

Radosław Pietrzyk

Akademia Ekonomiczna we Wrocławiu

**ESTYMACJA INDEKSU OGONA ROZKŁADU
STÓP ZWROTU NA POLSKIM RYNKU AKCJI****1. Wstęp**

Występowanie „grubych ogonów” rozkładów stóp zwrotu na rynkach finansowych skłania do poszukiwań rozkładów innych niż normalny, dokładniej opisujących zachowanie się stóp zwrotu. Do takich rozkładów należą rozkłady α -stabilne. Grube ogony rozkładów mogą być również modelowane za pomocą rozkładów wartości ekstremalnych czy też uogólnionego rozkładu Pareto.

Niech $X_1, \dots, X_n$ będzie daną próbą losową pochodzącą z rozkładu o dystrybucie F , którego prawy ogon (przy założeniu $x \rightarrow \infty$) spełnia warunek:

$$1 - F(x) = P(X > x) \approx x^{-\alpha} L(x), \quad (1)$$

gdzie $L(x)$ – wolno zmieniająca się funkcja (por. [Embrechts, Klüppelberg, Mikosch 1997]).

Do rozkładów spełniających ten warunek należą między innymi rozkłady α -stabilne, t -Studenta oraz Pareto, gdzie α oznacza odpowiedni indeks stabilności, liczbę stopni swobody oraz grubość ogona [Weron, Weron 1998].

Wykorzystanie rozkładów α -stabilnych do opisu rozkładu stóp zwrotu zostało zaproponowane m.in. w pracy Mandelbrota z 1963 r. Rozkłady te definiowane są najczęściej przez funkcję charakterystyczną, gdyż jawna postać dystrybucy i funkcji gęstości jest znana jedynie dla trzech przypadków rozkładów α -stabilnych: normalnego ($\alpha = 2$), Cauchy’ego ($\alpha = 1$) oraz Levy’ego ($\alpha = 1/2$) [Weron, Weron 1998]. Funkcja charakterystyczna przyjmuje postać:

$$\phi(t) = \begin{cases} \exp\left[(-\sigma^\alpha |t|^\alpha (1 - i\beta \operatorname{sgn}(t) \operatorname{tg} \frac{\pi\alpha}{2}) + i\mu t)\right], & \alpha \neq 1, \\ \exp\left[-\sigma |t| (1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \ln |t|) + i\mu t\right], & \alpha = 1, \end{cases} \quad (2)$$

gdzie: $\alpha \in (0; 2]$ – indeks stabilności, indeks ogona,

$\beta \in [-1; 1]$ – parametr skośności,

$\sigma > 0$ – parametr skali,

$\mu \in R$ – parametr położenia.

Dla $\alpha = 2$ zmienna losowa ma rozkład normalny, natomiast dla $0 < \alpha < 2$ otrzymujemy rozkład o ogonach grubszych od ogonów rozkładu normalnego. Im niższa wartość parametru α , tym ogon jest grubszy, co oznacza większe prawdopodobieństwo występowania obserwacji ekstremalnych.

Do takich analiz może zostać wykorzystany warunkowy rozkład przekroczenia (*conditional excess distribution*), który może być dobrze aproksymowany za pomocą uogólnionego rozkładu Pareto o dystrybuancie danej wzorem:

$$G_{\xi, \sigma}(y) = \begin{cases} 1 - (1 + \frac{\xi}{\sigma} y)^{-1/\xi} & \text{dla } \xi \neq 0, \\ 1 - e^{-y/\sigma} & \text{dla } \xi = 0, \end{cases} \quad (3)$$

dla $y \in [0, (x_F - u)]$, jeżeli $\xi \geq 0$, oraz $y \in [0, -\frac{\sigma}{\xi}]$, jeżeli $\xi \leq 0$.

Uogólniony rozkład Pareto jest charakteryzowany przez dwa parametry. Parametr $\sigma > 0$, zwany parametrem skali, oraz parametr ξ , zwany parametrem kształtu, definiujący zarazem poszczególne rozkłady należące do rodziny rozkładów Pareto ($\xi > 0$ – rozkład Pareto, $\xi = 0$ – rozkład wykładniczy, $\xi < 0$ – rozkład Pareto typu II).

