Report

Calculating quantile risk measures for financial return series using extreme value theory

Author(s):
McNeil, Alexander J.

Publication Date:
1998

Permanent Link:
https://doi.org/10.3929/ethz-a-004320029 →

Rights / License:

ETH Library

Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory

Alexander J. McNeil
Departement Mathematik
ETH Zentrum
CH-8092 Zürich
Tel: +41 1 632 61 62
Fax: +41 1 632 10 85
email: mcneil@math.ethz.ch

April 1998

Abstract

We consider the estimation of quantiles in the tail of the marginal distribution of financial return series, using extreme value statistical methods based on the limiting distribution for block maxima of stationary time series. A simple methodology for quantification of worst case scenarios, such as ten or twenty year losses is proposed. We validate methods on a simulated series from an ARCH(1) process showing some of the features of real financial data, such as fat tails and clustered extreme values; we then analyse daily log returns on a share price.

Keywords: Risk Measures, Quantile Estimation, Financial Time Series, ARCH models, Extreme Value Theory

1 Introduction

This paper is concerned with quantile measures of risk for financial return series and their calculation using extreme value theory (EVT). This is a topical issue which has given rise to several recent publications (Embrechts, Klüppelberg & Mikosch 1997, Reiss & Thomas 1997, Embrechts, Resnick & Samorodnitsky 1998, Longin 1997b, Longin 1997a, Danielsson & de Vries 1997a, Danielsson & de Vries 1997b).

For easy interpretation, we consider series of daily returns, although our methods extend to higher (or lower) frequency return series. Our basic idealisation is that returns follow a stationary time series model with conditional heteroscedasticity structure.

With such models there are two kinds of quantiles to be considered – conditional and unconditional – and both are of relevance to risk managers. By conditional we mean a quantile of the conditional distribution of the process, where the conditioning is on the current volatility. By unconditional we mean a quantile of the marginal or stationary distribution of the process.

The former quantity is essentially the object of interest in Value at Risk (VaR) calculation, where we are interested in knowing the possible extent of a loss caused by an adverse market movement over the next day (or next few days) against the current volatility background. An approach to estimating this requires us to estimate the current volatility, perhaps using GARCH modelling, and to make some assumption about the form of the conditional distribution. Whilst the Gaussian assumption is common, a heavy-tailed distribution like Student-t might be more realistic. In another recent paper we have looked at the combination of econometric modelling of volatility with extreme

value modelling of the tail of the conditional distribution to calculate VaR and related risk measures (McNeil 1998).

Calculation of unconditional quantiles provides different, but complementary information about risk. Here we take a long-term view and attempt to assign a magnitude to a specified rare adverse event, such as a 5-year or 10-year loss. This kind of information may be of interest to the risk manager who wishes to perform scenario analysis and get a feeling for the scale of worst case or stress losses.

In this paper we will look at methods for calculating such unconditional quantiles. In many ways this is the more difficult estimation problem, due to the fact that the marginal distribution of a conditionally heteroscedastic process is heavier-tailed than the conditional distribution. Normal approximations will generally fail and grossly underestimate the extent of the rare loss; methods from EVT explicitly address the heavy-tailed nature of the problem.

Although our emphasis is on quantile estimation, an important stage on the way to estimating extreme quantiles is the determination of the tail index of a return series. This has been a subject of some interest in the empirical finance literature; see papers by Koedijk, Schafgans & deVries (1990), Jansen & deVries (1991), Loretan & Phillips (1994), Lux (1996) and Longin (1996). A common theme in all of these papers is the use of semi-parametric estimation techniques based on higher order statistics, such as Hill's estimator (Hill 1975) and the moment estimator (Dekkers, Einmahl & de Haan 1989), to estimate the tail index ξ (or often $\alpha = 1/\xi$) of the series. There is general consent that financial time series are leptokurtic and there has been some interest in using estimates of ξ to decide how many finite moments the marginal distributions possess. In analyses of stock price returns, estimates of ξ are mostly less than 0.5, implying finite variance, and often less than 0.25, implying a finite fourth moment (Jansen & deVries 1991, Loretan & Phillips 1994, Lux 1996).

The Hill and moment estimators may be subsumed in a larger family of EVT methods known as threshold analyses or sometimes POT (peaks over thresholds). In this family there are also some fully parametric methods based on generalized Pareto approximations to excess losses over high thresholds, which have become popular in insurance applications (McNeil 1997, Rootzén & Tajvidi 1997) and which we use to calculate conditional risk measures in McNeil (1998). However, we note there are some practical problems with the use of threshold methods for unconditional inference about financial return series. The asymptotic properties of tail index estimators have not been extensively investigated for dependent samples of extreme values from dependent processes and estimates may be heavily biased unless the number of upper order statistics (or threshold) is chosen very carefully (see page 270 of Embrechts et al. (1997) for an example involving an ARMA process). The problem of sensitivity to the threshold is particularly pronounced for the Hill estimator, even in the case of iid data from certain distributions such as the alpha-stable family (Mittnik & Rachev 1996). There has been promising some recent work on using bootstrap methods to determine the optimal number of order statistics (Danielsson, de Haan, Peng & de Vries 1997), but further validation of such methods is still required. In addition to the problem of threshold choice, there is also the related unresolved problem of how to give reasonable errors and confidence intervals for point estimates of tail indices or quantiles based on dependent data.

In this paper we follow a different path which partially circumvents these dependence problems and which was also adopted in a study of an American stock index by Longin (1996). We estimate the tail index and calculate quantiles using the most traditional of extreme value methods, the fully parametric analysis of block maxima using the extreme value limiting distributions. This approach dates back to Gumbel (1958) and has a long history of application in hydrology in particular. It leads to very natural worst case quantile measures known as return levels, which have an easy interpretation in terms of

time horizons. For example, the ten year return level is that level which will be exceeded in one year every ten years, on average.

