Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 700 – 713
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Identifying multiple outliers in heavy-tailed distributions with an application to market crashes ☆

Christian Schluter [a,1], Mark Trede [b,*]

[a] *Department of Economics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK*
[b] *Institut für Ökonometrie und Wirtschaftsstatistik, Universität Münster, Am Stadtgraben 9, 48143 Münster, Germany*

Received 12 February 2007; received in revised form 17 October 2007; accepted 23 October 2007
Available online 12 November 2007

Abstract

Heavy-tailed distributions, such as the distribution of stock returns, are prone to generate large values. This renders difficult the detection of outliers. We propose a new outward testing procedure to identify multiple outliers in these distributions. A major virtue of the test is its simplicity. The performance of the test is investigated in several simulation studies. As a substantive empirical contribution we apply the test to Dow Jones Industrial Average return data and find that the Black Monday market crash was not a structurally unusual event.

JEL classification: C12; G10; C13; C14; G11
Keywords: Outliers; Outward testing; Masking

1. Introduction

Many distributions in economics are 'heavy' or 'long' tailed such as the distributions of earnings or stock returns. It is well known that such distributions are prone to generate extreme order statistics (such as the sample maximum) which are very large relative to the central body of the data. It is convenient at this stage to distinguish between two types of very large sample data. We refer to 'extremes' as large realizations of random variables which are generated by the population distribution of interest. By contrast, we refer to 'outliers' as large sample values which do not belong to the population of interest. The nature of such outliers depends on the context. In a standard cross-sectional setting in e.g. labour economics, outliers can be interpreted as measurement error. Survey data is rarely free of measurement error, this is an inevitable consequence of the tensions between the conflicting desires for accurate measurement and

☆ Many thanks to Kees Jan van Garderen, Johan Segers, Richard Smith, Maria-Pia Victoria-Feser and to our referees, whose constructive comments have helped to improve the paper.
* Corresponding author. Tel.: +49 251 83 25006; fax: +49 251 83 22012.
E-mail addresses: C.Schluter@soton.ac.uk (C. Schluter), mark.trede@uni-muenster.de (M. Trede).
URL: http://www.economics.soton.ac.uk/staff/schluter/ (C. Schluter).
[1] Tel.: +44 2380 59 5909; fax: +44 2380 59 3858.

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2007.10.003

large sample sizes in the face of limited survey budgets. Such contaminations might occur from misreporting, such as of annual for monthly income, or the presence of data miscoded by the data transcriber such as the classic decimal-point error. In a financial time-series setting, outliers may be the consequence of structurally unusual events such as a market crash. This paper is about the identification of outliers.

Identification of outliers is important in applications which focus on the tails of the distributions. A prominent class of examples is the measurement of risk since the tail of the distribution informs about how wrong things can go wrong if they go wrong. In finance, Value-at-Risk (VaR) has become, following the 1995 Basle Committee agreement, the standard risk measure adopted by financial institutions to define the risk exposure of a financial position, and thus determines capital requirements. Another measure focussing on the tail of the return distribution is the expected shortfall, being the expected return conditional on the return not exceeding the value-at-risk. Extreme market movements do occur from time to time for data described by heavy tailed distributions, and many distributions in finance exhibit heavy tails with tail indices estimated around 3. More precisely, Gonzalo and Olmo (2004) report estimates for log return data for Dax, Ibex, Nikkei and Dow Jones in the range from 2.4 and 3.2. Huisman, Koedijk, Kool, and Palm (2001) study exchange rate data and provide estimates ranging from 3.1 to 8.2. Quintos, Fan, and Phillips (2001) report estimates between 1.5 and 3 for Asian stock price indices. Danielsson and de Vries (1997) report estimates around 3 for the Olson data of forex exchange rate contracts.

Besides finance, heavy tailed distributions play an important role in other fields as well. A classical application is fitting the Pareto distribution to the upper tail of income or wealth distributions (Pareto, 1965; Klass et al., 2006). Piketty and Saez (2003) investigate the income shares of top quantile groups over time. Schluter and Trede (2002) establish statistical inference methods for the tails of Lorenz curves when the income distribution is heavy tailed. This empirical literature takes it invariably for granted that there are no outliers in the data.

Our test procedure, developed for heavy tailed distributions which decay like power functions, has the major virtue of being simple to implement. It is based on the ratio of successive extreme order statistics for which we derive the joint distribution in closed form. As multiple outliers may mask their presence, ours is an outward testing procedure: we start at the kth largest order statistic, test the ratio of this and the previous order statistic, and proceed to the next order statistic up the sample maximum unless an outlier is detected. Critical values for the test are easily calculated.

A substantive empirical contribution of this paper is the application of our test to the problem of market crashes: was Black Monday a structurally unusual event or just an extreme event consonant with the presumed return distribution?

Our contribution differs from recent papers in the literature by explicitly focussing on the identification of outliers rather than considering exclusively extremes. But by seeking to distinguish outliers from extremes we build, of course, on this literature. The study of extremes is the subject of extreme value theory. Given the sparseness of sample data for tails, the extreme quantiles of a distribution are approximated via tail indices, such as the well-known Hill estimator. The generic problem is to determine where the tail begins, i.e. which order statistics, viewed as exceedances over a threshold, to include in the estimation. Recently, Gonzalo and Olmo (2004) have proposed a bootstrap method to determine statistically “which extreme values are really extreme.” In contrast to Gonzalo and Olmo (2004) we do not seek to identify extreme values that can be described by a tail distribution, but rather to identify extreme values that do not belong to them. Another approach is taken by Huisman et al. (2001) who determine the threshold by resolving the mean-variance trade-off inherent in the Hill estimator. We apply their proposed procedure to a robust version of the Hill estimator.

For a comprehensive review of extreme value theory see e.g. Embrechts et al. (1997) or Beirlant et al. (2004). Our test statistics are ratios of extreme order statistics. Galambos (1978, 1984) reviews some relevant results for extreme order statistics relating to differences and ratio. In particular, the differences of iid exponential variates X are independent, and the marginal distributions are themselves exponential. It follows, using the transform $Z=\exp(X)$, that Z has a Pareto distribution, and that the ratios of its order statistics are independent with marginal distributions which are themselves Pareto. We show that similar results hold for distributions satisfying a more general domain-of-attraction assumption. The classic categorisation of tail behaviour, due to Schuster (1984) is based on extreme spacings, i.e. the difference between extreme order statistics. In particular long or heavy tailed distributions are prone to lead to samples which ‘often’ have extremes, and their extreme spacings diverge with sample size. Moreover, extreme spacings of Frechet variables are dependent. We therefore study the ratios which we show to converge and to be independent. The categorisation is further elucidated in e.g. Freimer, Muholkar, Kollia, and Lin (1989) by considering the rate of convergence of extreme spacings for specific short and medium tailed distributions via approximations to their quantile functions. The use of the gap between the largest and second largest order statistic is also discussed in Vandewalle, Beirlant, and Hubert (2004), who also propose a robust estimator for the tail index.

The economics literature treats the problem of outliers unsystematically. For instances, ad hoc trimming is the almost universal practice in labour economics. By contrast, the statistical literature has a tradition of seeking to identify outliers, see e.g. the surveys of Barnett and Lewis (1994), Beckman and Cook (1983), and Hawkins (1980). An important early contribution is Ferguson (1961). Davies and Gather (1993), and Rosner (1975, 1983) deal, among other, with the problem of identifying multiple outliers. The early contributions focus specifically on outliers in normal distributions. A recent contribution is Mittnik, Rachev, and Samorodnitsky (2001) who derive the limiting distribution of the normalised range for samples in the domain of attraction of a stable law. The approximate critical values for the range statistic are proposed as a basis for (single) outlier detection.

