# SAFETY FIRST PORTFOLIO SELECTION, EXTREME VALUE THEORY AND LONG RUN ASSET RISKS

Laurens de Haan (Erasmus University Rotterdam)

Dennis W. Jansen (Texas A&M University)

Kees Koedijk (University of Limburg)

Casper G. de Vries (Tinbergen Institute Rotterdam)

**Abstract**

The paper motivates the use of the statistical extreme value theory for the problem of portfolio selection in economics, both theoretically and empirically. It is shown that the conventional safety first criterion developed by Roy can be successfully improved upon by exploiting the fat tail property of asset returns. Extreme value theory is seen to provide a better bound than the Chebyshev bound. In the empirical application we calculate minimum threshold return levels given very low exceedance probabilities for bond and equity investors. A proof of a new quantile estimator is obtained in the appendix. The data cover at least a half-century of returns and allow for evaluation of investment risks in the long run.



*J. Galambos et al. (eds.), Extreme Value Theory and Applications,* 471–487.

## 1. Introduction

An important problem in financial analysis concerns the selection of a portfolio of assets which is optimal given some criterion. The most popular criterion of Markowitz (1952) is based on the tradeoff between mean and variance of the different asset return distributions. For an investor who is primarily concerned about the downside risk of an investment, however, the "safety first" criterion of Roy (1952) is more appropriate. Under the safety first rule an investor prespecifies a low threshold return level and selects the portfolio of assets which minimizes the probability of a return below this threshold or vice versa. An interesting discussion of the relationship between mean-variance analysis and safety first is provided by Bernstein (1992). The safety first rule is related to the mean variance criterion, see e.g. Levy and Sarnat (1972), but explicitly takes into account the probability of a highly negative return. Negative returns imply a loss of wealth and may lead to bankruptcy and even to a financial crisis. Therefore information about this downsided risk is important for economic decision making.

Empirical analysis of financial data such as equity and bond prices over the past 30 years has shown that nominal asset prices approximately follow a martingale, i.e. asset returns are (mean) unpredictable, and the distributions of asset returns tend to be heavily fat tailed. In the analysis below we use this latter fact to improve upon the Chebyshev bound which was used by Roy to make the safety first criterion operational. While the Chebyshev bound only presumes that the mean and variance are known, the other end of the information spectrum is to entertain a specific fat tailed distribution. For example, McCulloch (1981) uses the Cauchy distribution to calculate bankruptcy probabilities of commercial banks. This, however, may be assuming too much specific knowledge about the empirical probability law. An intermediate position is to base oneself on the fact that returns are fat tailed, and to rely on the limit laws of extreme value theory. Further motivation for using extreme value theory in economics is provided in Koedijk et al. (1990) and Hols and de Vries (1991).

In the empirical section we implement the revised safety criterion for two types of investment opportunities: a mutual fund of equities (equity index) and a mutual fund of bonds (bond index). These are the investment opportunities which are usually open to institutional investors like pension funds. Because of their actuarial goals, such investors are especially concerned about the downsided risks of their investments. For given low probabilities, we estimate the associated threshold return levels (quantiles) for the two asset indices upon which the investment decisions may be based. An earlier application to choosing between different equities within the equity index is contained in Jansen and de Vries (1991).

There is evidence, see e.g. Friedman and Laibson (1989) and the much discussed Peso problem, that agents tend to underestimate the probability of a financial crisis. This is not too surprising given the popularity of the mean variance criterion and the widely entertained hypothesis that returns are normally distributed. The index series are suited for addressing this issue. The series cover about a century of US financial history, and include several financial crashes. This enable us to provide fairly reliable estimates of the likelihood of such a financial crisis.

The methods we employ for estimation are based on the methods developed by Dekkers et al. (1989), and heavily rely on the idea that the distribution of the highest order statistics approximately follow the extreme value distribution. Given the fat tail property this involves using the Hill (1975) estimator, which can be interpreted as a moment estimator, see de Haan (1990), to estimate the index of the extreme value distribution. A corollary to the results in Dekkers et al. (1989) A proof of this new corollary is given in the appendix for the case where the index is positive is then used to obtain quantile estimates.

