A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 255, 2011

Wiesław Dziubdziela[*], *Michał Stachura*[**], *Barbara Wodecka*[***]

EXTREME VALUE INDEX OF LEFT AND RIGHT TAILS FOR FINANCIAL TIME SERIES

Abstract. The paper concerns possibility of tail heaviness degree identification with use of the extreme value index estimation.

For that purpose there are two methods employed and additionally the methods are compared. One of them is Pickands' estimator that is based on the k-th order statistics, and the other, proposed by Berred, is some kind of parallel of the former but based on the k-th record values.

Key words: k-th record values, k-th order statistics, extreme value index, estimator, log-returns, distribution tail.

I. INTRUDUCTION

One of the main problems of financial risk estimation is to recognize the degree of tail heaviness of log-returns of relevant time series, whereas the very choice of tail depends on an investor, since one may be interested in decreases or in increases or other features of a proper index. However, regardless of any investor preferences, long-term empirical research shows that in case of many time series there exists some disproportion between right and left distribution tails. Moreover, there is well-known fact that the tails of financial data strongly differ from the ones of the normal distribution.

Properties of estimators used in the paper are verified and compared by executing simulation research concerning some arbitrarily chosen types of distributions. However the main research is focused on the use of Berred's concept to the extreme value index estimation with respect to several stock indices (inter alia WIG, Dow Jones, Dax, FTSE). The estimation confirms a considerable degree of tail heaviness asymmetry of the log-return distributions for the examined indices.

---

[*] Professor, Department of Quantitative Methods in Economics, Bolesław Markowski Higher School of Commerce in Kielce.

[**] Ph.D., Institute of Mathematics, Jan Kochanowski University of Humanities and Sciences.

[***] M.S., Institute of Mathematics, Jan Kochanowski University of Humanities and Sciences.

II. THEORETICAL BACKGROUND

It is well known that all possible non-degenerate weak limit distribution of the normalized partial maxima $X_{n,n}$, of independent and identically distributed random variables X_1, ..., X_n, are extreme value distributions, i.e., if there exist constants $a_n > 0$, b_n, for $n \in \mathbb{N}_+$, and some non-degenerate distribution function G such that for all x holds

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{n,n} - b_n}{a_n} \leq x \right) = G(x) \tag{1}$$

then there exists a constant $\gamma \in \mathbb{R}$, such that the limit distribution $G(x)$ is of the form

$$G(x) = G_\gamma(x) = \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}), & 1+\gamma x > 0, \quad \text{for } \gamma \neq 0 \\ \exp(\exp(-x)), & -\infty < x < \infty, \quad \text{for } \gamma = 0 \end{cases} \tag{2}$$

The parameter γ is called the *extreme value index* and it is the primary parameter of interest in the extreme value analysis [de Haan, Ferreira (2006)].

This parameter influences on the asymptotic behaviour of the right tail of the common distribution F. When $\gamma > 0$, we say that the distribution has a heavy tail, like the Pareto, the Cauchy, and the Student's distribution. For $\gamma = 0$ the distribution is light-tailed, for example the normal, the exponential, the gamma distributions. And for $\gamma < 0$ the distribution has a short tail, like the uniform and the beta distributions.

Classical estimators of extreme value index are based on upper order statistics. Among wide variety of such estimators, only the Pickands' one is within our interest. The estimator has the functional form [Pickands III (1975), Gomes *et al*. (2008)]

$$\hat{\gamma}^{\mathrm{P}}_{n,k} = (\ln 2)^{-1} \ln\left(\frac{X_{n-[k/4]+1,n} - X_{n-[k/2]+1,n}}{X_{n-[k/2]+1,n} - X_{n-k+1,n}} \right) \tag{3}$$

for $\gamma \in \mathbb{R}$, where $[x]$ denotes the integer part of x, and $4 \leq k \leq n$.

From the theoretical point of view the main advantage of Pickads' estimator is its insensitivity for any linear transformation of data, providing positive slope, which is compatible with the very notion of extreme value index. Additionally, the estimator is convenient for any real γ. Nevertheless, from practical point of view the Pickads' estimator displays instability, and was criticized in the literature on the subject. There are many improvements in this area, however the straightforward generalizations the Pickads' estimator or other ideas based on the order statistics lead to estimators that are either adequate for some bounded range of γ's values (e.g. Hill's case) or sensitive for linear transformations of data (e.g. Dekkers-Einemahl-de Haan case).