This paper is organised in the following way: Section 2 establishes the notation and formalises the notion of heavy-tailedness. Section 3 presents the test procedures both for single and multiple outlier testing. In addition, we discuss the problem of robustly estimating the tail index in the presence of possible outliers. In Section 4 we simulate the behaviour of the test for single and multiple outliers in finite samples. It turns out that the tests are reliable in that their actual size is close to the nominal size. We also derive the power of the test and show its consistency. The empirical application is given in Section 5. Was the Black Monday market crash so extreme that it may be considered as an outlier not belonging to the heavy tailed return distribution governing the usual returns? Since return data are time-series data we pre-whiten the returns using an AR(1)-GARCH(1,1) model. We find evidence that the market crash was not a structurally unusual event. Section 6 concludes. The Appendix collects the proofs.

2. Preliminaries

We consider the class of distributions with tails which decay like power functions. These tails are heavy. More precisely, we assume that the cdf F satisfies, for some $\alpha > 0$ and sufficiently large x,

$$F(x) = 1 - L_0(x)x^{-\alpha} \tag{1}$$

where L_0 is a slowly varying function.[2] For sufficiently large x the distribution becomes Pareto-like, as L_0 then behaves almost like a constant. The parameter α is referred to as the tail index. This class of distributions is very broad. Many real world distributions are members of this class. We have enumerated in the introduction several examples in finance with estimates of the tail index around 3. Parametric examples include the generalised beta distribution (McDonald, 1984), nesting the Pareto, Singh-Maddala or Burr XII and the Dagum distributions as special cases, stable distributions with characteristic exponent in (0, 2), and Student's t-distributions with α equal to the degrees of freedom. Even medium tailed distributions can be approximated by setting the tail index sufficiently large. F lies in the domain of attraction of the Frechet distribution, being the only heavy-tailed limiting distribution of the sample maximum.

Next we define extremes and outliers. Consider a sample of size n and the associated order statistics $X_{1,n} \geq X_{2,n} \geq \ldots \geq X_{k,n} \geq \ldots \geq X_{n,n}$. The problem about which we want to make inferences is that the k extreme order statistics ($X_{k,n}$ to $X_{1,n}$) might, or might not, have been generated by F, the distribution of interest. For instance, the associated sample data might be contaminations, either measurement error or structurally unusual events. If the k extreme order statistics have not been generated by F we refer to them as 'outliers', otherwise they are 'extremes'. The problem of inlying contaminations is not considered. We specify a sequence of null hypotheses indexed by k: under $\mathbf{H}_{0,k}$ the kth order statistic is an extreme and not an outlier:

The null hypotheses $\mathbf{H}_{0,k}$, with $k = 1, 2, \ldots$: the kth extreme order statistic $X_{k,n}$ belongs to the population of interest F.

When $k = 1$ the test is for only a single outlier. When $k > 1$ we consider the possibility of multiple outliers, and we test the sequence of null hypotheses $\mathbf{H}_{0,k}$, $\mathbf{H}_{0,k-1}, \ldots$, $\mathbf{H}_{0,1}$. This is an outward testing procedure which addresses the masking problem stemming from multiple outliers, discussed in more detail below.

We base our testing procedure on ratios $R_\tau = X_{\tau,n} / X_{\tau+1,n}$ of order statistics, since these converge whereas it is well known that for heavy-tailed distributions extreme spacings diverge and are dependent.[3] It is therefore important to

[2] Consider $\lim_{x \to x0} g(tx)/g(x)$ with $t > 0$. The function g is regularly varying at x_0 if the limit equals t^θ with $\theta < 0$, and slowly varying if $\theta = 0$. We consider variation at infinity. $\bar{F}$ is regularly varying with $\theta = -a$.

[3] In fact, $R_1 = O_p(1)$ implies that $D_1 = X_{1,n} - X_{2,n} = O_p(n^{1/a})$. This is seen immediately by noting that $R_1 = D_1 / X_{2,n+1}$ and $c_{\mathrm{n}}^{-1}2\ X_{2,n}$ converging to the Frechet distribution where $c_{\mathrm{n}} = n^{1/a} L_0(n)$.

derive the marginal and the joint distribution of such ratios under the null hypothesis of no outliers $\mathbf{H}_{0,1}$. It turns out these have a simple form. First, we state a more general theorem about the marginal distribution of the ratios $X_{\tau,n}/X_{\tau+t,n}$.

Theorem 1. *Consider $X_{\tau,n}/X_{\tau+t,n}$ with τ and t fixed. Its asymptotic distribution is given by*

$$F_{X_{\tau,n}/X_{\tau+t,n}}(y) \xrightarrow{d} \left(1 - y^{-\alpha\tau}\sum_{j=0}^{k-1}[1-y^{-\alpha}]^j \frac{1}{j!}\frac{\Gamma(\tau+j)}{\Gamma(\tau)}\right),$$

where Γ denotes the Gamma function.

The asymptotic distribution of $R_\tau = X_{\tau,n}/X_{\tau+1,n}$ is a special case, whence

$$F_{R_\tau}(y) \xrightarrow{d} (1-y^{-\alpha\tau}) = (1-y^{-\alpha})\sum_{j=0}^{\tau-1}(y^{-\alpha})^j. \tag{2}$$

In particular, for $R_1 = X_{1,n}/X_{2,n}$ we have convergence to a Pareto distribution with parameter equal to the tail index α. Next we state the joint distribution of the ratios $R_1,\ldots, R_k$.

Theorem 2. *The asymptotic joint distribution of $R_1,\ldots, R_k$ is*

$$F_{R_1,\ldots,R_k}(r_1,\ldots,r_k) = \prod_{i=1}^{k}\left(1-r_i^{-\alpha}\right)\sum_{j=0}^{i-1}\left(r_i^{-\alpha}\right)^j = \prod_{i=1}^{k} F_{R_i}(r_i). \tag{3}$$

It is immediate that the test statistics $R_1,\ldots, R_k$ are independent, because the joint distribution factorises into the product of the marginal distributions. This independence will greatly simplify the computation of critical values. Theorems 1 and 2 follow from a generalisation of Renyi's exponential representation (see Renyi, 1970, chap. VIII, Section 9 or Beirlant et al., 2004, p.110) which states that $Z_j \equiv j\,(\log X_{j,n} - \log X_{j+1,n})$ with $j = 1,\ldots, k$ are asymptotically iid with the same exponential distribution.[4] We are now in a position to propose the outlier identification procedure.

3. Outlier identification

A test for a single outlier is a test of $\mathbf{H}_{0,1}$. The test procedure is immediate: in a test of size δ reject $\mathbf{H}_{0,1}$ if the realisation of R_1 exceeds the critical level $\delta^{-1/\alpha}$.

The presence of more than one outlier can mask their presence. Three outliers, for instance, could result in R_1 and R_2 being small but R_3 being large. This suggests the following outward testing procedure:

Proposition 3. (Outward testing for multiple outlier identification) Start at testing $\mathbf{H}_{0,k}$ using R_k, i.e. the hypothesis that the kth largest order statistic belongs to the population of interest. If $\mathbf{H}_{0,k}$ is not rejected, test the next null hypothesis $\mathbf{H}_{0,k-1}$ using R_{k-1}. This outward testing continues until the first outlier is detected or the sample maximum has been tested. For the overall test to have size δ, the individual critical levels can be set to

$$r_i = \left[1-(1-\delta)^{1/k}\right]^{-1/(\alpha i)}. \tag{4}$$

We discuss the estimation of the nuisance parameter α below. The choice of the critical levels to achieve the overall test to have size δ is justified as follows. The size of a joint test of $\mathbf{H}_{0,k},\ldots, \mathbf{H}_{0,1}$ for critical levels $(r_1,\ldots, r_k)$ is

$$\Pr\left\{\text{reject}\, \mathbf{H}_{0,k} \cup \ldots \cup \mathbf{H}_{0,1}\right\} = 1 - F_{R_1,\ldots,R_k}(r_1,\ldots r_k) = 1 - \prod_{i=1}^{k} F_{R_i}(r_i).$$

Setting the individual $F_{Ri}(r_i) = (1-\delta)^{1/k}$, it follows by independence that $1-\prod_{i=1}^{k}(1-\delta)^{1/k} = 1-(1-\delta) = \delta$. Also $F_{R_i}(r_i) = 1 - r_i^{-\alpha i}$, hence $1 - r_i^{-\alpha i} = (1-\delta)^{1/k}$ and the result (4) follows.