The paper is structured as follows. Section 2 contains the necessary theoretical background, with particular reference to extreme values in stationary series. EVT for independent, identically distributed series can be found in a number of sources, but the generalisation to stationary series is less well known. For financial applications it is essential to encounter this theory and, in particular, to understand the notion of the extremal index of a stationary series.
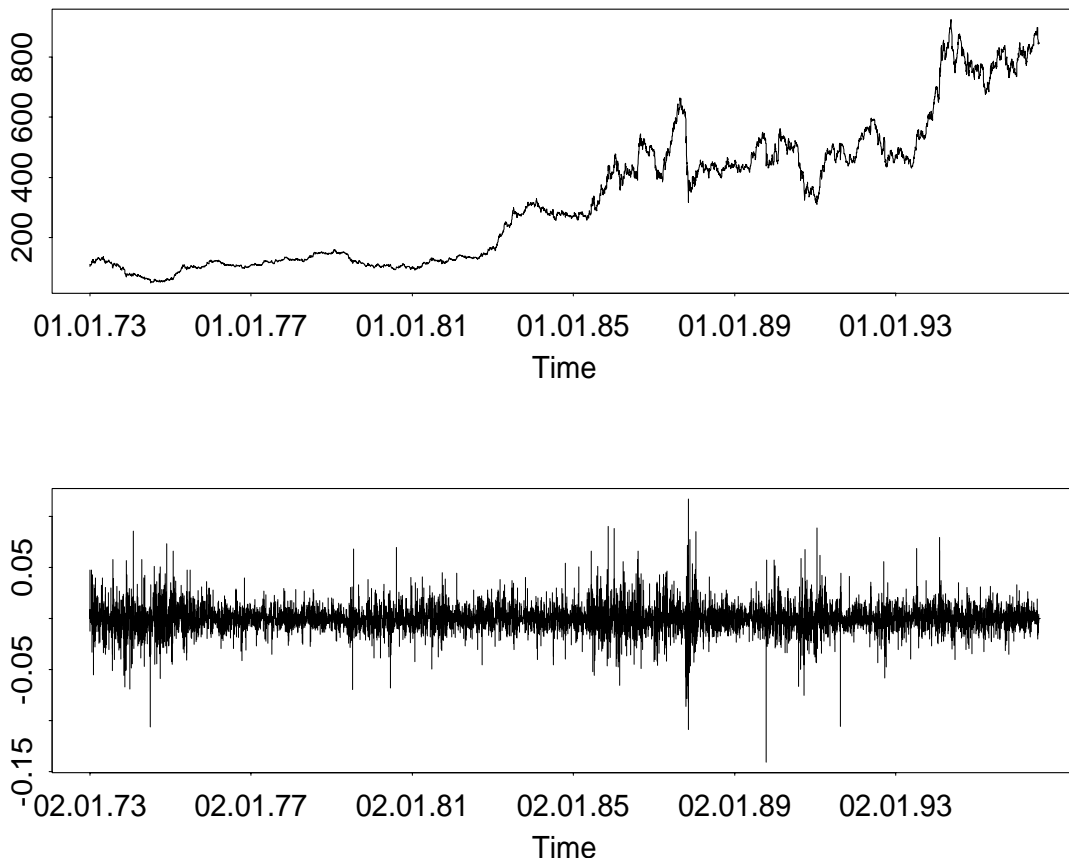

Figure 1: Daily closing prices of BMW shares from January 1973 to July 1996; raw d ata in upper picture, log returns in lower

In Section 3 the method for tail index and quantile estimation is demonstrated and verified on a simulated ARCH(1) process which mimics features of real return data (Figure 2). In section 4 the methods are applied to the series of 6146 daily logarithmic returns on the BMW share price from January 1973 to July 1996 (Figure 1), this particular return series being chosen purely for illustrative purposes. If the stock price at time t is S_t the log return is

$$r_t = \log\left(S_t/S_{t-1}\right),$$

which, for small price movements, is approximately equal to the relative return $(S_t - S_{t-1})/S_{t-1}$. We will be interested in the downward movements of the price represented by log returns less than zero. We analyse large values in the series of negative log returns since results in extreme value theory are most frequently expressed in terms of large values and maxima.
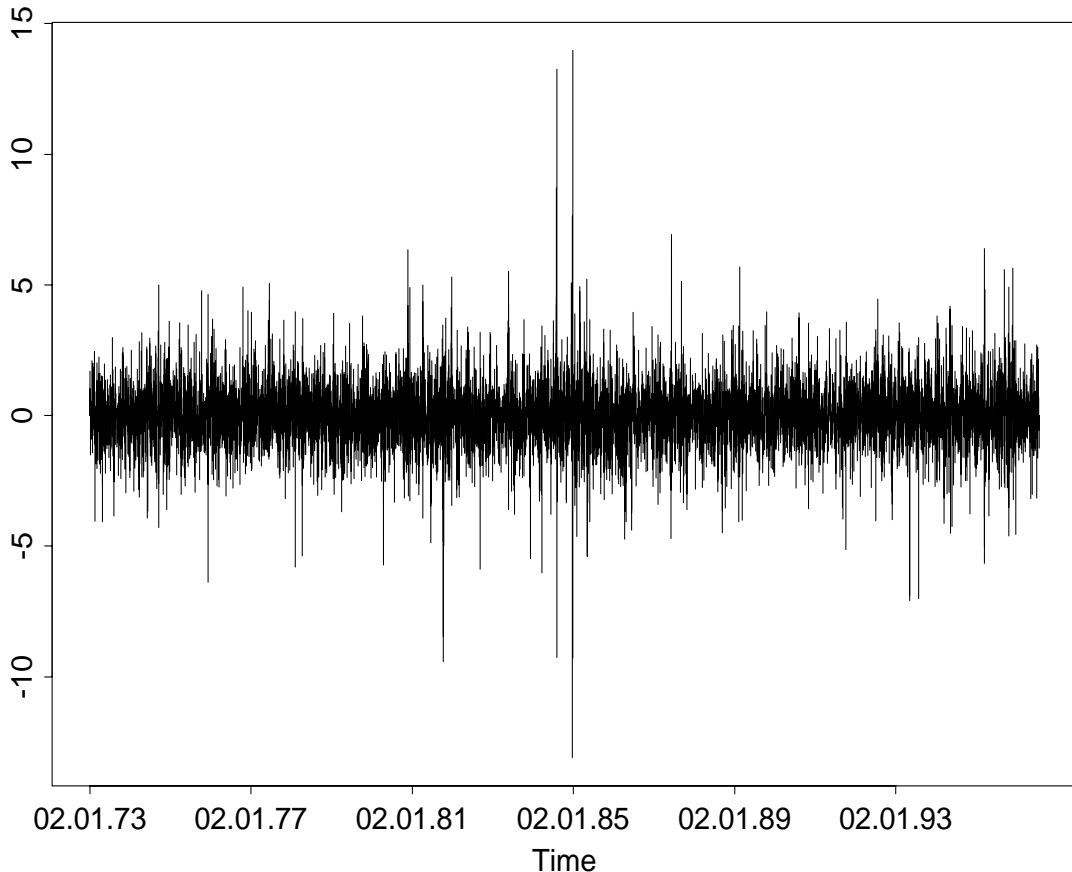
Figure 2: Simulated ARCH(1) process with same length as BMW log return series

2 Extreme Value Theory for Block Maxima

2.1 Extremal Types Theorem for iid Series

Let $(X_n, n \in \mathbb{Z})$ be a family of random variables representing daily observations of the log return on a stock price. We make the crude first assumption that these returns are independent, identically distributed (iid) from an underlying unknown distribution F; shortly we relax the assumption of independence. We are interested in extrema over longer periods such as quarters, semesters or years; we call such periods blocks.

Assume that we have data which may be divided into m such blocks with approximately n days in each block. Clearly, if we aggregate a dataset by calendar quarters or other blocks, the exact number of days in each block will differ slightly, but we ignore this feature for simplicity of notation. For an arbitrary block we denote the daily data by $X_1, \ldots, X_n$ and we let $M_n = \max(X_1, \ldots, X_n)$ denote the maximum observation in the block. Extreme Value Theory suggests fitting the generalized extreme value distribution to our m independent realisations of the block maxima random variable M_n. This is due to the celebrated extremal types theorem of Fisher & Tippett (1928).