An alternative idea, proposed by Berred [Berred (1995)], is based on the *k*-th record values instead of the *k*-th order statistics. Dziubdziela & Kopociński defined the notion of the *k*-th record values as follows [Dziubdziela, Kopociński (1976)].

Let *k* be an integer. The sequences of the *k*-th times and the *k*-th values are defined by

$$L(1,k) = k, \quad \text{for } n = 1, \tag{4a}$$

$$L(n,k) = \min\{\ j > L(n-1,k) \ :\ X_j > X_{L(n-1,k)-k+1,\ L(n-1,k)}\ \} \text{ for } n > 1\,, \tag{4b}$$

$$X^{(k)}(n) = X_{L(n,k)-k+1,L(n,k)} \quad \text{for } n \geq 1\,, \tag{5}$$

In other words, by eliminating repetitions in the non-decreasing sequence of *k*-th order statistics a strictly increasing subsequence is obtained and it is called a sequence of *k*-th record values

$$X_{1,k} < X_{L(2,k)-k+1,L(2,k)} < X_{L(3,k)-k+1,L(3,k)} < \ldots \tag{6}$$

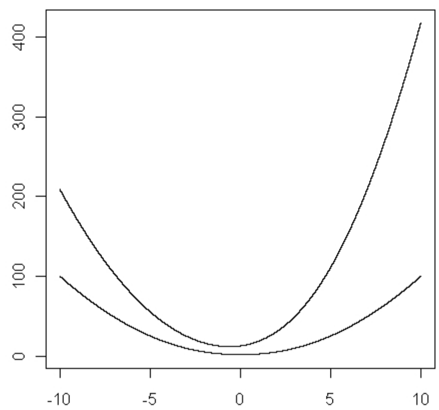
The estimator based on the *k*-th record values, introduced by Berred, is of the form

$$\hat{\gamma}^{B}_{n,k} = \ln\left(\frac{X^{(k)}(n) - X^{(k)}(n-k)}{X^{(k)}(n-k) - X^{(k)}(n-2k)} \right) \tag{7}$$

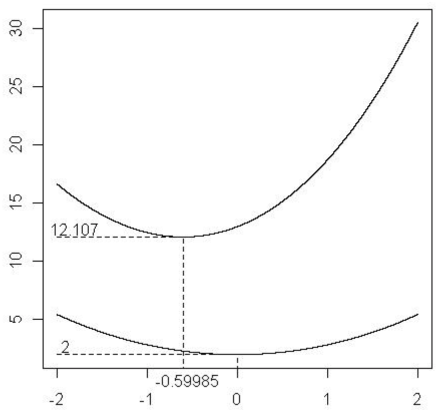
where $k = k_n$ is a sequence of positive integers satisfying $k_n / \ln n \rightarrow \infty$ and $k_n / n \rightarrow 0$ for $n \rightarrow \infty$. Such a choice of the sequence k_n guaranties almost sure convergence of the estimator. Among many possibilities the sequence of the type $k_n = n^{\lambda}$, for $\lambda \in (0,1)$, satisfies the above conditions and it is employed in empirical research, in continuation of this paper.

Limit laws of the considered estimators are known and their theoretical properties are widely discussed in the literature on the subject. The estimators $(\hat{\gamma}^{\mathrm{P}}_{n,k_n} - \gamma)\sqrt{k_n}$, $(\hat{\gamma}^{\mathrm{B}}_{n,k_n} - \gamma)\sqrt{k_n}$ are asymptotically zero-mean normal, with variances equal to $\dfrac{\gamma^2(2^{2\gamma+1}+1)}{((2^{\gamma}-1)\ln 2)^2}$, and $\dfrac{\gamma^2(e^{2\gamma}+1)}{(e^{\gamma}-1)^2}$, respectively, provided that $\gamma \neq 0$.

For $\gamma = 0$ analogous theoretical results are unknown. Nevertheless both the expressions $\dfrac{\gamma^2(2^{2\gamma+1}+1)}{((2^{\gamma}-1)\ln 2)^2}$, $\dfrac{\gamma^2(e^{2\gamma}+1)}{(e^{\gamma}-1)^2}$ may be regarded as well defined and continuous functions for all real numbers γ by setting proper limits (that are equal to $\dfrac{3}{\ln^4 2}$ and 2, respectively) as values for the argument $\gamma = 0$.