[4] We are grateful to a referee for pointing out this result. For an alternative direct proof see the web Appendix.

We proceed to discuss some theoretical properties of the test of Proposition 3. Below we examine its actual performance in a simulation study. Obviously, the test has size δ. To measure the performance of the test we consider next its conditional power, i.e. "the probability that the test is significant given that the contaminant is the extreme value tested" (Hawkins, 1980). The alternative 'not $\mathbf{H}_{0,1}$' is thus 'the maximum z is an outlier'. Note that we are thus considering a general alternative hypothesis. An alternative approach is derive the power under specific alternatives as e.g. in Pagurova (1996), who considers for distributions $F((x-\theta_{1i})/\theta_2)$ with $i=1,\ldots,n$ the null hypotheses $\theta_{1i}=\theta$ for all i, against the specific alternative $\theta_i>\theta$ for some i.

The conditional power of the test for a single outlier at level δ is

$$\begin{aligned}\pi_1(z) &= \Pr\left\{R_1>r_1|\text{not } \mathbf{H}_{0,1},\ z \text{ is an outlier}\right\}\\ &= \frac{\Pr\left\{z/X_{2,n}>r_1\right\}}{\Pr\left\{X_{2,n}<z\right\}}\\ &= \exp\left\{c_{n-1}^{\alpha} z^{-\alpha}\left[1-\delta^{-1}\right]\right\}\end{aligned} \tag{5}$$

with fixed z and $c_{n-1}=F^{-1}(1-(n-1)^{-1})$ and r_1 as defined in Eq. (4). We utilized the fact that the normalised maximum of heavy tailed distributions converges in distribution to the Frechet distribution. Note that $X_{2,n}$ is the maximum of the population of interest when z is an outlier.

The conditional power of the multiple outlier test is the probability that the outward testing procedure starting with $\mathbf{H}_{0,k}$ detects outliers, given that the m largest values $(1\le m\le k)$ are contaminants. We assume for simplicity that the contaminating m outliers are all equal to z, so that their ratios $R_j=1<r_j$ for $j=1,\ldots,m-1$ by construction (hence the test will not indicate the presence of outliers within the group of outliers). The conditional power is

$$\begin{aligned}\pi_{k,m}(z) &= \Pr\left\{R_1>r_1 \text{ or}\ldots\text{or } R_k>r_k|\text{there are } m \text{ outliers } z\right\}\\ &= 1-\frac{\Pr\left\{1\le R_m\le r_m,\ldots,R_k\le r_k\right\}}{\Pr\left\{X_{m+1,n}<z\right\}}.\end{aligned} \tag{6}$$

The probability in the denominator can be calculated using the fact that $X_{m+1,n}$, being the maximum of the population of interest, follows a Frechet law. The probability in the numerator is (see the Appendix for the derivation)

$$\begin{aligned}&P(1\le R_m\le r_1,\ldots,R_k\le r_k)\\ &= c_n^{(k-m+1)\alpha}\sum_{n=0}^{k-m}\sum_{\{i_1,\ldots,i_n\}}\left[(-1)^n r_{i_1}^{-(i_1-m)^{\alpha}}\cdots r_{i_n}^{-(i_n-m)\alpha}\times\left[\exp\left(-z^{-\alpha}c_n^{\alpha}r_{i_1}^{\alpha}\cdots r_{i_n}^{\alpha}\right)-\exp\left(-z^{-\alpha}c_n^{\alpha}r_{i_1}^{\alpha}\cdots r_{i_n}^{\alpha}r_1^{\alpha}\right)\right]\right],\end{aligned} \tag{7}$$

where the inner sum runs over all n-combinations $\{i_1,\ldots,i_n\}$ out of the set $\{m+1,\ldots,k\}$.

For sufficiently large outliers, the conditional power π approaches 1, and these outliers are detected with probability one. The test is thus consistent as $z\to\infty$. This is the appropriate notion of consistency in the current context since $c_n=O(n^{1/\alpha})$ but the conditioning event must be guaranteed to hold. This condition would not be met if z were fixed and $n\to\infty$ (as in the standard definition of consistency) since the extreme quantile of the heavy-tailed F then grows without bound.

The implementation of the test requires an estimate of the unknown nuisance parameter α. For distributions satisfying Eq. (1) the most popular estimator in the literature is the Hill (1975) estimator

$$\hat{\alpha}(\kappa)=\left[\frac{1}{\kappa}\sum_{i=1}^{\kappa}\ln X_{i,n}-\ln X_{\kappa+1,n}\right]^{-1} \tag{8}$$

where κ is the number of extremes to be included for estimation.

The estimator (8) is well behaved, e.g. has an asymptotic normal distribution (Embrechts et al., 1997). However, the estimator requires choosing the threshold κ, i.e. determining where the tail of the distribution begins. It is well known that κ gives rise to a mean-variance trade-off. In particular, for the distributions which satisfy $L_0(x)=g_1[1+g_2x^{-\rho}+O(x^{-2\rho})]$ with $\rho>0$ and $g_1>0$, Hall (1982) has shown that the asymptotic bias of the Hill estimator is $g_2g_1^{-\rho/\alpha}(\kappa/n)^{\rho/\alpha}\alpha\rho/(\alpha+\rho)$, whereas the asymptotic variance is α^2/κ.

We apply a procedure proposed by Huisman et al. (2001) in order to resolve the mean-variance trade-off. The procedure is as follows. The asymptotic bias is exactly linear in κ for $\alpha=\rho$ or approximately linear for sufficiently small κ. This suggests writing $\hat{\alpha}(\kappa)=\beta_0+\beta_1\kappa+\varepsilon(\kappa)$ with $\kappa=1,\dots,N$ for some threshold N which we set to $N=0.1n$. The error term $\varepsilon(\kappa)$ is heteroscedastic since the asymptotic variance is not constant for different κ. This is a weighted least squares (WLS) problem, and the regression constant β_0 is an unbiased estimator for α. The WLS-estimator thus obtained is, of course, a weighted sum of the Hill-type estimators $\hat{\alpha}(\kappa)$, $\kappa=1,\dots,N$.

We need to adapt the Hill estimator to our setting. In particular, the Hill estimator is consistent for extremes (i.e. under $\mathbf{H}_{0,1}$), but not in the presence of outliers. In order to test $\mathbf{H}_{0,k}$ it is therefore advisable to exclude the k extreme order statistics. An estimator excluding the potential outliers is the modified Hill-type estimator

$$\widehat{\alpha_{-k}}(\kappa)=\left[\frac{k}{\kappa-k+1}\ln X_{k+1,n}-\frac{\kappa}{\kappa-k+1}\ln X_{\kappa+1,n}+\frac{1}{\kappa-k+1}\sum_{i=k+1}^{\kappa}\ln X_{i,n}\right]^{-1}. \tag{9}$$

For $k=0$ we arrive at the usual Hill estimator $\hat{\alpha}(\kappa)$. The modified Hill-type estimator (9) turns out to inherit the nice properties which the Hill estimator has under $\mathbf{H}_{0,1}$, such as the asymptotic normality. Lemma 4 in the Appendix makes this more precise.

We apply the Huisman et al. (2001) procedure to the modified Hill estimator in all cases but one. It is well known that for stable distributions, the Hill estimator substantially overestimates the tail index (McCulloch, 1997). A robust (because quantile-based) maximum likelihood estimator is proposed by McCulloch (1986). A preliminary data analysis using diagnostics proposed by Nolan (1997) should typically help in determining whether the sample is reasonably well described by a stable distribution.

We also note that there are more sophisticated methods to robustly estimate the tail parameter (Brazauskas and Serfling, 2000; Peng and Welsh, 2001; Vandewalle et al., 2004; Vandewalle et al., 2007). However, in the framework of this paper the tail index is just a nuisance parameter, and consistency is all that is required. In order to keep the test procedure as simple to implement as possible we prefer (9) as an estimator.