## 2. Economic Theory

By definition the one period rate of return $R_{t+1}$ on an investment in a (risky) asset with current price $P_t$ is

$$R_{t+1} = (P_{t+1} + d_{t+1} - P_t)/P_t.$$

Under continuous compounding (e.g. $d_{t+1} = 0$, and $P_{t+1} = P_t \exp(tQ_{t+1})$, with $t=1$), or as an approximation, the gross return is often expressed as a logarithmic price change

$$\text{(1)} \qquad Q_{t+1} = \log((d_{t+1}+P_{t+1})/P_t).$$

At the time t of the investment decision, $P_{t+1}$ and $d_{t+1}$ are usually unknown, and hence $Q_{t+1}$ is viewed as a random variable. Arbitrage ensures that $Q_{t+1}$ satisfies the fair game property, i.e. $\{Q_t\}$ is a martingale. For example prior knowledge that $E[P_{t+1}] > P_t$ leads to a buying wave, thereby raising $P_t$ and quickly eliminating any riskless profit opportunities. But for certain asset prices the submartingale model is more appropriate due to the natural growth of the economy.

Suppose there exist several different investment opportunities which satisfy $E[Q_{t+1}] = Q$, how does one choose between these alternatives? Without any further assumptions, not much can be said. But suppose the investor is especially concerned about the downside risk of the investment. To meet this concern, Roy (1952) developed a two step portfolio selection procedure in which the investor first specifies a critical threshold return level q below which Q should fall with the smallest possible probability. In the second step the investment is selected which minimizes this probability. Alternatively, a very low probability on excess losses is prespecified by the investor, and the investment opportunity is selected which yields the highest safety level q. This latter procedure will be implemented in the empirical section, but here we will go the other way.

More formally, given some q and n investment alternatives the problem is to

$$(2) \qquad \min_{Q_i} P\{Q_i \leq q\}, \ i - 1, \ldots, n.$$

To solve this problem, Roy (1952) assumed that $E[Q] = \varepsilon$ and $\mathrm{Var}[Q] = \sigma^2$ are finite and can be estimated for the various investment opportunities i. For the sake of robustness no further assumptions about the d.f. of Q were made. In order to capture $P\{Q \leq q\}$, Roy employed the Chebyshev bound

$$(3) \qquad P\{(Q - \varepsilon)^2 \geq (\varepsilon - q)^2\} \leq \frac{\sigma^2}{(\varepsilon - q)^2},$$

for $\varepsilon > q$, and where q can be negative. By the following manipulations

$$P\{(\varepsilon - Q)^2 \geq (\varepsilon - q)^2\} \geq P\{\varepsilon - Q \geq \varepsilon - q\} = P\{q \geq Q\},$$

if follows that

$$(4) \qquad P\{Q \leq q\} \leq \frac{\sigma^2}{(\varepsilon - q)^2}.$$

Instead of solving problem (2), the second step of the safety first criterion proposes to minimize the RHS of eq. (4). This amounts to maximization of $(\varepsilon - q)/\sigma$. (The efficient frontier is the locus generated by portfolios i which for given risk levels $\sigma_i$ generate the highest expected return $\varepsilon_i$. The typical concave shape of the frontier is derived in e.g. Copeland and Weston (1983).) Given a mean variance frontier as drawn in Figure 1, the portfolio is selected which lies on the straight line which is tangent to the mean variance frontier and has intercept q. To show this, note that all lines through (0,q) can be parameterized as follows:

$$(5) \qquad \varepsilon = q + \frac{\bar{\varepsilon}-q}{\bar{\sigma}}\sigma .$$

Hence the inverse of the RHS of eq. (4) is maximized if $(\bar{\sigma}, \bar{\varepsilon})$ are chosen on the frontier, i.e. by rotating the lines (5) around (0,q) upwards. Note that if there exists an asset with risk free return $\bar{q}$, then choosing $q = \bar{q}$ generates the same unlevered choice of portfolio under the mean-variance criterion and the safety first criterion; see Levy and Sarnat (1972). Moreover, if $q < \bar{q}$ no investment in the risky asset ensues under the safety first criterion.

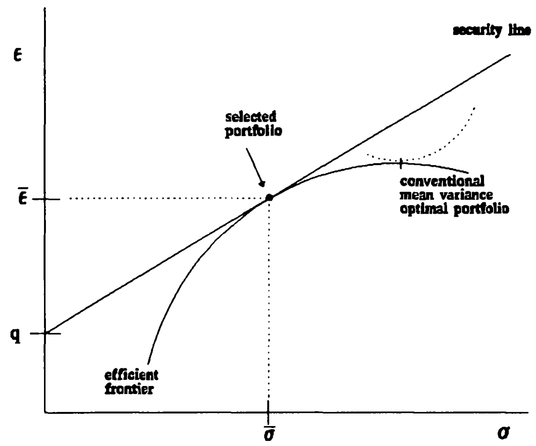


**Figure 1: Safety First Criterion**

Instead of the one period analysis, it is of interest to consider the multiperiod portfolio selection problem. For example, transactions costs preclude a continuous adjustment of the portfolios, and multiperiod investments better reflect the likelihood of a financial crisis. To develop the multiperiod safety first criterion we employ the multivariate Chebyshev bound, cf. Roy (1952). Let A be a positive semidefinite n x n matrix, and let Q be a n x 1 vector of r.v's such that $E[Q^T A Q]$ exists. Then