Problemem pojawiającym się przy wykorzystaniu warunkowego rozkładu przekroczenia jest właściwy wybór progu u . W niniejszym artykule zastosowano metodę graficzną opisaną między innymi w pracy [Këllezi, Gilli 2000]. W metodzie tej próg przyjmowany jest na poziomie, na którym rozpoczyna się liniowa zależność między funkcją wartości oczekiwanej $e(u)$ a u .

2. Podejście nieparametryczne do estymacji indeksu ogona

Jednym z podstawowych problemów pojawiających się przy rozpatrywaniu uogólnionego rozkładu Pareto jest estymacja parametrów rozkładu, a w szczególności parametru α , zwanego indeksem ogona. Estymacji można dokonać za pomo-

ca estymatorów Pickandsa, Hilla, Dekkersa–Einmahla–de Haana. Estymatory te pozwalają na estymację parametru $\hat{\xi} = \frac{1}{\hat{\alpha}}$.

Zastosowanie estymatorów wymaga utworzenia z obserwacji ciągu statystyk pozycyjnych $X_{(1)}, \dots, X_{(n)}$, czyli uporządkowania wartości zmiennych $X_1, \dots, X_n$, takiego, że $X_{(i)} \geq X_{(i+1)}$ (n – wielkość próby).

Estymator Pickandsa wyrażony jest wzorem:

$$\hat{\xi}_{k,n}^{(P)} = \frac{1}{\ln 2} \ln \frac{X_k - X_{2k}}{X_{2k} - X_{4k}}, \quad (4)$$

gdzie: $X_1, \dots, X_n$ – niezależne zmienne pochodzące z tego samego rozkładu (uporządkowane malejąco),

n – liczebność próby,

k – parametr estymatora przy czym $4k \leq n$.

Estymator ten charakteryzuje się największą wariancją z przedstawionych estymatorów, co może utrudniać interpretacje otrzymanych wyników [Weron, Weron 1998].

Estymator Hilla jest opisany wzorem:

$$\hat{\xi}_{k,n}^{(H)} = \frac{1}{k} \sum_{j=1}^k \ln X_j - \ln X_k, \quad (5)$$

gdzie $k \leq n$ – liczba obserwacji z ogona rozkładu.

Estymator Hilla jest obciążony, a jego obciążenie rośnie w miarę wzrostu k , jednocześnie dla małych k wariancja estymatora jest duża, co powoduje, że podstawowym problemem staje się dobór takiej wartości zmiennej k , aby otrzymane wyniki nie różniły się znacząco od wartości rzeczywistej i jednocześnie wariancja nie była zbyt duża.

Kolejnym estymatorem jest estymator Dekkersa–Einmahla–de Haana, określony wzorem:

$$\hat{\xi}_{k,n} = 1 + H_n^{(1)} + \frac{1}{2} \left(\frac{(H_n^{(1)})^2}{H_n^{(2)}} - 1 \right)^{-1}, \quad (6)$$

gdzie:

$$H_n^{(1)} = \frac{1}{k} \sum_{j=1}^k \ln X_j - \ln X_k, \quad (7)$$

$$H_n^{(2)} = \frac{1}{k} \sum_{j=1}^k (\ln X_j - \ln X_{k+1})^2. \quad (8)$$

Estymator Dekkersa–Einmahla–de Haana jest rozszerzeniem estymatora Hilla i podobnie jak estymator Hilla jest estymatorem zgodnym.

3. Podejście parametryczne

Jednym z podstawowych i najczęściej stosowanych podejść do estymacji parametru α jest wykorzystanie metody największej wiarygodności. Metoda ta zakłada, że mając n niezależnych obserwacji $X_1, X_2, \dots, X_n$ zmiennej losowej X o rozkładzie zależnym od wektora parametrów θ , a f_θ jest funkcją gęstości rozkładu prawdopodobieństwa zmiennej losowej X , funkcję

$$L(X_1, \dots, X_n; \theta) = f_\theta(X_1) \dots f_\theta(X_n) \quad (9)$$

nazywamy funkcją wiarygodności rozważanych parametrów (por. [Theil 1979, Gajek, Kałuska 2000]).