Suppose that the block maxima M_n show regular limiting behaviour, in the sense that there exist sequences of real constants b_n and $a_n > 0$ such that

$$\lim_{n \to \infty} P\left\{(M_n - b_n)/a_n \le x\right\} = \lim_{n \to \infty} F^n(a_n x + b_n) = H(x), \tag{1}$$

for a non-degenerate distribution function $H(x)$. If this condition holds F is said to be in

4

the maximum domain of attraction of H, written $F \in \text{MDA}(H)$.

The generalized extreme value distribution (GEV distribution) is defined to have the distribution function

$$H_\xi(x) = \begin{cases} \exp(-(1 + \xi x)^{-1/\xi}) & \text{if } \xi \neq 0, \\ \exp(-e^{-x}) & \text{if } \xi = 0, \end{cases} \tag{2}$$

where x is such that $1 + \xi x > 0$, and where the case $\xi = 0$ should be viewed as the limit of the distribution function as $\xi \to 0$. The three cases $\xi < 0$, $\xi = 0$ and $\xi > 0$ are sometimes referred to as the Weibull, Gumbel and Fréchet distributions respectively.

Theorem 1 (Fisher-Tippett)

$$F \in MDA(H) \implies H \text{ is of the type } H_\xi \text{ for some } \xi.$$

Thus, if condition (1) is fulfilled for a non-degenerate $H(x)$ then, up to changes in scale and location, this limit distribution is a GEV.

2.2 The Domains of Attraction

Essentially all the common, continuous distributions of statistics are in $\text{MDA}(H_\xi)$ for some value of ξ. The normal distribution, commonly assumed for log returns in finance despite its deficiencies in the tail area, is a member of $\text{MDA}(H_0)$ (the Gumbel case). Distributions in this class we shall call thin-tailed.

Distributions in $\text{MDA}(H_\xi)$ for $\xi > 0$ (the Fréchet case) we shall call heavy-tailed. This class includes the Pareto, the Student-t and the alpha-stable distributions with characteristic exponent in $(0, 2)$; the marginal distribution of a stationary ARCH process is also in $\text{MDA}(H_\xi)$ for $\xi > 0$. The following characterization was given by Gnedenko (1943).

Theorem 2 *For $\xi > 0$,*

$$F \in MDA(H_\xi) \iff 1 - F(x) = x^{-1/\xi} L(x),$$

for some slowly varying function $L(x)$.

Thus distributions giving rise to the Fréchet case are those whose tails decay essentially like a power function. For a distribution in $\text{MDA}(H_\xi)$ with $\xi > 0$, we will call ξ the tail index; such a distribution has a finite jth moment if $j < 1/\xi$. (Note that some authors reserve the tail index label for $\alpha = 1/\xi$.)

We do not know the underlying distribution of our log returns but believe it to be heavy-tailed so that the Fréchet limit will be the relevant case. We exploit the extremal types theorem by fitting a three parameter form of the GEV distribution $H_{\xi,\mu,\sigma}(x) = H_\xi((x - \mu)/\sigma)$ to our data on block maxima. The location parameter μ and the positive scale parameter σ take care of the unknown sequences of normalizing constants a_n and b_n in (1). The GEV distribution can be fitted using various methods. We choose to use maximum likelihood (Hosking 1985); the asymptotic likelihood theory was worked out for the regular case ($\xi > -0.5$) in Smith (1985).

2.3 Alternative Condition for Convergence of Maxima

By taking logarithms of both sides, it may be shown that an alternative condition to (1) is that $a_n > 0$ and b_n exist so that

$$\lim_{n \to \infty} n \left(1 - F(a_n x + b_n) \right) = -\log H_\xi(x).$$

5

In fact, we have the following more general equivalence. Let u_n be any non-decreasing sequence of real numbers, and not necessarily a linear function such as $a_n x + b_n$. For $0 \leq \tau \leq \infty$

$$\lim_{n \to \infty} n\left(1 - F(u_n)\right) = \tau$$

if and only if

$$\lim_{n \to \infty} P\left\{M_n \leq u_n\right\} = \exp(-\tau). \tag{3}$$

These equivalent statements will provide a contrast with the corresponding results for stationary series in the next sections.

2.4 Strictly Stationary Series

We now consider what happens if we relax our iid assumption to consider strictly stationary time series with serial dependencies. Stationary time series should provide more realistic models for log returns, at least over time periods which are not too long.

Of course, there is nothing to prevent us simply fitting the GEV to block maxima as we have described; what we are seeking is some justification for choosing this particular limiting distribution in the stationary case, as we had in the iid case. Under certain additional assumptions it can indeed be shown that appropriately normalized block maxima follow the GEV asymptotically. The relevant results are given in detail in Chapter 3 of Leadbetter, Lindgren & Rootzén (1983) and Section 4.4 of Embrechts et al. (1997) and summarised only briefly here.

We need the following notation. (X_n) is now our stationary time series and, as before, the marginal distribution of an observation X_i is F and $M_n = \max(X_1, \ldots, X_n)$. We denote by $(\tilde{X}_n)$ an associated iid series with the same marginal distribution F and we let $\tilde{M}_n = \max(\tilde{X}_1, \ldots, \tilde{X}_n)$.

The first natural question to look at is, under which conditions do maxima of (X_n) have exactly the same limiting behaviour as maxima of $(\tilde{X}_n)$. In Leadbetter et al. (1983) we find two technical conditions for this to happen. If the stationary series shows only weak long-range dependence (made precise by the mixing condition $D(u_n)$) and if it shows no tendency to form clusters of large values (made precise by the anti-cluster condition $D'(u_n)$) then maxima of the two series have identical limiting behaviour. If $F \in \mathrm{MDA}(H_\xi)$ for some ξ, then H_ξ will also be the natural limiting distribution for normalized block maxima of the stationary series; moreover the normalizing sequences will be the same for both the stationary and the associated iid series.

Whilst the weak long-range dependence condition may be tenable for a financial series, the anti-clustering condition is not. Financial time series show clusters of volatility which lead to clusters of large values. A stationary ARCH(1) process satisfies $D(u_n)$ but not $D'(u_n)$. We cannot apply the above theory but must look more carefully into the issue of clustering. To do this we define the extremal index of a stationary process.

2.5 Extremal Index of a Stationary Series

First we give a formal definition before offering a more intuitive heuristic definition.

Let $0 \leq \theta \leq 1$ be a real number and suppose for every $\tau > 0$ we can find a sequence $u_n(\tau)$ such that the following hold

$$\lim_{n \to \infty} n\left(1 - F(u_n(\tau))\right) = \tau \tag{4}$$
$$\lim_{n \to \infty} P\left\{M_n \leq u_n(\tau)\right\} = \exp(-\theta\tau)$$

Then we say that (X_n) has extremal index θ.