4. Simulation evidence

We proceed to examine the performance of the outlier tests in several simulations. In particular, we take a close look at the empirical size in small samples of the single and multiple outlier test procedures. If the empirical size deviates substantially from the nominal size our test would be uninformative. The sample sizes we consider are in the range usually available in finance.

The data are generated by four heavy-tailed distributions, and we have chosen tail indices reported in the finance literature.

1. Student's t-distribution with 3 degrees of freedom.
2. Stable distribution with characteristic exponent $\alpha=1.5$.
3. Pareto distribution with tail index $\alpha=3$.
4. Burr XII or Singh-Maddala distribution with density

$$f(x;a,b,c)=\frac{abcx^{b-1}}{(1+ax^b)^{c+1}},$$

a member of the generalised beta distributions (McDonald, 1984). This distribution is frequently used not to model return distributions but rather to model income or wealth distributions (Singh and Maddala, 1976), and the parametrisation f(.; 100, 2.8, 1.7) fits real world income data well up to scale (Brachmann et al., 1996). The tail index is $\alpha=bc=4.76$.

Random numbers were generated by the inversion method, except for the stable distribution where we used the stable random number generator proposed by Chambers, Mallows, and Stuck (1976), see also Weron (1996). In all experiments, we generated 100,000 random samples. When considering actual test sizes, we have used a nominal value of $\delta=0.05$.

Table 1
Simulated empirical size of the single outlier test for various sample sizes

	Student's t	Stable	Pareto	Singh-Maddala
n	$\alpha=3$	$\alpha=1.5$	$\alpha=3$	$\alpha=4.76$
100	0.242	0.044	0.071	0.104
500	0.076	0.049	0.051	0.048
1000	0.066	0.049	0.051	0.046
1500	0.064	0.049	0.049	0.045
2000	0.061	0.051	0.050	0.046
2500	0.059	0.051	0.050	0.045
5000	0.060	0.049	0.051	0.044
7500	0.060	0.050	0.051	0.043
10000	0.058	0.050	0.050	0.044

The first simulation study investigates the error probability of the first kind of the single outlier test. In each iteration we generate a sample of size n from the distribution under consideration. As we are concerned with the error probability of the first kind, no outliers are inserted into the sample. The tail index α is estimated using the methods of Section 3. In this setting, if the realisation of test statistic R_1 is larger than the critical level $r_1=\delta^{-1/\hat{\alpha}}-1$ we (incorrectly) reject the null hypothesis. The proportion of rejections of $\mathbf{H}_{0,1}$ for the nominal size $\delta=0.05$ is reported in Table 1 for various sample sizes n. We conclude that the empirical size is mostly close to the nominal size, even for small samples. Only if the sample size gets as small as $n<1500$ the proportion of wrong rejections increases above the nominal size, especially for the t-distribution and (to a lesser extent) for the Singh-Maddala distribution. For sample sizes of $n\geq 1500$ the test keeps its nominal size virtually exactly for the stable distribution and the Pareto distribution, it is slightly conservative for the t-distribution and the Singh-Maddala distribution.

We turn to the test for multiple outliers, and investigate the error probability of the first kind. The sample size is set to $n=2000$, and the nominal significance level is $\delta=0.05$. We consider the actual size of the multiple test as a function of the number of potential outliers $k=1,\ldots, 10$. As before, the sample does in fact not contain any outliers. For $k>1$ we test the multiple null hypotheses $\mathbf{H}_{0,i}$ for $i=k,\ldots, 1$ by computing the test statistics $R_i=X_{i,n}/X_{i+1,n}$ and the tail index estimates $\widehat{\alpha_{-i}}$. Following Section 3, the null hypothesis $\mathbf{H}_{0,i}$ is rejected if the realisation of R_i is larger than the critical level (4). If any one of $\mathbf{H}_{0,k}$ to $\mathbf{H}_{0,1}$ is rejected the overall null hypothesis of no outliers is (wrongly) rejected. In practice the test procedure is sequential, starting with $i=k$ and decrementing i until rejection or $i=1$.

The results, given in Table 2, reveal that nominal and actual size for all distributions are in close agreement. The empirical size tends to decrease slightly with increasing k (and constant n), making the test more conservative. An exception is the Student's t-distribution where the empirical size is greater than the nominal size. The discrepancy is, however, small and may be neglected in applications.

Table 2
Simulated empirical size of the multiple outlier test for various numbers of potential outliers k and constant sample size $n=2000$

	Student's t	Stable	Pareto	Singh-Maddala
k	$\alpha=3$	$\alpha=1.5$	$\alpha=3$	$\alpha=4.76$
1	0.063	0.050	0.051	0.045
2	0.064	0.050	0.050	0.043
3	0.065	0.049	0.049	0.041
4	0.069	0.049	0.049	0.041
5	0.071	0.049	0.048	0.041
6	0.070	0.049	0.048	0.041
7	0.070	0.048	0.047	0.041
8	0.072	0.047	0.047	0.041
9	0.073	0.047	0.046	0.040
10	0.074	0.047	0.045	0.040

In summary, we find that the test procedures are not only valid asymptotically, but also in finite samples.

The same qualitative results are obtained concerning the power of the tests. We ran the same simulations ($n=2000$, $\delta=0.05$) with a single outlier z (or m multiple outliers) inserted into the data. Fig. 1 displays the power of the single outlier test for the four distributions. The dotted line shows the power as calculated by Eq. (5) with the true tail index α assumed to be known.

The solid line gives the simulated power when the tail index is estimated by Eq. (9) with $k=1$. In general, the simulated small sample power is close to $\pi_1(z)$. Taking the sampling error of α into account we find that $\pi_1(z)$ slightly understates the power for smaller outliers and overstates the power for larger outliers, the reason being that $\hat{\alpha}_{-1}$ is a biased estimator of α under the alternative hypothesis. In the presence of an outlier $X_{2,n}$ is the maximum of the population of interest while the estimator assumes $X_{2,n}$ to be the second-largest value. An exception is the stable distribution where the tail index is estimated by a quantile based estimator.

Fig. 2 shows the power in the presence of multiple outliers. We set the number of outliers to $m=3$ and apply the outward testing procedure starting with $k=5$. For clarity we assume that all $m=3$ outliers are equal to z. The dotted lines give the power $\pi_{k,m}(z)$ as a function of z (setting the tail index α to its true value). The solid lines show the simulated power when the tail index is estimated by $\hat{\alpha}_{-k}$. The results are similar to the single outlier case. The bias of the tail index estimator $\hat{\alpha}_{-k}$ under the alternative induces the power to be somewhat smaller than in the case where the true tail index is known. Again, the stable distribution is an exception because of its quantile based estimator. Comparing the single outlier test with the outward testing procedure in the presence of multiple outliers we find that the chances of detecting multiple outliers are larger in the latter case.

