PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU
Nr 18 — 2008
Inwestycje finansowe i ubezpieczenia
– tendencje światowe a polski rynek

Bartosz Kaszuba

Uniwersytet Ekonomiczny we Wrocławiu

ODPORNE METODY ESTYMACJI WSPÓŁCZYNNIKA BETA NA PRZYKŁADZIE POLSKIEGO RYNKU AKCJI

1. Wstęp

Ryzyko jest jednym z najważniejszych pojęć teoretycznych w naukach ekonomicznych, a zwłaszcza w naukach o finansach [Jajuga (red.) 2007]. Jednym z częściej stosowanych mierników ryzyka jest współczynnik beta zawarty w jednowskaźnikowym modelu Sharpe'a. Najczęściej wykorzystywaną metodą do wyliczenia współczynnika beta jest klasyczna metoda najmniejszych kwadratów (KMNK), która ma dużą wrażliwość na obserwacje nietypowe. Wykorzystywanie metody KMNK na polskim rynku akcji, który cechują duża zmienność i częste występowanie obserwacji nietypowych, może spowodować błędne decyzje inwestycyjne.

Celem pracy jest zaimplementowanie wybranych metod odpornych do wyznaczania współczynnika beta dla najdłużej notowanych spółek o wysokiej kapitalizacji, wchodzących w skład indeksu WIG20 na Giełdzie Papierów Wartościowych w Warszawie. Opierając się na rzeczywistych danych historycznych, wyliczono współczynniki beta i zbadano ich stabilność w okresie od 1.07.1998 do 20.02.2008. W artykule podjęto próbę odpowiedzi na pytanie, czy dla polskiego rynku akcji stosowanie metod odpornych pozwoli uzyskać bardziej stabilne estymatory parametrów beta niż te wyznaczane metodą KMNK. Do porównania współczynników beta użyto metod opartych na M-estymatorach z funkcjami Hubera i podwójnie kwadratową, metodę najmniejszej mediany kwadratów oraz metodę najmniejszych uciętych kwadratów.

2. Podstawowe pojęcia odporności

Głównym celem stosowania metod odpornych jest poprawa wyników estymacji parametrów służących do budowy modelu [Ostasiewicz 1998]. W wielu modelach przyjmuje się założenia dotyczące rozkładu badanej próby. Jednak w praktyce model dany pewnym rozkładem F często jest zakłócany przez obserwacje nietypowe, które pochodzą z nieznanego rozkładu G. Zatem, badając pewne właściwości modelu o dystrybuancie F, powinniśmy rozważać model ε-zaburzony:

$$F_{\varepsilon}(G) = (1-\varepsilon)F + \varepsilon G \text{, gdzie } \varepsilon \in [0;1].$$

Jako miarę globalnej odporności estymatorów możemy użyć punktu załamania ε^* (*break down point*), który można interpretować następująco: jest to najmniejszy udział danych zakłócających, dla których estymator może przyjmować dowolnie duże wartości.

Estymatory odporne posiadające punkt załamania bliski 0,5 nazywane są estymatorami o wysokim punkcie załamania.

Kolejnym zagadnieniem mającym znaczenie przy wyborze odpowiednich estymatorów jest ich efektywność (*efficiency*). Efektywność estymatora $\hat{\theta}$ parametru θ może być zdefiniowana następująco:

$$eff(\hat{\theta}) = \frac{v_0}{v},$$

gdzie v_0 jest asymptotyczną wariancją estymatora największej wiarygodności parametru θ, natomiast v jest asymptotyczną wariancją odpornego estymatora parametru θ. Zatem w przypadku, gdy liczba zaburzeń jest niewielka, preferowane są estymatory o wysokiej efektywności.

W dalszej części pracy podawana efektywność będzie tzw. efektywnością gaussowską, czyli obliczaną przy założeniu, że próba pochodzi z rozkładu normalnego.

3. Metody estymacji parametrów w modelu regresji liniowej

Rozważmy następujący model liniowy:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon,$$

gdzie $\mathbf{Y}' = (Y_1,\dots,Y_n)$, $\mathbf{X}$ jest $n \times p$ wymiarową macierzą, $\beta' = (\beta_1,\dots,\beta_p)$ jest wektorem nieznanych parametrów oraz $\varepsilon' = (\varepsilon_1,\dots,\varepsilon_n)$ jest wektorem błędów losowych o takim samym rozkładzie F, mającym średnią 0 i nieznaną wariancję σ^2.