Metoda ta określa, że estymatorem parametru θ jest takie $\hat{\theta}$, dla którego funkcja wiarygodności $L(X_1, \dots, X_n; \theta)$ osiąga największą wartość. Otrzymany tą metodą estymator $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ jest nazywany estymatorem największej wiarygodności. Ważnym faktem jest, że maksymalizacja funkcji L jest równoważna maksymalizacji logarytmu naturalnego tej funkcji $\ln L$. Więcej na temat samej metody największej wiarygodności można odnaleźć między innymi w pracy Theila [1979].

Rozważania w tym artykule będą dotyczyły estymacji indeksu ogona dla uogólnionego rozkładu Pareto. Przez $\ell(\theta; X) = \ln L(X; \theta)$ oznaczmy logarytm naturalny funkcji wiarygodności. Estymator największej wiarygodności (*maximum likelihood estimator* – MLE) parametru θ przyjmuje postać:

$$\ell((\xi, \sigma; X)) = \begin{cases} -n \ln \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \ln \left(1 + \frac{\xi}{\sigma} X_i\right) & \text{dla } \xi \neq 0, \\ -n \ln \sigma - \frac{1}{\sigma} \sum_{i=1}^n \ln X_i & \text{dla } \xi = 0. \end{cases} \quad (10)$$

Podczas estymacji parametrów rozkładu metodą największej wiarygodności konieczne jest zastosowanie metod numerycznych. Takie rozwiązania zaproponowali w swoich pracach między innymi Jenkinson [1969] oraz Prescott i Walden [1980; 1983], a także Hosking [1985] i Macleod [1989] (por. [Embrechts, Klüppelberg, Mikosch 1997, s. 318]).

4. Badania empiryczne

W niniejszym artykule, przy zastosowaniu estymatorów nieparametrycznych i estymatora największej wiarygodności, zostaną zbadane rozkłady dziennych logarytmicznych stóp zwrotu akcji oraz indeksów giełdowych. Pozwoli to na określenie wartości indeksów ogonów rozkładów stóp zwrotu na polskim rynku akcji. Przedmiotem badań będą zarówno indeksy giełdowe, jak i spółki notowane na Giełdzie Papierów Wartościowych w Warszawie. Jednym z podstawowych problemów jest

wielkość próby, jaką należy wykorzystać w badaniach. Dane z polskiego rynku mają niespełna piętnastoletnią historię, a do badań mogą być zastosowane dane od połowy lat dziewięćdziesiątych, kiedy na polskiej giełdzie rozpoczęto notowania ciągle. Dostępne dane nie pozwalają na zbudowanie długich serii danych, z kolei ich nadmierne wydłużanie powoduje, że obserwacje pochodzące z odległej przeszłości mogą zbyt silnie wpływać na postać rozkładu.

W celu estymacji indeksów ogonów rozkładów stóp zwrotu na rynku polskim wybrano 8 spółek i 2 indeksy (WIG i WIG20), które charakteryzują się długimi seriami danych (2741 notowań od 4 października 1994 r. do 16 września 2005 r.). Dla porównania do badań wybrano również 6 dużych spółek pochodzących z indeksu WIG20 (Agora, TP SA, Prokom, Computerland, PKN Orlen, KGHM), charakteryzujących się dużą płynnością na rynku, ale posiadających krótsze serie danych (od 1609 do 2475 notowań).