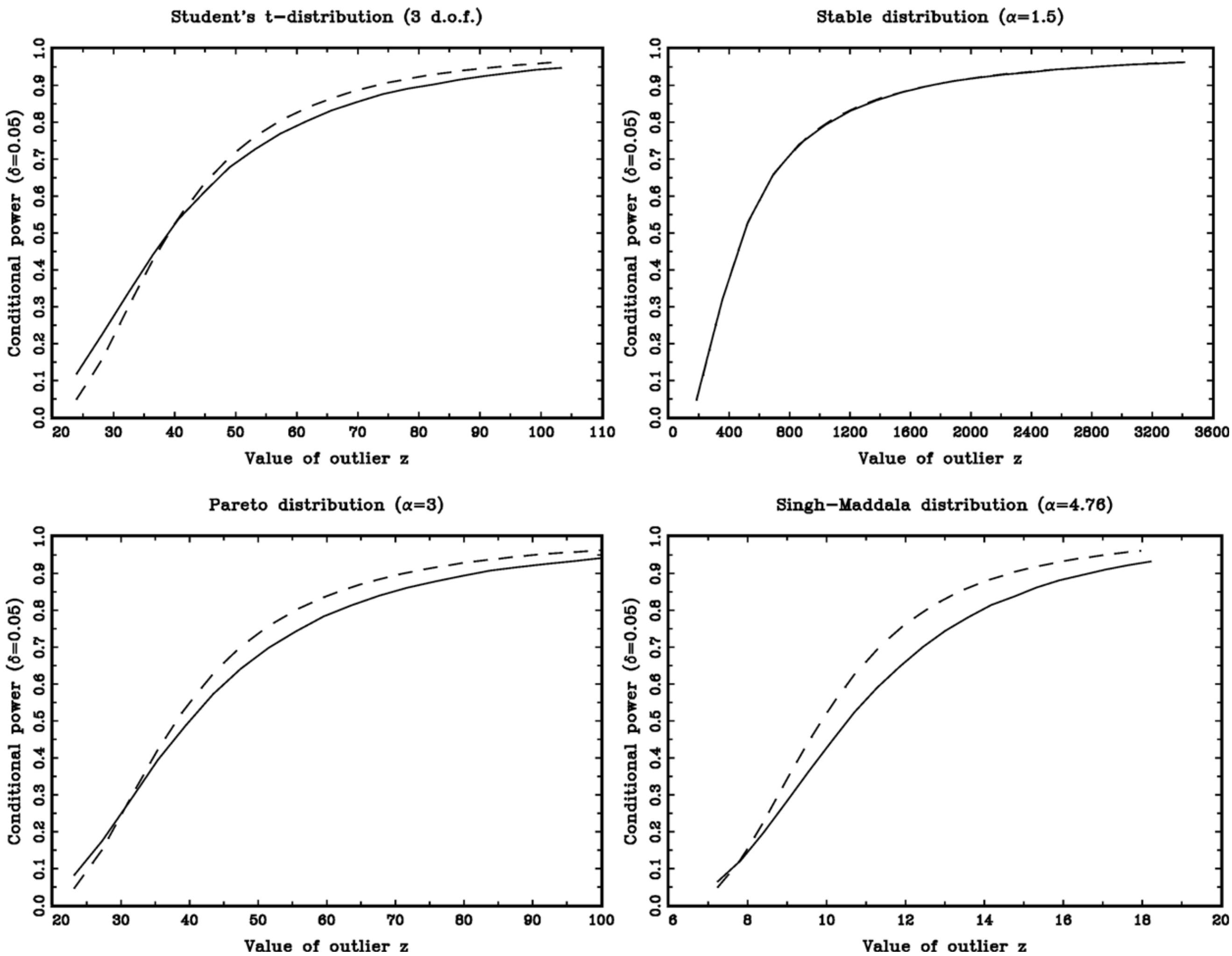

Fig. 1. Conditional power of the single outlier test as a function of the value z of the outlier.

5. Application to market crashes

We now apply the test to answer the question: Was the 1987 stock market crash (Black Monday) a structurally unusual event, i.e. an outlier? There is no consensus on this issue in the literature. McNeil (1998) concludes that the magnitude of the market crash could have been predicted using extreme value theory. However, his analysis is based on the worst daily loss in each year from 1960 until Black Monday, resulting in just 28 observations and hence prone to large sampling error. In contrast, Matthys and Beirlant (2001) use a bias-corrected estimator of the tail index and find that the market crash is unlikely to belong to the same return distribution as the rest of the observations. In a more recent paper, Novak and Beirlant (2006) estimate the tail index of the S&P500 to be $\hat{\alpha}=3.88$ implying that the market crash might in fact belong to the usual return distribution.

The data set consists of daily log-returns $X_1,\ldots, X_n$ of the Dow Jones Industrial Index from 3 January 1977 to 31 January 2005. The number of returns is $n=7326$. The mean and standard deviation of daily returns are $\bar{x}=0.032$ and $s=1.03$. The crash occurred on 19 October 1987 with a loss of 25.63% (in log-returns) on a single day. The second largest loss is much smaller (in absolute value): 8.38%, occurring one week after Black Monday. Fig. 3 depicts the logarithm of the Dow Jones Industrial Average from 1977 until 2005. The Black Monday crash is obviously the most extreme event during the observation period. Since tails of return distributions are known to be heavy, rather than thin, tailed it is not surprising to find returns that are (very) large in absolute value. It is evidently a very important question whether extreme events such as the 1987 crash are to be interpreted as large losses within the usual return distribution or rather as outliers not belonging to it. If a heavy tailed distribution fitted to the returns does not assign enough probability to crashes such as the Black Monday, any attempt to successfully manage portfolio risks is bound to fail.

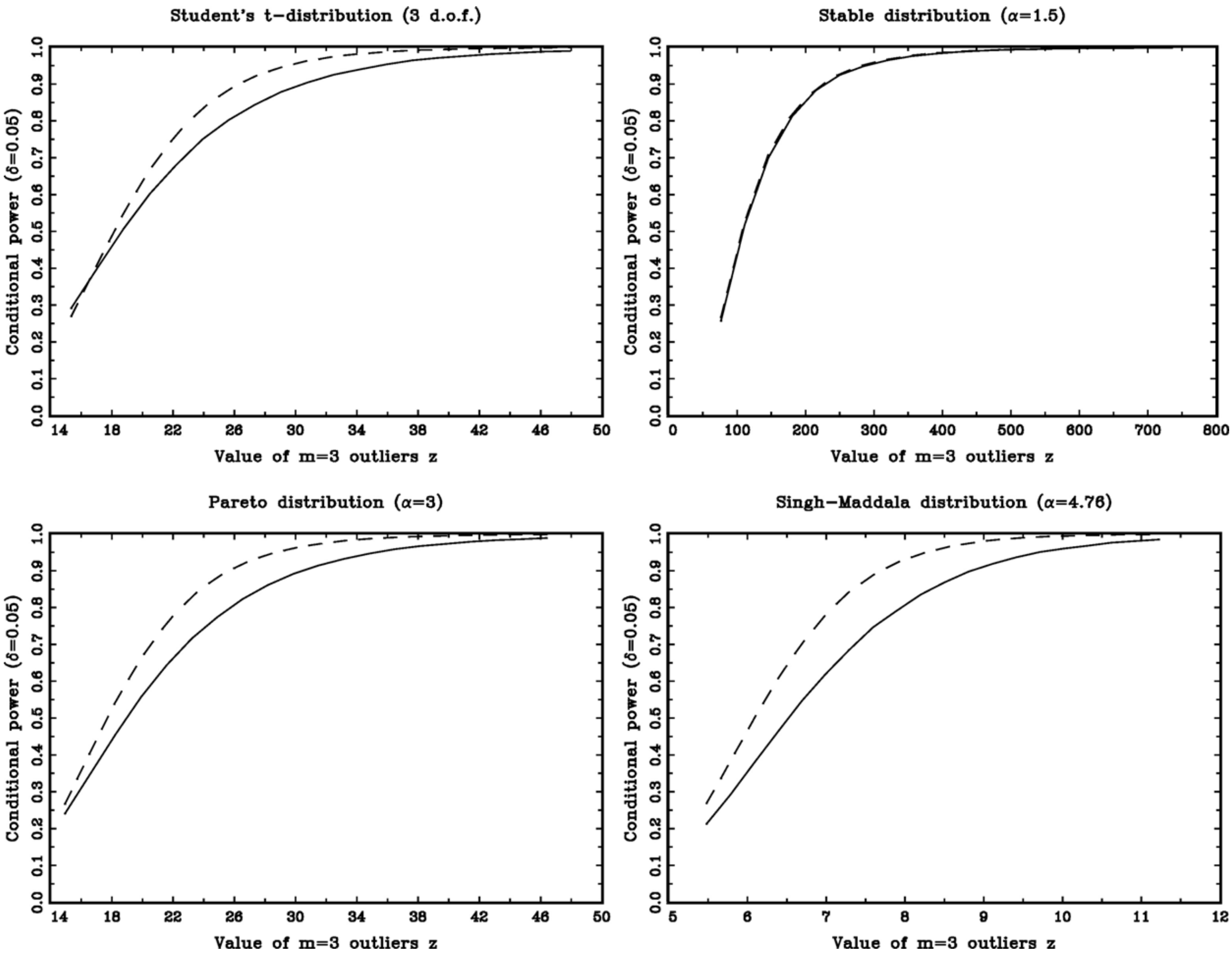

Fig. 2. Conditional power of the multiple outlier test as a function of the value z of the $m=3$ multiple outliers.