$$(6) \qquad P\{Q^T A Q \geq q^2\} \leq \frac{\operatorname{tr} A\, E[Q Q^T]}{q^2}.$$

This follows directly from the univariate bound by realizing that $Q^T A Q$ is just a univariate r.v. and the fact that $E[Q^T A Q] = \operatorname{tr} A\, E[Q Q^T]$.

To further the analysis it is helpful to exploit the characteristic properties of asset returns. Firstly, recall that the net returns $(Q-\varepsilon)$ satisfy the fair game property so that all covariances are zero, and hence the off-diagonal elements in the A-matrix are zero after correction for the mean (the natural growth rate of the economy). For example, this is the case for the popular ARCH process, see Baillie and McMahon (1989).

Secondly, returns are additive, i.e. the n-period return is just the sum of the n single period returns. Thus the diagonal elements are all 1, and hence A = I the identity matrix. Hence, in analogy with e.g. (3) we get

$$\text{(7)} \qquad \frac{n\,\sigma^2}{(\varepsilon - q)^2} \geq P\left\{\sum_{t=1}^{n} (Q_t - \varepsilon)^2 \geq (\varepsilon - q)^2\right\}.$$

Let $m_a = \min_t\{Q_1,\dots,Q_n\} = -\max_t\{-Q_1,\dots,-Q_n\}$ be the minimum, and manipulate the RHS of eq. (7):

$$\begin{aligned} P\left\{\sum_{t=1}^{n} (\varepsilon - Q_t)^2 \geq (\varepsilon - q)^2\right\} &\geq P\left\{\max_t(\varepsilon - Q_t)^2 \geq (\varepsilon - q)^2\right\} \\ &\geq P\left\{\max_t(\varepsilon - Q_t) \geq \varepsilon - q\right\} \\ &= P\left\{-\max_t(-Q_t) \leq q\right\} \\ &= P\left\{\min_t(Q_t) \leq q\right\}. \end{aligned}$$

Hence, we have established the multiperiod analogue of eq. (4):

$$\text{(8)} \qquad P\{m_n \leq q\} \leq \frac{n\,\sigma^2}{(\varepsilon - q)^2}.$$

The multiperiod safety first criterion involves the following two steps. First establish a quantile level $q < \varepsilon$ and then maximize $(\varepsilon - q)^2/n\,\sigma^2$. Note that this is tantamount to the one period problem as the RHS of (4) and (8) differ only by the scale factor n.

Suppose the r.v.'s $Q_t$ are i.i.d. with common d.f. F(q). Then the exact bound for eq. (8) is well know from the theory of order statistics:

$$\text{(9)} \qquad P\{m_n \leq q\} = 1 - [1 - F(q)]^n.$$

One wonders whether something in between (8) and (9) can be obtained by using some information about the d.f. of $Q_t$ and still retain some of the robustness associated with the Chebyshev bound. As is well known, stock returns do exhibit the fat tail property, and hence it seems reasonable to assume the $Q_t$ vary regularly at infinity, see e.g. Jansen and de Vries (1991). In fact this is a sufficient condition for the limit law

$$\text{(10)} \qquad P\{b_n m_n \leq q\} \to 1 - \exp(-(-q)^{-\alpha}),\ q < 0, \text{ as } n \to \infty,$$

to apply. In (10) $b_n > 0$ is a scaling constant dependent on n, $\alpha > 0$ is called the tail index and the convergence is weak. The tail index characterizes the tail fatness as it is 1 - 1 with the number of moments that exists for $Q_t$. E.g. $\alpha$ equals the degrees of freedom if F(q) is a Student-t d.f.

Below we provide a numerical example which shows that eq. (10) may constitute a considerable improvement over eq. (8). Apart from this gain at least one other potential advantage needs to be mentioned. Several empirical studies indicate that only the first few moments of $Q_t$ are well defined. Eq. (10) has the advantage that it holds regardless of the number of moments, except if all moments exist. But the latter would violate the fat tail property. Eq. (8) holds as long as the mean and variance exist. Thus for the Cauchy d.f., which has no moments, eq. (10) applies with $\alpha = 1$ and $b_n = \pi/n$, while eq. (8) cannot be used.