Najpopularniejszą metodą estymacji parametrów β jest klasyczna metoda najmniejszych kwadratów pochodząca od Gaussa lub Legendre'a (do dziś prowadzone są dyskusje historyczne dotyczące twórcy tej metody [Stigler 1981]). KMNK może być opisana następująco:

$$\min_{\beta} \sum_{i=1}^{n} e_i^2,$$

gdzie $e_i = y_i - (\hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_p x_{ip})$ oraz y_i jest wartością zmiennej losowej Y_i.

Mimo że powyższa metoda odgrywa ważną rolę w statystyce, to jednak jest często krytykowana za duże braki w odporności [Rousseeuw 1984]. Jedną z przyczyn braku odporności estymatora KMNK są założenia dotyczące rozkładu składnika losowego [Chan, Lakonishok 1992]. Gdy jednak składnik losowy ma rozkład normalny, wtedy estymator parametrów otrzymany metodą najmniejszych kwadratów ma najmniejszą wariancję w klasie estymatorów nieobciążonych [Rao 1973].

Ponadto, gdy korzysta się z nierówności Jensena, procedura optymalizacyjna KMNK (przy założeniu gaussowskich rozkładów składnika losowego) może być rozszerzona do każdej wypukłej funkcji straty [Rao 1973]. Jeśli natomiast rozkład błędów losowych ma grubsze ogony niż rozkład normalny, to estymatory MNK są nieefektywne oraz mają wysoką wrażliwość na obserwacje odstające i nawet jedna obserwacja odstająca może mieć duży wpływ na postać estymatora.

Obecność obserwacji odstających może być modelowana przez założenie, że rozkład błędów losowych F jest mieszanką dwóch różnych rozkładów, gdzie jeden z rozkładów reprezentuje „dobre" dane, np. standaryzowany rozkład normalny, drugi zaś reprezentuje błędne dane, np. rozkład normalny o wariancji przekraczającej 1. Wtedy rozkład F ma grubsze ogony niż standaryzowany rozkład normalny. W takim przypadku lepiej sprawdzają się metody odporne. Jest to związane z tym, że estymatory odporne nadają mniejsze wagi obserwacjom odstającym, np. przez minimalizację sumy bezwzględnych odchyleń (L1), zamiast minimalizować sumę kwadratów (KMNK). Ponadto można pokazać, że w przypadku estymatora KMNK punkt załamania wynosi $\varepsilon^* = 0$. Zatem wskazane jest użycie bardziej odpornych estymatorów, takich jak estymator otrzymany **metodą najmniejszych bezwzględnych odchyleń** (L1), który dany jest wzorem:

$$\min_{\beta} \sum_{i=1}^{n} |e_i| .$$

Mimo że estymatory L1 są odporne na obserwacje odstające y_i, to jednak nie są odporne na duże odchylenia w próbie $x_i = (x_{i1}, \ldots, x_{in})$, które mają duży wpływ na dopasowanie prostej regresji. Ponadto można pokazać, że punkt załamania L1, tak jak w przypadku MNK, wynosi $\varepsilon^* = 0$.

Kolejnym etapem tworzenia metod odpornych w regresji liniowej są **M-estymatory** [Huber 1973], które otrzymuje się przez rozwiązanie następującego wyrażenia:

$$\min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}}\right),$$

gdzie $\rho(x)$ jest pewną regularną funkcją. Różniczkując względem $\hat{\beta}$, otrzymujemy:

$$\sum_{i=1}^{n} \psi\left(\frac{e_i}{\hat{\sigma}}\right) x_i = 0,$$

gdzie $\psi = \rho'$ (w praktyce rozwiązanie powyższego równania możemy otrzymać, stosując metodę Newtona-Raphsona lub iteracyjną metodę zmiennych wag najmniejszych kwadratów [Rousseeuw 1984]). Jako funkcję ρ możemy użyć funkcję Hubera:

$$\rho(x) = \begin{cases} x^2 & \text{jeżeli } |x| \le \text{k} \\ 2k|x| - k^2 & \text{jeżeli } |x| > \text{k} \end{cases},$$

gdzie k jest pewną stałą. Z postaci funkcji Hubera wynika, że dla $k = \infty$ otrzymany estymator będzie estymatorem KMNK.