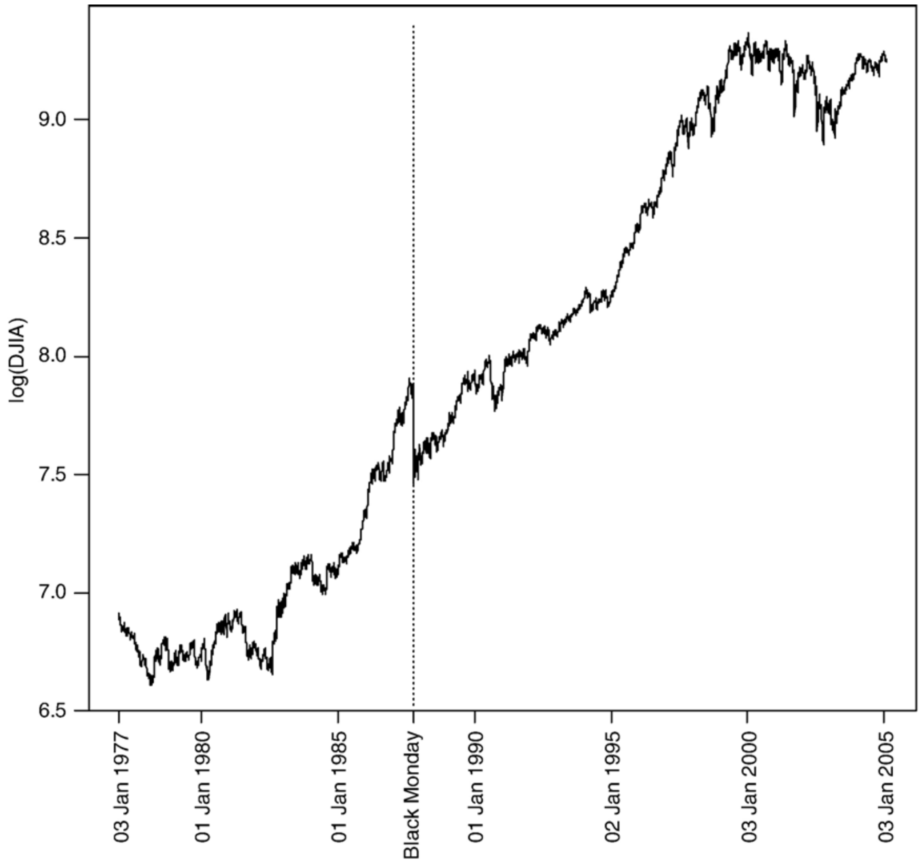

Fig. 3. Logarithm of the Dow Jones Industrial Average.

The distribution theory for our test is based on the assumption that the underlying data are iid. Yet, it is well known that financial data such as the Dow Jones Industrial Index are dependent. In order to be able to apply our test, we therefore need to pre-whiten the returns first. To this end we follow McNeil and Frey (2000) and estimate a standard AR (1)-GARCH(1,1) model

$$\begin{aligned} X_t &= \mu_t + \sigma_t \varepsilon_t \\ \mu_t &= \mu + \phi(X_{t-1} - \mu) \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{aligned}$$

by pseudolikelihood.[5] The fitted standardized residuals are $\hat{\varepsilon}_t = (X_t - \hat{\mu}_t)/\hat{\sigma}_t$, and depicted in Fig. 4. Next, we apply a battery of tests (Ljung-Box tests on the residuals and the squared residuals) to verify that these residuals are approximately independent. Hence we proceed to apply our test procedure to the residuals (i.e. the pre-whitened returns).[6] Since we do not make any symmetry assumption about the innovations we concentrate on the loss distribution and test whether there are outliers in the upper tail of the negative standardized residuals. Fig. 4 suggests that there might be two outliers, the first dated by the market crash, the second dated 13 October 1989. Hence we test for two and one outlier ($k=2$).

[5] For the purpose of estimation the innovations are assumed to be normally distributed. The point estimates are $\hat{\mu}=0.0452$, $\hat{\phi}=0.0338$, $\hat{\alpha}_0=0.0143$, $\hat{\alpha}_1=0.0680$, $\hat{\beta}_1=0.9200$.

[6] We have verified the robustness of the pre-whitening approach against misspecifications by a simulation study, based on a misspecified stochastic volatility model. The details, not reported here for the sake of brevity, can be obtained from the authors.

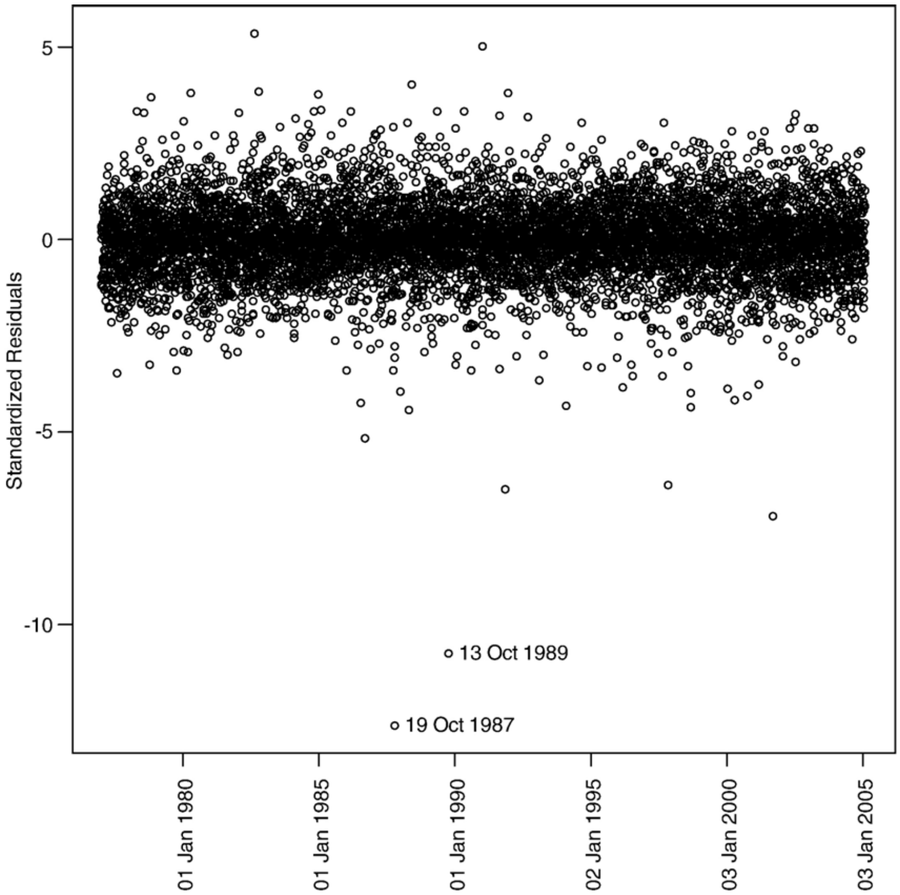

Fig. 4. Standardized residuals (pre-whitened returns).

The results are as follows: The ratio of the second and third largest pre-whitened negative return is 1.490. Our estimate of the tail index of the negative returns excluding the two potential outliers is $\widehat{\alpha_{-2}}=4.152$ where we have set $N=n/10=732$.[7] According to Eq. (4) the critical value for a test of size $\delta=0.05$ is

$$r_2 = \left[1-(1-0.05)^{1/2}\right]^{-1/(2\cdot 4.152)} = 1.557$$

which is larger than the test statistic. We conclude that the second largest pre-whitened loss is not an outlier. Next, we proceed to consider Black Monday. The ratio of the largest and second largest pre-whitened return is 1.176 and the re-estimated tail index $\widehat{\alpha_{-1}}=4.078$ leading to a critical value of $r_1=(1-(1-0.05)^{1/2})^{-1/4.078}=2.463$ for a test of size $\delta=0.05$ — again we conclude that the pre-whitened negative return of Black Monday is not an outlier.[8]

6. Conclusions

We have proposed a new testing procedure to identify multiple outliers, defined as large contaminations, in heavy-tailed distributions. The masking problem stemming from the presence of multiple outliers is solved by outward testing. Our test statistics are ratios of order statistics, and we have derived their asymptotic joint and marginal distributions. These outlier test statistics are shown to be independent, and therefore make the test simple to implement. The good performance of the test is demonstrated in several simulation studies. A substantive empirical contribution is the application of our test to Dow Jones Industrial Average return data and we find that the Black Monday market crash in October 1987 was not a structurally unusual event if the intertemporal dependence of the returns is taken into account.