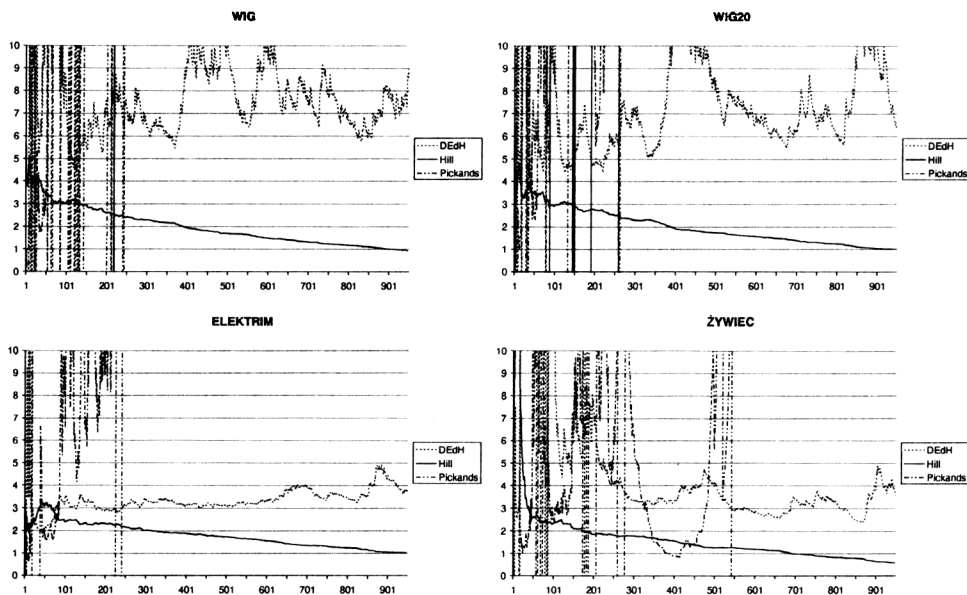
Tabela 1. Podstawowe statystyki indeksów i akcji

Spółki	Liczba notowań	Odchylenie standardowe	Średnia (w %)	Skośność	Kurtoza	Minimum (w %)	Maksimum (w %)
WIG	2741	0,016506	0,05	-0,11086	6,27332	-10,29	7,89
WIG20	2741	0,019461	0,04	-0,07331	6,836551	-14,16	13,71
ING	2741	0,025265	0,04	0,013344	8,364776	-15,46	14,01
BRE	2741	0,027495	0,05	-0,03851	6,36731	-15,93	15,93
Millenium	2741	0,032767	0,02	0,360642	10,6661	-21,67	25,66
Elektrum	2741	0,042829	-0,04	-0,87344	21,11901	-49,64	28,51
Irena	2741	0,028386	-0,02	0,244786	7,348733	-15,27	21,50
Krosno	2741	0,027584	0,03	-0,07044	6,190097	-11,71	12,45
Wólczanka	2741	0,033182	0,00	0,27263	9,27888	-27,51	24,69
Żywiec	2741	0,023871	0,03	0,133276	11,41527	-14,01	21,07
Agora	1609	0,0250437	-0,03	0,2639936	7,3951374	-14,57	15,72
TPSA	1713	0,024121	0,02	0,274729	4,606483	-10,18	12,78
Prokom	1907	0,0298922	0,01	0,2458866	6,8834953	-14,52	20,00
Computerland	2475	0,036492	0,12	0,179353	5,541283	-17,62	17,91
PKN Orlen	2457	0,018434	0,07	0,203877	0,854569	-7,16	8,46
KGHM	2052	0,028442	0,03	0,179836	6,794244	-17,50	17,96

Źródło: opracowanie własne.

W pierwszej kolejności przeprowadzono badania z zastosowaniem estymatorów nieparametrycznych. Dokonano próby estymacji indeksu ogona dla ujemnych stóp zwrotu. W tym celu stopy zwrotu zostały przemnożone przez (-1) .

Dla przedstawionych powyżej danych zostały obliczone estymatory indeksu ogona $\hat{\alpha}$. Rysunek 1 przedstawia estymatory Pickandsa, Hilla oraz Dekkersa–Einmahla–de Haana $\hat{\alpha}$ obliczone dla indeksów WIG i WIG20 oraz spółek Elektrim i Żywiec. Jak widać, trudno mówić w tych przypadkach o obszarze stabilizowania się estymatorów. Estymator Pickandsa charakteryzuje się zbyt dużą wariancją, aby mógł być uznany za użyteczny.