It can be shown that θ is well defined in this way and not dependent on the particular choice of sequence $u_n(\tau)$. In fact, if (X_n) has extremal index θ, then for general u_n and $\tau > 0$ the following can be demonstrated to be equivalent

$$\lim_{n \to \infty} n \left(1 - F(u_n) \right) = \tau$$
$$\lim_{n \to \infty} P \left\{ \tilde{M}_n \leq u_n \right\} = \exp(-\tau)$$
$$\lim_{n \to \infty} P \left\{ M_n \leq u_n \right\} = \exp(-\theta\tau).$$

These three equivalent statements for a stationary time series with extremal index θ may be compared with the two equivalent statements in (3) for iid series. We infer that for large n

$$P \left\{ M_n \leq u_n \right\} \approx P^\theta \left\{ \tilde{M}_n \leq u_n \right\} = F^{n\theta}(u_n). \tag{5}$$

This will be our heuristic definition of θ. A stationary series showing property (5) has extremal index θ.

This definition gives some intuition for the effect of clustering and the nature of θ. We see that the maximum of n observations from the stationary series with extremal index θ behaves like the maximum of $n\theta$ observations from the associated iid series. $n\theta$ can thus be thought of as counting the number of pseudo-independent clusters in n observations so that θ has an interpretation as the reciprocal of the mean cluster size.

We now state the main result for stationary series.

Theorem 3 *If (X_n) is stationary with extremal index $\theta > 0$ then*

$$\lim_{n \to \infty} P\{(\tilde{M}_n - b_n)/a_n \leq x\} = H(x),$$

for a non-degenerate $H(x)$ if and only if

$$\lim_{n \to \infty} P\{(M_n - b_n)/a_n \leq x\} = H^\theta(x), \tag{6}$$

with $H^\theta(x)$ also non-degenerate.

Thus, for $F \in \mathrm{MDA}(H_\xi)$, the asymptotic distribution of normalized maxima of the stationary series (X_n) with extremal index θ is also an extreme value distribution. The shape parameter is ξ as in the iid case, since raising the df to the power θ only affects location and scaling parameters.

This is the justification we require for fitting the GEV to the block maxima of a stationary time series which shows a tendency to form clusters of large values. The dependence of the X_i does have the effect that convergence to the GEV is slower, since the effective sample size is $n\theta$ which is smaller than n. Thus more care should be taken that blocks are sufficiently long to admit a GEV model for their maxima. Also, blocks should be sufficiently long so that maxima in different blocks can be assumed to be effectively independent observations for the purposes of statistical estimation.

3 Analysis of an ARCH(1) Process

3.1 Properties of the Simulated Process

The autoregressive conditional heteroscedasticity (ARCH) model of Engle (1982) reproduces the kind of stochastic volatility changes observed in real financial time series. For a review article on ARCH and other related processes see, for example, Bollerslev, Chou & Kroner (1992). We consider an ARCH(1) process defined by

$$X_i = Z_i(\beta + \lambda X_{i-1}^2)^{1/2}, \ i \geq 1,$$

where the Z_i are iid standard Normal variates and β and λ are parameters to be set. The extremal properties of this process have been analysed in de Haan, Resnick, Rootzén & de Vries (1989).

Although it is possible to fit this model to our BMW data by maximum likelihood, we choose not to simulate from the fitted model thus obtained. The ARCH(1) process with only two parameters and normal innovations is too simplistic a model for real financial time series. In particular, the fitted model fails to explain the extreme movements observed in the data, which are the focus of this paper. The marginal distribution is thin-tailed and high quantiles are considerably smaller than the empirical quantiles of the data; the fitted model also shows no tendency to cluster. It seems likely we would have to fit GARCH or ARCH processes of high order with heavy-tailed innovations to begin to reproduce the kind of extremal behaviour seen in the original data. Since the extremal properties of such processes are less easy to derive analytically, we choose not to follow this path. Instead we simulate an ARCH(1) process with parameters $\beta = 1$, $\lambda = 0.5$ and starting value $X_0 = 0$. β is a scaling parameter and not important as far as extremal properties are concerned but the reasons for choosing $\lambda = 0.5$ are as follows.

First, the resulting process is stationary. Second, de Haan et al. (1989), using results for stochastic difference equations from Kesten (1973), have shown that for x large

$$1 - F(x) \sim cx^{-4.73},$$

for some value of c. Thus, by the result of Gnedenko (1943), the marginal distribution of this series is in the domain of attraction of the Fréchet with shape parameter $\xi = 0.21$, which is the kind of value estimated in empirical studies of real stock market returns. Finally de Haan et al. (1989) show that the extremal index θ exists and is equal to 0.835 so that this process forms clusters and the average cluster size is 1.20. This is enough information for us to know that the asymptotic distribution for maxima of this process is a Fréchet distribution with $\xi = 0.21$.

3.2 Fitting the GEV Distribution

	Quarter		Semester		Year	
$\approx$ n	65		130		261	
m	95		48		24	
ξ	0.22	(0.06-0.38)	0.24	(0.00-0.49)	0.26	(-0.11-0.64)
σ	0.92	(0.75-1.09)	1.04	(0.76-1.31)	1.11	(0.68-1.54)
μ	2.96	(2.74-3.16)	3.48	(3.14-3.82))	4.26	(3.74-4.78)
$R_{n,20}$	6.83	(5.88-8.64)	8.00	(6.60-11.82)	9.29	(7.30-19.26)
R_{20}	9.78	(7.71-14.67)	9.64	(7.58-16.31)	9.29	(7.30-19.26)

Table 1: Parameter estimates and estimated return levels for GEV models fitted to block maxima of ARCH(1) data; confidence intervals in brackets

We simulated an ARCH series of the same length as the BMW series and attached similar dates to each observation (Figure 2). We then analysed quarterly, semesterly and yearly maxima as shown in Table 1. The fit of the Fréchet model to these block maxima was investigated using crude residuals in the sense of Cox & Snell (1968); for each block maximum $M_{ni}, i = 1, \ldots, m$ the corresponding residual is defined to be

$$W_i = \left(1 + \xi \frac{M_{ni} - \mu}{\sigma}\right)^{-1/\xi}.$$

These should be iid unit exponentially distributed and this hypothesis can be checked using graphical diagnostics such as a QQplot. The exponential assumption seemed tenable for all models, suggesting the block maxima were adequately modelled by their respective Fréchet distributions.

The estimated ξ parameters were 0.22, 0.24 and 0.26 for quarterly, semesterly and annual maxima respectively. In no case does the estimated confidence interval allow rejection of the true asymptotic value 0.21, so that our models are consistent with asymptotic theory. On the other hand, using our models for quarterly and semesterly maxima we have enough data to exclude values less than 0 and greater than 0.5 from our confidence intervals; in these models, we conclude that our underlying marginal distribution is heavy-tailed but reject the infinite variance hypothesis. In the analysis of yearly maxima the data quantity (24 maxima) is too meagre and the confidence interval too wide to permit such statements.