[7] These results are consistent with McNeil and Frey (2000) using S&P500 data.
[8] The p-value of the joint test is 0.071.

Appendix A

A. Proofs

Proof of Theorems 1 and 2. Assume that Eq. (1) holds. The generalised Renyi result is that

$$Z_j \equiv j(\log X_{j,n} - \log X_{j+1,n}) \overset{d}{\simeq} \left(\alpha^{-1} + b\left(\frac{n+1}{j+1}\right)\right) Ej$$

with $j=1,\ldots,k$ where E_j are iid standard exponential random variables and b is a regularly varying function with some index $-\beta<0$ (see Beirlant et al., 2004, p.110). Theorems 1 and 2 are now immediate.

Derivation of the joint probability (7): For simplicity we assume that there is just one single outlier ($m=1$) and the outward testing procedure starts at the kth ratio of order statistics. The value of the outlier is z. Consider the transformation

$$T(x_1,\ldots,x_k) = \left(\frac{z}{x_1}, \frac{x_1}{x_2}, \ldots, \frac{x_{k-1}}{x_k}\right)$$

with inverse

$$T^{-1}(r_1, r_2, \ldots, r_k) = \left(\frac{z}{r_1}, \frac{z}{r_1 r_2}, \ldots, \frac{z}{r_1 r_2 \cdots r_k}\right).$$

The Jacobian is

$$\frac{\partial T^{-1}}{\partial r} = \begin{bmatrix} -\frac{z}{r_1^2} & 0 & \cdots & \cdots & 0 \\ -\frac{z}{r_1^2 r_2} & -\frac{z}{r_1 r_2^2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & -\frac{z}{r_1 r_2 \cdots r_{k-1}^2} & 0 \\ -\frac{z}{r_1^2 r_2 \cdots r_k} & -\frac{z}{r_1 r_2^2 \cdots r_k} & \cdots & \cdots & -\frac{z}{r_1 r_2 \cdots r_k^2} \end{bmatrix}.$$

Since all entries above the diagonal disappear the determinant is

$$|\det\left(\frac{\partial T^{-1}}{\partial r}\right)| = \prod_{i=1}^{k} \frac{z}{r_1 \cdots r_i^2} = z^k r_1^{-(k+1)} r_2^{-k} r_3^{-(k-1)} \cdots r_k^{-2}.$$

The asymptotic joint density of the first k order statistics is (Embrechts et al., 1997, p. 201)

$$\varphi_\alpha(x_1,\ldots,x_k) = \alpha^k \exp\left(-\left(x_k c_n^{-1}\right)^{-\alpha}\right) \prod_{j=1}^{k} \left(x_j c_n^{-1}\right)^{-\alpha-1}$$

with $c_n = F^{-1}(1-n^{-1})$. Hence, the joint density of the ratios $(R_1, R_2,\ldots, R_k) = (z/X_{1,n}, X_{1,n}/X_{2,n},\ldots, X_{k-1,n}/X_{k,n})$ is

$$\begin{aligned}
g_{R_1,\ldots,R_k}(r_1,\ldots r_k) &= \varphi_\alpha\left(\frac{z}{r_1}, \frac{z}{r_1 r_2}, \ldots, \frac{z}{r_1 r_2 \cdots r_k}\right) |\det\left(\frac{\partial T^{-1}}{\partial r}\right)| \\
&= c_n^{-k} \alpha^k \exp\left(-\left(\frac{z}{c_n r_1 r_2 \cdots r_k}\right)^{-\alpha}\right) \left(\frac{z}{c_n r_1}\right)^{-\alpha-1} \left(\frac{z}{c_n r_1 r_2}\right)^{-\alpha-1} \\
&\quad \times \ldots. \left(\frac{z}{c_n r_1 r_2 \cdots r_k}\right)^{-\alpha-1} z^k r_1^{-(k+1)} r_2^{-k} r_3^{-(-k-1)} \cdots r_k^{-2} \\
&= \alpha^k c_n^{k\alpha} z^{-k\alpha} \exp\left(-z^{-\alpha} c_n^{\alpha} r_1^{\alpha} r_2^{\alpha} \cdots r_k^{\alpha}\right) \cdot r_1^{k\alpha-1} r_2^{(k-1)\alpha-1} \ldots.. r_k^{\alpha-1}.
\end{aligned}$$

The required joint probability is

$$P(1\leq R_1\leq r_1, R_2\leq r_2, \ldots, R_k\leq rk) = \int_1^{r_1} \ldots \int_1^{r_{k1}} g_{R_1}, \ldots, R_k(x_1, \ldots, x_k) dx_k \ldots dx_1$$
$$= \alpha^k c_n^{k\alpha} z^{-k\alpha} \int_1^{r_1} x_1^{k\alpha-1} \int_1^{r_2} x_2^{(k-1)\alpha-1} \ldots \int_1^{r_k} x_k^{\alpha-1} \times \exp\left(-z^{-\alpha} c_n^{\alpha} x_1^{\alpha} x_2^{\alpha} \cdots x_k^{\alpha}\right) dx_k \ldots dx_2 dx_1$$
$$= c_n^{k\alpha} \sum_{n=0}^{k-m} \sum_{\{i_1,\ldots,i_n\}} \left\{ (-1)^n r_{i_1}^{-(i_1-1)\alpha} \cdots r_{i_n}^{-(i_n-1)\alpha} \times \left[\exp\left(-z^{-\alpha} c_n^{\alpha} r_{i_1}^{\alpha} \cdots r_{i_n}^{\alpha}\right) - \exp\left(-z^{-\alpha} c_n^{\alpha} r_{i_1}^{\alpha} \cdots r_{i_n}^{\alpha} r_1^{\alpha}\right)\right]\right\}$$

where the inner sum runs over all n-combinations $\{i_1,\ldots, i_n\}$ of the set $\{2,\ldots, k\}$. If $m>1$ the indices are simply shifted.

Lemma 4. (Properties of the modified Hill estimator) *Under Assumption (1) and the null hypothesis* $\mathbf{H}_{0,k}$ *the modified estimator has the following desirable properties:*

(i) *The estimator (9) is the maximum likelihood estimator of* α *of a Pareto distribution if the k largest observations are not included.*
(ii) *As* $n \to \infty$ *the estimators (8) and (9) converge almost surely:* $\left(\hat{\alpha} - \widehat{\alpha_{-k}}\right) \xrightarrow{a.s.} 0$.
(iii) *The estimator (9) is consistent for* α*: if* $\kappa \to \infty$, $\kappa/n \to 0$, *as* $n \to \infty$ *then* $\widehat{\alpha_{-k}} \xrightarrow{p} \alpha$.
(iv) *If* $\kappa/n \to 0$, $\kappa/\ln\ln n \to \infty$, *as* $n \to \infty$ *then* $\widehat{\alpha_{-k}} \xrightarrow{a.s.} \alpha$.
(v) *The estimator (9) is asymptotically normally distributed. Assume* $\lim_{x\to\infty} [\bar{F}(tx)/\bar{F}(x) - t^{-\alpha}]/\gamma(x) = t^{-\alpha}[t^{-\rho}-1]/(-\rho)$, $t>0$ *exists where* $\gamma(x)$ *is a measurable function of constant sign.* $-\rho$ *is the "second order parameter of regular variation". Let* $U(t) = F^{-1}(1-t^{-1})$, *and* $A(x) = \alpha^{-2}\gamma(U(x))$ *and* $\kappa \to \infty$ *but* $\kappa/n \to 0$. *If* $\lim_{n\to\infty} \sqrt{\kappa} A(n/\kappa) = \lambda \in \mathbb{R}$ *then, as* $n \to \infty$, *the estimator* $\widehat{\alpha_{-k}}$ *is consistent and asymptotically normal with* $\sqrt{\kappa}\left(\widehat{\alpha_{-k}} - \alpha\right) \xrightarrow{d} N(\lambda\alpha^3/(-\rho-\alpha), \alpha^2)$.