(a)



(b)

Figure 1. Panels (a) and (b) show plots of limit variances of both estimators in different scales. The limit variance of Berred's estimator is lower than the limit variance Pickands' estimator against any γ (going on horizontal axes)

Source: self-study.

Therefore one of the purposes of the simulation research, hereinafter described, is to detect differences between limit behaviour of the analysed estimators with $\gamma \neq 0$ vs. $\gamma = 0$.

The other purpose is to examine the degree of divergence between variances for simulated sequences of data and variances for the adequate theoretical limit distributions, which additionally enables to compare the convergence rapidity of both analysed estimators with respect to *k* (introduced before).

III. SIMULATION RESEARCH

All the simulation research is carried out in R environment.[1] The choice of analysed distribution types is made arbitrarily to include distributions of positive, negative and equal to zero extreme value indices. Thus the standard log-normal, the standard Cauchy, the standard exponential, the standard uniform distributions are considered as their extreme value indices are independent of their parameter changes. Additionally the Burr and the Pareto distributions are considered for proper choices of parameters, i.e. the Burr distribution for (*k*, *c*) = (0.2, 0.5), (0.5, 0.2), (0.5, 0.5), (1, 1), (2, 2), and the Pareto distribution for α = 0.05, 0.1, 2, 5, 10, 20. Total number of distribution types is therefore equal to 15.

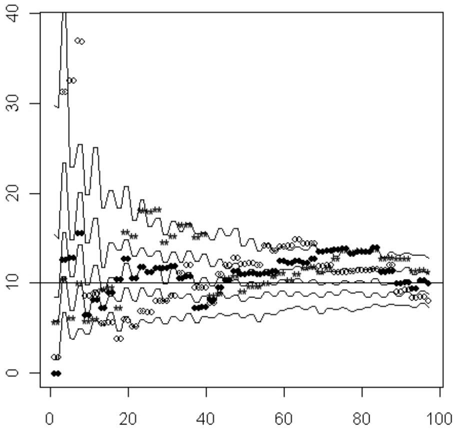
Firstly, there are simulated 100 series of the length 1000 for each of the distribution. Then the series of extreme value index estimators of both types are calculated for all k in range from 4 to 100 with respect to each simulated series. Values of *k* in Pickands' estimator correspond to the *k*-th order statistics, while in Berred's estimator to the *k*-th records.

And finally, quantile lines of orders 0.1, 0.3, 0.5, 0.7, 0.9 are determined for each hundred of computed estimator sequences of the both types.

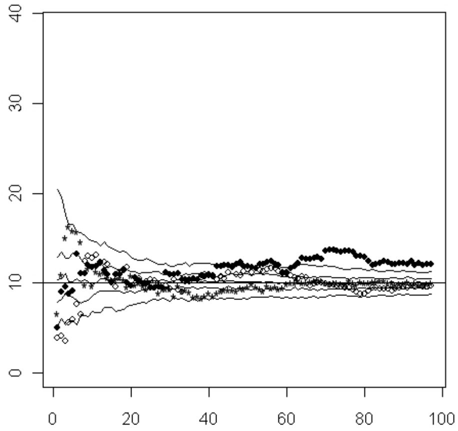
Figure 2. and figure 3. present simulation research results in graphical form for several selected distributions that represent different values of parameter γ. The other distributions lead to analogous plots and conclusions as well.

The carried simulation study leads to the following conclusions. The quantile lines in Berred's estimator case are more fast convergent to the theoretical limit law quantiles, and they exhibit less degree of volatility, in comparison with Pickand's case. The same refers to the variances and their limit values. Moreover, there are no substantial differences detected for $\gamma = 0$ against $\gamma \neq 0$ in view of quantile lines and selected series plots for both the estimators.