Our example is based on the "negative" Pareto d.f. $F(x) = (-x)^{-3}$, for $x \leq -1$. The density reads $f(x) = 3(-x)^{-4}$ and the first two moments are E[X] = -3/2, Var[X] = 3/4, and no other (integer) moments do exist. For this case the exact expression for the distribution of the minimum is found from eq. (9) as $P\{m_n \leq x\} = P\{m_n \leq x\} = 1 - (1 - (-x)^{-3})^n$. From Leadbetter et al. (1983, Ch. 1) in (10) the scale parameter is $b_n = n^{-1/3}$ and tail index reads $\alpha$=3 (corroborating the fat tail property). Rewriting eq. (10) gives $P\{m_n \leq x\} \to 1 - \exp(-n(-x)^{-3})$. The Chebyshev bound for the Pareto d.f. is found from eq. (8) as $P\{m_n \leq x\} \leq 3/4\ n\ (-3/2 - x)^{-2}$. For a sample with size n = 100 and for different quantiles x the probabilities are presented in Table 1.

**Table 1: $P\{m_n \leq x\}$**

| quantile | x = -10 | x = -50 | x = -100 |
|---|---|---|---|
| exact | .09521 | .0008 | .0001 |
| extreme value | .09516 | .0008 | .0001 |
| Chebyshev bound | 1.0000 | .0319 | .0077 |

This table gives the probability that the minimum from a sample of size n = 100 from the "negative" Pareto d.f. is below the quantile x.

Note that the in-sample probabilities can always be readily estimated from the empirical d.f., due to the M.S.E. consistency of this procedure. As this can be done for all $P \geq 1/n = 0.1$, the two columns with x = -50, -100 are more informative. The example reveals the extreme value approach constitutes a considerable improvement over the Chebyshev bound for intermediate values of x. The differences can be studied analytically by taking: $-\log[1 - P\{m_n \leq x\}]$ and letting both x and n get large. For the exact and the extreme value approach this gives both times $n(-x)^{-3}$, while the Chebyshev bound leads to $3/4n(-3/2 - x)^{-2} \approx 3/4nx^{-2}$. Thus while both approximations are of the order n, they differ by a scale factor $-4/(3x)$. Also note that the Chebyshev bound (8) is uninformative for $x = -10$. For small $|x|$, and assuming independence, a better Chebyshev bound is $1 - [1 - \sigma^2(\varepsilon-x)^{-2}]^n$, which yields for $x = -10$: $P = 0.6478$, instead of the 1.0000 entry in the table. For $x = -100$, the two approaches are indistinguishable, i.e. (8) is the first order Taylor expansion to the latter bound. A final word on the use of the normal distribution commonly employed in the finance literature. If the above mean and variance would have been used in calculating the P-values on basis of the normal d.f., then for $x = -10$, one already finds $P < 10^{-22}$, which is a gross underestimation of the true $10^{-1}$ value.

## 3. Empirical Results

The improved safety first criterion is now implemented for a two point investment problem. Either all funds are allocated towards investing in a mutual bond fund, or in a mutual equity fund. To this end the investor is asked to specify a maximal acceptable downsided risk level (P-value) for which the threshold return level q is to be maximized by choosing the appropriate investment. (In Jansen and de Vries (1991) some results for the inverse problem, choosing q and minimizing p, are presented.) To implement this choice process we need estimates of the threshold level $q_p$, given a certain p-value. The estimation procedure is based on the limit law in (10), and consists of 2 steps. First the tail index $\alpha$ is calculated by means of the now popular moment estimator of Hill, as in de Haan (1990):

$$\widehat{1/\alpha} = \frac{1}{m} \sum_{i=1}^{m} [\log(X_{(n+i-t)}/X_{(n-m)}], \tag{10}$$

where $(\widehat{1/\alpha} - 1/\alpha)\sqrt{m}$ is asymptotically normal $N(0,1/\alpha^2)$, and where the $X_{(i)}$ is the i-th descending order statistic. Subsequently the quantiles $q_p$ are estimated on the basis of a result presented in the appendix:

$$\hat{q}_p = X_{(n-m)} \left[\frac{m}{pn}\right]^{\widehat{1/\alpha}}, \tag{11}$$

and where

$$\frac{\sqrt{m}}{\log(\frac{m}{pn})}\left[\frac{\hat{q}_p}{q_p} - 1\right] \tag{12}$$

is asymptotically normal $N(0, 1/\alpha^2)$. In both steps we need to choose the number of lowest order statistics m. We follow the bootstrap procedure suggested by Hall (1990) in choosing m, as this method is less ad hoc than other procedures and may correct for possible bias.

