Academy of Economic Studies Bucharest
Doctoral School of Finance and Banking

Dissertation Paper

Measuring market risk: a copula and extreme value approach

Supervisor
PhD. Professor Moisă Altăr

M. Sc. Student
Stângă Alexandru Leonard

July 2007

Abstract

This paper presents a methodology for measuring the risk of a portfolio composed of assets with heteroscedastic return series. In order to obtain good estimates for Value-at-Risk and Expected Shortfall, the model tries to capture as realistically as possible the data generating process for each return series and also the dependence structure that exists at the portfolio level. For this purpose, the individual return series are modelled using GARCH methods with semi-parametric innovations and the dependence structure is defined with the help of a Student t copula. The model built with these techniques is then used for the simulation of a portfolio return distribution that allows the estimation of the risk measures. This methodology is applied to a portfolio of five Romanian stocks and the accuracy of the risk measures is then tested using a backtesting procedure.

Contents

A. Introduction

The financial institutions with significant amounts of trading activity proved to be very vulnerable to extreme market movements and, in time, the measurement of market risk became a primary concern for regulators and also for internal risk control.

For example, U.S. banks and bank holding companies with an important trading portfolio are subject to market risk requirements. They have been required to hold capital against their defined market risk exposures, and, the necessary capital is a function of banks' own risk estimates.

In this context, Value-at-risk (VaR) has emerged as one of the most used risk measure in the financial industry, mostly because of its simplicity and intuitive interpretation. Value at Risk measures the worst loss to be expected of a portfolio over a given time horizon at a given confidence level.

Although a clear definition of VaR may be given, this measure of risk doesn’t have a unique method of estimation because its accuracy highly depends on the ability to identify the true portfolio loss distribution. Simple models of estimation, like Historical Simulation or Variance-Covariance failed to give accurate high confidence level estimates but are used often because of their low computing power demands. More complex models based on Monte Carlo simulation have the advantage of flexibility in modelling the loss distribution and the potential of being more accurate but they are difficult to compute for very complex portfolios with a high number of risk factors.

Although VaR offers a simple and intuitive way of evaluating market risk, Artzner et al. (1997, 1998) have criticized it as a measure of risk for two main reasons. First they proved that VaR is not necessarily subadditive[1] and secondly, this measure gives only an upper limit on the losses given a confidence level, but it tells nothing about the potential size of the loss if this upper limit is exceeded. In order to solve these two issues, they propose the use of the so-called expected shortfall or tail conditional expectation instead of VaR. Expected shortfall measures the expected loss given that the loss exceeds VaR; in mathematical terms it can be written as $E[L \mid L > VaR]$.

[1] This property represents the benefits of diversification for a portfolio, the risk derived from the portfolio (x + y) is lower than (or equal to) the risk derived from the sum of the risk of the individual securities x , y

This paper aims at computing accurate estimates for both measures of risk by using a flexible modelling technique in order to build the loss distribution of the portfolio. The first stage of this process represents the modelling of the return series for each individual stock.

One possible solution for this issue is based on non-parametric methods (empirical methods), that make no assumptions concerning the nature of the empirical distribution function. However, these techniques have several drawbacks, for example they cannot be used to estimate out of sample quantiles and the kernel based estimators usually perform poorly in the smoothing of tails (Silverman, 1986).

Another possible option would be the use of parametric methods for describing the entire distribution of the series. Empirical evidences have shown that the distributions of financial returns series exhibit fat tails and sometimes negative skews (Zangari 1996). For this reason, the normal distribution, in spite of its popularity, it is not considered a good choice as its symmetry and exponentially decaying tail doesn't seem to be supported by data. An alternative to the normal distribution may be considered the Student-t distribution as it displays polynomial decay in the tails and thus having heavier tails than the normal one. Hence, it may be able to capture the observed excess kurtosis although it maintains the hypothesis of symmetry.

A third possibility would be the use of extreme value parametric methods for describing the tails of the distribution and parametric (ex. gaussian, student-t) or non-parametric methods (ex. kernel smoothing) for the interior of the distribution. These types of tools permit a high flexibility because the parameters for each tail can be estimated separately and thus, both the excess kurtosis and the skewness of the financial series can be incorporated into the model.

The methodology used in this paper takes advantage of the flexibility provided by the third method in order to capture the data generating process for each financial series. More specifically it uses extreme value theory for the estimation of the tail parameters and a kernel smoothing technique for building the interior of the distribution.

Once the tools used for the construction of the returns series distribution are defined, the next step consists of choosing the distribution that should be analyzed. In the context of market risk management, the analysis of both the conditional and unconditional returns distribution provides useful information. However, the conditional returns distribution takes into account the current volatility background and forms the basis for short term risk evaluation, thus being the main interest of market risk management.

The analysis of unconditional tails provides additional information about risk and can be used for the estimation of the magnitude of a rare adverse event that can lead to an important loss. This kind of information may be used for stress testing scenarios and long – term risk estimation.

Because short-term risk evaluation is a primary concern of market risk management, the analysis of the conditional return distribution is the main focus of this paper and it is based on the assumption that returns follow a stationary time series process with stochastic volatility structure. This premise is supported by empirical evidence regarding the presence of stochastic volatility in the financial time series (Frey, 1997) and implies that returns are not necessarily independent over time.

The specific methodology used for describing volatility dynamics is based on the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) methods and has an additional advantage of providing an iid innovations series that can be directly modelled by using a semi-parametric distribution with tails described by extreme value theory.

The estimation of the GARCH models for each returns series and the construction of the semi-parametric distributions based on the innovations represents the first stage of the portfolio risk evaluation. At the second stage of the process the dependence structure between assets is defined with the help of copula methods.

A joint distribution function for risk factors contains a description of the marginal distribution for each individual factor and also a description of their dependence structure. The copula methods provide a mechanism for isolating the dependence structure of the portfolio from the individual margins of the assets and as a consequence it provides flexibility in modelling the portfolio as a whole.

The composition of the portfolio has to be taken into consideration in order to choose a specific copula for modelling. There is empirical evidence that equity markets tend to be more correlated in volatile times (Longin, 2000) and this implies that the dependence structure should allow for high tail dependence among assets.

Unfortunately the Gaussian dependence structure, relying only on the notion of correlation, doesn't allow for extreme co-movements regardless of potentially large magnitudes in correlation between the underlying individual assets.

Because of these drawbacks, the dependence structure used in this paper is based on the Student-t copula that relies on the notion of correlation but, in addition it is also characterized by a parameter (Degrees of Freedom – DoF) that controls the tail dependence (extreme co-movements) of marginal distributions.

The final stage of the risk evaluation process is based on the parameters estimated in the previous stages and consists of the simulation of a portfolio conditional returns distribution that can be used for Value-at-Risk and Expected Shortfall estimation.

In the following sections of this paper, the methodology of the risk evaluation process is presented in detail and the results of each intermediary stage are displayed. The final part of this analysis presents an evaluation of the methodology by using a backtesting procedure in order to test the accuracy of the risk measures.

B. Literature Review

The modelling of financial return distributions using extreme value theory is applied and tested in several research papers, both from a general market risk perspective and from a more specific financial sector perspective.

Danielsson and De Vries (1997b) propose a semi-parametric method for VaR estimation, where the unconditional return distribution is described by a combination of non-parametric historical simulation and extreme value theory. The authors build their model based on the assumption that extreme returns occur infrequently, and do not appear to be related to a particular level of volatility or exhibit time dependence. As a consequence they propose the unconditional loss distribution as a base for Value-at-Risk estimation. However, the research conducted by McNeil and Frey (2000) in their 2000 article 'Estimation of Tail-Related Risk Measures for Heteroskedastic Financial Time Series: an Extreme Value Approach' contradicts the assumption of Danielsson and De Vries regarding the superiority of the VaR estimates obtained from the unconditional distribution and prove that a conditional approach against the current volatility background is better suited for VaR estimation.

Kaj Nystrom and Jimmy Skoglund, agreed on this matter in their 2001 article on univariate Extreme Value Theory, GARCH and Measures of risk and believe that in order to measure portfolio risks it is important to correctly identify a model for the risk factors.

Although they both combine GARCH models to estimate the current volatility and the Extreme Value Theory for estimating the tail of the innovation distribution of the GARCH model, Nystrom and Skoglund introduce the assumption of asymmetry. This concept is important in the context of the GARCH model as it no longer assumes that negative and positive shocks have the same impact on volatility.

One important similarity between the two articles is given by the assumption of a Student's t distribution for the innovation series. Although the Student-t distribution continues to assume the symmetry hypothesis, it also makes it possible to capture the observed kurtosis. This is an important part of the model, as Nystrom ang Skoglund emphasize, because the distributions of the financial series are often characterized by excess kurtosis and negative skewness. So by using the normal distribution approximation, the risk of high quantiles is severely underestimated and it is for this reason that the authors chose an alternative to this distribution in the form of a Student-t distribution.

After having looked at empirical evidence, Nystrom and Skoglund came to the conclusion that, in the case of daily risk measurement, while the normal model indeed tends to underestimate the lower tail and overestimate the upper one, the t distribution also has its flaws, in the sense that it actually tends to overestimate both tails.

In relation to the distribution of residuals, McNeil and Frey, while agreeing in favour of a t distribution instead of a normal one, they believe GPD-approximation to be a much better model for when the tails are asymmetric. If the tails of the distribution of residuals were symmetric then the t distribution, they argue, is a good alternative.

Embrechts et al (1999) in their article "Extreme Value Theory as a Risk Measurement Tool" pronounced themselves in favour of using a parametric estimation technique which is based on a limit result for the excess distribution over high threshold, a technique also preferred by McNeil and Frey. This is a technique which will be explained in greater detail later in the methodology of this paper.

C. Methodology

I. GARCH models

In a generalized autoregressive conditional heteroscedasticity (GARCH) model, returns are assumed to be generated by a stochastic process with time-varying volatility. This implies that the conditional distributions change over time in an autocorrelated way and the conditional variance is an autoregressive process

The GARCH model was introduced by Bollerslev (1986) and it consists of two equations, the conditional mean equation that explains how the expected value of the return changes over time and the conditional variance equation that describes the evolution of the conditional variance of the unexpected return process.

An ARMA(m,n) model describes how the return changes over time (the conditional mean equation):

$$y_t = c + \sum_{i=1}^{m} \phi_i y_{t-i} + \sum_{j=1}^{n} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

In this model it is assumed that ε_t is independent and identically distributed with mean zero and variance σ^2. GARCH extends the ARMA model by assuming that $\varepsilon_t = z_t \sigma_t$, where z_t is independent and identically distributed with mean zero and unit variance and $z_t \sigma_t$ are stochastically independent.

The dynamics of σ_t^2, the conditional variance at time t, is described by the second equation of GARCH model, and the representation of this equation for GARCH(p,q) is:

$$\sigma_t^2 = k + \sum_{i=1}^{p} G_i \sigma_{t-i}^2 + \sum_{j=1}^{q} A_j \varepsilon_{t-j}^2$$

The coefficients of the GARCH model must respect some constraints in order to avoid the possibility that the volatility becomes negative or the process non-stationary.

$$\sum_{j=1}^{q} A_j + \sum_{i=1}^{p} G_i < 1$$; for a stationary volatility process

$$A_j \geq 0$$
$$G_j \geq 0 \qquad \text{; for positive volatility}$$
$$k > 0$$

In some equity markets it can be observed that volatility is higher if the market is falling than if the market is rising. The volatility response to a large negative return is often greater than it is to a large positive return of the same magnitude. One possible reason for this effect may be explained by the debt/equity ratio. When the equity price falls the debt remains constant in the short term, so the debt/equity ratio increases, the company becomes more highly leveraged and so the future of the firm becomes more uncertain.

The asymmetry in volatility clustering caused by the leverage effect can be captured with asymmetric GARCH models like GJR-GARCH, introduced by Glosten, Jagannathan and Runkle (1993).

$$\sigma_t^2 = k + \sum_{i=1}^{p} G_i \sigma_{t-i}^2 + \sum_{j=1}^{q} A_j \varepsilon_{t-j}^2 + \sum_{j=1}^{q} L_j Sgn_{t-j} \varepsilon_{t-j}^2$$

$$Sgn_{t-j} = \begin{cases} 1 & \varepsilon_{t-j} < 0 \\ 0 & \varepsilon_{t-j} \geq 0 \end{cases}$$

$$\sum_{j=1}^{q} A_j + \sum_{i=1}^{p} G_i + \frac{1}{2}\sum_{j=1}^{q} L_j < 1 \qquad \text{; for a stationary volatility process}$$

$$A_j \geq 0$$
$$G_j \geq 0$$
$$k > 0 \qquad \text{; for positive volatility}$$
$$A_j + L_j \geq 0$$

In estimating the return series with GARCH models it is commonly assumed that the innovation series (z_t) has a standard normal distribution. This premise relies on the fact that the excess kurtosis of the return distribution can be partially captured by the GARCH model. However, it is possible that some of the excess kurtosis to remain unexplained and as a consequence the assumption of normality for the innovations might not be valid (McNeil, 2000). Bollerslev (1986) proposed the use of the t-distribution for the innovation series in order to better explain the excess kurtosis of the financial series.

II. Extreme Value Theory (EVT) models

Extreme Value Theory (EVT) was conceived as the probabilistic theory for studying rare events (i.e. realizations from the tails of a distribution) and it is mainly used for the parametric modelling of the tails of a distribution. Because EVT needs information only about extreme events in order to model the tails, it is not necessary to make a particular assumption about the shape of the entire distribution in order to use the theory. Furthermore, because EVT is a parametric technique it can be used to estimate out of sample quantiles by extrapolation.

The EVT uses two approaches in order to study the extreme events. The first method (block maxima) is used to describe the distribution of minimum or maximum realizations of a process. In order to apply this method the data sample is divided into blocks and the maximum value from each block is considered an extreme event. These values are then extracted from the sample data and modelled separately by fitting them to a Generalized Extreme Value (GEV) distribution.

The second method (peak-over-threshold) is used for modelling the distribution of exceedances over a particular threshold. In order to identify the extreme values, this method sets a threshold over which all realizations of the process are considered extreme. After setting this critical value, the observations of the sample data that are larger than the threshold are extracted and the exceedances are computed (exceedance = extreme value – threshold). Finally, in order to describe the extreme events, the exceedances are fitted to a Generalized Pareto Distribution.

Both EVT methods have a parameter that is used for identifying the extreme values of the process that is analyzed. This parameter has to be fixed before the extreme data can be fitted to a certain distribution. In the case of block maxima methods this parameter is the size of the block and in the case of peak-over-threshold methods the value of the threshold has this role.

1. Limiting distributions of the maxima.

The extreme value theory is applied in order to describe the limiting distributions of the sample maxima. This concept is similar to the central limit theorem that sets the normal distribution as the limiting distribution of sample averages. The EVT describes this family of limiting distributions under a single parameterization known as the generalized extreme value (GEV) distribution.

If r_t, t = 1, 2, . . . , n, is an uncorrelated sample of returns with the distribution function $F(x) = P\{r_t \leq x\}$, which has variance σ^2 and mean μ [2], we denote the sample maxima[3] of r_t by $M_1 = r_1$, $M_2 = \max(r_1, r_2)$,…, $M_n = \max(r_1, ..., r_n)$, where $n \geq 2$. Let R denote the real line, if there exists a sequence $c_n > 0$, $d_n \in R$ and some non-degenerate distribution function H such that

$$\left(\frac{M_n - d_n}{c_n}\right) \xrightarrow{d} H$$

then H belongs to one of the following three families of distributions:

Gumbel:

$$\Lambda(x) = e^{-e^{-x}} \text{, } x \in R$$

Fréchet:

$$\Phi_{\alpha}(x) = \begin{cases} 0, x \leq 0 \\ e^{-x^{-\alpha}}, x > 0, \alpha > 0 \end{cases}$$

Weibull:

$$\Psi_{\alpha}(x) = \begin{cases} e^{-\left(-x^{\alpha}\right)}, x \leq 0, \alpha < 0 \\ 1, x > 0 \end{cases}$$

[2] We assume for convenience that $\mu = 0$ and $\sigma^2 = 1$.
[3] The sample maxima is $\min(r_1, ..., r_n) = -\max(-r_1, ..., -r_n)$.

The Fisher and Tippett (1928) theorem suggests that the limiting distribution of the maxima belongs to one of the three distributions above, regardless of the original distribution of the observed data.

If we consider **ξ= 1/α** (von Mises, 1936) and Jenkinson, 1955), Fréchet, Weibull and Gumbel distributions can be expressed as a unified model with a single parameter. This representation is known as the generalized extreme value distribution (GEV):

$$H_{\xi}(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0,\ 1+\xi x > 0 \\ e^{-e^{-x}} & \text{if } \xi = 0 \end{cases}$$

where **ξ= 1/α** is a shape parameter and **α** is the tail index.