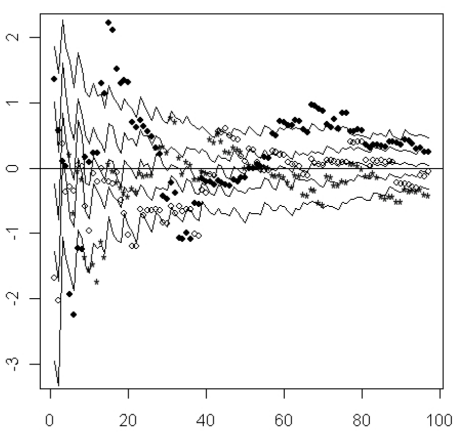
[1] The simulations are exercised with use of "Mersenne-Twister" RNG, i.e. R's default Random Number Generation implementation that has generalized feedback shift register (GFSR) with period equal to $2^{19937}-1$.



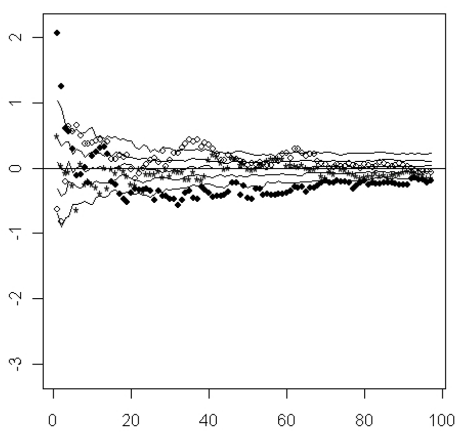
(1a)



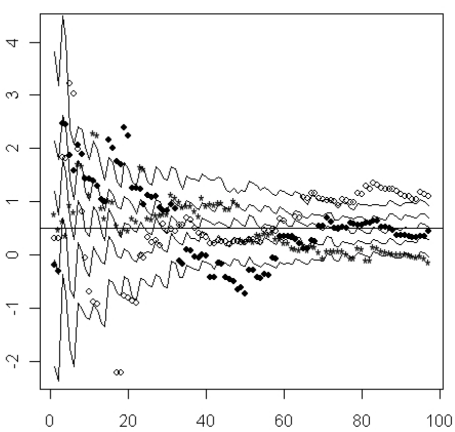
(1b)



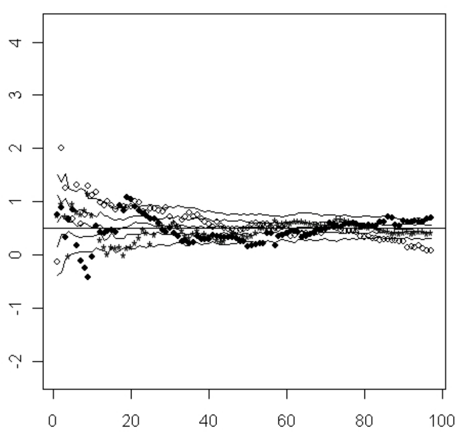
(2a)



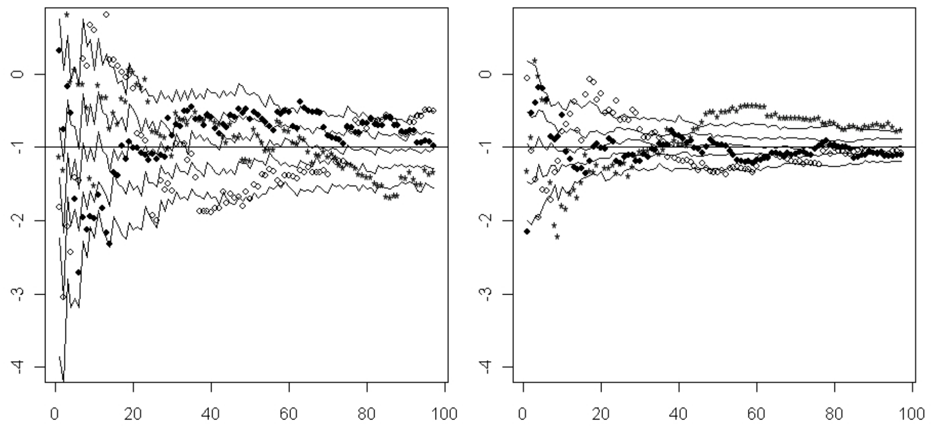
(2b)



(3a)