Proof of Lemma 4. (i) The joint density of $X_{j,n},\ldots, X_{\kappa,n}$ with $\kappa \leq n$ is

$$f_{X_{j,n},\ldots,X_{\kappa,n}}(x_j,\ldots,x_k) = \frac{n!}{(n-\kappa)!(j-1)!} F^{n-\kappa}(x_\kappa) \bar{F}^{j-1}(x_j) \prod_{i=j}^{\kappa} f(x_i)$$

for $x_j \geq \ldots \geq x_\kappa$ (David, 1981, p. 10). The proof then simply follows the usual derivation for the Hill estimator (see e.g. Embrechts et al., 1997). One immediately obtains $\widehat{\alpha_{-k}}(\kappa)$. (ii) is immediate. (iii)–(v) This theorem has been proved for the usual Hill estimator $\hat{\alpha}$, but by (ii) also apply to $\widehat{\alpha_{-k}}$. See, for instance Embrechts, Kluppelberg, and Mikosch (1997, chap. 6.4). Lemma 4(iii) is due to de Haan and Peng (1999). Another version is given in Haeusler and Teugels (1985). See also Weissman (1978). Asymptotic normality in the submodel with $L_0(x) = g_1[1 + g_2 x^{-\rho} + O(x^{-2\rho})]$ is also shown by Hall (1982). □

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jempfin.2007.10.003.

References

Barnett, V., Lewis, T., 1994. Outliers in Statistical Data, 3 ed. Wiley, Chichester.
Beckman, R., Cook, R., 1983. Outlier...s. Technometrics 25 (2), 119–149.
Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J. (Eds.), 2004. Statistics of Extremes: Theory and Applications. Wiley and Sons, Chichester.
Brachmann, K., Stich, A., Trede, M., 1996. Evaluating parametric income distribution models. Allgemeines Statistisches Archiv 80, 285–298.
Brazauskas, V., Serfling, R., 2000. Robust estimation of tail parameters for two-parameter Pareto and exponential models via generalized quantile statistics. Extremes 3 (3), 231–249.
Chambers, J., Mallows, C., Stuck, B., 1976. A method for simulating stable random variables. Journal of the American Statistical Association 71 (354), 340–344.
Danielsson, J., de Vries, C.G., 1997. Tail index and quantile estimation with very high frequency data. Journal of Empirical Finance 4, 241–257.
David, H., 1981. Order Statistics, 2 ed. Wiley, New York.
Davies, L., Gather, U., 1993. The identification of multiple outliers. Journal of the American Statistical Association 88, 782–792.

de Haan, L., Peng, L., 1999. Exact rates of convergence to a stable law. Journal of the London Mathematical Society 59, 1134–1152.
Embrechts, P., Kluppelberg, C., Mikosch, T., 1997. Modelling Extremal Events. Springer, Berlin.
Ferguson, T.S., 1961. On the rejection of outliers. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, pp. 253–287.
Freimer, M., Muholkar, G., Kollia, G., Lin, C., 1989. Extremes, extreme spacings and outliers in the Tukey and Weibull families. Communications in Statistics. Theory and Methods 18, 4261–4274.
Galambos, J., 1978. The Asymptotic Theory of Extreme Order Statistics. Wiley, New York.
Galambos, J., 1984. Introduction, Order Statistics, Exceedances, Laws of Large Numbers. D. Reidel Publishing Company, Dordrecht.
Gonzalo, J., Olmo, J., 2004. Which extremes are really extreme? Journal of Financial Economics 2 (3), 349–369.
Haeusler, E., Teugels, J., 1985. On asymptotic normality of Hill's estimator for the exponent of regular variation. Annals of Statistics 13, 743–756.
Hall, P., 1982. On some simple estimates of an exponent of regular variation. Journal of the Royal Statistical Society, Series B, 44, 37–42.
Hawkins, D., 1980. Identification of Outliers. Chapman and Hall, London.
Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. Annals of Statistics 3, 1163–1174.
Huisman, R., Koedijk, K.G., Kool, C.J., Palm, F., 2001. Tail-index estimates in small samples. Journal of Business and Economic Statistics 19, 208–216.
Klass, O.S., Biham, O., Levy, M., Malcai, O., Solomon, S., 2006. The Forbes 400 and the Pareto wealth distribution. Economics Letters 90 (2), 290–295.
Matthys, G., Beirlant, J., 2001. Extreme quantile estimation for heavy-tailed distributions. Working Paper.
McCulloch, J., 1986. Simple consistent estimators of stable distribution parameters. Communications in Statistics. Simulation and Computation 15, 1109–1136.
McCulloch, J.H., 1997. Measuring tail thickness to estimate the stable index α: a critique. Journal of Business and Economic Statistics 15, 74–81.
McDonald, J., 1984. Some generalized functions for the size distribution of income. Econometrica 52, 647–663.
McNeil, A., 1998. On extremes and crashes. Risk 11, 99.
McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7 (3–4), 271–300.
Mittnik, S., Rachev, S., Samorodnitsky, G., 2001. The distribution of test statistics for outlier detection in heavy-tailed samples. Mathematical and Computer Modelling 34, 1171–1183.
Nolan, J., 1997. Maximum likelihood estimation and diagnostics for stable distributions. American University Working Paper, Washington D.C.
Novak, S., Beirlant, J., 2006. The magnitude of a market crash can be predicted. Journal of Banking and Finance 30, 453–462.
Pagurova, V., 1996. On the asymptotic power of a test for detecting outliers. Theory of Probability and its Applications 42, 433–443.
Pareto, V., 1965. Ecrits sur la Courbe de la Repartition de la Richesse. Oeuvres completes, vol. III. Librairie Droz, Geneva.
Peng, L., Welsh, A., 2001. Robust estimation of the generalized Pareto distribution. Extremes 4 (1), 53–65.
Piketty, T., Saez, E., 2003. Income inequality in the United States, 1913–1998. Quarterly Journal of Economics 98 (1), 1–39.
Quintos, C.E., Fan, Z., Phillips, P.C., 2001. Structural change in tail behavior and the Asian financial crisis. Review of Economic Studies 68 (3), 633–663.
Renyi, A., 1970. Probability Theory. North-Holland, Amsterdam.
Rosner, B., 1975. On the detection of many outliers. Technometrics 17, 221–227.
Rosner, B., 1983. Percentage points for a generalized ESD many-outlier procedure. Technometrics 25 (2), 165–172.
Schluter, C., Trede, M., 2002. Tails of Lorenz curves. Journal of Econometrics 109, 151–166.
Schuster, E., 1984. Classification of probability laws by tail behaviour. Journal of the American Statistical Association 79, 936–939.
Singh, S., Maddala, G., 1976. A function for size distribution of incomes. Econometrica 44, 963–970.
Vandewalle, B., Beirlant, J., Hubert, M., 2004. A robust estimator of the tail index based on an exponential regression model. In: Hubert, M., Pison, G., Struyf, A., Van Aelst, S. (Eds.), Theory and Applications of Recent Robust Methods. Birkhauser, Basel, pp. 367–376.
Vandewalle, B., Beirlant, J., Christmann, A., Hubert, M., 2007. A robust estimator for the tail index of Pareto-type distributions. Computational Statistics & Data Analysis 51 (12), 6242–6251.
Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. Journal of the American Statistical Association 73, 812–815.
Weron, R., 1996. On the Chambers–Mallows–Stuck method for simulating skewed stable random variables. Statistics & Probability Letters 28, 165–171.