The class of distributions of F(x) where the Fisher-Tippett theorem holds is quite large. One of the conditions (Falk, Hüssler, and Reiss 1994) is that F(x) has to be in the domain of attraction of the Frechet distribution (ξ > 0), which in general is true for the financial time series. Gnedenko (1943) shows that if the tail of F(x) decays like a power function (heavy-tailed distributions like Pareto, Cauchy, Student-t), then it is in the domain of attraction of the Fréchet distribution.

2. Limiting distribution of exceedances over a threshold

The limiting distribution of exceedances over a threshold is a member of the family of extreme value distributions. In order to estimate the parameters of this limiting distribution, first we have to identify the extreme values of the sample data. If we take a sample of observations, r_t, t = 1, 2, . . . , n with a distribution function $F(x) = \Pr\{r_t \leq x\}$ and we set a high-threshold u, then the exceedances over this threshold occur when $r_t > u$ for any t in t = 1, 2, . . . , n. An excess over u is defined by $y = r_i - u$ (peak-over-threshold method).

For a high threshold u, the probability distribution of excess values of r over threshold u is defined as:

$$F_u(y) = \Pr\{r - u \le y \mid r > u\}$$

Given that r exceeds the threshold u, this represents the probability that the value of r exceeds the threshold u by at most an amount y. This conditional probability may be written as:

$$F_u(y) = \frac{\Pr\{r \mid -u \le y, r > u\}}{\Pr(r > u)} = \frac{F(y+u) - F(u)}{1 - F(u)}$$

Because r > u and x = y + u , we can also write the following expression:

$$F(x) = [1 - F(u)]F_u(y) + F(u)$$

The theorem of Balkema and de Haan (1974) and Pickands (1975) shows that for sufficiently high threshold u, the distribution function of the excess may be approximated by the Generalized Pareto Distribution (GPD)

The GPD can be defined as:

$$G_{\xi,\sigma,\upsilon} = \begin{cases} 1 - \left(1 + \xi \dfrac{x - \upsilon}{\sigma}\right)^{-\frac{1}{\xi}} & ,if \quad \xi \ne 0 \\ 1 - e^{-\frac{(x-\upsilon)}{\sigma}} & ,if \quad \xi = 0 \end{cases}$$

$$x \in \begin{cases} [\upsilon, \infty] & ,if \quad \xi \ge 0 \\ \left[\upsilon, \upsilon - \dfrac{\sigma}{\xi}\right] & ,if \quad \xi < 0 \end{cases}$$

where $\boldsymbol{\xi} = \mathbf{1}/\boldsymbol{\alpha}$ is a shape parameter and $\boldsymbol{\alpha}$ is the tail index, $\boldsymbol{\sigma}$ is the scale parameter, and $\boldsymbol{\upsilon}$ is the location parameter.

When $\boldsymbol{\upsilon} = 0$ and $\boldsymbol{\sigma} = 1$, the representation is known as the standard GPD. The relationship between the limiting distribution of exceedances (standard GPD) and the limiting distributions of the sample maxima (GEV) is:

$$G_\xi(x) = 1 + \log H_\xi(x) \qquad if \quad \log H_\xi(x) > -1$$

When $\xi > 0$, GDP takes the form of the ordinary Pareto distribution which is a heavy-tail distribution, and as a consequence it is very useful for the analysis of financial series. If $\xi > 0$, $E[X^k]$ is infinite for $k \geq 1/\xi$ and in order to have a financial series with finite variance, ξ must be less than 0.5. When $\xi = 0$, the GPD corresponds to the thin-tailed distributions and for $\xi < 0$ it corresponds to finite-tailed distributions.

For $\xi > -0.5$ the GPD model can be estimated with the maximum-likelihood method because in this case maximum-likelihood regularity conditions are fulfilled and the maximum-likelihood estimates are asymptotically normally distributed (Hosking and Wallis 1987).

One important aspect when applying EVT is the choice of the threshold value. This value must be set low enough in order to have a sufficient number of exceedances for computing accurate estimates of the tail parameters with the ML method. At the same time the threshold must be set high enough in order to have the GPD as the limiting distribution of the exceedances. Unfortunately there is no natural estimator of the threshold and thus, its value must be set more or less arbitrarily. In practice, instead of assuming that the tail of the underlying distribution begins at the threshold *u*, we can choose a fraction *k/n* of the sample data which is considered to be the tail of the distribution, hence implicitly choosing also the threshold value.

McNeil and Frey (2000) and Nyström and Skoglund (2001) conducted Monte-Carlo experiments in order to evaluate the properties of the ML estimator for various distributions and sample sizes. The results show that the ML estimator is almost invariant to the threshold value if k is set between 5-13% of the sample data.

III. Copula models

The essential idea behind the copula approach is that a joint distribution can be decomposed into marginal distributions and a dependence structure represented by a function called copula. Using a copula, marginal distributions that are estimated separately can be combined in a joint risk distribution that preserves the original characteristics of the marginals.

A real advantage of using copula functions for the description of dependence structures consists in the ability to combine different types of marginal distributions (parametric or non-parametric) into a joint risk distribution. At the same time, the joint distributions created using copulas can have a dependence structure described by more than a simple correlation matrix (e.g. the t-copula has an additional tail dependence parameter - degrees of freedom).

Definition (Copula) A function $C:[0,1]^n \to [0,1]$ is a n-dimensional copula if it satisfies the following properties:

a) For all $u_i \in [0,1], C(1,...,1,u_i,1...,1) = u_i$

b) For all $u_i \in [0,1]^n, C(u_1...,u_n) = 0$ if at least one $u_i = 0$

c) C is grounded and n-increasing

Sklar's theorem: Given a d-dimensional distribution function G with continuous marginal cumulative distributions F_1, . . . , F_d, then there exists a unique n-dimensional copula $C:[0,1]^d \to [0,1]$ such that for $x \in \Re^n$:

$$G(x_1,....,x_n) = C(F_1(x_1),....,F_n(x_n)) \qquad \text{(c1)}$$

Moreover, if $F_1, F_2,\ldots,F_n$ are continuous, then C is unique.

Sklar's Theorem is a fundamental result concerning copula functions and basically it states that any joint distribution can be written in terms of a copula and marginal distribution functions.

If F is a univariate distribution function then the generalized inverse of F is defined as

$$F^{-1}(t) = \inf\{x \in R : F(x) \geq t\}$$

for all $t \in [0,1]$ and using the convention $\inf\{\phi\} = \infty$

Corollary Let G be an n-dimensional distribution function with continuous marginals $F_1, \ldots,$ F_d and an n-dimensional copula C. Then for any $u \in [0,1]^n$,

$$C(u_1,...,u_n) = G(F_1^{-1}(u_1),..., F_n^{-1}(u_n)) \qquad \text{(c2)}$$

Note: without the continuity assumption, this relation may not hold (Nelsen 1999).

The copula links the quantiles of the two distributions rather than the original variables, so one of the key properties of a copula is that the dependence structure is unaffected by a monotonically increasing transformation of the variables.

Theorem (copula invariance) Consider n continuous random variables $(X_1, \ldots, X_n)$ with copula C. If $g_1, \ldots, g_n : R \rightarrow R$ are strictly increasing on the range of $X_1, \ldots, X_n$, then $(g_1(X_1), \ldots, g_n(X_n))$ also have C as their copula.

Remark By applying Sklar's theorem and by exploiting the relation between the distribution and the density function, we can easily derive the multivariate copula density $c(F_1(x_1),\ldots,F_n(x_n))$ associated with a copula function $C(F_1(x_1),\ldots,F_n(x_n))$

$$f(x_1,.....,x_n) = \frac{\partial^n [C(F_1(x_1),....,F_n(x_n))]}{\partial F_1(x_1)....\partial F_n(x_n)} * \prod_{i=1}^{n} f_i(x_i) = c(F_1(x_1),.....,F_n(x_n)) * \prod_{i=1}^{n} f_i(x_i)$$

(c3)

where we define:

$$c(F_1(x_1),.....,F_n(x_n)) = \frac{f(x_1,......,x_n)}{\prod_{i=1}^{n} f_i(x_i)}$$

(c4)

Definition (Normal-copula) Let R be a symmetric, positive definite matrix with diag(R) = 1 and let Φ_R denote the standard multivariate normal distribution with correlation matrix R[4]. Then the Multivariate Gaussian Copula is defined as:

$$C(u_1, u_2,....,u_n; R) = \Phi_R\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2),....., \Phi^{-1}(u_n)\right)$$

(c5)

4 Given a random vector $X = (X_1,....,X_n)'$ we define the standardized normal joint density function f(x) with corrwlation matrix R, as follows:

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} x' R^{-1} x\right)$$

where $\Phi^{-1}(u)$ denotes the inverse of the normal cumulative distribution function. The associated multinormal copula density is obtained by applying equation (c4):

$$c(\Phi(x_1),....,\Phi(x_n)) = \frac{f^{gaussian}(x_1,....,x_n)}{\prod_{i=1}^{n} f_i^{gaussian}(x_i)} = \frac{\frac{1}{(2\pi)^{\frac{n}{2}}|R|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}x'R^{-1}x\right)}{\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x_i^2\right)}$$

and hence, fixing $u_i = \Phi(x_i)$, and denoting with $\zeta = (\Phi^{-1}(u_1),\ldots, \Phi^{-1}(u_n))'$ the vector of the gaussian univariate inverse distribution functions, we have

$$c(u_1,u_2,....,u_n;R) = \frac{1}{|R|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}\varsigma'(R^{-1}-I)\varsigma\right]$$

(c6)

Definition (Student *t*-copula) Let R be a symmetric, positive definite matrix with diag(R) = 1 and let $T_{R,v}$ denote the standard multivariate Student's t distribution with correlation matrix R and v degrees of freedom [5]. Then the multivariate Student's t copula is defined as follows:

$$C(u_{1,}u_2,....,u_n;R,\nu) = T_{R.\nu}\left(t_\nu^{-1}(u_1),t_\nu^{-1}(u_2),....,t_\nu^{-1}(u_n)\right)$$

(c7)

where $t_v^{-1}(u)$ denotes the inverse of the Student's t cumulative distribution function. The associated Student's t copula density is obtained by applying equation (4)

[5] Given a random vector X=(X1,....,Xn)' with a joint standardized multinormal distribution with correlation matrix R and a χ_v^2 – distributed random variable S, independent from X, we define the standardized multivariate Student's t joint density function with correlation matrix R and v degrees of freedom, as the joint distribution function of the random vector EMBED Equation.3 $Y = (\frac{X_1}{S/\sqrt{\nu}},.....,\frac{X_n}{S/\sqrt{\nu}})$:

$$f(y) = \frac{\Gamma\left[\frac{1}{2}(\nu+n)\right]}{\Gamma\left(\frac{1}{2}\nu\right)} * \frac{1}{(\pi\nu)^{\frac{n}{2}}|R|^{\frac{1}{2}}} * \left(1+\frac{y'R^{-1}y}{\nu}\right)^{-\frac{1}{2}(\nu+n)}$$

$$c(u_1,u_2,....,u_n,R,\nu)=\frac{f^{Student}(x_1,....,x_n)}{\prod_{i=1}^{n} f_i^{Student}(x_i)}=|R|^{-\frac{1}{2}}*\frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left[\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\right]^n*\frac{\left(1+\frac{\varsigma' R^{-1}\varsigma}{\nu}\right)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^{n}\left(1+\frac{\varsigma_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \tag{c8}$$

where $\zeta = (t_v^{-1}(u_1),\ldots, t_v^{-1}(u_n))'$.

IV. Measures of risk

The market risk represents the uncertainty of observing an event in the future that could lead to an important portfolio loss. In this context, a measure of risk is a function that takes as an argument the distribution that characterizes the risk factor and returns a scalar value that describes the potential risk implied. The key aspect of measuring risk resides in the ability to correctly identify the distributions of the risk factors.

Even thought the correspondence between the distribution of the risk factor and a scalar could be expressed in different ways, only a part of all these potential functions are appropriate indicators of risk. Artzner et al. (1997, 1998) proposed the theory of coherent risk that captures the desired properties of a risk measure.

If x is a set of real-valued random variables (e.g. the loss distribution of an equity) and the function ω is a real-valued risk measure, then ω should respect the following properties in order to be considered coherent:

Positive homogeneity.

This property basically states that if we increase the quantity of a certain equity in our portfolio we should also have a linear increase in the risk involved and not a diversification effect.

$\omega(\lambda x) = \lambda\omega(x)$.

Subadditivity.

This property represents the advantages of portfolio diversification. The risk of a portfolio (x + y) should be lower than (or equal to) the sum of the risk of the individual securities (x,y)

$$\boldsymbol{\omega(x + y) \leq \omega(x) + \omega(y).}$$

Monotonicity.

This property implies that a higher risk should be considered for a higher return.

$$\boldsymbol{x \leq y \rightarrow \omega(x) \leq \omega(y).}$$

Translational invariance.

This property states that the inclusion of n units of a risk-free asset with returns r in the portfolio should lower the risk of the portfolio.

$$\boldsymbol{\omega(x + nr) = \omega(x) - n.}$$

Value-at-Risk

Value at Risk measures the worst loss to be expected of a portfolio over a given time horizon at a given confidence level. If we mark losses with a positive sign and gains with a negative sign we can estimate Value at Risk by taking the relevant quantile q_α of the conditional distribution.

$$VaR_\alpha = q_\alpha$$

Although VaR offers a simple and intuitive way of evaluating risk Artzner et al. (1997, 1998) have criticized it as a measure of risk for two main reasons. Firstly they showed that VaR is not necessarily subadditive and as a consequence it is not a coherent measure of risk and secondly, this measure gives only an upper limit on the losses given a confidence level, but it tells nothing about the potential size of the loss if this upper limit is exceeded.

Expected Shortfall

The Expected Shortfall (ES) of an asset or a portfolio is the average loss given that VaR has been exceeded.

$$ES_t(\alpha) = E[r_t \mid r_t > VaR_t(\alpha)]$$

where r_t is the return at time t

Although ES is a coherent measure of risk, its accuracy also depends on the ability to identify the true loss distribution of the portfolio.

D. Application

I. Data

The risk evaluation techniques described earlier in the paper are applied to a portfolio of five Romanian equities traded on the Bucharest Stock Exchange (symbols: SIF1, SIF2, SIF3, SIF4, SIF5). These particular stocks were selected due to their high market liquidity, a long time series with very few missing values and high volatility periods that can help evaluate the accuracy of the model with backtesting procedures. The companies are part of the financial sector and their primary activity is the investment in Romanian firms.

The price series covers the period 04/01/2001 – 05/06/2007, has a total of 1564 observations and is adjusted for corporate events. The missing data was replaced by the previous value of the series (or the next value if data at the beginning of the series is missing).

The original series of prices was transformed into return series with the help of the logarithmic formula:

$$\mathbf{rt = \log(p_t/p_{t-1})}$$

II. Estimation and results

1. GARCH models

Before the estimation of the GARCH model, an analysis of the data is made in order to verify if the returns are autocorrelated and the volatility clustering effect is present in the series[6]. The analysis can be made both visually by studying the plot of the autocorrelation functions (*Appendix I, Figure 1 and Figure 2*) and statistically by using a Ljung-Box test for randomness (*Appendix I, Table 1*).

From the visual analysis we can conclude that the series SIF1, SIF2 and SIF4 have a strong first order autocorrelation and the series SIF3 and SIF5 display a weak at most of autocorrelation, or even a non-existent degree of autocorelation. At the same time the visual analysis of the squared returns suggests that a strong autocorrelation is present in all five series.

The results of the Ljung-Box test confirm the outcome of the visual analysis. For the return series, the null hypothesis of the test (where data is random) is rejected at a 5% significance level for all series except SIF3. Similar results are obtained for the squared returns series where the null hypothesis of the test is rejected at a 5% significance level for all equities.

These results confirm the assumptions that the return series are autocorelated and have a time-varying volatility, thus a GARCH model should be appropriate for explaining the data generating process of each series.

In order to find the best GARCH model for each series, a GJR-GARCH model is estimated at first and the coefficients that are not statistically significant are removed, then the model is estimate again in a simpler form. In addition to the verification of significance for each coefficient, the Akaike criterion is also used for model selection. The coefficients are estimated with the maximum – likelihood method and an assumption of t-distributed innovations.

