VaR	-0.017142	-0.019209	-0.015791	-0.016383	-0.017286	-0.019537

Table 8 - Backtesting results for VaR forecasts (NASDAQ index, cl = 0.99, period 22.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	HS 500	VCV	Risk Metrics	GARCH RM	HHS
Number of failures	15	17	20	17	13	5
Frequency of failures	0.015	0.017	0.02	0.017	0.013	0.005
Kupiec test	0.047871	0.013833	0.0014965	0.013833	0.13444**	0.93386**
Christoffersen UC test	0.13898	0.043113	0.0051465	0.043113	0.36211	0.078594
Christoffersen IND test	0.49887**	0.44295**	0.366**	0.44295**	0.55821**	0.82254**
Christoffersen CC test	0.26626	0.096347	0.01327	0.096347	0.55617	0.20763
Lopez test	5.0699	7.064	10.096	7.0791	3.0631	-4.9712
Blanco-Ihle test	3.3692	2.8606	4.7917	4.9517	3.3408	1.3432
RMSE	0.020071	0.023546	0.020269	0.020252	0.021057	0.02641
MAPE	1.0166	1.4107	1.8335	0.93008	0.98002	0.69589
Average VaR	-0.021891	-0.024722	-0.021931	-0.021535	-0.022588	-0.027502

Table 9 - Backtesting results for VaR forecasts (S&P500 index, cl = 0.99, period 22.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	HS 500	VCV	Risk Metrics	GARCH RM	HHS
Number of failures	21	20	29	25	20	14
Frequency of failures	0.021	0.02	0.029	0.025	0.02	0.014
Kupiec test	0.0006518	0.0014965	2.06E-07	1.559E-05	0.0014965	0.082412*
Christoffersen UC test	0.0023116	0.0051465	9.051E-07	6.192E-05	0.0051465	0.23056
Christoffersen IND test	0.45862**	0.366**	0.86291**	0.65394**	0.366**	0.52813**
Christoffersen CC test	0.0073236	0.01327	5.7E-06	0.0002969	0.01327	0.39944
Lopez test	11.076	10.094	19.143	15.092	10.07	4.0508
Blanco-Ihle test	4.5406	5.2283	8.8611	7.2991	4.2648	3.1555
RMSE	0.016322	0.018924	0.015127	0.016409	0.017001	0.017984
MAPE	1.5327	1.5683	2.1321	1.7469	1.1376	1.1077
Average VaR	-0.017388	-0.019405	-0.01633	-0.016992	-0.017985	-0.018847

Table 10 - Backtesting results for VaR forecasts (CAC index, cl = 0.99, period 21.4.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	19	15	30	23	17	10
Frequency of failures	0.019	0.015	0.03	0.023	0.017	0.01
Kupiec test	0.0032884	0.047871	6.42E-08	0.0001091	0.013833	0.41696**
Christoffersen UC test	0.010956	0.13898	2.888E-07	0.0004102	0.043113	1
Christoffersen IND test	0.049257	0.49887**	0.067832*	0.10865**	0.28968**	0.65293**
Christoffersen CC test	0.0056872	0.26626	3.629E-07	0.000537	0.073827	0.90383
Lopez test	9.1363	5.1303	20.21	13.14	7.107	0.074238
Blanco-Ihle test	5.8632	5.2447	10.05	7.6256	4.9213	3.0033
RMSE	0.023047	0.027756	0.01876	0.020346	0.020674	0.023136
MAPE	1.7059	1.5361	2.182	1.4173	0.93341	0.70699
Average VaR	-0.024122	-0.027796	-0.02002	-0.020685	-0.02152	-0.023783