3.3 Simulation Study

We performed a small simulation study to examine the properties of the GEV estimator for ξ. We generated 1000 similar realisations from the ARCH(1) process and estimated ξ each time. We used these replications to estimate the bias, standard deviation and mean squared error of the estimator. We also calculated coverage percentages for symmetric 95% confidence intervals calculated from the estimated standard errors for ξ.

Criterion	Quarter	Semester	Year
m	95	48	24
Bias	-0.052 (25%)	-0.052 (25%)	-0.032 (15%)
Standard Deviation	0.088	0.138	0.216
RMSE	0.102	0.147	0.218
Coverage	87%	87%	92%

Table 2: Simulation Study of the properties of the GEV estimator for ξ

Table 2 shows that the bias of the procedure is smallest for year blocks, as one would expect since the procedure is based on an asymptotic argument. However, the variability of the estimates and the root mean squared error is high so that detailed inference about the shape parameter is difficult, as one would also expect from an analyses of only 24 data points. On the other hand, quarterly blocks, whilst giving a somewhat higher bias, give a much smaller RMSE, so that more confident statements about the range in which ξ lies are possible. The coverage percentages are somewhat smaller than the 95% value we would expect, but this is most likely due to the symmetric nature of our confidence intervals; asymmetric intervals might be more realistic as we shall later discuss.

We compared the results for quarterly blocking with threshold analyses using the generalized Pareto distribution based on the 95 upper order statistics of the data; i.e. we used a similar quantity of data. The bias, RMSE and coverage properties of the block maxima method were all superior to those of the threshold method. This may be due to the performance degradation of the threshold methods caused by the use of dependent data, but this awaits a proper large scale analysis.

3.4 Quantiles and Return Levels

The usual definition of the qth quantile x_q of a continuous distribution with df F is

$$x_q = F^{-1}(q), \tag{7}$$

where F^{-1} is the inverse of the distribution function. We consider an alternative type of quantile which is particularly easy to estimate from our fitted model, the so-called return level. This is a concept borrowed from hydrology which also seems well-suited to risk management.

Suppose we consider our model for annual (261 day) maxima in the last column. We define the k–year return level $R_{261,k}$ to be given by

$$P\{M_{261} > R_{261,k}\} = 1/k, \tag{8}$$

for $k > 1$. This is a level we would expect to be exceeded only in one year out of every k years, on average. Similarly, $R_{n,k}$ is a level we expect to be exceeded in one n–block every k n–blocks, on average. We shall call the n–block in which the return level is exceeded a stress period.

If we believe that maxima in blocks of length n follow the generalized extreme value distribution with df $H_{\xi,\mu,\sigma}$, then $R_{n,k}$ is a quantile of this distribution and a simple function of its parameters

$$R_{n,k} = H_{\xi,\mu,\sigma}^{-1}(1 - 1/k) = \mu - \frac{\sigma}{\xi}\left(1 - (-\log(1 - 1/k))^{-\xi}\right). \tag{9}$$

We can estimate $R_{n,k}$ using our maximum likelihood parameter estimates of (ξ, μ, σ) and it is also possible to construct asymmetric confidence intervals using the profile likelihood method which is described in the Appendix.

Using our annual maxima model the maximum likelihood estimate of the 20-year return level $R_{261,20}$ is 9.29 (see Table 1) and a 95% confidence interval is (7.30, 19.3). Point estimates of the 20-quarter and 20-semester return levels $R_{65,20}$ and $R_{130,20}$ are 6.83 and 8.00 respectively. Clearly if the models for maxima in different blocks are consistent then the 20-year return level should correspond roughly with the 40-semester and 80-quarter return levels. In the table we confirm this to be the case by calculating $R_{261,20}$, $R_{130,40}$ and $R_{65,80}$. These are shown in the row marked R_{20}. Note that the confidence intervals based on semesterly and quarterly blocking are smaller, as we would expect.

The return level is a particular quantile of the marginal distribution associated with a particular probability. In the case where data are iid it is easy to calculate this probability since from (7) and (8) we can deduce that

$$(1 - 1/k)^{1/n} = P^{1/n}\{M_n \leq R_{n,k}\} = F(R_{n,k}). \tag{10}$$

so that $R_{n,k} = x_{(1-1/k)^{1/n}}$. The 20-year return level would be the 0.99980 quantile or the 10-semester return level would be the 0.99919 quantile. Alternatively, to calculate the 0.999 quantile we could calculate the 8.2–semester return level or the 15.9–quarter return level.

However, our data are not iid and, instead of the identity (10), we can use our informal definition of extremal index (5) to obtain the asymptotic approximation

$$(1 - 1/k)^{1/(n\theta)} = P^{1/(n\theta)}\{M_n \leq R_{n,k}\} \approx F(R_{n,k}),$$

so that

$$R_{n,k} \approx x_{(1-1/k)^{1/(n\theta)}}. \tag{11}$$

Thus, in the case of stationary series it is less easy to say exactly what the probability is associated with the return level. If we wish to answer this question we need to know, or be able to estimate accurately, the extremal index θ of the series.

If we do not know θ there is some vagueness in the statements we can make about extreme returns. We know the return level $R_{n,k}$ is exceeded in one n–block out of every k, on average. However, we do not know how many exceedances will occur in that particular

extreme n–block or stress period, since this depends on the propensity of the series to form clusters. In this sense, $R_{n,k}$ gives us an idea of the frequency of stress periods but not of the true frequency of extreme returns averaged out over the long term. If we want more precise information we need θ.