Rys. 1. Wykres wartości estymatorów Pickandsa, Hilla i Dekkersa–Einmahla–de Haana dla indeksów WIG i WIG20 oraz spółek Elektrim i Żywiec
Źródło: opracowanie własne.

Inne wyniki otrzymujemy po zastosowaniu estymatora Hilla. Jak pokazuje rys. 1, możemy wyróżnić obszary stabilizacji tych estymatorów. W przypadku estymatora dla indeksu WIG można zaobserwować stabilizowanie się wartości α na poziomie 2,5-3 dla $k \in [50; 200]$; w przypadku indeksu WIG20 stabilizację można zaobserwować na poziomie 2-3 dla $k \in [90; 350]$. W miarę wzrostu wartości k rośnie obciążenie estymatora, a jego wartość zaczyna maleć. Ponieważ estymatory Hilla stabilizują się na poziomie zbliżonym do liczby 2, nie można wykluczyć, że rozkład stóp zwrotu indeksu można by przybliżyć rozkładem normalnym (α -stabilnym dla α bliskiemu liczbie 2).

Ostatnim z zaprezentowanych estymatorów jest estymator Dekkersa–Einmahla–de Haana (DEdH), który charakteryzuje się wariancją większą od estymatora Hilla, a określanie obszaru jego stabilizacji na poziomie powyżej 2 nie daje oczekiwanych wyników. Jedynie w przypadku akcji Elektrimu można mówić o wyraźnym stabilizowaniu się wartości estymatora na poziomach 3-4.

Brak możliwości określenia obszarów stabilizowania się estymatorów może być spowodowany przede wszystkim zbyt krótkimi seriami danych. Może to również oznaczać, że grube ogony nie występują, na co wskazywałoby stabilizowanie się estymatora wokół liczby 2 (co ma miejsce w przypadku estymatora Hilla). W literaturze znane są również przykłady, gdy estymowana wartość parametru α jest

większa od 2, mimo że jego faktyczna wartość należy do przedziału od 1 do 2 (por. [Weron 2001]).

Kolejnym etapem badań była próba estymacji indeksów ogona metodą największej wiarygodności. W tym celu konieczne stało się wyznaczenie progów, powyżej których obserwacje uznaje się za należące do ogonów i wykorzystuje przy estymacji parametrów uogólnionego rozkładu Pareto. Standardowo dla wszystkich indeksów i akcji przyjęto, że obserwacje przekraczające próg u będą stanowiły odpowiednio 5 i 1% wszystkich obserwacji. Dodatkowo za pomocą wykresu zależności funkcji wartości oczekiwanej $e(u)$ od u wyznaczono progi u dla indeksów WIG i WIG20 oraz akcji odpowiednio na poziomach przedstawionych w tab. 2. Wykresy zależności dla przykładowych indeksów i akcji przedstawia rys. 2.

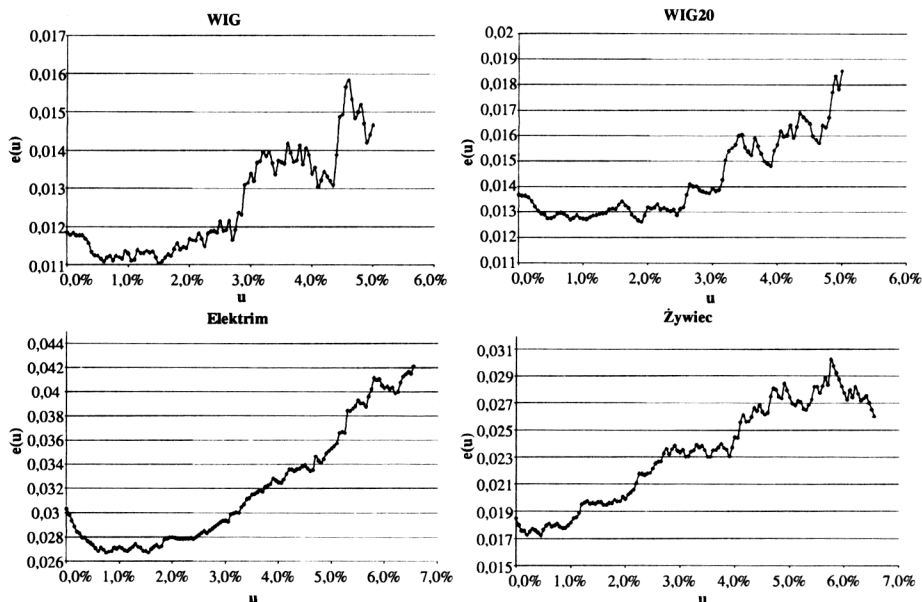
Tabela 2. Wartości parametru α dla wybranych progów u