Turning now to the analysis of the data, we consider the problem of choosing between investing in a mutual fund of bonds or a mutual fund of stocks. To this end we employ a 60-year US bond index and William Schwert's 130-year US-stock index. This stock index has been analyzed before by Schwert (1989), Pagan and Schwert (1990), and Loretan and Phillips (1992). These papers show that the stock index returns are clearly heteroskedastic, i.e. volatility is related to the phase of the business cycle.

Interestingly, Jansen and de Vries (1991), and Loretan and Phillips (1992) also report tail index estimates $\hat{\alpha}$ for the upper and lower tails of the stock index series and for subperiods, in these articles the number of order statistics m is not selected on basis of Hall's (1990) bootstrap method. Equality of $\alpha$ across tails and over time cannot be rejected. The stock index estimates for $1/\hat{\alpha}$ reported in Table 2 compare to the previous results, and equality of the tail indexes can again not be rejected at the 5% level. In addition to the stock index, we now also have another important wealth index available. In contradistinction to the stock index, the point estimate $\hat{\alpha}$ for the lower tail of the bond index is much lower than the $\hat{\alpha}$ for the upper tail, albeit not significantly. The upper tail parameter of the stock and bond indexes are comparable, but the lower tail $\hat{\alpha}$'s do differ (again, however, not significantly). The height of the lower tail bond index parameter points towards the, not completely surprising, limited downsided risk of a bond investment.

This brings us to the investment decision problem. Suppose we measure the probabilities on excessively low returns, i.e. the risk factor, along the x-axis; and the threshold low return levels (quantiles), i.e. the return factor, along the y-axis. The risk-return properties of the bond and equity investments are then plotted in this (x,y)

**Table 2: Tail Indices and Quantile Estimates**

| Asset and sample period | M,n | 1/α | Quantile Estimates* | | |
|---|---|---|---|---|---|
| | | | 1/n | 1/(1.5n) | 1/(2n) |
| Stocks 1854-1987 lower tail | 40, 1608 | .3646 (± .1103) | -.3874 (-.2743, -.6589) | -.4491 (-.3083, -.8267) | -.4987 (-.3350, -.9750) |
| Stocks 1854-1987 upper tail | 27, 1608 | .2680 (± .1011) | .2723 (.2042, .4083) | .3035 (.2209, .4850) | .3279 (.2337, .5493) |
| Stocks 1926-1987 lower tail | 17, 744 | .4047 (± .1924) | -.3677 (-.2380, -.8082) | -.4333 (-.2670, -1.149) | -.4868 (-.2901, -1.5134) |
| Stocks 1926-1987 upper tail | 58, 744 | .3769 (± .0970) | .3060 (.2196, .5050) | .3566 (.2488, .6292) | .3974 (.2720, .7376) |
| Govt. Bonds 1926-1987 lower tail | 12, 744 | .1674 (± .0952) | -.0723 (-.0585, -.0947) | -.0774 (-.0607, -.1068) | -.0813 (-.0624, -.1165) |
| Govt. Bonds 1926-1987 upper tail | 17, 744 | .3503 (± .1665) | .1459 (.0991, .2762) | .1682 (.1092, .3650) | .1860 (.1172, .4506) |
| Corp. Bonds 1926 - 1991 lower tail | 22, 792 | .3878 (± .1620) | -.1061 (-.0707, -.2126) | -.1242 (-.0793, -.2865) | -.1388 (-.0861, -.3589) |
| Corp. Bonds 1926-1991 upper tail | 25, 792 | .2982 (± .1246) | .1324 (.0962, .2123) | .1494 (.1049, .2592) | .1628 (.1117, .3000) |

*1/n, t/(1.5n), 1/(2n) are the imputed exceedence probabilities, p.

space. The investment decision is then to select the fund which for given risk (P-level) generates the highest return (q-level), or vice versa to select the investment which generates the lowest P given a certain q-level. Looking at the quantile estimates in Table 2, we see that with each small P-level (note that n is the sample size) the government bond investment generates much higher threshold return levels than the stock investment. The corporate bonds fall between the U.S. government bonds and stocks. Hence, on the basis of the revised safety first criterion the government bond investment seems preferable. On the other hand, risk preferring agents might go for the stock investment. Because, even though the tail parameters of the upper tail bond and the stock index are about equal, the highest quantiles which are rarely exceeded (1-P) are much higher for the stock investment. An interesting feature of the upper tail is that corporate bonds have (marginally) lower estimated quantiles than government bonds. Thus, corporate and government bonds seem to have about the same chance of giving large returns, but government bonds have less chance of giving large (in absolute value) negative returns.