Figure 3 is intended to give some more intuition for the problem. We have simulated a new realisation of our ARCH process over an entire century of trading days and marked with horizontal lines point estimates for the 20-year and 100-year return levels derived from the model for annual maxima; the vertical gridlines mark each of the 100 years. The 20-year level of 9.29 is exceeded in exactly 5 years which is in line with our expectation. In some years more than one observation exceeds the level, which is again what we expect, since our ARCH process gives rise to clusters of extremes. Thus, although there are five extreme years, there are eight extreme observations. The 100-year level of 14.26 is exceeded in two years, which is only once more than our average long-run expectation; in those two extreme years there are three extreme observations.
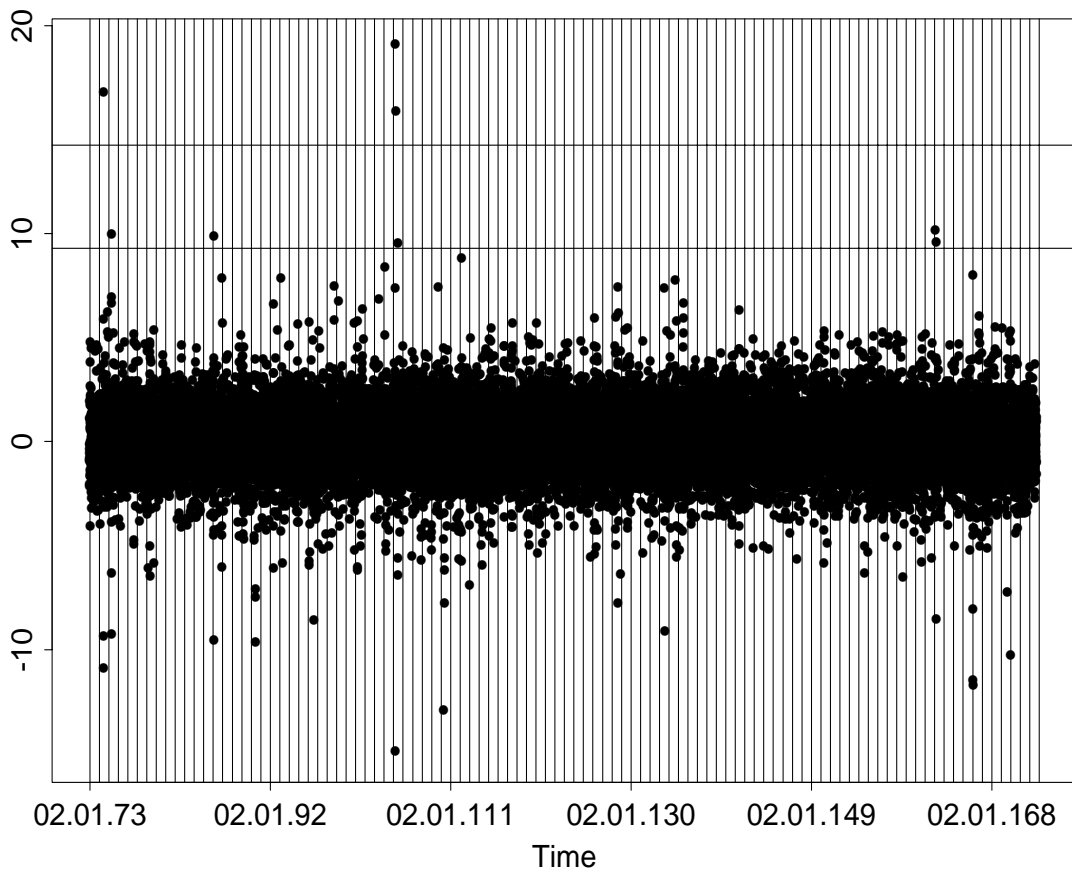

Figure 3: A simulated century of daily ARCH(1) values

3.5 Estimating the Extremal Index

Figure 3 also suggests a rough estimator for the extremal index. We can divide the number of yearly blocks in which the maxima exceed a high level by the total number of exceedances in those blocks. In our figure, using the R_{20} level, our estimate is $5/8 = 0.625$ which can be compared with the true value for the ARCH process of 0.835.

This is a crude example of the application of the so-called blocks method of calculating

the extremal index. The method, which is supported by asymptotic arguments given in Chapter 8 of Embrechts et al. (1997), involves dividing the data into m blocks of observations of length n, as we have done to fit the GEV distribution, and setting high thresholds, at u say. The natural asymptotic estimator of θ is

$$\hat{\theta} = n^{-1} \frac{\log(1 - K_u/m)}{\log(1 - N_u/(mn))} \tag{12}$$

where N_u is the number of exceedances of the threshold and K_u is the number of blocks in which the threshold is exceeded. For K_u/m and $N_u/(mn)$ small, this estimator reduces to K_u/N_u, as applied above.

The statistical properties of such estimators are investigated in Smith & Weissman (1994) as is the problem of choosing n (or m) and u. The asymptotic derivation of (12) suggests we should attempt to keep both m and n large. With this in mind we have applied the method using quarterly and semesterly blocks in Table 3. We choose thresholds such that they are exceeded by between 15 and 200 observations; that is, instead of setting u we set N_u (see Table 3). We have used the estimator as defined in (12) rather than the simplified approximate form K_u/N_u, which can only perform well if we have many more data.

(m, n)	N_u	15	20	25	30	40	50	100	200
quarter	K_u	12	15	18	21	27	34	57	81
(95, 65)	$\hat{\theta}$	0.85	0.81	0.79	0.79	0.79	0.83	0.86	0.89
semester	K_u	11	13	16	18	22	25	39	46
(47, 130)	$\hat{\theta}$	0.84	0.76	0.79	0.76	0.74	0.71	0.83	0.90

Table 3: Estimates of Extremal Index using Blocks Method for simulated ARCH process.

In fact, our estimates are quite close to the true value of 0.835, particularly for quarterly blocks. The mean estimate over the various thresholds is 0.83; for semesterly blocks it is 0.79. With such estimates we can return to the asymptotic relationship (11) and estimate the quantile probability associated with the return period; we shall do this in the next section.

4   Analysis of BMW Log returns

As in the case of the ARCH "returns", we analysed quarterly, semesterly and yearly maxima as shown in Table 4. The fit of the Fréchet model to these block maxima was again investigated using exponential residuals and found to be adequate.

The estimated ξ parameters are 0.24, 0.27 and 0.21 for analyses of quarterly, semesterly and annual maxima respectively. In the analysis of quarterly maxima, which we consider to be most reliable on the basis of our simulation study, our confidence interval allows us to reject the hypothesis of a thin-tailed marginal distribution and the hypothesis of infinite variance. In the analysis of semesterly maxima we reject the thin-tailed hypothesis but cannot formally rule out infinite variance. For yearly maxima the data quantity (24 maxima) is again too meagre to make such statements.

We show estimates of the location and shape parameters μ and σ and various return levels multiplied by 100. For values up to about 10 or 15, these may be interpreted as percentage changes. Consider the analysis of quarterly maxima. A scaled estimate of the 20–quarter return level is 7.6 and a 95% confidence interval is (6.3, 9.9); thus in one quarter block every 5 years we expect to see daily falls of a magnitude of up to 10%. A scaled estimate of the 80–quarter return level is 11.5 with confidence interval (8.9, 17.9),

	Quarter		Semester		Year	
$\approx$ n	65		130		261	
m	95		48		24	
ξ	0.24	(0.08-0.40)	0.27	(0.13-0.52)	0.21	(-0.16-0.58)
$\sigma \times 100$	1.12	(0.11-1.34)	1.37	(0.18-1.74)	1.87	(1.17-2.56)
$\mu \times 100$	2.70	(0.13-2.96)	3.34	(0.22-3.78))	4.36	(3.50-5.22)
$R_{n,20} \times 100$	7.6	(6.3-9.9)	9.6	(7.6-15.0)	12.0	(9.1-24.3)
$R_{20} \times 100$	11.5	(8.9-17.9)	12.0	(8.9-21.7)	12.0	(9.1-24.3)

Table 4: Parameter estimates for GEV model fitted to BMW negative log returns

so that every 20 years on average we expect a stress period (meaning a volatile quarter) where log returns may reach levels as low as -0.179.