The initial model has the following form:

$$r_t = c + \phi r_{t-1} + \varepsilon_t$$
$$\sigma_t^2 = k + G\sigma_{t-1}^2 + A\varepsilon_{t-1}^2 + LSgn_{t-1}\varepsilon_{t-1}^2$$

[6] The volatility clustering effect can be detected by calculating the degree of autocorrelation for the squared returns.

The results of the initial estimation (*Appendix I, Table 2)* show that the coefficient of leverage is not significant or has a wrong sign for all the five series. This implies that the volatility of the series is influenced in an equal way by a negative or positive return of the same magnitude and the assumption that a negative return has a higher impact doesn't seem to be supported by the data.

Because the results do not support the assumption of a leverage effect, this parameter is removed and the model becomes a GARCH(1,1). In addition, the coefficient for the AR(1) parameter is not significant for all the series, thus this parameter is also removed where it is found to be irrelevant.

The results of the final estimation (*Appendix I, Table 3)* show that the coefficients are all significant at a 5% with the exception of the coefficient of the AR(1) parameter of the SIF2 series which is significant only at a 10% level.

In order to evaluate the outcomes of the GARCH modelling, the residuals are first standardized and then the Ljung-Box test is applied to the standardized residuals.

The assumption of the GARCH model is that $\varepsilon_t = z_t\sigma_t$, where z_t (standardized residuals) is independent and follows a Student's t - distribution. The z_t series is obtained by dividing the residuals(ε_t) at the conditional standard deviation (σ_t).

The Ljung-Box Test applied to the standardized residual (*Appendix I, Table 4)* shows that the null hypothesis cannot be rejected at a 5% significance level for SIF1, SIF3, SIF4 and at 1% significance level for SIF2 and SIF5. In the case of the squared standardised residuals the results are even more relevant, the null hypothesis being accepted at a 5% significance level for all five series.

These results prove that the GARCH models accurately describe the time series and that the standardized residuals fulfil the independence criteria that is necessary in order to use the extreme value theory.

2. Extreme Value Theory (EVT) models

Even thought the estimation of the GARCH model with the assumption of t-distributed innovations (standardized residuals) can explain a large degree of the excess kurtosis found in the return series, it still cannot capture its asymmetry because the distribution is presumed to be symmetric. At the same time, the correct parameterization of the innovation series is very important because it is later used for the simulation of the portfolio loss distribution.

In order to better describe the innovation series resulted from the GARCH models, extreme value theory is used to estimate each tail of the distribution and a kernel smoothing method is used for the interior.

The tail parameters can be estimated using one of the methods described by the extreme value theory. However, for this study, the peak-over-threshold method is used mainly because it needs a smaller data sample compared to the block-maxima method.

In order to identify the tails of the innovation series we sort the values of the series in ascending order and we consider the first 10% of the values to be the lower tail and the last 10% of the values to be the upper tail. By using this method the threshold value is implicitly determined for the lower and upper tail. The 10% fraction of the distribution is selected by taking into consideration the simulations performed by McNeil and Frey (2000) and Nyström and Skoglund (2001) that showed that the ML estimator is almost invariant to the threshold value if the tail is considered between 5-13% of the sample data.

The parameters of the GPD distribution that are estimated using the maximum-likelihood method are displayed in (*Appendix II, Table 1.*). By studying the results it can be observed that the coefficient that gives the heaviness of the tail (tail index) is statistically different for the upper compared with the lower tail, thus confirming the assumption of an asymmetric distribution of innovations. Furthermore the coefficients of the tail index for the lower tails are all statistically different from zero, thus giving an indication that the lower tails are ”heavier” then the tails of a normal distribution.

The estimated value of the tail index for the upper tail is not statistically different from zero for neither of the series, thus we can draw the conclusion that the shape of these tails resembles the tail of a normal distribution.

By comparing the shape of the tails estimated using the extreme value method with the shape of a Student’s t-distribution that has the same degrees of freedom as the parameter estimated in the GARCH model (*Appendix II, Figure 1 a,b*), it can be observed that the extreme value method describes much better the distributions of the innovations.

In order to have a complete semi-parametrical distribution for each series of innovations, a pseudo CDF (cumulative distribution function) and ICDF (inverse cumulative distribution function) are built. The pseudo CDF function receives a value, it identifies where the value is situated in the estimated semi-parametrical distribution (in one of the tails or in the centre) by comparing it with the thresholds and based on this information it computes a corresponding cumulated probability. The pseudo ICDF function is built on the same principle, the cumulative probability that is given as an input is mapped to one area of the semi-parametrical distribution and a corresponding quantile is computed and returned based on that information. With the help of these pseudo functions a representation of each semi-parametric distribution can be build (*Appendix II, Figure 1, c)*.

3. Copula models

Once the model for each time series is defined, the dependence structure of the portfolio can be estimated by linking together the semi-parametrical distributions with copula methods. The Student's t copula is selected for this task because in addition to a correlation matrix it is also characterized by the degrees of freedom parameter, which defines the amount of tail dependence between the series.

The copula is calibrated by using the Canonical Maximum Likelihood (CML) method because this method allows an estimation of the copula parameters without making any assumption about the marginal distributions. The CML method[7] can be implemented in two stages.

First we transform the initial data set $X = (X_{1t,\dots,} X_{nt})_{t=1}^{T}$ into uniform variates using the marginal distribution function, that is, for t=1,….T , let $u_t = (u^t_{1,\dots\dots,} u^t_n) = [F_1(X_{1t}),\dots, F_n(X_{nt})]$. In this case, for equity i, X_i represents the innovation series and F_i represents the pseudo-cumulative distribution.

Secondly we estimate the vector of copula parameters α, via the following relation:

$$\hat{\alpha}_{\mathrm{CML}} = \arg\max_{\alpha} \sum_{t=1}^{T} \ln c(\mathrm{u}_1^{\mathrm{t}}, \dots\dots, \mathrm{u}_{\mathrm{n}}^{\mathrm{t}}; \alpha)$$

where c is the copula density function, in this case the density of the Student's t copula.

The actual estimation of the copula parameters is done in two steps; the first step maximizes the log-likelihood function with respect to the linear correlation matrix, given a fixed value for the degrees of freedom. The second step uses the results from the first optimization in order to maximize the function with respect to the degrees of freedom, thus maximizing the log-likelihood over all parameters. The function that is maximized in the second step is called the profile log-likelihood function for the degrees of freedom.

The estimated correlation matrix (*Appendix III, Table1*) shows a positive correlation between all five series, while the relative small value of the degrees of freedom parameter (*Appendix III, Table2*) confirms the presence of a strong tail dependence. The standard error

[7] See Mashal and Zeevi, p 25. for more details.

of the degrees of freedom parameter (0.328318) was obtained using a simple bootstrap method.

4. Portfolio simulation

Once the models for the marginal distributions and the dependence structure are estimated, we can simulate the conditional loss distribution of the portfolio for the next period and compute the risk measures of interest.

The first stage of the simulation process represents the generation of dependent series by using the dependence structure given by the t-copula, that is, for each series, for a horizon of h days, n trials are generated from a multivariate Student's t distribution that has the same correlation matrix and degrees of freedom parameters as those estimated with the t-copula.

The result of this step is a collection of (no of equities x horizon) distributions that have the same dependence structure as the original data. However these distributions were generated using a multivariate Student's t distribution, so they must be transformed in order to follow the semi-parametrical distributions used by the GARCH model.

The transformation of each distribution is done in two steps, first the distribution is shaped into a uniform variate, by using the cumulative distribution function of the Student's t distribution, and secondly these uniform variates are converted into the semi-parametrical distributions by using the pseudo-inverse cumulative distribution function of the corresponding semi-parametrical distribution. A visual example of a simulation for two correlated series can be found in (*Appendix IV, Figure 1*).

At the second stage of the simulation process, the semi-parametric distributions are given as an input to the GARCH model that reintroduces the volatility into the series and gives as an output conditional series of returns.

At this stage we have a conditional distribution of returns for each equity in the portfolio and for each day of the horizon (h). These conditional distributions can be cumulated in order to build the loss distribution of the entire portfolio for a horizon of up to h days.

5. Measures of risk

If we mark losses with a positive sign and gains with a negative sign we can estimate Value at Risk by taking the relevant quantile q_α of the conditional distribution.

$$VaR_\alpha = q_\alpha$$

Also, Expected Shortfall is estimated by using the following formula:

$$ES = E(X \mid X > VaR) = (\sum_{i=[n\alpha]}^{n} X_{n(i)})/(n-[na])$$

where n represents the number of trials.

Both measures of risk are applied to the individual conditional distributions and the conditional portfolio loss distribution (computed with an equal weight for each asset) and the results are displayed in (*Appendix IV, Table 1-4*). The estimated risk is comparable between the equities, SIF3 displaying the highest level of risk while SIF5 the lowest. Furthermore, the high correlation between the assets reduces the benefit of diversification to a minimum, thus making the risk of the portfolio comparable to any of the individual equities.

6. Backtesting

In order to evaluate the accuracy of the methodology used for the estimation of risk, a backtest is applied for each individual return series and also for the portfolio. The test implies the estimation of Value-at-Risk for a number of days for which we already know the actual returns. By comparing the estimated Value-at-Risk with the actual returns we can observe if the confidence levels of the risk measure are indeed respected.

The tests are performed with a 1 day horizon, for the last 500 days of the series and with a fixed data sample of 1000 observations. For each day the methodology is applied from the beginning and all the parameters are reestimated.

Plots of these tests can be seen in (*Appendix IV, Figure 2*) and the number of violations for each series is displayed in (*Appendix IV, Table 5*). The backtesting results are not very clear, firstly because the evaluation at 90% and 95% confidence levels gave mixed results and secondly because the accuracy of the risk measure at 99% confidence level cannot be test properly due to the small number of back-testing days (500).

For the individual series, the results at a 90% confidence level show that the risk is being slightly underestimated for two of the series (SIF1, SIF2) and slightly overestimated for the other three (SIF3, SIF4, SIF5). The situation is different at a 95% confidence level, where we can see a higher underestimation of the risk for the majority of the series. At a 99% confidence level the risk appears to be overestimated for the majority of the series, perhaps due to the lack of sufficient observations leading up to unclear conclusions.

At the portfolio level, the results show an underestimation of the risk at all three confidence levels, although the degree of underestimation is rather small for the 90% and 99% levels and more significant for the 95% level.

E. Conclusions.

This paper aims at computing accurate estimates for the risk of a portfolio by constructing its conditional loss distribution with a flexible methodology that separates the description of the marginal distributions from the dependence structure. The return series for each of the equities was modelled using GARCH methods in order to explain the autocorrelation and time-varying volatility. Then, the innovation series resulted from the GARCH model is described as a semi-parametrical distribution with GPD tails and a kernel-smoothed interior that captures the asymmetry and excess-kurtosis often found in these series. The link between the semi-parametrical distributions is then explained by a Student's t copula that gives the dependence structure of the entire portfolio.

The estimated parameters of the marginal distributions and the dependence structure serve as a base for the simulation of a conditional portfolio distribution and implicitly for the estimation of the risk measures.

By analyzing the intermediary results of this methodology the following conclusions can be drawn:

- the GARCH models explain very well the autocorrelation found in the return series and the volatility clustering effect
- the distributions of the innovations are asymmetric with heavy lower tails and thin upper tails
- the GPD description for the tails of the innovation series is more accurate compared to the description given by the t-distribution estimated by the GARCH models.
- the backtesting results for Value-at-Risk are not conclusive but give an indication of a possible underestimation of the risk at 95% confidence level

Further research can be done in two main directions; first this methodology could be applied for portfolios with different risk factors in order to evaluate its accuracy for a larger collection of assets. Secondly the methodology can be improved by using better measures of risk or more flexible tools for describing the data generating process for the returns of the portfolio. This implies for example, the estimation of the GARCH models without any assumption about the distribution of the innovations, the use of different copulas that might describe better the dependence structure of the assets or the use of spectral measures of risk that take into account the risk aversion of the risk manager.

F. References

1. Alexander, C. (2001), „Market Models: A guide to Financial Data Analysis”

2. Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath (1998), „Coherent Measures Of Risk”, Universite Louis Pasteur, EidgenÄossische Technische Hochschule, Societe Generale, Carnegie Mellon University, Pittsburgh

3. Bao, Y., T.H. Lee, and B. Saltoglu (2004), „Evaluating Predictive Performance of Value-at-Risk Models in Emerging Markets: A Reality Check”, UT San Antonio, UC Riverside, Marmara University

4. Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli (2000), „Copulas for Finance: A Reading Guide and Some Applications”, Financial Econometrics Research Centre City University Business School London

5. Clemente, A. and C. Romano (2004a), „Measuring and optimizing portofolio credit risk: A Copula-Based Approach**”,** Working Paper n.1 - Centro Interdipartimentale sul Diritto e l'Economia dei Mercati

(2003b) „Measuring portofolio value-at- risk by a Copula-Evt based approach”

6. Cotter, J. and K. Dowd (2005), "Extreme Spectral Risk Measures: An Application to Futures Clearinghouse Margin”, University College Dublin, Nottingham University Business School

7. Danielsson, J. And C. G. de Vries (1997), „Value-at-Risk and Extreme Returns”, London School of Economics and Institute of Economic Studies at University of Iceland, Tinbergen Institute and Erasmus University

8. Demarta, S. and A. J. McNeil (2004), „The t Copula and Related Copulas”, Department of Mathematics Federal Institute of Technology ETH Zentrum

9. Dias, A. and P. Embrechts, „Dynamic copula models for multivariate high-frequency data in Finance”, Warwick Business School, - Finance Group, Department of Mathematics, ETH Zurich

10. Diebold, F. X. , T. Schuermann, and J. D. Stroughair (1998), „Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management”, The Wharton Financial Institutions Center

11. García, A. and R. Gençay (2006), „Risk-Cost Frontier and Collateral Valuation in Securities Settlement Systems for Extreme Market Events”, Bank of Canada Working Paper 2006-17

12. Hotta, L.K. , E.C. Lucas, and H.P. Palaro, „Estimation of VaR Using Copula and Extreme Value Theory”, State University of Campinas, Department of Statistics, Campinas SP, Brazil. ESAMC, Campinas SP

13. Joe, H. and J. J. Xu , „The Estimation Method of Inference Functions for Margins for Multivariate Models", Department of Statistics, University of British Columbia

14. Longin, F. M. (2000), „From value at risk to stress testing: The extreme value approach", Journal of Banking & Finance 24, 1097-1130

15. Mashal, R. and A. Zeevi (2002), „Beyond Correlation: Extreme Co-movements Between Financial Assets", Columbia University

16. McNeil, A.J. (1996a), „Estimating the Tails of Loss Severity Distributions using Extreme Value Theory", Departement Mathematik ETH Zentrum

17. (1998b), „Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory", Departement Mathematik ETH Zentrum

18. McNeil, A.J. and R.Frey (2000), „Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach", Departement Mathematik ETH Zentrum

19. McNeil, A.J. and T. Saladin (1997), „The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions", Departement Mathematik ETH Zentrum

20. Nyström, K. and J. Skoglund (2002a), „A Framework for Scenariobased Risk Management", Swedbank, Group Financial Risk Control

21. (2002b), „Univariate Extreme Value Theory, GARCH and Measures of Risk", Swedbank, Group Financial Risk Control

22. Rosenberg, J. V. and T. Schuermann (2005), „A General Approach to Integrated Risk Management with Skewed, Fat-tailed Risks", Federal Reserve Bank of New York