Spółka	u (w)	N_u	α
WIG	2,60	136	6,71919
	2,70	127	4,95124
WIG20	1,90	344	7,19393
	2,45	221	5,29493
	3,02	136	6,6886
	3,10	128	4,94182
	3,90	71	4,88118
	4,65	42	3,624
Żywiec	2,00	328	4,82
	3,64	136	36,802
	3,90	123	11,928
Wólczanka	2,10	486	20,8904
	2,80	366	19,4585
	3,70	250	38,4657
	5,25	136	58,2827
Krosno	1,80	482	8,6626
	4,25	153	-2,3985
	4,47	136	-1,9983
Irena	2,35	403	7,3461
	2,95	285	8,3238
	3,75	189	7,3306
	4,26	136	-9,9732
Elektrim	2,40	579	4,2676
	4,65	233	3,1390
	4,80	219	3,1247
	5,97	136	3,70

Spółka	u (w)	N_u	α
Millennium	3,50	269	5,981
	4,89	136	7,222
BRE	2,70	323	9,405
	4,10	154	9,970
	4,26	136	88,515
	5,30	81	-11,038
ING	1,80	415	5,737
	3,50	163	549
	3,79	136	-11,703
Agora	2,10	251	5,136
	2,95	133	4,734
	3,75	79	4,288
	4,80	42	2,97
TPSA	3,82	85	-10,758
	4,85	41	151,43
Prokom	3,95	140	5,193
	4,59	95	6,514
	5,00	75	7,588
Computerland	2,20	423	6,005
	4,63	123	5,359
Orlen	2,86	72	-46,872
	4,20	18	3,193
KGHM	4,10	126	3,869
	4,32	102	7,922

Źródło: opracowanie własne.

Na tej podstawie oszacowano wartości parametru α dla rozkładów stóp zwrotu akcji i indeksów giełdowych.



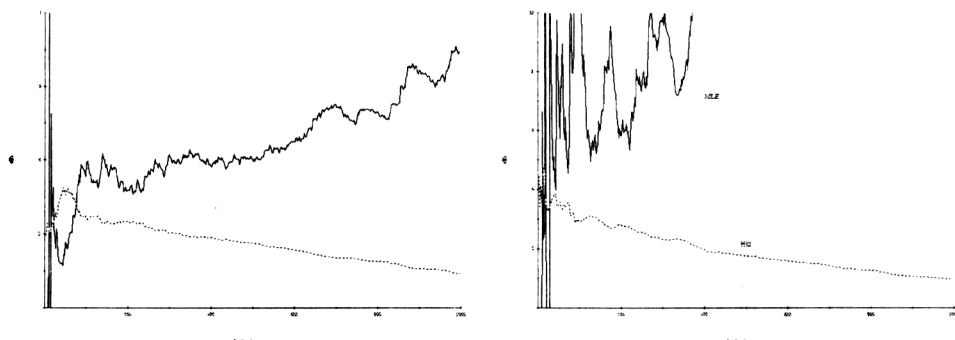
Rys. 2. Wykresy zależności $e(u)$ od u

Źródło: opracowanie własne.