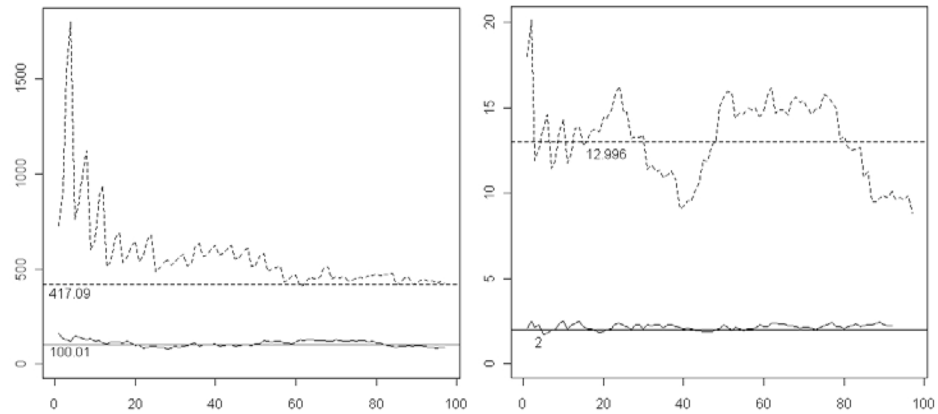
(3b)



(4a) **(4b)**

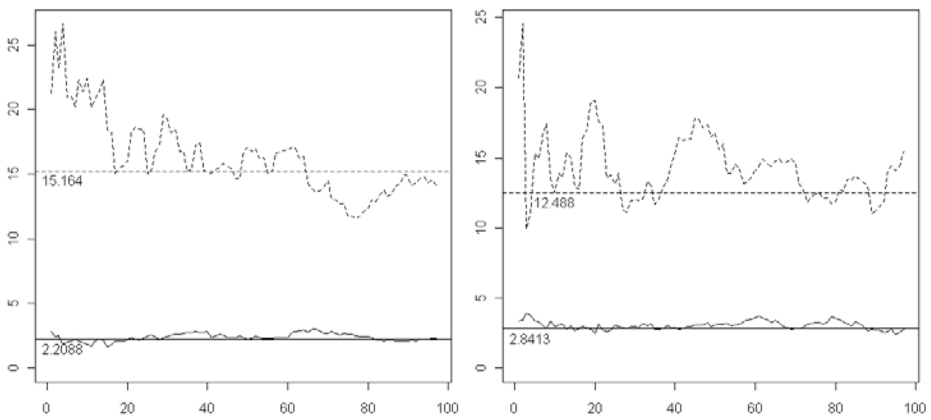
Figure 2. All panels present quantile lines (continuous lines) and 3 randomly selected series (points of 3 different patterns) for the exemplary distributions and for all k going on horizontal axes. The column (a) refers to Pickands' estimator, while the column (b) refers to Berred's estimator. The rows [(1) , (2), (3), (4)] comprise plots according to the Burr (k = 0.2, c = 0.5), the exponential, the Pareto (α = 2), and the uniform distribution, respectively.

Source: self-study.



(1) **(2)**

Figure 3a. The panels [(1) , (2), (3), (4)] comprise plots according to the Burr (k = 0.2, c = 0.5), the exponential, the Pareto (α = 2), and the uniform distribution, respectively. Rough lines represent empirical variances and straight horizontal lines represent the corresponding limit law variances. Dashed lines stand for Pickands' case, while continuous lines stand for Berred's case



(3) **(4)**

Figure 3b. The panels [(1) , (2), (3), (4)] comprise plots according to the Burr (k = 0.2, c = 0.5), the exponential, the Pareto (α = 2), and the uniform distribution, respectively. Rough lines represent empirical variances and straight horizontal lines represent the corresponding limit law variances. Dashed lines stand for Pickands' case, while continuous lines stand for Berred's case

Source: self-study.

IV. EMPIRICAL DATA ANALYSIS

The empirical research concerns log-returns of daily closure quotations of 8 stock indices (BUX, DAX, Dow Jones in the sequel in short DOW, FTSE – FTS, Hang-Seng – HAN, Nikkei225 – NIK, S&P500 – SAP, WIG). Each of them generates 1000 observations from about October or November, 2005 till October 30th, 2009 (depending on different numbers of working days of every stocks).[2]

All of the analysed time series exhibit specific behaviour during the global economic crisis of 2008, which can be observed on figure 3.

[2] BUX – 2005 10 26, DAX – 2005 11 18, DOW – 2005 11 09, FTS – 2005 11 16, HAN – 2005 10 10, NIK – 2005 10 04, SAP – 2005 11 08, WIG – 2005 11 07.