Appendix I – GARCH

Figure 1. Correlograms for the return series

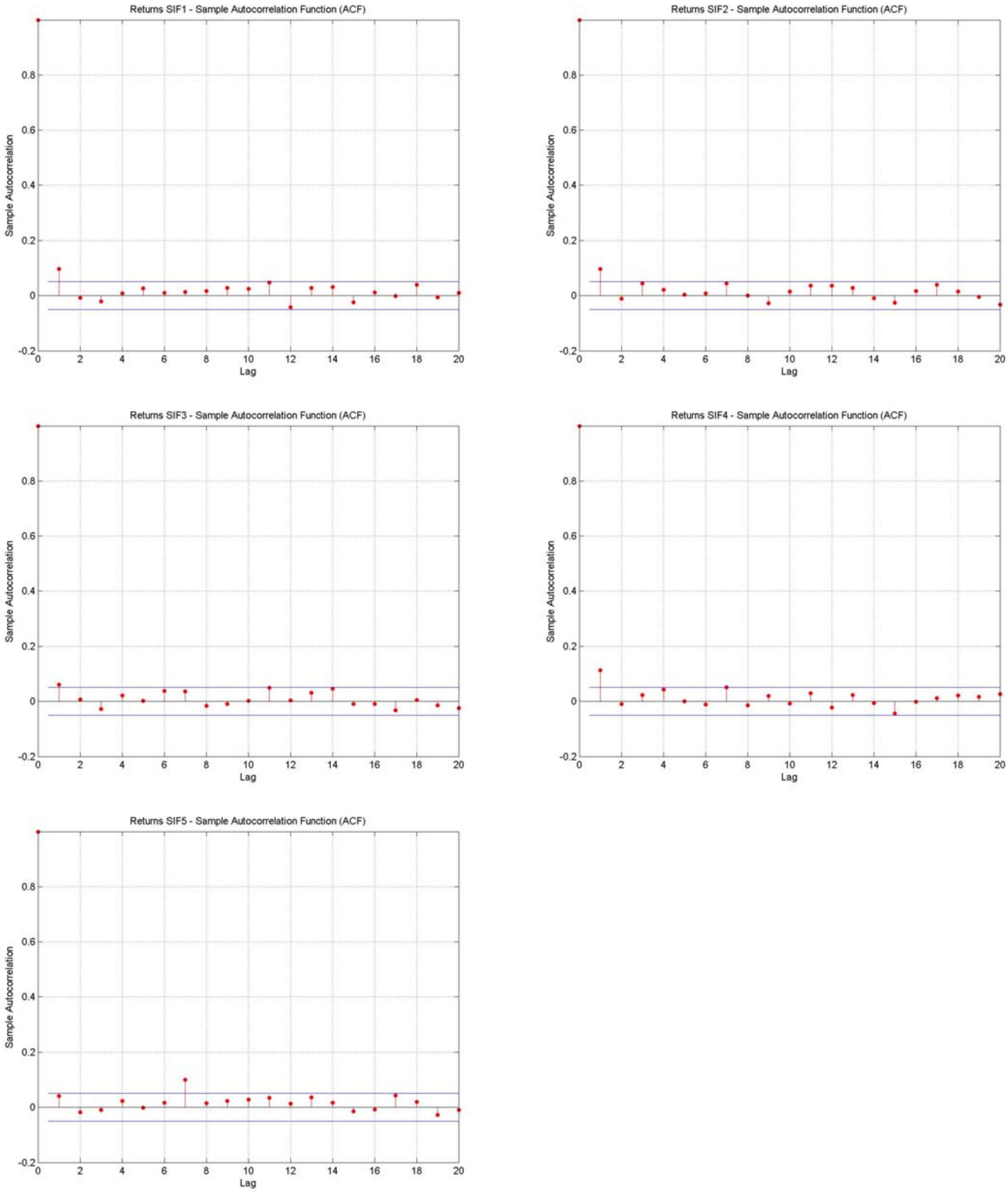


Figure 2. Correlograms for the squared return series

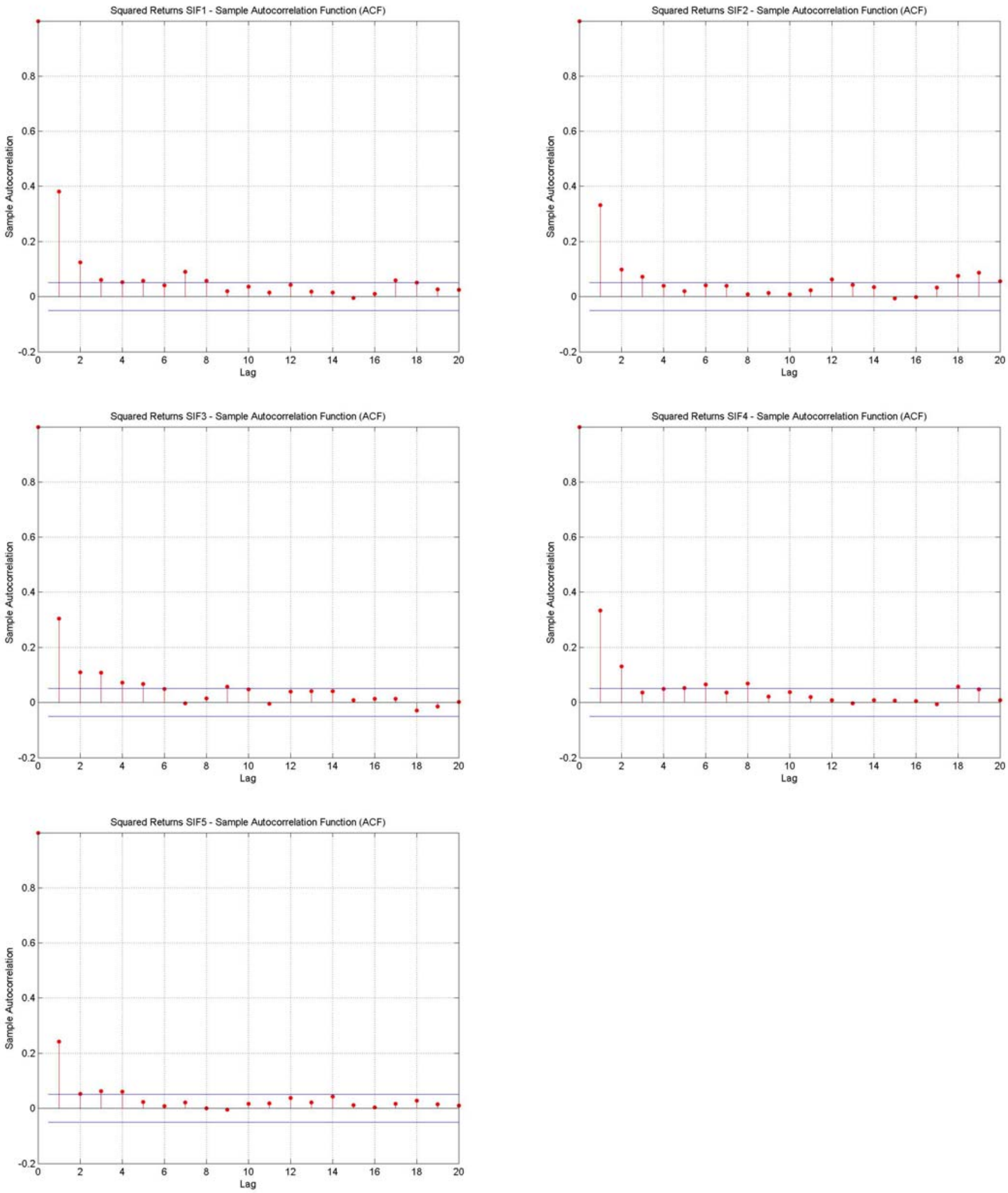


Table 1.Ljung-Box test results for the return series and the squared return series.

LjungBoxTest for returns @ 20 lags, @ 0.05 confidence level

Null Hypotesis: Data is random (no serial correlation)

Null Hypotesis Accepted -> H = 0 ; Null Hypotesis Rejected -> H = 1

Series	H	pValue	Statistic	Critical Value
SIF1	1	0.031399	33.284	31.4104
SIF2	1	0.019863	35.045	31.4104
SIF3	0	0.19554	25.156	31.4104
SIF4	1	0.0087669	38.036	31.4104
SIF5	1	0.031252	33.302	31.4104

LjungBoxTest for squared returns @ 20 lags, @ 0.05 confidence level

Null hypotesis: Data is random (no serial correlation)

Null Hypotesis Accepted -> H = 0 ; Null Hypotesis Rejected -> H = 1

Series	H	pValue	Statistic	Critical Value
SIF1	1	0	300.04	31.4104
SIF2	1	0	248.48	31.4104
SIF3	1	0	221.78	31.4104
SIF4	1	0	245.23	31.4104
SIF5	1	3.7748e-015	113.9	31.4104

Table 2. Results for the first estimation of the GARCH models (GJR-GARCH).

Initial estimation - GARCH: SIF1

Mean: ARMAX(1,0,0); Variance: GJR(1,1)

Conditional Probability Distribution: T

Number of Model Parameters Estimated: 7

Parameter	Value	Standard Error	T Statistic
C	0.0013624	0.0004647	2.9317
AR(1)	0.0061472	0.026416	0.2327
K	4.6024e-005	1.1727e-005	3.9245
GARCH(1)	0.72498	0.033663	21.5365
ARCH(1)	0.30953	0.059267	5.2226
Leverage(1)	-0.077775	0.060984	-1.2753
DoF	3.6836	0.39126	9.4147

Initial estimation - GARCH: SIF2

Mean: ARMAX(1,0,0); Variance: GJR(1,1)

Conditional Probability Distribution: T

Number of Model Parameters Estimated: 7

Parameter	Value	Standard Error	T Statistic
C	0.0019917	0.00053772	3.7040
AR(1)	0.040896	0.026685	1.5326
K	8.9754e-005	2.1894e-005	4.0994
GARCH(1)	0.68387	0.043946	15.5614
ARCH(1)	0.3013	0.06246	4.8239
Leverage(1)	-0.11796	0.0608	-1.9402
DoF	4.0824	0.45051	9.0616

Initial estimation - GARCH: SIF3

Mean: ARMAX(1,0,0); Variance: GJR(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 7

Parameter	Value	Standard Error	T Statistic
C	0.001685	0.00047358	3.5581
AR(1)	-0.023526	0.026172	-0.8989
K	5.3273e-005	1.4384e-005	3.7037
GARCH(1)	0.73855	0.03692	20.0037
ARCH(1)	0.24825	0.053417	4.6474
Leverage(1)	-0.037494	0.055747	-0.6726
DoF	3.6475	0.38714	9.4218

Initial estimation - GARCH: SIF4

Mean: ARMAX(1,0,0); Variance: GJR(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 7

Parameter	Value	Standard Error	T Statistic
C	0.0015988	0.00050452	3.1690
AR(1)	0.060422	0.025919	2.3312
K	6.8256e-005	1.713e-005	3.9845
GARCH(1)	0.71545	0.040135	17.8260
ARCH(1)	0.2714	0.057564	4.7147
Leverage(1)	-0.080635	0.05978	-1.3489
DoF	3.8946	0.41809	9.3153

Initial estimation - GARCH: SIF5

Mean: ARMAX(1,0,0); Variance: GJR(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 7

Parameter	Value	Standard Error	T Statistic
C	0.0019634	0.00051959	3.7788
AR(1)	0.0074924	0.025493	0.2939
K	9.2278e-005	2.3704e-005	3.8930
GARCH(1)	0.71299	0.046853	15.2174
ARCH(1)	0.26942	0.063197	4.2631
Leverage(1)	-0.12281	0.059662	-2.0584
DoF	3.5881	0.34257	10.4741

Table 3. Results for the final estimation of the GARCH models.

Final estimation - GARCH: SIF1

Mean: ARMAX(0,0,0); Variance: GARCH(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 5

Parameter	Value	Standard Error	T Statistic
C	0.0012395	0.00046064	2.6908
K	4.5935e-005	1.172e-005	3.9195
GARCH(1)	0.72085	0.033774	21.3436
ARCH(1)	0.27915	0.048701	5.7319
DoF	3.6882	0.39026	9.4506

Final estimation - GARCH: SIF2

Mean: ARMAX(1,0,0); Variance: GARCH(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 6

Parameter	Value	Standard Error	T Statistic
C	0.0017968	0.00053892	3.3341
AR(1)	0.034957	0.026676	1.3104
K	8.6898e-005	2.1328e-005	4.0744
GARCH(1)	0.68862	0.042992	16.0174
ARCH(1)	0.24198	0.043716	5.5352
DoF	4.1105	0.45594	9.0155

Final estimation - GARCH: SIF3

Mean: ARMAX(0,0,0); Variance: GARCH(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 5

Parameter	Value	Standard Error	T Statistic
C	0.0015991	0.0004655	3.4353
K	5.3986e-005	1.4579e-005	3.7031
GARCH(1)	0.73498	0.037461	19.6198
ARCH(1)	0.23442	0.045878	5.1097
DoF	3.6559	0.38638	9.4617

Final estimation - GARCH: SIF4

Mean: ARMAX(1,0,0); Variance: GARCH(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 6

Parameter	Value	Standard Error	T Statistic
C	0.0014639	0.00050067	2.9238
AR(1)	0.058049	0.026091	2.2248
K	7.0242e-005	1.7569e-005	3.9980
GARCH(1)	0.70692	0.040694	17.3718
ARCH(1)	0.2431	0.047402	5.1285
DoF	3.8641	0.41209	9.3770

Final estimation - GARCH: SIF5

Mean: ARMAX(0,0,0); Variance: GARCH(1,1)

Conditional Probability Distribution: T
Number of Model Parameters Estimated: 5

Parameter	Value	Standard Error	T Statistic
C	0.001778	0.00051569	3.4479
K	8.361e-005	2.1759e-005	3.8425
GARCH(1)	0.72883	0.044148	16.5090
ARCH(1)	0.20305	0.043688	4.6478
DoF	3.5868	0.34209	10.4850

Table 4. Ljung-Box test results for the standardized residuals and squared residuals.

LjungBoxTest for returns @ 20 lags, @ 0.05 confidence level

Null Hypotesis: Data is random (no serial correlation)
Null Hypotesis Accepted -> H = 0 ; Null Hypotesis Rejected -> H = 1

Serie	H	pValue	Statistic	Critical Value
SIF1	0	0.10089	28.371	31.4104
SIF2	1	0.015045	36.082	31.4104
SIF3	0	0.77333	15.054	31.4104
SIF4	0	0.36896	21.487	31.4104
SIF5	1	0.043424	31.988	31.4104

LjungBoxTest for squared returns @ 20 lags, @ 0.05 confidence level

Null hypotesis: Data is random (no serial correlation)
Null Hypotesis Accepted -> H = 0 ; Null Hypotesis Rejected -> H = 1

Serie	H	pValue	Statistic	Critical Value
SIF1	0	0.99529	7.3687	31.4104
SIF2	0	0.84678	13.671	31.4104
SIF3	0	0.8839	12.846	31.4104
SIF4	0	0.95327	10.716	31.4104
SIF5	0	0.99314	7.7947	31.4104

Appendix II – EVT

Table 1. Estimated GPD parameters for the tails of the standardized residuals distributions.

SIF1						
Tail	**Tail Index**	**Std Error**	**T-Stat**	**Sigma**	**Std Error**	**T-Stat**
Lower	**0.2022**	0.088745	2.2785	**0.52983**	0.062743	8.4444
Upper	**0.012265**	0.094734	0.12947	**0.68873**	0.085424	8.0625

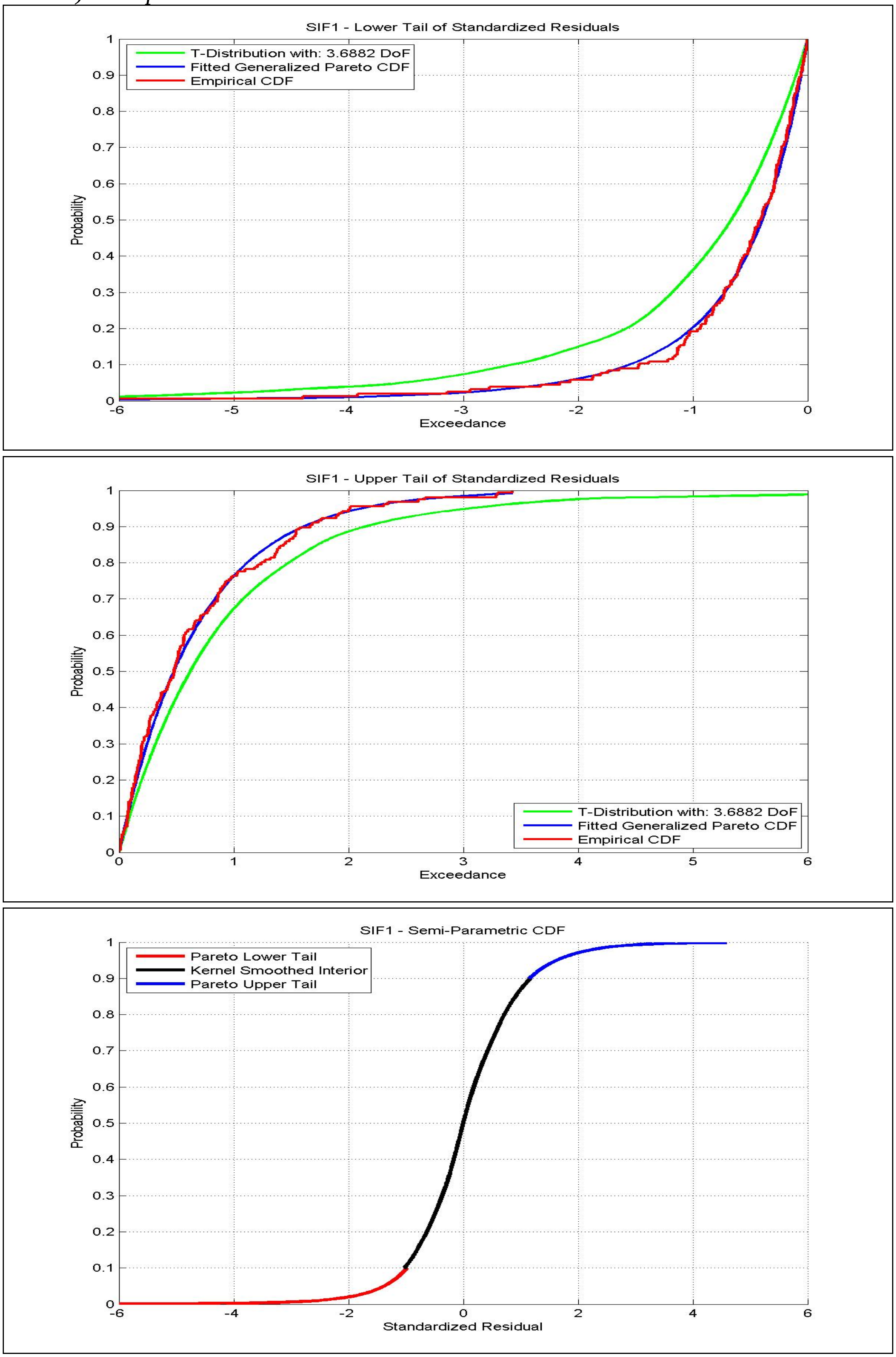
SIF2						
Tail	**Tail Index**	**Std Error**	**T-Stat**	**Sigma**	**Std Error**	**T-Stat**
Lower	**0.21153**	0.089722	2.3576	**0.5177**	0.061633	8.3997
Upper	**-8.11E+00**	05 0.084005	-0.00097	**0.67981**	0.078891	8.6171