Otrzymane wyniki pokazują, że również estymacja przeprowadzona metodą największej wiarygodności nie dała zadowalających rezultatów. Otrzymane wyniki są w znacznym stopniu zależne od wyboru progu, a wariancja estymatora w większości przypadków jest zbyt wysoka, aby można było zaakceptować otrzymane wyniki. Jedynie w przypadku Elektrimu możemy mówić o przedziale (3,5-4,5), w którym mieści się wartość parametru α . W pozostałych przypadkach możemy stwierdzić jedynie, że obszary stabilizacji (aczkolwiek nie tak wyraźne jak w przypadku Elektrimu) możemy zaobserwować dla wartości powyżej 2.

Elektrim

WIG20



Rys. 3. Estymator największej wiarygodności (MLE) i estymator Hilla dla Elektrimu i WIG20

Źródło: opracowanie własne.

Wyniki uzyskane metodą największej wiarygodności w większości przypadków wskazują na dość wysokie wartości parametru α , który dla poszczególnych akcji i indeksów wahał się w granicach 2-10. Dyskusyjne pozostają oszacowania indeksu ogona powyżej wartości 2, a nie – jak można się było spodziewać – w przedziale (0, 2). Może to jednak wskazywać, że rzeczywista jego wartość jest zbliżona do 2. Taką możliwość wskazuje również w swojej pracy Weron [2001].

5. Podsumowanie

Uzyskane wyniki potwierdzają, że badanie grubości ogonów rozkładów nie jest rzeczą łatwą, a otrzymane wyniki często są niejednoznaczne i bardzo wrażliwe na wybór liczby obserwacji wykorzystywanych do estymacji indeksu ogona. Wpływ na to ma głównie zbyt mała liczba danych z rynku polskiego. Z estymatorów nieparametrycznych jedynie estymator Hilla okazał się dość stabilny, co pozwalało na określenie jego obszaru stabilizowania się. Pozostałe dwa parametry w większości przypadków nie pozwoliły na określenie nawet przybliżonej wartości indeksu ogona α . Dla danych pochodzących z polskiego rynku również estymator największej wiarygodności dawał często rozbieżne wyniki, bardzo wrażliwe na wielkość próby.

Literatura

- Embrechts P., Klüppelberg C., Mikosch T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.
- Fofack H., Nolan J. (2000), *Tail Behavior, Modes and Other Characteristics of Stable Distributions*, maszynopis, American University, Washington.
- Gajek L., Kałuska M. (2000), *Wnioskowanie statystyczne*, WNT, Warszawa.
- Gilli M., Këllezi E. (2003), *An Application of Extreme Value Theory for Measuring Risk*, <http://www.gloriamundi.org>.
- Jajuga K. (2001), *Podstawy analizy wartości ekstremalnych na rynkach finansowych*, „Rynek Terminowy” nr 11, s. 123-127.
- Jajuga K. (red.) (2000), *Metody ekonometryczne i statystyczne w analizie rynku kapitałowego*, AE, Wrocław.
- Këllezi E., Gilli M. (2000), *Extreme Value Theory for Tail-Related Risk Measures*, maszynopis, <http://www.unige.ch/ses/metri/gilli/publicat.html>.
- McNeil A. (1999), *Extreme Value Theory for Risk Managers*, maszynopis, ETHZ, Zurich.
- Pickands J. (1995), *Statistical Inference Using Extreme Order Statistics*, „The Annals of Statistics” 3, s. 119-131.
- Theil H. (1979), *Zasady ekonometrii*, PWN, Warszawa.
- Weron A., Weron R. (1998), *Inżynieria finansowa*, WNT, Warszawa.

Weron R. (2001), *Levy-stable Distribution Revisited: Tail Index > 2 Does Not Exclude the Levy-stable Regime*, „International Journal of Modern Physics C” vol. 12, nr 2, s. 209-223.

TAIL INDEX ESTIMATION OF RETURNS DISTRIBUTION ON POLISH STOCK EXCHANGE

Summary

In this paper a study on tail index estimation. Some data sets requiring heavy tailed models necessitates reliable methods of estimation α parameter governing the degree of tail heaviness In this paper three estimators Hill, Pickands and Dekkers-Einmahl-de Haan and the method of maximum likelihood estimation have been verified on Polish Stock Exchange.