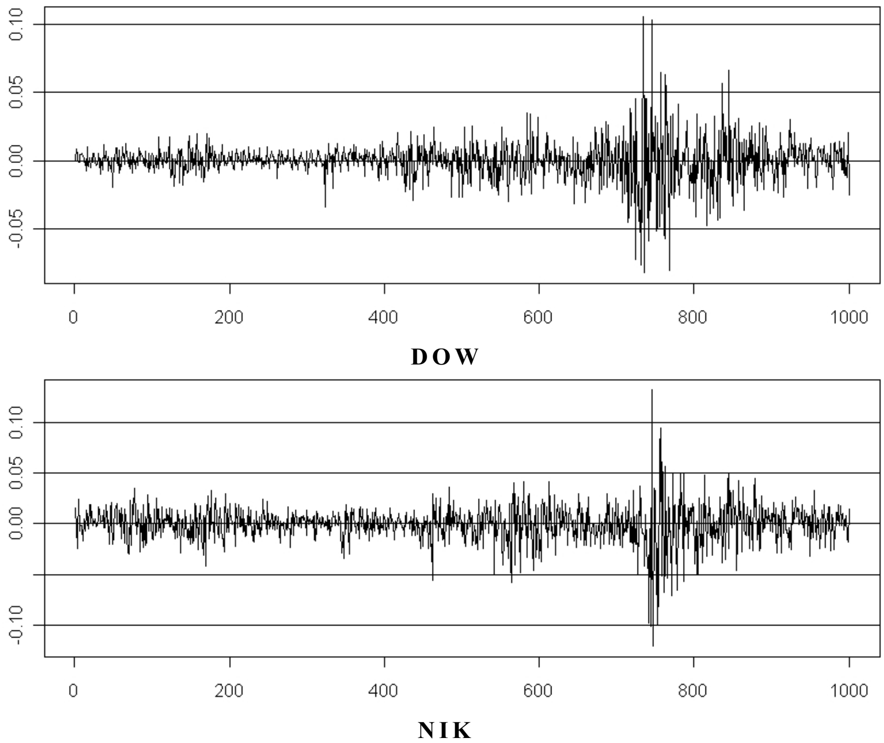


Figure 4. Panels present log-returns of two exemplary stock indices

Source: self-study.

A noticeable increase of volatility emerges after about 700th observation (September, 2008) with respect to all of the log-returns. Such an increase is synchronized with the crisis appearance.

In order to detect some qualitative change in the analysed time series distributions during the crisis, the extreme value indices are calculated in a way as follows.

First, with respect to every stock index, their log-returns series are partitioned into 2 subseries compound of the first 700 records, and the last 300 records, respectively.

Second, the extreme value indices of the right and the left tail[3] of both subseries are evaluated with use of the Berred's estimator. The arithmetic average of extreme value indices estimated for k's from properly chosen range

[3] From computational point of view the left tail of a variable X is the right tail of the variable $-X$.

($k = 6, \dots, 31$)[4] is treated as the empirical value of γ instead of single estimate (for a fixed k). Thus the risk of over- or underestimation may be reduced, since each of extreme value index is estimated only for a single empirical series.

Third, shift of disproportion between right and left tails for log-return series of the analysed stock indices is detected during the crisis. This disproportion is meant as the difference of extreme value index estimates of right and left tails of distribution, which reveals dominance of the right tail (if positive), or of the left one (if negative).

Table 1. The left part contains extreme value index estimators for right and left tails during and before the crisis [columns labelled by (1) and (2) respectively]. The right part consists of disproportion values (dv) and points out heavier tail (R, L). Bold fonts (**R**, **L**) designate disproportions that are arbitrarily admitted as significant (> 0.15), while contour fonts (ℝ, 𝕃) designate insignificant ones

Stock index	right tail (R)		left tail (L)		during crisis (1)		before crisis (2)	
	(1)	(2)	(1)	(2)	dv	tail	dv	tail
BUX	–0.0647	–0.3562	–0.1138	–0.0686	0.0490	ℝ	–0.2876	**L**
DAX	–0.1483	–0.1459	0.3538	–0.3192	–0.5021	**L**	0.1733	**R**
DOW	–0.1934	0.3986	–0.1875	0.2050	–0.0059	𝕃	0.1936	**R**
FTS	–0.1652	0.4644	0.1036	0.2653	–0.2687	**L**	0.1991	**R**
HAN	–0.5585	0.5585	–0.0258	0.4054	–0.5327	**L**	0.1530	**R**
NIK	0.2415	0.0669	0.7829	0.0957	–0.5414	**L**	–0.0289	𝕃
SAP	–0.0676	0.3911	–0.1599	0.3333	0.0923	ℝ	0.0578	ℝ
WIG	–0.4369	–0.0667	–0.0588	0.0971	–0.3781	**L**	–0.1638	**L**

Source: self-study.