SIF3						
Tail	**Tail Index**	**Std Error**	**T-Stat**	**Sigma**	**Std Error**	**T-Stat**
Lower	**0.13105**	0.093611	1.4	**0.52577**	0.064529	8.1478
Upper	**0.11445**	0.079746	1.4352	**0.65982**	0.074316	8.8786

SIF4						
Tail	**Tail Index**	**Std Error**	**T-Stat**	**Sigma**	**Std Error**	**T-Stat**
Lower	**0.30175**	0.10018	3.012	**0.44784**	0.056405	7.9397
Upper	**0.10545**	0.096637	1.0912	**0.59198**	0.074122	7.9865

SIF5						
Tail	**Tail Index**	**Std Error**	**T-Stat**	**Sigma**	**Std Error**	**T-Stat**
Lower	**0.44799**	0.12018	3.7278	**0.39466**	0.055205	7.149
Upper	**0.076981**	0.088776	0.86714	**0.61118**	0.072967	8.3761

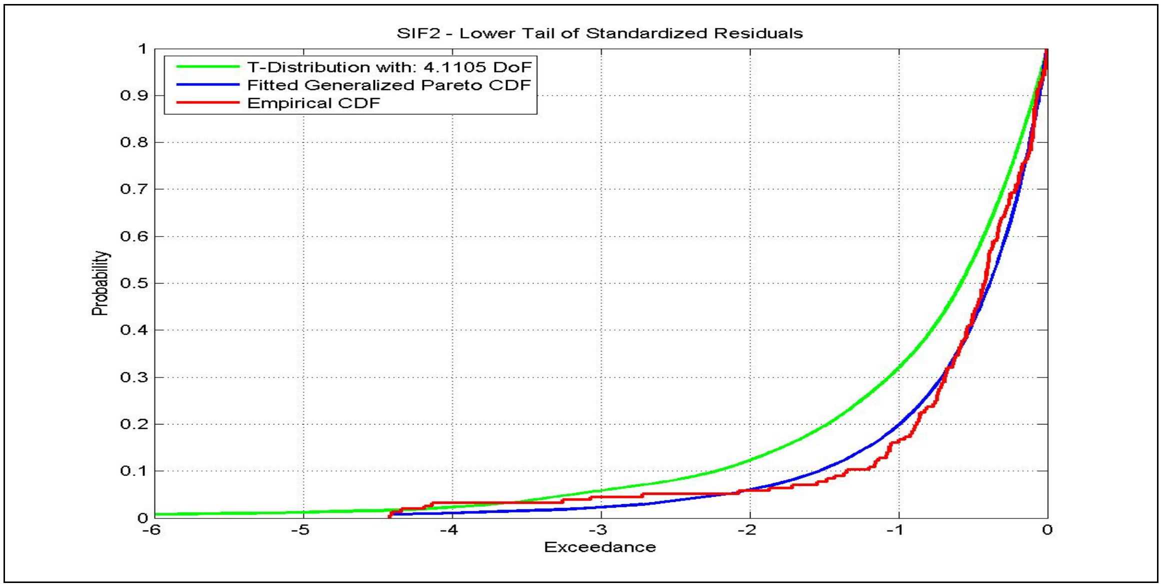
Figure 1. a) CDF for the lower tail of the standardized residuals distribution: GPD vs. Student's t vs empirical
b) CDF for the upper tail of the standardized residuals distribution: GPD vs. Student's t vs empirical
b) Semi-parametric CDF with GPD tails and kernel smoothed interior



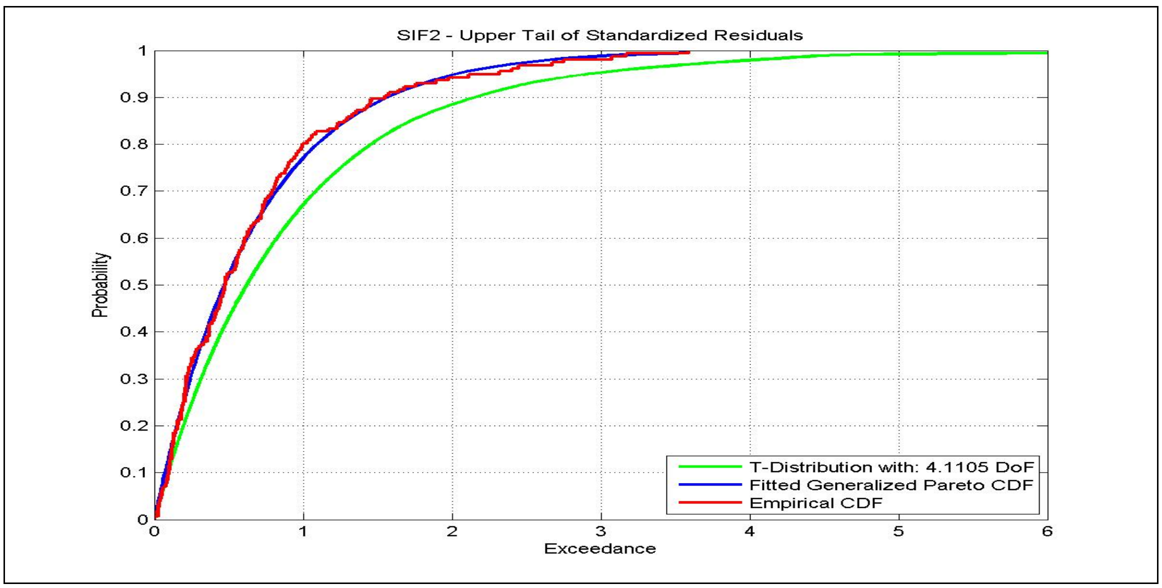
(a)

(b)

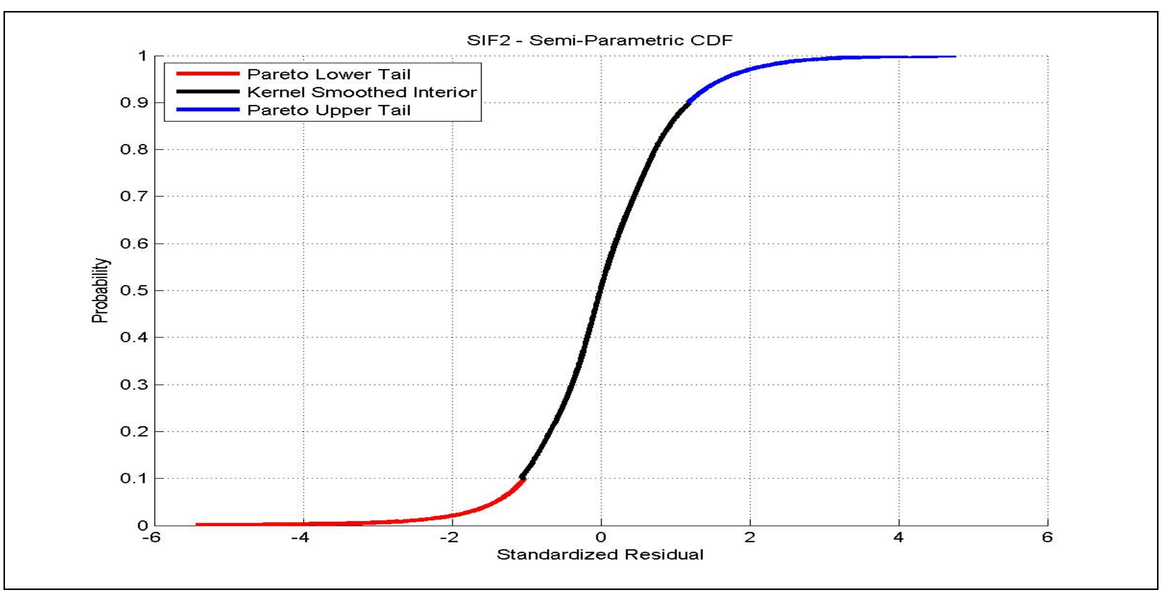
(c)


SIF2 - Lower Tail of Standardized Residuals
T-Distribution with: 4.1105 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

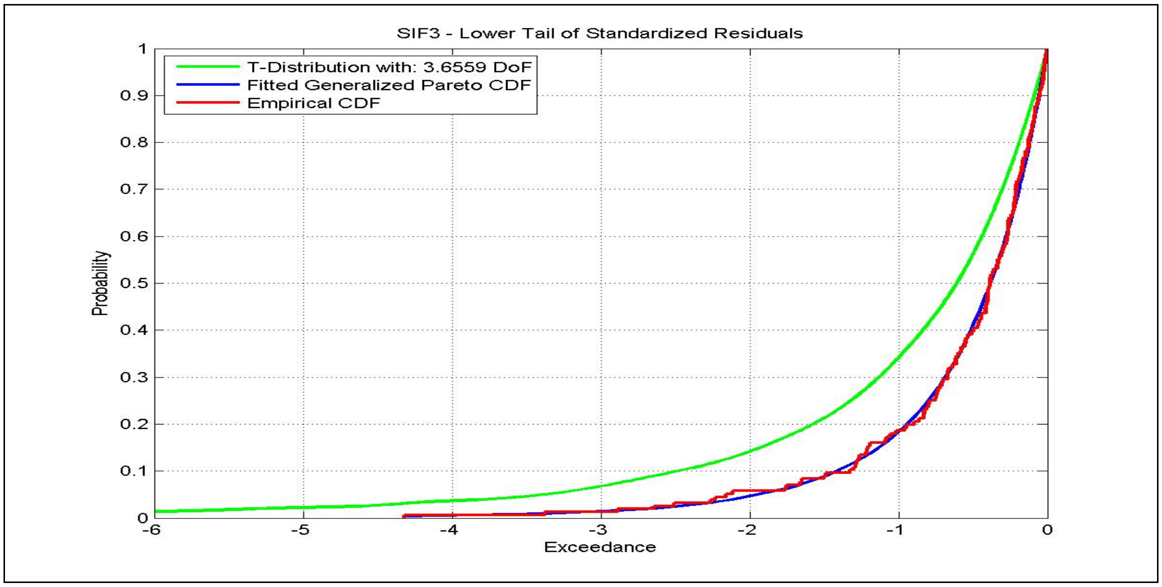
(a)


SIF2 - Upper Tail of Standardized Residuals
T-Distribution with: 4.1105 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

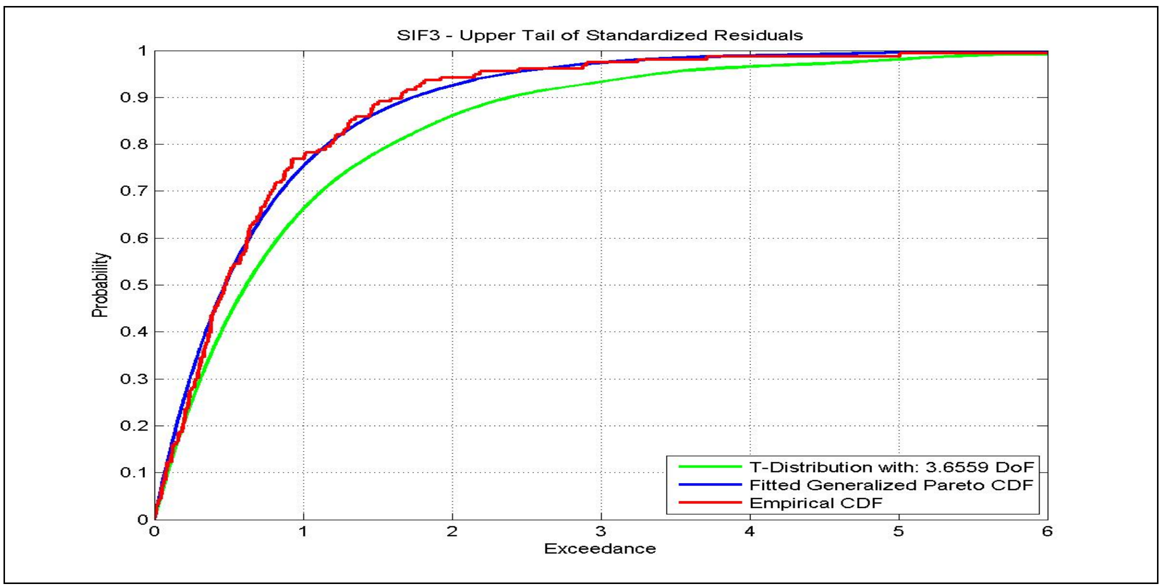
(b)


SIF2 - Semi-Parametric CDF
Pareto Lower Tail
Kernel Smoothed Interior
Pareto Upper Tail
Probability
Standardized Residual

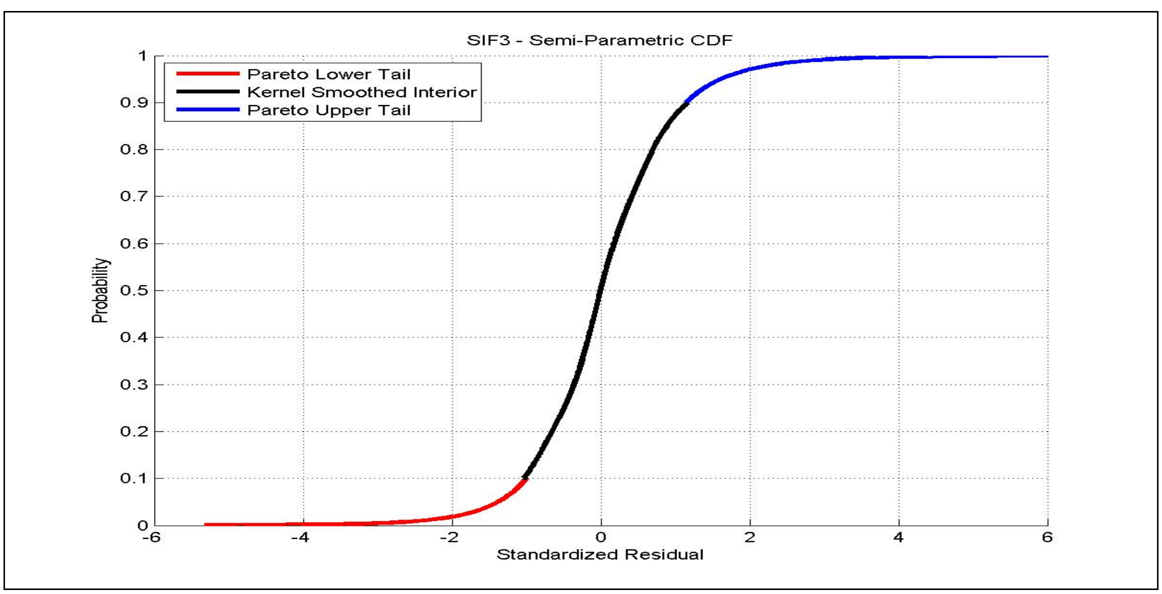
(c)