Table 11 - Backtesting results for VaR forecasts (DAX index, cl = 0.99, period 14.4.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	16	13	24	23	19	13
Frequency of failures	0.016	0.013	0.024	0.023	0.019	0.013
Kupiec test	0.026391	0.13444**	4.203E-05	0.0001091	0.0032884	0.13444**
Christoffersen UC test	0.079429	0.36211	0.0001625	0.0004102	0.010956	0.36211
Christoffersen IND test	0.47047**	0.55821**	0.60331**	0.10865**	0.37015**	0.15701**
Christoffersen CC test	0.16551	0.55617	0.0007132	0.000537	0.026309	0.24252
Lopez test	6.129	3.1247	14.175	13.157	9.1104	3.0804
Blanco-Ihle test	5.5367	5.3097	8.3212	9.1349	5.0871	3.2936
RMSE	0.023363	0.030537	0.02018	0.020661	0.021443	0.023791
MAPE	1.1676	1.374	1.9822	1.1987	0.85572	0.88679
Average VaR	-0.024563	-0.030006	-0.02148	-0.021246	-0.022405	-0.024519

Table 12 - Backtesting results for VaR forecasts (FTSE index, cl = 0.99, period 29.3.2004 - 12.3.2008). Symbols * and ** denote significance at 5 and 10% levels.

	HS 250	**HS 500**	**VCV**	**Risk Metrics**	**GARCH RM**	**HHS**
Number of failures	19	15	32	25	17	14
Frequency of failures	0.019	0.015	0.032	0.025	0.017	0.014
Kupiec test	0.0032884	0.047871	5.674E-09	1.559E-05	0.013833	0.082412*
Christoffersen UC test	0.010956	0.13898	2.669E-08	6.192E-05	0.043113	0.23056
Christoffersen IND test	0.39069**	0.49887**	0.37506**	0.15123**	0.44295**	0.52813**
Christoffersen CC test	0.027197	0.26626	1.294E-07	0.0001172	0.096347	0.39944
Lopez test	9.1265	5.1385	22.202	15.138	7.0805	4.0471
Blanco-Ihle test	6.4493	6.9151	11.775	7.8833	3.9671	2.2519
RMSE	0.019465	0.023463	0.015917	0.018223	0.018082	0.020077
MAPE	1.5339	1.3973	2.6848	1.7037	1.1232	0.95006
Average VaR	-0.020054	-0.02309	-0.016766	-0.017972	-0.018391	-0.020045

Table 13 - Backtesting results for ETL forecasts (DOW JONES index, ξ = 0.2, cl = 0.95, 0.99, period 22.3.2004 - 12.3.2008)

Cl	**95%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0296	0.0133	**0.0116**	0.0273	0.0129	0.0124
RMSE2	0.0177	0.0058	**0.0052**	0.0164	0.0055	0.0058
MAE	0.0161	0.0045	**0.0035**	0.0155	0.0047	0.0042
MAPE	1.0222	0.2771	**0.1846**	1.0865	0.3167	0.2179
Cl	**99%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0526	0.0196	**0.0109**	0.0512	0.0189	0.0141
RMSE2	0.0358	0.0082	**0.0049**	0.0350	0.0081	0.0057
MAE	0.0336	0.0073	**0.0026**	0.0338	0.0073	0.0043
MAPE	1.6505	0.3498	**0.1038**	1.7272	0.3816	0.1679

Table 14 - Backtesting results for ETL forecasts (NASDAQ index, ξ = 0.31, cl = 0.95, 0.99, period 22.3.2004 - 12.3.2008)

Cl	**95%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0577	0.0178	0.0137	0.0555	0.0165	**0.0132**
RMSE2	0.0411	0.0074	0.0054	0.0391	0.0067	**0.0054**
MAE	0.0396	0.0066	**0.0037**	0.0381	0.0058	0.0038
MAPE	2.0284	0.3368	**0.1485**	1.9568	0.3046	0.1623
Cl	**99%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.1023	0.0246	0.0162	0.1052	0.0242	**0.0133**
RMSE2	0.0807	0.0107	0.0072	0.0832	0.0104	**0.0049**
MAE	0.0787	0.0102	0.0042	0.0823	0.0099	**0.0031**
MAPE	3.3442	0.4395	0.1198	3.4052	0.4190	**0.0999**