These long run time series can also be used to evaluate the likelihood of a financial crisis, as these indices represent major parts of financial wealth. For example, from the first row of the table we see that about once every 250 years (1/2n) there will be a month in which stocks fall by about 40% (≈-.5 on the log scale). Even though the accompanying fall in bonds would be less, such a loss of wealth would probably constitute a financial crisis (however defined). But the rareness of the event, see Friedman and Laibson (1989), makes that people tend to discount the possibility of a crisis too much.

As a word of caution, the quantile estimates for the longer and shorter period stock index are somewhat discomforting. At the lower tail we do find -.4987 for P = 1/2n and n = 1608, while for the shorter sample n = 744, 1/2n has a quantile estimate of -.4868. These are virtually the same, but should be further apart. Something we hope to investigate in the future.

## Appendix

Let $U_1$, $U_2$,..., be i.i.d. random variables with a uniform (0,1) distribution. Let

$\{U_{k,n}\}^{n}_{k=1}$ be the n-th order statistics. Then the stochastic process

$$\sqrt{k}\left\{\frac{n}{k}U_{[kt],n}-t\right\}$$

converges in $D(0,\infty)$ to Brownian motion $W(t)$ (cf. for example Einmahl [1992]). Hence for i.i.d. random variables $Y_i$ defined by

$$Y_i := \frac{1}{U_i}(i=1,2,...)$$

(hence $P\{Y_i > x\} = 1/x$ for $x > 1$) with n order statistics $\{Y_{k,n}\}^{n}_{k=1}$ we have

$$t^2\sqrt{k}\left\{\frac{k}{n}Y_{n-[kt],n}-t^{-1}\right\}\rightarrow -W(t) \tag{1}$$

in $D(0,\infty)$. Finally consider i.i.d. random variables $X_1$, $X_2$, ... with distribution function F and let $\{X_{k,n}\}^{n}_{k=1}$ be the n-th order statistics. Define

$$U := \left(\frac{1}{1-F}\right)^{\leftarrow}$$

and suppose

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)} = x^{-\alpha}$$

for all $x>0$ and some parameter $\alpha>0$. Then

$$\lim_{t\to\infty}\frac{U(tx)}{U(t)} = x^{1/\alpha}$$

for $x>0$. Write

$$\gamma := \frac{1}{\alpha}.$$

We assume a second order refinement of this limit relation, i.e. suppose there exists a positive function A such that

$$\lim_{t\to\infty}\left\{\frac{U(tx)}{U(t)} - x^{\gamma}\right\} /A(t) = \pm x^{\gamma}\log x \tag{2}$$

for $x>0$.

**Lemma 1** Suppose (2) and let the sequence of integers $k = k(n)$ satisfy: $\lim_{n\to\infty} k(n) = \infty$ and

$$\lim_{n\to\infty}\sqrt{k}\,A\left(\frac{n}{k}\right) = 0. \tag{3}$$

Then the stochastic process

$$t^{\gamma+1}\sqrt{k}\left\{\frac{X_{n-[kt],n}}{U\left(\frac{n}{k}\right)} - t^{-\gamma}\right\} \to -\gamma W(t) \tag{4}$$

in $D(0,\infty)$.

**Proof** For $n \to \infty$ (cf. (1))

$$t^{\gamma+1}\sqrt{k}\left\{\left(\frac{k}{n}Y_{n-[kt],n}\right)^{-\gamma}\right\} - t^{-\gamma} \sim \sqrt{k}(-\gamma)\left\{\frac{k}{n}Y_{n-[kt],n} - t^{-1}\right\} \to -\gamma W(t) \tag{5}$$

in $D(0,\infty)$. One can write $X_i = U(Y_i)$; $i = 1, 2, \ldots$ and hence

$$t^{\gamma+1}\sqrt{k}\left[\frac{X_{n-[kt],n}}{U\left(\frac{n}{k}\right)} - \left(\frac{k}{n}Y_{n-[kt],n}\right)^{-\gamma}\right] = \tag{6}$$

$$t^{\gamma+1}\sqrt{k}\,A\left(\frac{n}{k}\right)\left\{\frac{U\left(\frac{n}{k}\left(\frac{k}{n}Y_{n-[kt],n}\right)\right)}{U\left(\frac{n}{k}\right)} - \left(\frac{k}{n}Y_{n-[kt],n}\right)^{-\gamma}\right\}, /A\left(\frac{n}{k}\right) \to 0$$

in $D(0,\infty)$ by (2) and (3). The result follows by combining (5) and (6).