SIF3 - Lower Tail of Standardized Residuals
T-Distribution with: 3.6559 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

(a)

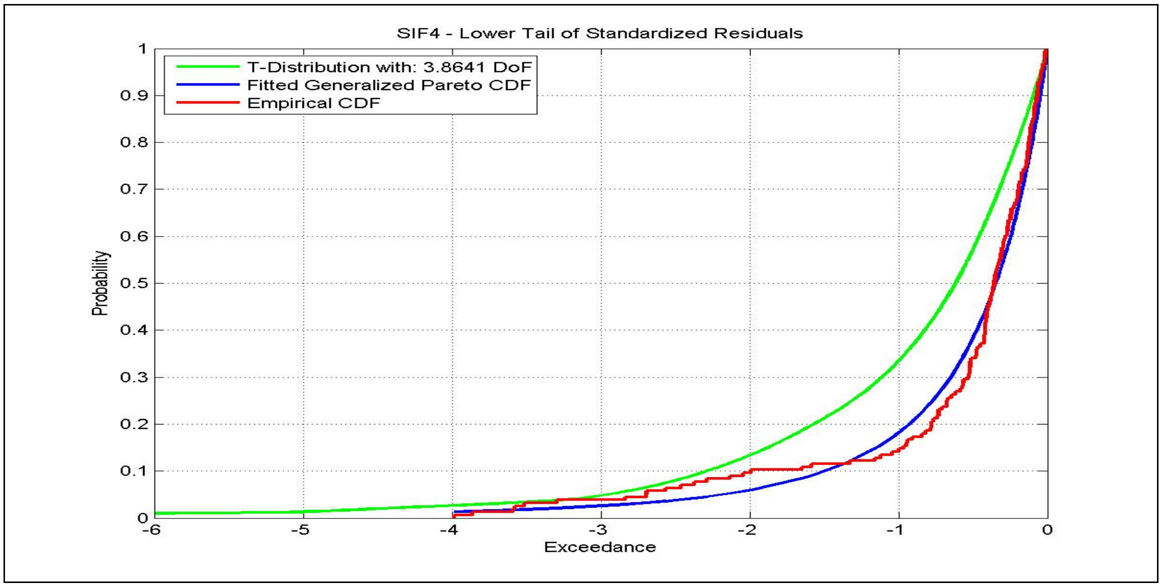

SIF3 - Upper Tail of Standardized Residuals
T-Distribution with: 3.6559 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

(b)

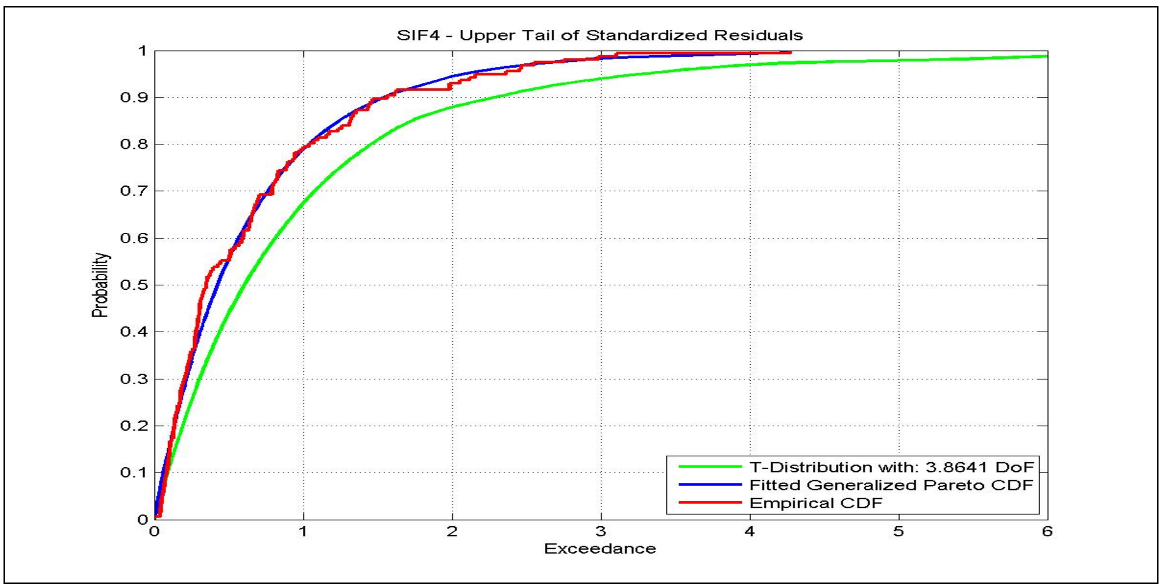

SIF3 - Semi-Parametric CDF
Pareto Lower Tail
Kernel Smoothed Interior
Pareto Upper Tail
Probability
Standardized Residual

(c)

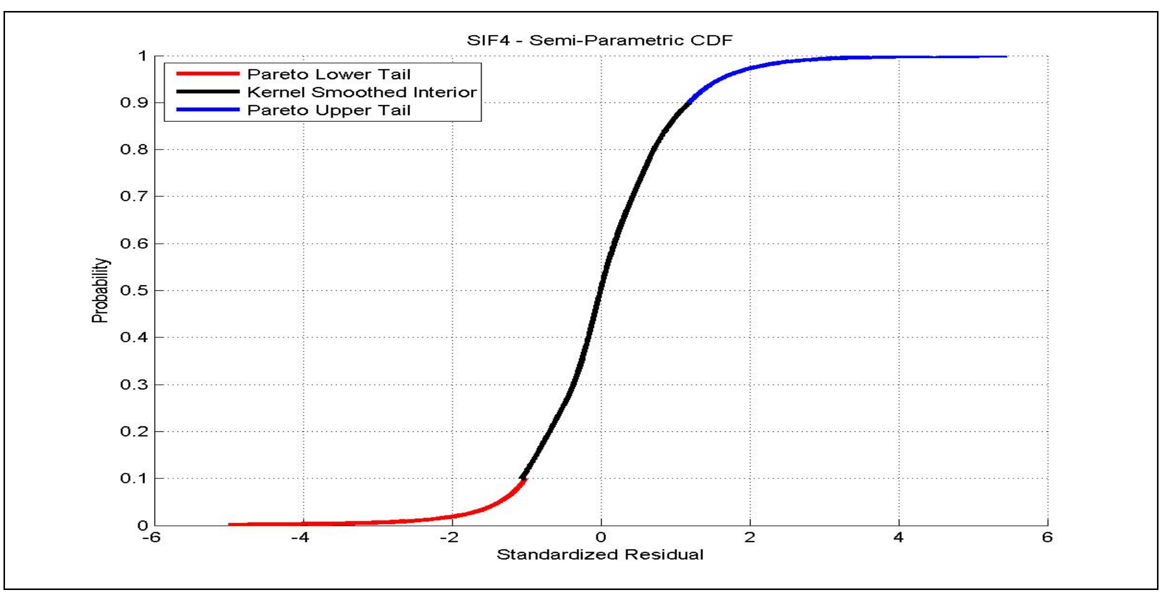
(a)

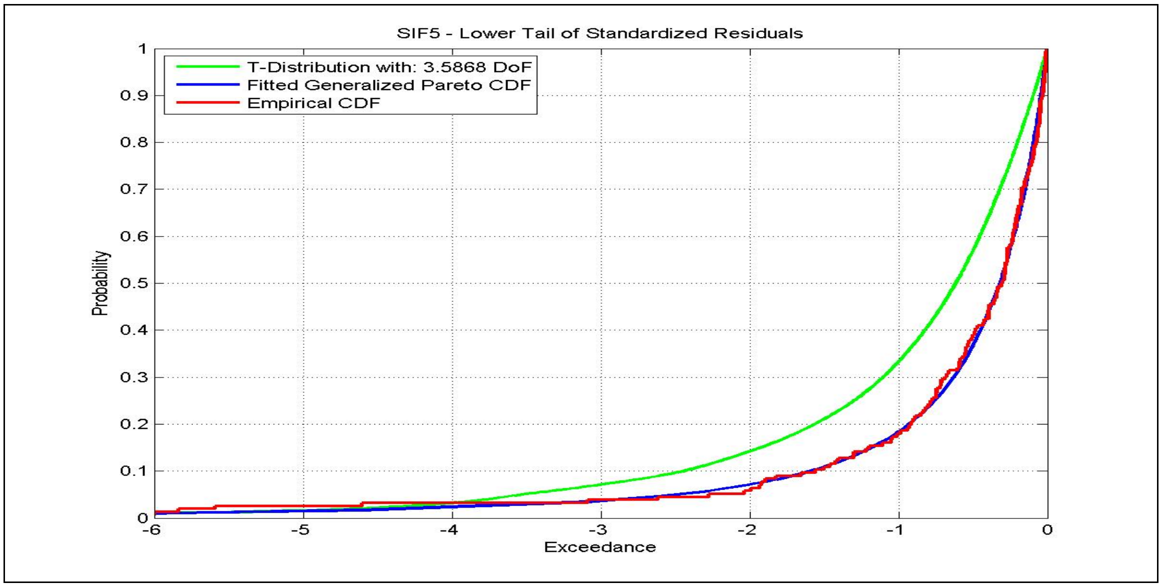

SIF4 - Lower Tail of Standardized Residuals
T-Distribution with: 3.8641 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

(b)


SIF4 - Upper Tail of Standardized Residuals
T-Distribution with: 3.8641 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

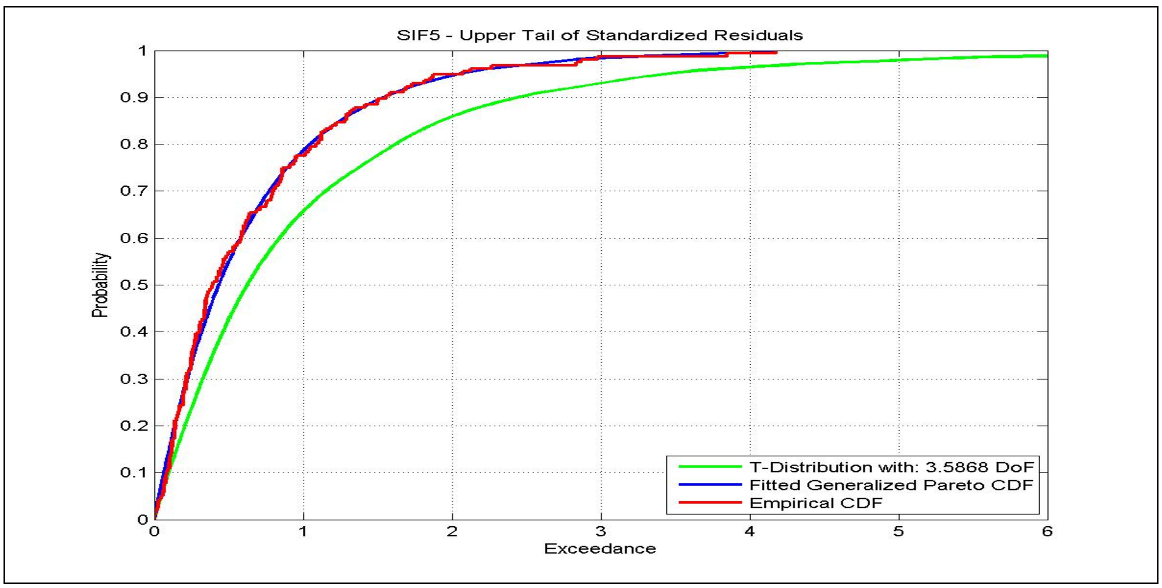
(c)


SIF4 - Semi-Parametric CDF
Pareto Lower Tail
Kernel Smoothed Interior
Pareto Upper Tail
Probability
Standardized Residual



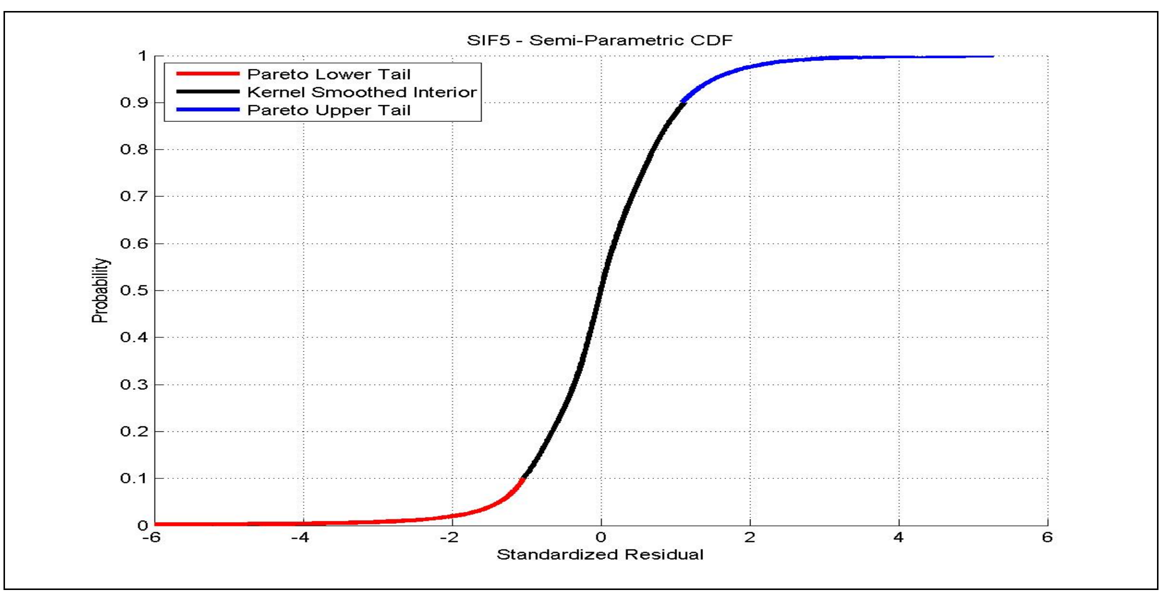
SIF5 - Lower Tail of Standardized Residuals
T-Distribution with: 3.5868 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

(a)



SIF5 - Upper Tail of Standardized Residuals
T-Distribution with: 3.5868 DoF
Fitted Generalized Pareto CDF
Empirical CDF
Probability
Exceedance

(b)



SIF5 - Semi-Parametric CDF
Pareto Lower Tail
Kernel Smoothed Interior
Pareto Upper Tail
Probability
Standardized Residual

(c)

Appendix III – Copula

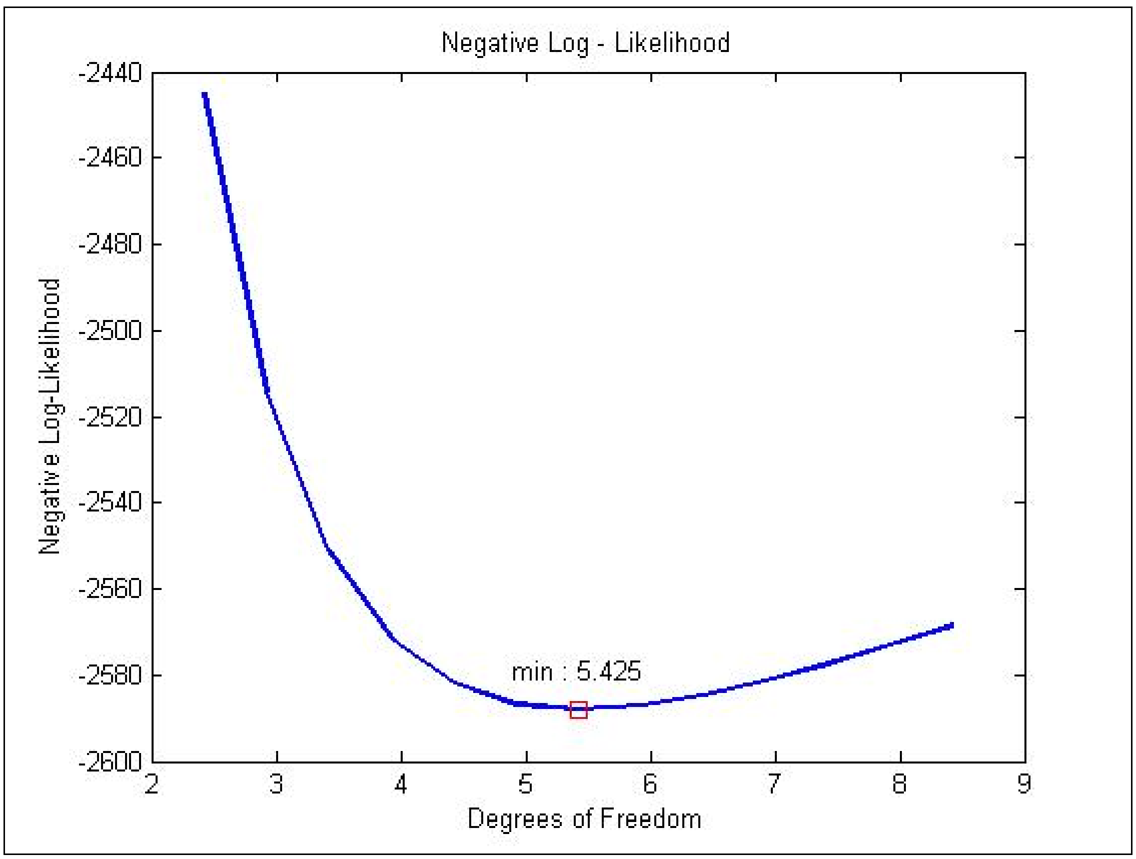
Table 1. Estimated correlation matrix

Correlation Matrix					
	SIF1	**SIF2**	**SIF3**	**SIF4**	**SIF5**
SIF1	1	0.7118	0.6822	0.6673	0.6994
SIF2	0.7118	1	0.6615	0.6693	0.7701
SIF3	0.6822	0.6615	1	0.6469	0.6408
SIF4	0.6673	0.6693	0.6469	1	0.6798
SIF5	0.6994	0.7701	0.6408	0.6798	1

Table 2. Estimated degrees of freedom and the equivalent standard error.