It is of interest now to estimate the quantile probabilities associated with such return levels. Thus we attempt to estimate the extremal index θ using quarterly and semesterly blocks as we did for the ARCH data. In Table 5 we see that quarterly blocks yield estimates between 0.51 and 0.59 for an average of 0.55. Semesterly blocks yield estimates between 0.42 and 0.61 for an average of 0.54. Again the former analysis gives more stable estimates, as was the case for the ARCH data; both m and n are reasonably large.

(m, n)	N_u	15	20	25	30	40	50	100	200
quarter	K_u	8	10	13	15	21	25	40	65
(95, 65)	$\hat{\theta}$	0.55	0.52	0.56	0.54	0.59	0.58	0.51	0.53
semester	K_u	8	10	13	14	17	21	28	42
(47, 130)	$\hat{\theta}$	0.58	0.56	0.61	0.56	0.53	0.56	0.42	0.52

Table 5: Estimates of Extremal Index using Blocks Method for BMW negative log returns

We assume that the BMW negative log returns have extremal index 0.55 so that the average cluster size is about 1.8. This means that our 20–quarter return level corresponds roughly to the 0.9986 quantile using the approximate equation (11). Thus, the marginal probability of exceeding the return level is estimated to be 0.0014.

Now we can phrase our results in terms of frequency of both extreme blocks and extreme returns. In five years we expect one stress quarter when daily falls as high as 10% could occur. When falls of this magnitude occur we expect to see clusters of large returns with an average size of 1.8. Moreover, in these five years there are 1305 trading days so that we expect roughly $0.0014 \times 1305 \approx 1.9$ actual days when returns of this magnitude occur.

Conversely, suppose we require an estimate of the 0.999 quantile of the marginal distribution of the negative log returns. Again using (11) and knowledge of θ we judge that this is equivalent to estimating the 28.5 quarter return level. We choose a round number and estimate the 28 quarter return level. Our point estimate is 0.084 and a 95% confidence interval is (0.070, 0.115). A rough check that such values are plausible is provided in Figure 4. Superimposed on the BMW negative log returns we have marked the point estimate of the 0.999 quantile or 28–quarter return level with a solid horizontal line and the asymmetric confidence interval with dotted lines. 6 returns exceed the lower band of the confidence interval; 5 returns exceed the point estimate and 1 return exceeds the upper band. The total length of the BMW return series is 6146 so that we expect about 6.1 returns to exceed the 0.999 quantile.
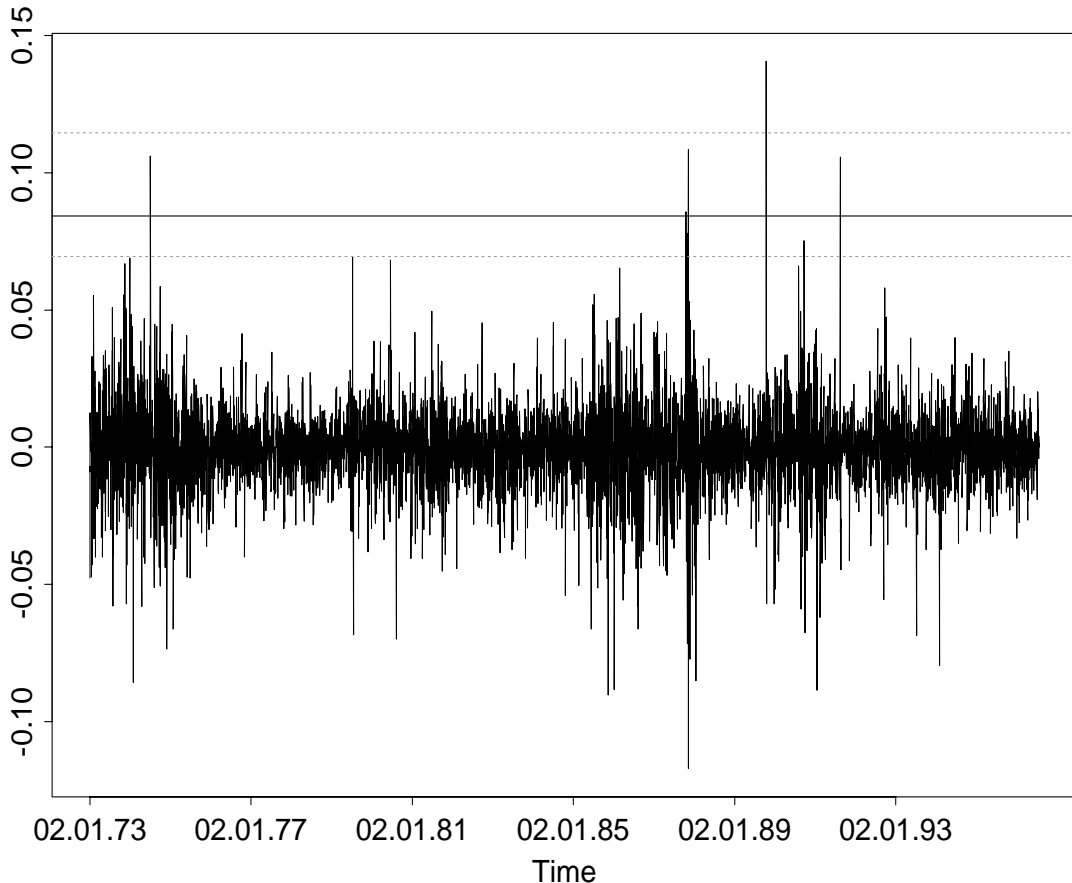
Figure 4: The BMW negative log returns. Solid horizontal line is maximum likelihood estimate of the 0.999 quantile or 28–quarter return level; dotted lines provide confidence interval.

5 Conclusion

We hope to have demonstrated the simplicity of calculating useful estimates of longer term extreme losses using the blocking approach to extreme value theory. This approach does require alot of data as we have seen - over 6000 daily observations to obtain reasonable results for quarterly blocks and more to obtain comparable accuracy with yearly blocks. However, the main advantage is that the problems of dependency which greatly complicate the use of threshold methods are largely circumvented by the use of block maxima methods. We are able to calculate useful risk measures in the form of return levels and to give errors for these estimates.

Depending on the nature of the statements we wish to make, the concepts of the return level and the traditional quantile are both useful. To pass between the concepts we need the extremal index θ. For a full understanding of the unconditional extremal behaviour of a financial return series there are two estimation problems to be addressed. First, we need to know the weight of the tail of the marginal distribution (ξ) and the associated extreme value limit distribution for maxima of the process in large blocks. If we know this we can make statements about the magnitude of returns in stress periods, which are defined in terms of specified time horizons. We can calculate quantities like the 20–quarter return level.

Second, we need to know the propensity of the series to form clusters of large values

(as summarized by θ) so that we can get a proper understanding of how often extreme levels might be hit in such stress periods and what the true long-term frequency of extreme returns might be.