Table 15 - Backtesting results for ETL forecasts (S&P500 index, ξ = 0.24, cl = 0.95, 0.99, period 22.3.2004 - 12.3.2008)

Cl	**95%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0367	0.0127	**0.0112**	0.0329	0.0130	0.0134
RMSE2	0.0228	0.0052	**0.0048**	0.0206	0.0055	0.0063
MAE	0.0218	0.0040	**0.0030**	0.0197	0.0046	0.0045
MAPE	1.2980	0.2468	**0.1428**	1.3009	0.2963	0.2137
Cl	**99%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0754	0.0253	**0.0125**	0.0679	0.0214	0.0143
RMSE2	0.0560	0.0115	**0.0057**	0.0494	0.0093	0.0057
MAE	0.0532	0.0109	**0.0031**	0.0476	0.0084	0.0042
MAPE	2.3979	0.4834	**0.1146**	2.2176	0.3976	0.1542

Table 16 - Backtesting results for ETL forecasts (CAC index, ξ = 0.23, cl = 0.95, 0.99, period 21.4.2004 - 12.3.2008)

Cl	**95%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0435	0.0183	**0.0153**	0.0362	0.0166	0.0169
RMSE2	0.0273	0.0083	**0.0083**	0.0217	0.0077	0.0095
MAE	0.0250	0.0058	**0.0040**	0.0208	0.0054	0.0056
MAPE	1.2463	0.2677	**0.1360**	1.1310	0.2595	0.2222
Cl	**99%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0723	0.0261	**0.0232**	0.0732	0.0243	0.0261
RMSE2	0.0511	0.0113	**0.0131**	0.0515	0.0106	0.0146
MAE	0.0500	0.0103	**0.0067**	0.0504	0.0092	0.0100
MAPE	2.0705	0.4046	**0.1544**	2.0357	0.3711	0.2571

Table 17 - Backtesting results for ETL forecasts (DAX index, $\xi = 0.3$, cl = 0.95, 0.99, period 14.4.2004 - 12.3.2008)

Cl	**95%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0548	0.0186	0.0182	0.0501	0.0176	**0.0174**
RMSE2	0.0378	0.0091	0.0106	0.0345	0.0087	**0.0103**
MAE	0.0349	0.0060	0.0054	0.0336	0.0061	**0.0053**
MAPE	1.7558	0.2702	**0.1730**	1.8551	0.3105	0.1901
Cl	**99%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.1007	0.0273	**0.0219**	0.0984	0.0263	0.0266
RMSE2	0.0781	0.0124	**0.0132**	0.0758	0.0120	0.0166
MAE	0.0765	0.0110	**0.0059**	0.0750	0.0105	0.0093
MAPE	3.1525	0.4231	**0.1357**	3.1049	0.4156	0.2041

Table 18 - Backtesting results for ETL forecasts (FTSE index, $\xi = 0.22$, cl = 0.95, 0.99, period 29.3.2004 - 12.3.2008)

Cl	**95%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0352	**0.0140**	0.0143	0.0306	0.0154	0.0156
RMSE2	0.0215	**0.0055**	0.0065	0.0183	0.0071	0.0086
MAE	0.0198	0.0045	**0.0041**	0.0165	0.0051	0.0055
MAPE	1.2424	0.2967	**0.1674**	1.0248	0.2653	0.2370
Cl	**99%**					
	Frechet GARCH RM	Gumbel GARCH RM	Bootstrap HHS	Frechet VCV	Gumbel VCV	Bootstrap HS
RMSE1	0.0799	0.0266	**0.0156**	0.0590	0.0209	0.0227
RMSE2	0.0567	0.0115	**0.0060**	0.0407	0.0088	0.0115
MAE	0.0523	0.0099	**0.0039**	0.0392	0.0081	0.0089
MAPE	2.0076	0.3885	**0.1362**	1.9004	0.3924	0.2696