Information contained in the table 1. discloses a tendency for log-returns tails of the analysed indices toward switching their dominance during the crisis. In general it may be inferred that before the crisis right tails seem to be heavier than the left ones, whilst after the crisis emerged the situation seems to be conversed.

[4] $\lambda_1 = 0.3$ and $\lambda_2 = 0.6$ provide $k_n = \lceil n^{0.3} \rceil, \dots, \lceil n^{0.6} \rceil$, where $\lceil x \rceil$ denotes the ceiling of x. For $n = 300$, $\lceil n^{0.3} \rceil = 6$, $\lceil n^{0.6} \rceil = 31$.

V. CONCLUSIONS

The presentation of the analysed estimators and the executed simulation research have demonstrated some advantages of Berred's estimator over the Pickands' one. Moreover, substituting order statistics by record values is proposed as an alternative approach for estimating the extreme value index in more precise and simple way than classical approach based on k-th order statistics.

The exemplary data analysis concerning the log-returns of several stock indices has disclosed that during the global crisis of 2008 there emerged not only substantial change in the log-returns volatility, but also a significant switch of log-return tail dominance.

REFERENCES

Arnold B.C., Balakrishnan N., H.N. Nagaraja H.N. (1998), *Records*, Wiley, New York.

Arrouchi M. El, Imlahi A. (2005), *Optimal choice of k_n-records in the extreme value index estimation*, Statistics & Decision 23, 101–115.Beirlant J., Goegebeur Y., Segers J., Teugles J. (2004), *Statistics of Extremes. Theory and Applications*, Wiley, Chichester.

Berred M. (1995), *K-record values and the extreme-value index*, J. Stat. Plan. Inference 45, 49–63.

Chandler K.N. (1952), *The distribution and frequency of record values*, J. R. Statist. Soc., Ser. B 14, 220–228.DeFusco R.A., McLeavey D.W., Pinto J. E., Runkle D.E. (2004), *Quantitative investment analysis*, Wiley, Hoboken.Dziubdziela W., Kopociński B. (1976), *Limiting properties of the k-th record values*, Zastosowania Matematyki 15, 187–190.

Dziubdziela W., Stachura M., Wodecka B. (2009), *Estymacja indeksu ekstremalnego szeregów finansowych w oparciu o rekordowe zwroty*, Rola informatyki w naukach ekonomicznych i społecznych 2/2009, tom 2, 327–337 Wydawnictwo Wyższej Szkoły Handlowej, Kielce.

Gomes M.I., e Castro L.C., Fraga Alves M.I., Pestana D. (2008), *Statistics of extremes for IID data and breakthroughs in the estimation of the extreme value index*: *Laurens de Haan leading contributions*, Extremes 11, 3–34.

de Haan L., Ferreira A. (2006), *Extreme Value Theory. An Introduction*, Springer, New York.

Nevzorov V.B. (2001), *Records: Mathematical Theory.* Transl. Math. Monographs 194, Amer. Math. Soc., Providence RI.

Pickands III J. (1975), *Statistical inference using extreme order statistics*, Ann. Statist. 3, 119–131.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Wiesław Dziubdziela, *Michał Stachura*, *Barbara Wodecka*

INDEKS EKSTREMALNY PRAWYCH I LEWYCH OGONÓW ROZKŁADÓW DLA SZEREGÓW FINANSOWYCH

Praca dotyczy możliwości identyfikacji stopnia grubości ogonów rozkładu poprzez estymację indeksu ekstremalnego.

W tym celu stosowane, i dodatkowo porównywane między sobą, są dwie metody. Jedną z nich jest estymator Pickandsa oparty o k-te statystyki pozycyjne, drugą zaś jest alternatywna metoda zaproponowana przez Berreda oparta o k-te wartości.