It is important to point out that quantification of these kinds of worst case losses is an inherently difficult problem. We are working in the tail of the marginal distribution of the process and we have only a limited amount of data which can help us. The uncertainty in our analyses is often high, as reflected by large confidence intervals for return levels. However, if we wish to quantify rare events we are better off using the theoretically supported methods of EVT than other ad hoc approaches. EVT represents an honest approach to measuring the uncertainty inherent in the problem.

Appendix

The asymptotic confidence interval for the return level $R_{n,k}$ is calculated using the profile likelihood method.

If we assume the exactness of the GEV distribution for block maxima in blocks of size n we can use (9) to write

$$\mu = R_{n,k} + \frac{\sigma}{\xi}\left(1 - (-\log(1 - 1/k))^{-\xi}\right),$$

and thus we can eliminate μ from the GEV log-likelihood by reparametrizing in terms of ξ, σ and $R_{n,k}$. Suppose we do this to obtain the log-likelihood $L(\xi, \sigma, R_{n,k})$ from m observed block maxima.

Now consider a test of the null hypothesis $R_{n,k} = r$ using an asymptotic likelihood ratio test. Let $L(\hat{\hat{\xi}}_r, \hat{\sigma}_r, r)$ be the maximum log-likelihood under the null hypothesis and let $L(\hat{\xi}, \hat{\sigma}, \hat{R}_{n,k})$ be the unconstrained maximum log-likelihood. Asymptotically (as $m \to \infty$) under the null hypothesis

$$-2\left(L(\hat{\hat{\xi}}_r, \hat{\sigma}_r, r) - L(\hat{\xi}, \hat{\sigma}, \hat{R}_{n,k})\right) \sim \chi_1^2,$$

so that the hypothesis is tested using a chi-squared distribution with one degree of freedom.

An $\alpha\%$ asymptotic confidence interval for $R_{n,k}$ is the set of r for which the null hypothesis is not rejected at the $\alpha\%$ level. That is, the set

$$\left\{r : L(\hat{\hat{\xi}}_r, \hat{\sigma}_r, r) \geq L(\hat{\xi}, \hat{\sigma}, \hat{R}_{n,k}) - 0.5\chi_1^2(\alpha)\right\},$$

where $\chi_1^2(\alpha)$ is the $\alpha\%$ point of the chi-squared distribution on one degree of freedom. The curve $(r, L(\hat{\hat{\xi}}_r, \hat{\sigma}_r, r))$ is known as the profile log-likelihood. In general, this curve is not symmetrical about its maximum, so that confidence intervals are not symmetrical about the maximum likelihood estimate $\hat{R}_{n,k}$. This asymmetry reflects the asymmetry of our uncertainty concerning extreme events. It is easier to bound the confidence interval below than to bound it above.

Software

The analyses in this paper were carried out in S-Plus using the author's own EVIS functions. These functions together with documentation and some datasets are available to S-Plus users over the World Wide Web at `http://www.math.ethz.ch/~mcneil`

Acknowledgement

References

Bollerslev, T., Chou, R. & Kroner, K. (1992), 'ARCH modeling in finance', *Journal of Econometrics* **52**, 5–59.

Cox, D. & Snell, E. (1968), 'A general definition of residuals (with discussion)', *Journal of the Royal Statistical Society, Series B* **30**, 248–275.

Danielsson, J., de Haan, L., Peng, L. & de Vries, C. (1997), Using a bootstrap method to choose the sample fraction in tail index estimation. Erasmus University, Rotterdam, preprint.

Danielsson, J. & de Vries, C. (1997*a*), 'Tail index and quantile estimation with very high frequency data', *Journal of Empirical Finance* **4**, 241–257.

Danielsson, J. & de Vries, C. (1997*b*), 'Value-at-risk and extreme returns', FMG-Discussion Paper NO 273, Financial Markets Group, London School of Economics.

de Haan, L., Resnick, S., Rootzén, H. & de Vries, C. (1989), 'Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes', *Stochastic Processes and their Applications* **32**, 213–224.

Dekkers, A., Einmahl, J. & de Haan, L. (1989), 'A moment estimator for the index of an extreme-value distribution', *Annals of Statistics* **17**, 1833–1855.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997), *Modelling extremal events for insurance and finance*, Springer, Berlin.

Embrechts, P., Resnick, S. & Samorodnitsky, G. (1998), 'Living on the edge', *RISK Magazine* **11**(1), 96–100.

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

Fisher, R. & Tippett, L. (1928), 'Limiting forms of the frequency distribution of the largest or smallest member of a sample', *Proceedings of the Cambridge Philosophical Society* **24**, 180–190.

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

Gumbel, E. (1958), *Statistics of Extremes*, Columbia University Press, New York.

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

Hosking, J. (1985), 'Maximum-likeihood estimation of the parameters of the generalized extreme-value distribution', *Applied Statistics* **34**, 301–310.

Jansen, D. & deVries, C. (1991), 'On the frequency of large stock returns: putting booms and busts into persepctive', *The Review of Economics and Statistics* pp. 18–24.

Kesten, H. (1973), 'Random difference equations and renewal theory for products of random matrices', *Acta Mathematica* **131**, 207–248.

Koedijk, K., Schafgans, M. & deVries, C. (1990), 'The tail index of exchange rate returns', *Jourmal of International Economics* **29**, 93–108.

Leadbetter, M., Lindgren, G. & Rootzén, H. (1983), *Extremes and related properties of random sequences and processes*, Springer, Berlin.

Longin, F. (1996), 'The asymptotic distribution of extreme stock market returns', *Journal of Business* **69**, 383–408.

Longin, F. (1997*a*), Beyond the VaR, Technical Report 97-011, CERESSEC.

Longin, F. (1997*b*), From value at risk to stress testing, the extreme value approach, Technical Report 97-004, CERESSEC.

Loretan, M. & Phillips, P. (1994), 'Testing the covariance stationarity of heavy-tailed time series', *Journal of Empirical Finance* **1**, 211–248.

Lux, T. (1996), 'The stable Paretian hypothesis and the frequency of large returns: an analysis of major German stocks', *Applied Financial Economics* **6**, 463–475.

McNeil, A. (1997), 'Estimating the tails of loss severity distributions using extreme value theory', *ASTIN Bulletin* **27**, 117–137.

McNeil, A.J. amd frey, R. (1998), 'Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach', preprint, ETH Zürich.

Mittnik, S. & Rachev, S. (1996), 'Tail estimation of the stable index α', *Appl. Math. Lett.* **9**, 53–56.

Reiss, R. & Thomas, M. (1997), *Statistical Analysis of Extreme Values*, Birkhäuser, Basel.

Rootzén, H. & Tajvidi, N. (1997), 'Extreme value statistics and wind storm losses: a case study', *Scandinavian Actuarial Journal* pp. 70–94.

Smith, R. (1985), 'Maximum likelihood estimation in a class of nonregular cases', *Biometrika* **72**, 67–92.

Smith, R. & Weissman, I. (1994), 'Estimating the extremal index', *Journal of the Royal Statistical Society, Series B* **56**, 515–528.