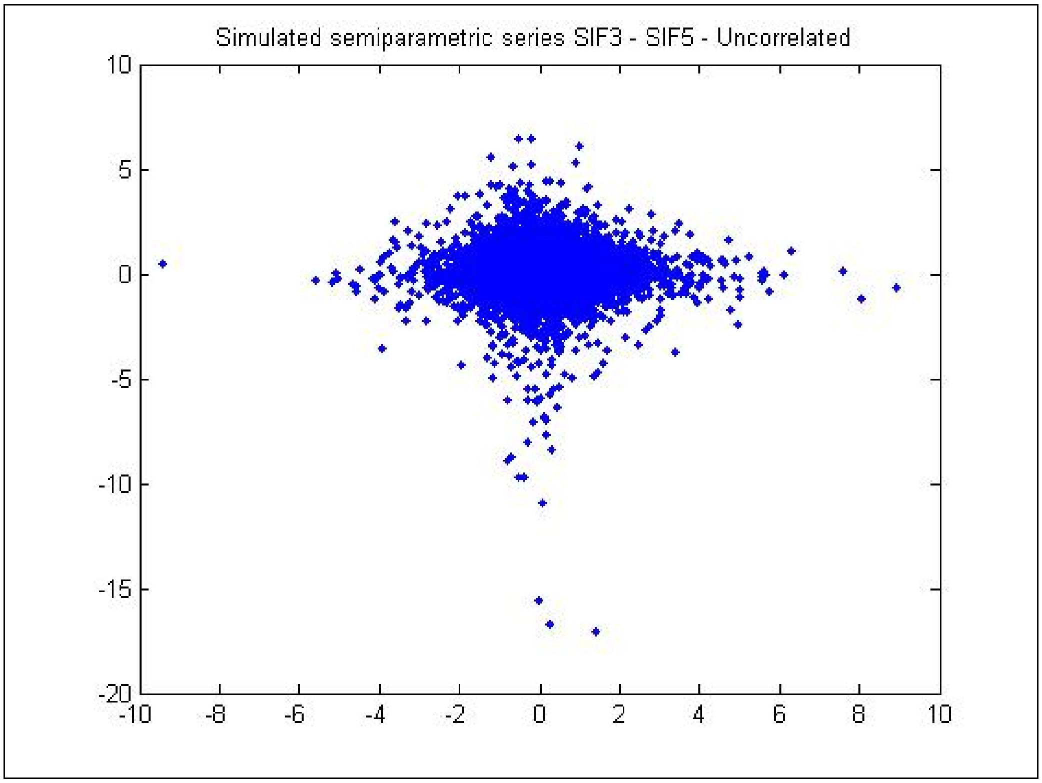
DoF	Std Error
5.425141	0.328318

Figure 1. The negative log-likelihood function of the t-copula

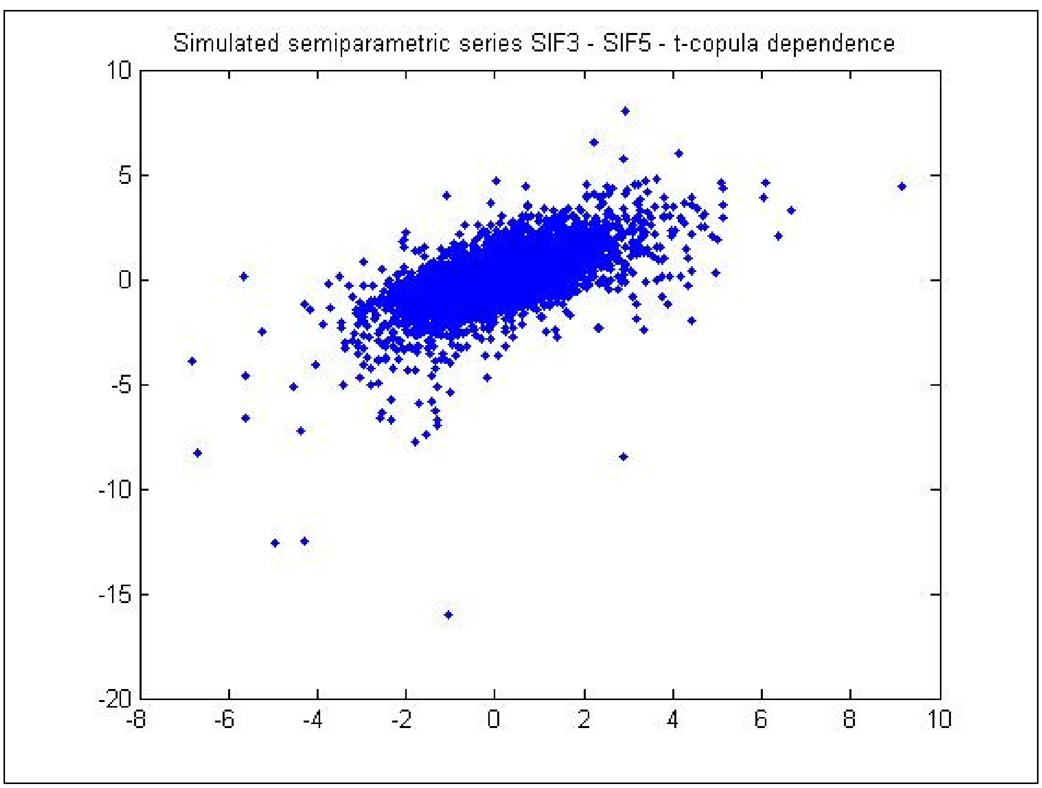


Appendix IV – Simulation and measures of risk

Figure 1. a) Simulated semi-parametric series, SIF3 vs SIF5 under the assumption of independence (b) Simulated semi-parametric series, SIF3 vs SIF5 with the dependence structure given by the t-copula



(a)



(b)

Table 1. Estimated Value-at-Risk and Expected Shortfall for a 1-day horizon – individual series

SIF1					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-1.86%	-2.58%	-5.40%	-3.22%	-4.30%	-7.78%

SIF2					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-2.02%	-2.86%	-5.68%	-3.55%	-4.73%	-8.58%

SIF3					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-2.50%	-3.49%	-6.39%	-4.11%	-5.37%	-8.63%

SIF4					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-2.40%	-3.09%	-6.08%	-3.80%	-4.92%	-8.33%

SIF5					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-1.83%	-2.41%	-4.31%	-2.94%	-3.78%	-6.65%

Table 2. Estimated Value-at-Risk and Expected Shortfall for a 10-day horizon – individual series

SIF1					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-5.57%	-8.51%	-15.75%	-10.36%	-13.87%	-24.54%

SIF2					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-6.85%	-9.98%	-16.88%	-11.47%	-14.75%	-23.91%

SIF3					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-6.80%	-10.12%	-18.72%	-12.19%	-16.08%	-27.26%

SIF4					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-6.23%	-9.52%	-16.99%	-11.07%	-14.37%	-22.77%

SIF5					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-5.75%	-8.69%	-20.48%	-12.24%	-17.45%	-37.68%

Table 3. Estimated Value-at-Risk and Expected Shortfall for a 1-day horizon – portfolio

PORTFOLIO					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-1.81%	-2.52%	-4.87%	-3.03%	-3.93%	-6.40%

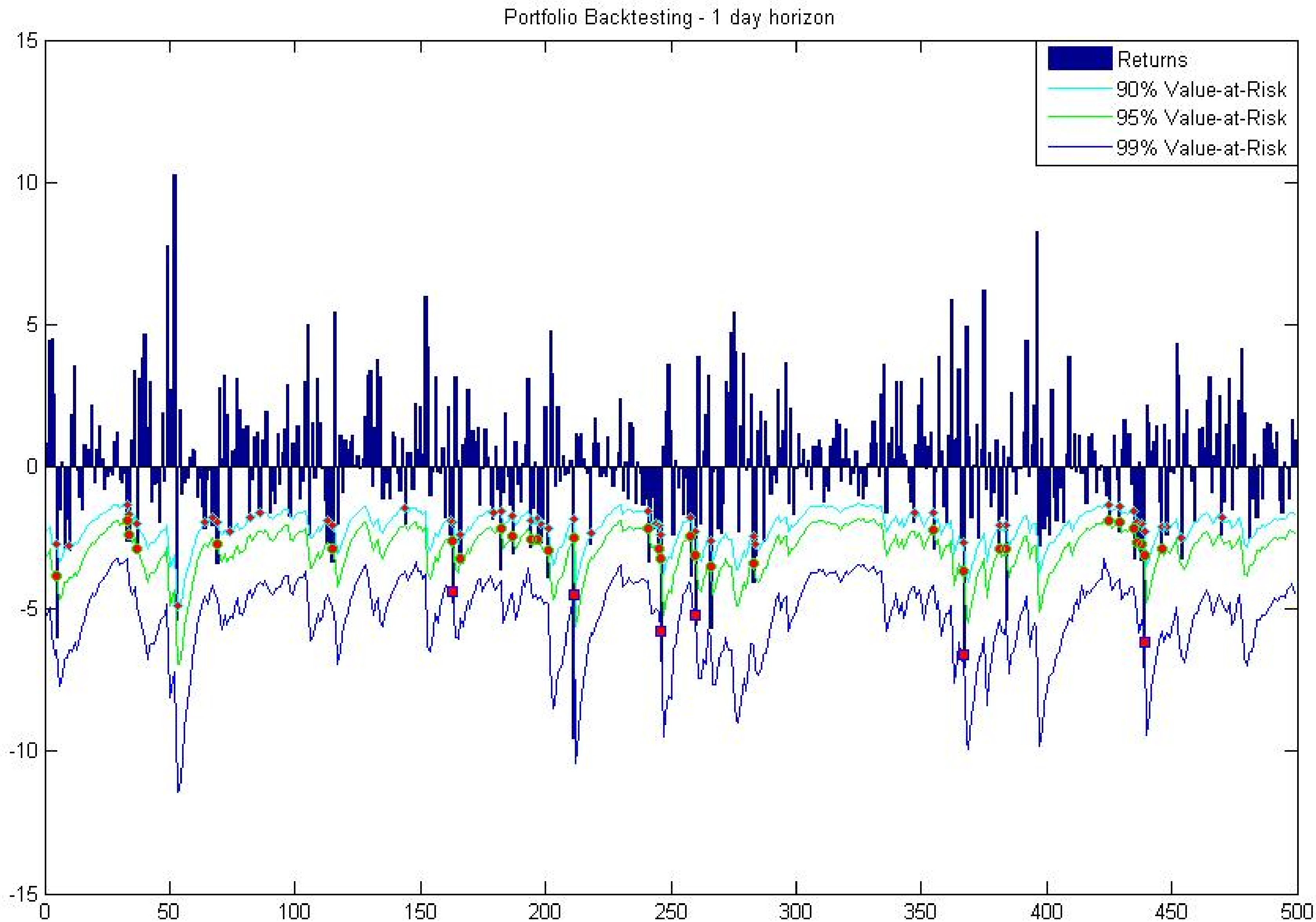
Table 4. Estimated Value-at-Risk and Expected Shortfall for a 1-day horizon – portfolio

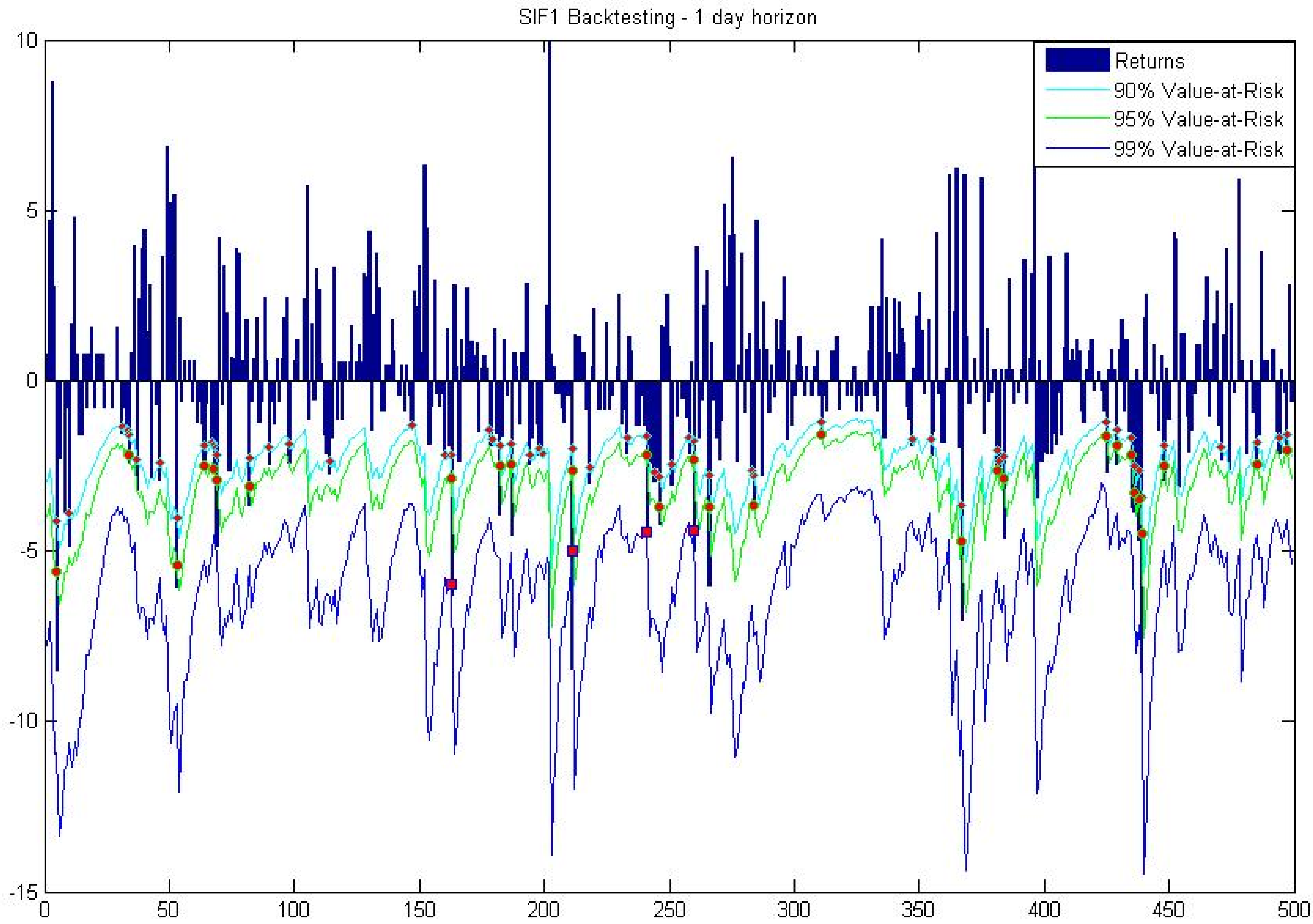
PORTFOLIO					
VaR			**ES**		
90%	95%	99%	90%	95%	99%
-5.03%	-7.54%	-15.74%	-9.58%	-13.03%	-24.14%