**Corollary 1**

$$t\sqrt{k}\left\{\log X_{n-[kt],n} - \log U\left(\frac{n}{k}\right) + \gamma \log t\right\} \to -\gamma W(t)$$

in $D(0,\infty)$.

**Corollary 2**

$$t\sqrt{k}\{\log X_{n-[kt],n} - \log X_{n-k,n} + \gamma \log t\} \to -\gamma\ \{W(t) - tW(1)\}$$

in $D(0,\infty)$.

Next define for a fixed sequence $k = k(n)$

$$\hat{\gamma}\ := \frac{1}{k}\sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n}$$

**Corollary 3** Under conditions (2) and (3)

$$\sqrt{k}\,(\hat{\gamma} - \gamma) \to -\gamma\left\{\int_0^1 \frac{w(t)}{t}dt - W(1)\right\}$$

in distribution.

**Sketch of Proof**

Note that

$$(\hat{\gamma} - \gamma) = \int_0^1 (\log X_{n-[kt],n} - \log X_{n-k,n} + \gamma \log t)dt.$$

so we have to prove

$$\int_0^1 (\log X_{n-[kt],n} - \log X_{n-k,n} + \gamma \log t)dt \to \int_0^1 -\gamma \left\{ \frac{W(t)}{t} - W(1) \right\} dt \tag{7}$$

in distribution. Clearly (2) implies

$$\int_\epsilon^1 (\log X_{n-[kt],n} - \log X_{n-k,n} + \gamma \log t)dt \to \int_\epsilon^1 -\gamma \left\{ \frac{W(t)}{t} - W(1) \right\} dt$$

in distribution for each $\epsilon > 0$. A result of Einmahl (1992) on the behaviour of log $X_{n-[kt],n}$ - log $X_{n-k,n}$ + $\gamma$log t near zero permits us to prove (7) itself.

We assemble the results obtained so far.

**Theorem 1** Under conditions (2) and (3)

$$\sqrt{k}\left( \frac{X_{n-k,n}}{U\left(\frac{n}{k}\right)} - 1, \hat{\gamma} - \gamma \right) \to -\gamma \left( \int_0^1 \frac{W(t)}{t} dt - W(1), W(1) \right) \tag{8}$$

in distribution.

Next we consider a sequence of probabilities $p_n$ as in the paper with

$$\lim_{n\to\infty} np_n = 0$$

and define for n = 1, 2, ...

$$a_n := \frac{k}{np_n}.$$

We shall require a sharpening of conditions (2) related to the sequence $\{a_n\}$:

$$\lim_{n\to\infty} \left\{ \frac{U\left(\frac{n}{k}a_n\right)}{U\left(\frac{n}{k}\right)} - a_n^\gamma \right\} / \left\{ A\left(\frac{n}{k}\right) a_n^\gamma \log a_n \right\} = 1. \tag{9}$$

Write

$$\hat{q}_p := X_{n-k,n} \left( \frac{k}{np} \right)^{\gamma_n}$$

and

$$q_p := U\left(\frac{1}{p}\right)$$

with $p = p_n$.

**Theorem 2** Under conditions (9) and (3) the sequence of random variables

$$\frac{\sqrt{k}\left(\hat{q}_p - q_p\right)}{X_{n-k,n}\, a_n^{\gamma} \log a_n}$$

is asymptotically normal with mean zero and variance $\gamma^2$, provided

$$\lim_{t\to\infty} (\log a_n)/\sqrt{k} = 0. \tag{10}$$

**Proof**

$$\frac{\sqrt{k}\left(\hat{q}_p - q_p\right)}{X_{n-k,n}\, a_n^{\gamma} \log a_n} = \frac{\sqrt{k}\left\{a_n^{\hat{\gamma}} - a_n^{\gamma}\right\}}{a_n^{\gamma}\log a_n} + \frac{\sqrt{k}}{\log a_n}\,\frac{U\left(\frac{n}{k}\right)}{X_{n-k,n}}\left\{\frac{X_{n-k,n}}{U\left(\frac{n}{k}\right)} - 1\right\}$$

$$- \frac{\sqrt{k}}{a_n^{\gamma}\log a_n}\left\{\frac{U\left(\frac{n}{k}a_n\right)}{U\left(\frac{n}{k}\right)} - a_n^{\gamma}\right\}\frac{U\left(\frac{n}{k}\right)}{X_{n-k,n}}.$$

The third part goes to zero by (9) and (8). The second part goes to zero by (8) since $a_n \to \infty (n \to \infty)$. The first part equals

$$\frac{\sqrt{k}}{\log a_n}\left(e^{(\hat{\gamma}_n - \gamma)\log a_n} - 1\right) = \frac{\sqrt{k}}{\log a_n}\left\{e^{(\Gamma_n \log a_n)/\sqrt{k}} - 1\right\}.$$

Since $\Gamma_n$ is asymptotically normal, this expression is asymptotically equivalent (cf.(10)) to $\frac{\sqrt{k}}{\log a_n}\,\frac{\Gamma_n}{\log a_n}\sqrt{k} = \Gamma_n$. The result follows since

$$\operatorname{var}\left(\int_0^1 \frac{W(t)}{t}\,dt - W(1)\right) = 1.$$

## REFERENCES

Baillie R. and McMahon P. (1989) The Foreign Exchange Market, Theory and Econometric Evidence, Cambridge Univerity Press, Cambridge.

Bernstein P. (1992) Capital Ideas, Free Press, New York, NY.

Copeland T.E. and Weston J.F. (1983) Financial Theory and Corporate Policy, Addison-Wesley, Reading.

Dekkers A.L.M., Einmahl J.H.J. and de Haan L. (1989) A moment estimator for the index of an extreme value distribution, The Annals of Statistics, 1833-1855.

Einmahl J.H.J. (1992) The a.s. behaviour of the weighted empirical process and the LIL for the weighted tail empirical process, Annals Probab., 20, 781-695.

Friedman B.M. and Laibson D.I. (1989) Economic implications of extraordinary movements in stock returns, Brookings Papers on Economic Activity, 2, 137-172.

Haan L., de (1990) Fighting the arch-enemy with mathematics, Statistica Neerlandica, 45-68.

Hall P. (1990) Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems, Journal of Multivariate Analysis, 177-203.

Hill B.M. (1975) A simple general approach to inference about the tail of a distribution, The Annals of Statistics, 1163-1173.

Hols M., de Vries C.A.B. and C.G. (1991) The limiting distribution of extremal exchange rate returns, Journal of Applied Econometrics, 287-302.

Jansen D. and de Vries C.G. (1991) On the frequency of large stock returns: putting booms and busts into perspective, The Review of Econ. and Stat., 18-24.

Koedijk K.G., Schafgans M.M.A. and de Vries C.G. (1990) The tail index of exchange rate returns, Journal of International Economics, 93-108.

Leadbetter M.R., Lindgren G. and Rootzen H. (1983) Extremes and Related Properties of Random Sequences and Processes, Springer-Verlag, Berlin.

Levy H. and Sarnat M. (1972) Safety first--an expected utility principle, Journal of Financial and Quantitative Analysis, 1829-1834.

Loretan M. and Phillips P.C.B. (1992) Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets, SSRI Working Paper 9208, University of Wisconsin-Madison.

Markowitz H.M. (1959) Portfolio Selection, Wiley, New York.

McCulloch J.H. (1981) Interest rate risk and capital adequacy for traditional bank and financial intermediaries, in S.Y. Maisel (ed.), Risk and Capital Adequacy in Commercial Banks, University of Chicago Press, Chicago, 223-248.

Pagan A.R. and Schwert G.W. (1990) Alternative models for conditional stock volatility, Journal of Econometrics, 267-290.

Roy A.D. (1952) Safety first and the holding of assets, Econometrica, 431-449.

Schwert W.G. (1989) Business cycles, financial crisis and stock volatility, Carnegie-Rochester Conference Series on Public Policy, 83-126.

Schwert W.G. (1990) Indexes of U.S. stock prices from 1802 to 1987, Journal of Business, 399-427.

Professor L. de Haan; Erasmus University Rotterdam; 3000 DR Rotterdam PB 1738; The Netherlands

Professor D.W. Jansen; Texas A&M University; Department of Economics; College Station, TX 77845

Professor K.G. Koedijk; University of Limburg; P.O. Box 616; Maastricht 6200 MD; The Netherlands

Professor C.G. de Vries; Tinbergen Institute; Oostmaaslaan 950; 3063 DM Rotterdam; 3063 DM; The Netherlands