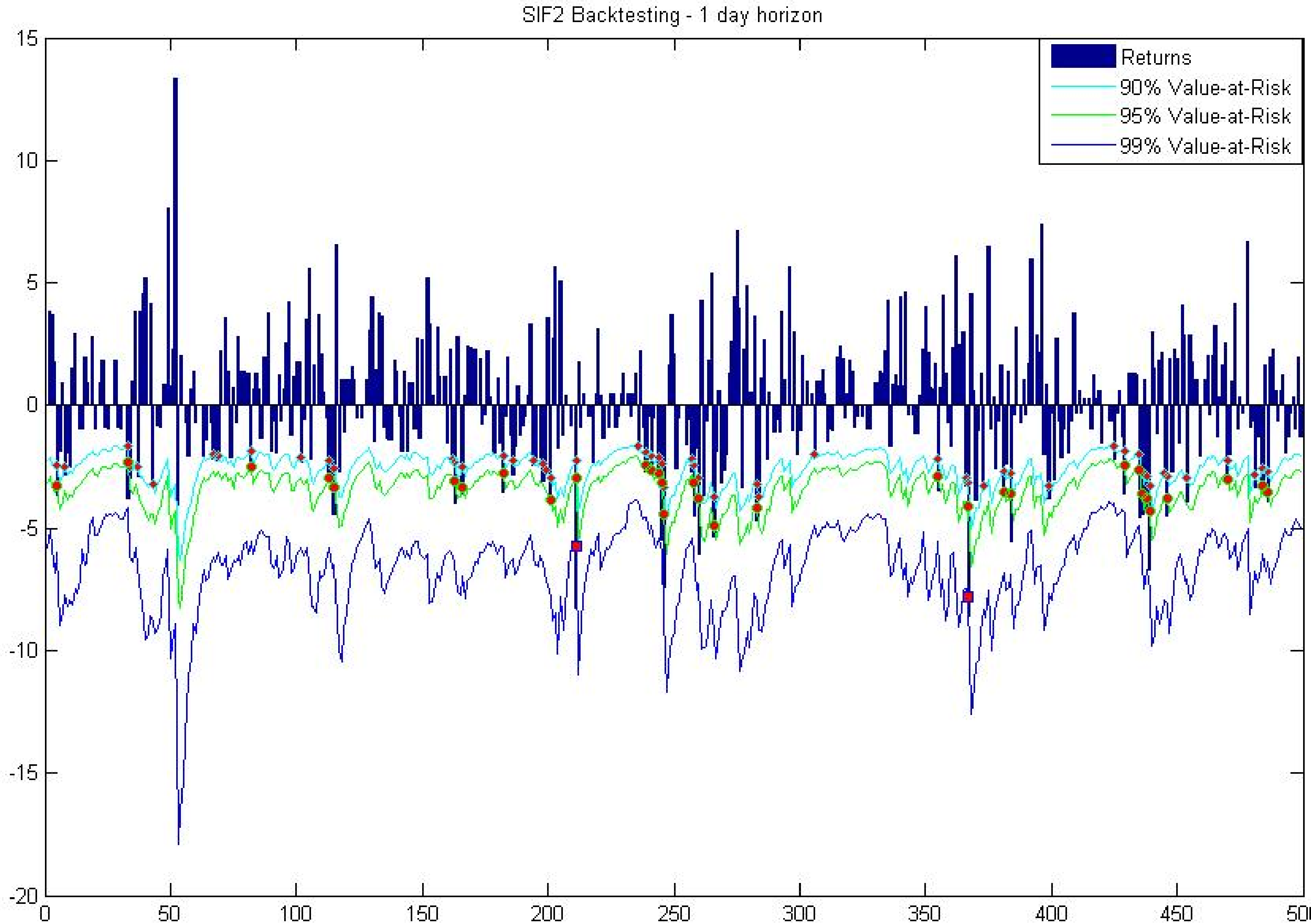
Table 5.Backtesting results –actual vs expected number of exceedances for Value-at-Risk

Backtesting Results			
	VaR - 90%	**VaR - 95%**	**VaR - 99%**
Expected	50	25	5
SIF1	56	29	4
SIF2	55	32	2
SIF3	47	28	3
SIF4	46	32	6
SIF5	45	24	4
Portfolio	51	32	6

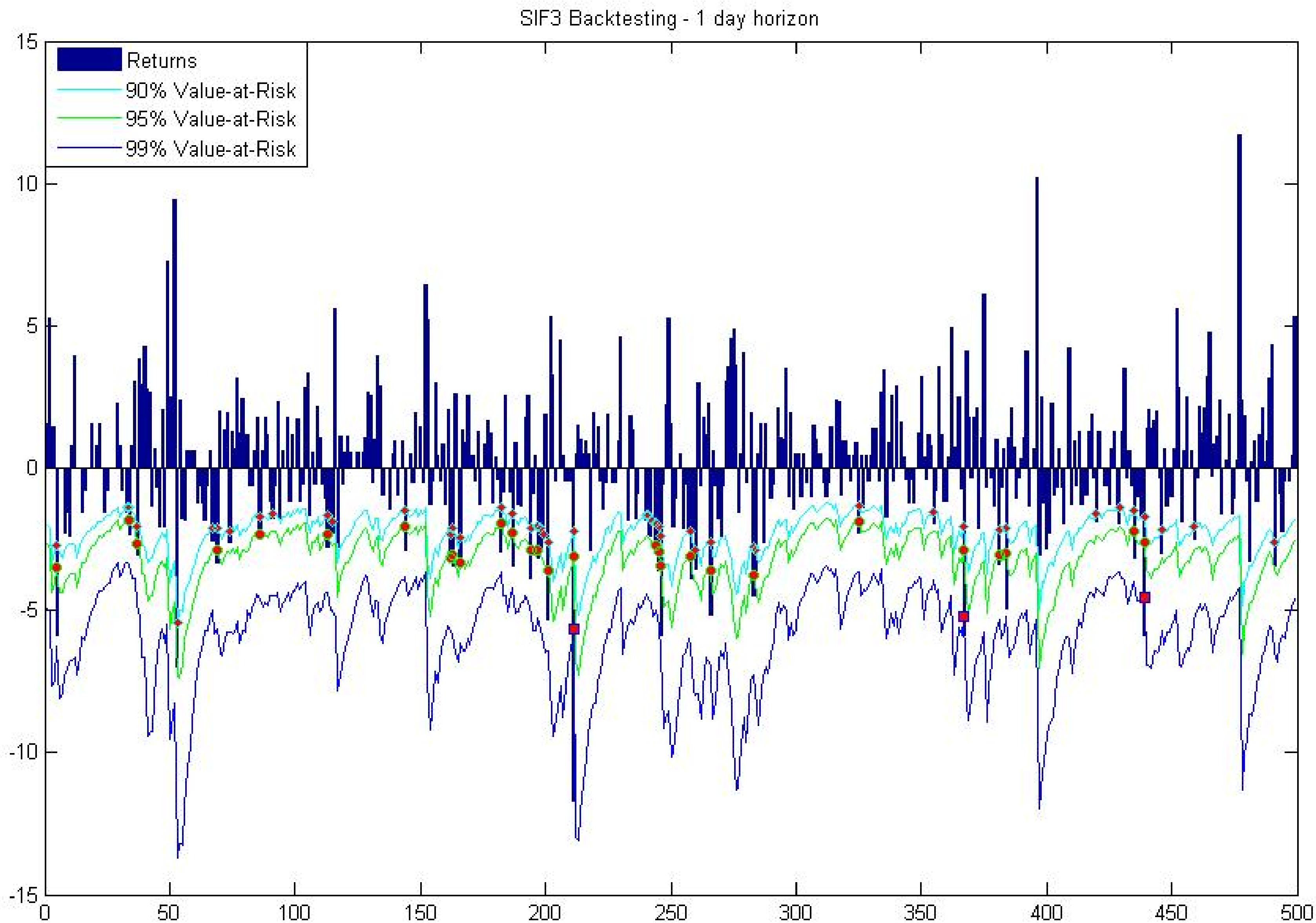
Figure 2. A display of the backtesting results for the individual series and for the portfolio. The exceedances are marked with red.



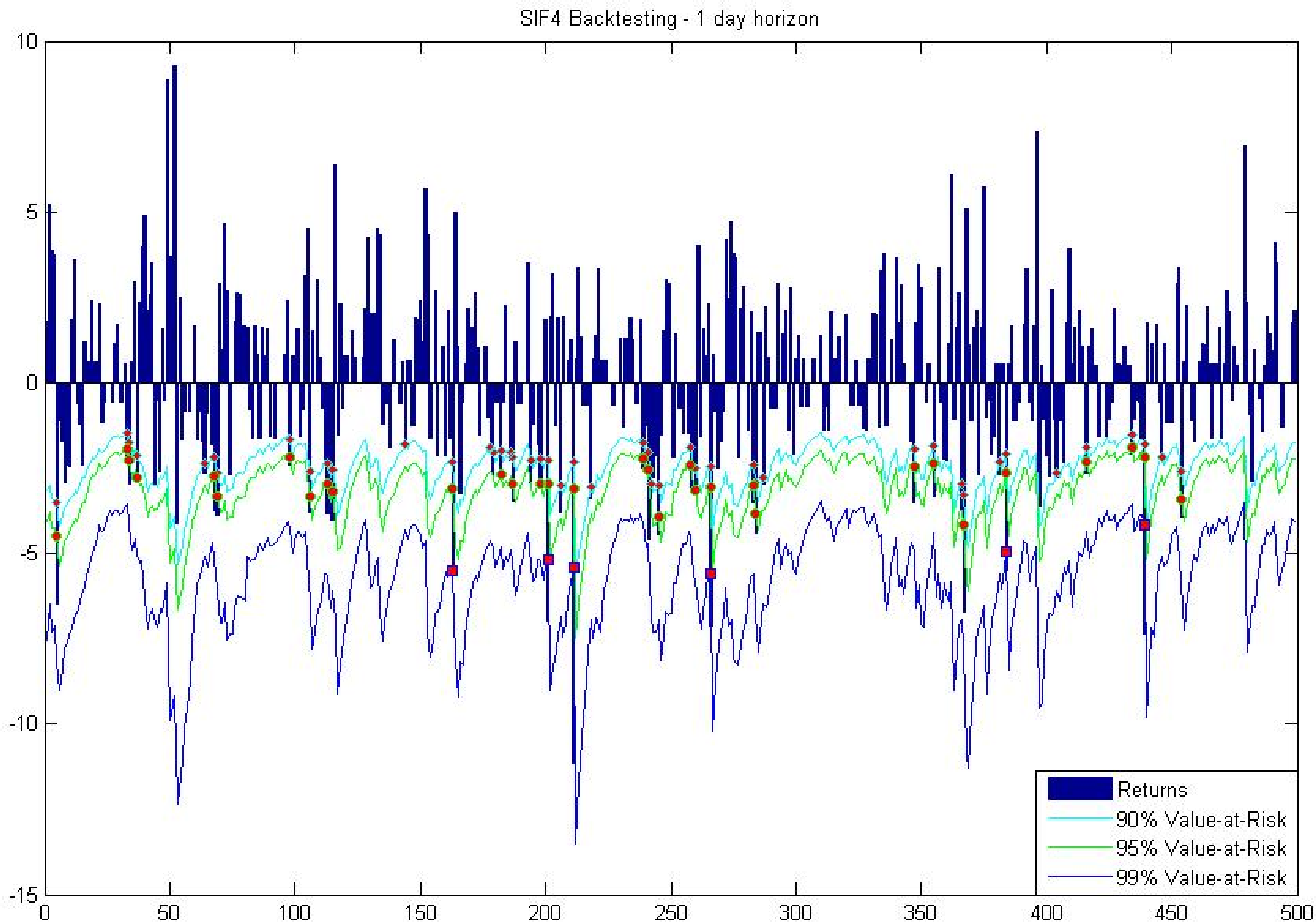

SIF1 Backtesting - 1 day horizon
Returns
90% Value-at-Risk
95% Value-at-Risk
99% Value-at-Risk
10
5
0
-5
-10
-15
0
50
100
150
200
250
300
350
400
450
500

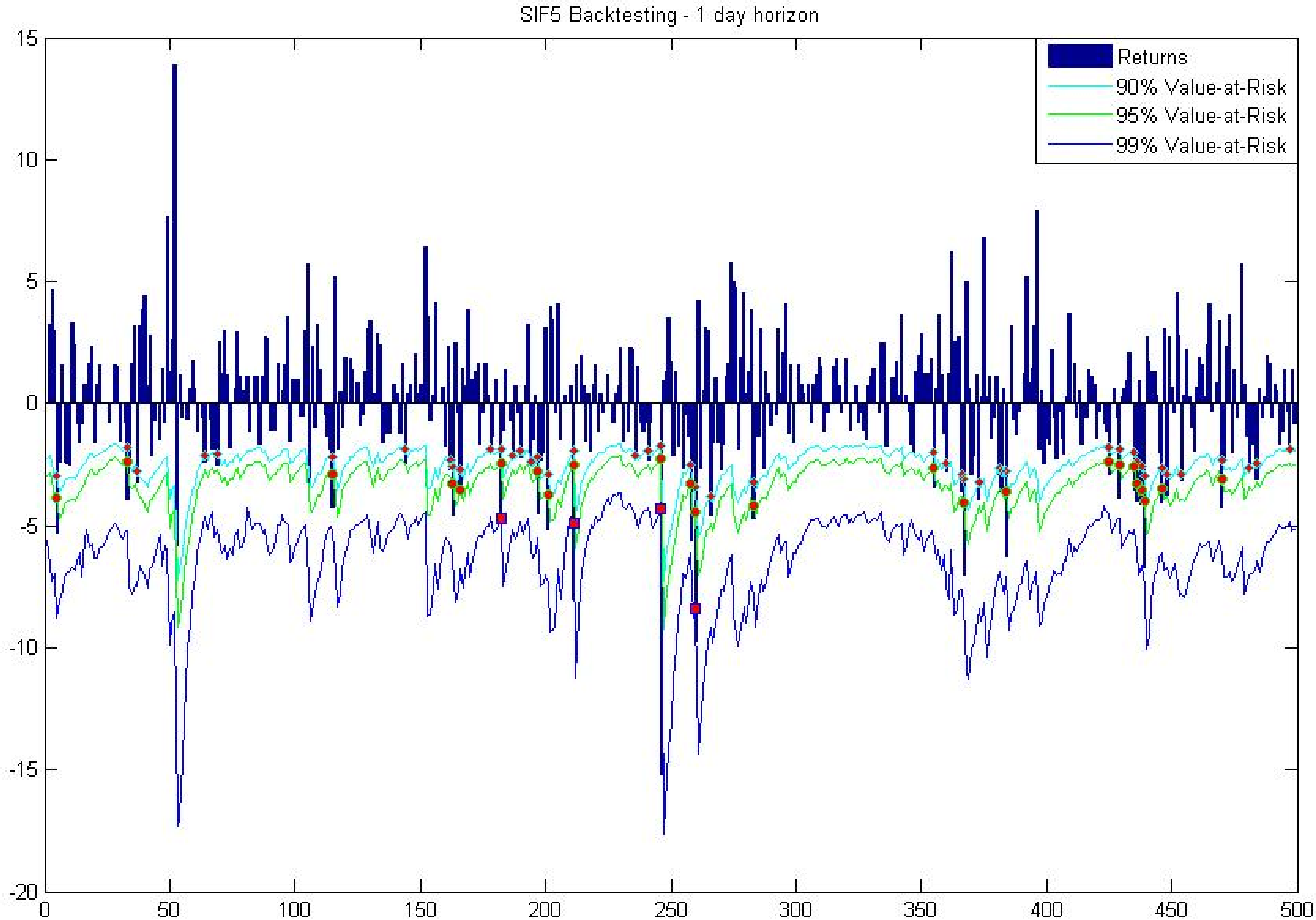


SIF2 Backtesting - 1 day horizon
Returns
90% Value-at-Risk
95% Value-at-Risk
99% Value-at-Risk
15
10
5
0
-5
-10
-15
-20
0
50
100
150
200
250
300
350
400
450
500



SIF3 Backtesting - 1 day horizon
Returns
90% Value-at-Risk
95% Value-at-Risk
99% Value-at-Risk
15
10
5
0
-5
-10
-15
0
50
100
150
200
250
300
350
400
450
500


SIF4 Backtesting - 1 day horizon
10
5
0
-5
-10
-15
0
50
100
150
200
250
300
350
400
450
500
Returns
90% Value-at-Risk
95% Value-at-Risk
99% Value-at-Risk


SIF5 Backtesting - 1 day horizon
Returns
90% Value-at-Risk
95% Value-at-Risk
99% Value-at-Risk
15
10
5
0
-5
-10
-15
-20
0
50
100
150
200
250
300
350
400
450
500