Jako ρ można również użyć funkcję „opadającą" (*redescending*), której wagi (dla dużych wartości) zbliżają się do 0. Wynika stąd, że na estymację parametrów nie mają wpływu reszty przekraczające pewien ustalony zakres. Tego typu funkcją jest następująca funkcja podwójnie kwadratowa (*bisquare, biweight*):

$$\rho(x) = \begin{cases} 1 - \left[1 - \left(\dfrac{x}{k}\right)^2\right]^2 & \text{jeżeli } |x| \le \text{k} \\ 1 & \text{jeżeli } |x| > \text{k} \end{cases},$$

stałą k wybiera się w zależności od preferowanej efektywności.

Stałe k są często nazywane stałymi regulującymi (*tuning constant*); mniejsze wartości k zwiększają odporność estymatorów, kosztem jest jednak strata efektywności, gdy rozkład błędów losowych jest rozkładem normalnym. Stałe regulujące są na ogół wybierane tak, aby estymatory dawały wysoką efektywność w przypadku normalnego rozkładu błędów. W praktyce $k = 1{,}345\sigma$ dla funkcji Hubera oraz $k = 4{,}685\sigma$ dla funkcji podwójnie kwadratowej (gdzie σ jest odchyleniem standardowym błędów). Taki wybór stałych daje 95% efektywność oraz wciąż chroni przed obserwacjami odstającymi [Chen, Hong 2005]. Pomimo jednak, że M-estymatory są bardziej odporne, to również i w tym przypadku $\varepsilon^* = 0$.

Bardziej odporna metoda estymacji parametrów w modelu liniowym polega na minimalizacji miary parametru skali reszt, która jest niewrażliwa na duże wartości. Taką miarą może być mediana.

Powyższe rozumowanie prowadzi nas do metody **najmniejszej mediany kwadratów** (LMS – *least median of squares*):

$$\min_{\beta} med\{e_i^2 : i = 1, \ldots, n\}.$$

Punkt załamania estymatora LMS wynosi $\varepsilon^* \approx 0{,}5$, jednak kosztem wysokiej odporności jest on bardzo nieefektywny dla dużych n [Maronna, Martin, Yohai 2006; Rousseeuw 1984].

Inną odporną metodą jest metoda **najmniejszych uciętych kwadratów** (LTS – *least trimmed squares*) zdefiniowana następująco:

$$\min_{\hat{\beta}} \sum_{i=1}^{h} e_{i:n}^2 ,$$

gdzie $e_{1:n}^2 \leq e_{2:n}^2 \leq \ldots \leq e_{n:n}^2$. Natomiast h jest pewną ustaloną stałą. Gdy $h = [n/2]+1$, wtedy punkt załamania wynosi $\varepsilon^* \approx 0{,}5$. Jednak również i w tym przypadku asymptotyczna efektywność jest bardzo niska kosztem wysokiej odporności estymatora [Maronna, Martin, Yohai 2006]. W teorii możemy znaleźć różne modyfikacje LTS, takie jak metoda najmniejszych α-uciętych kwadratów, która minimalizuje sumę kwadratów reszt, pomijając 0,5α% najmniejszych i 0,5α% największych wartości. Pokazano, że efektywność estymatora LTS wynosi ok. 7% [Maronna, Martin, Yohai 2006].

4. Wyniki empiryczne

W badaniach za reprezentanta rynku przyjęto indeks WIG, który obejmuje wszystkie spółki giełdowe, z wyjątkiem tych, które nie spełniają minimalnych wymogów płynności. Natomiast spółkami, które uwzględniono w badaniach, były BRE Bank, BUDIMEX, BZWBK, CERSANIT, KGHM, PEKAO, PROKOM. Na wybór spółek miała wpływ zarówno płynność danej spółki, jak również data debiutu, która determinuje długość okresu badań.

W pracy rozważano dzienne, logarytmiczne stopy zwrotu dla cen zamknięcia z poszczególnych spółek w okresie od 1.07.1998 do 20.02.2008, co daje 2420 notowań.

Na wstępie testem Dickeya-Fullera zbadano stacjonarność szeregów czasowych stóp zwrotu omawianych spółek. Otrzymane wyniki były statystycznie istotne na poziomie 0,01, zatem szeregi stóp zwrotu rozważanych spółek są stacjonarne. Szczegółowe wyniki znajdują się w tab. 1.

Tabela 1. Wartości statystyk testu Dickeya-Fullera wybranych indeksów w okresie od 1.07.1998 do 20.02.2008

Okres	BRE	BUDIMEX	BZWBK	CERSANIT	KGHM	PEKAO	PROKOM	WIG
1998-2007	-11,738	-13,348	-14,187	-12,305	-13,519	-14,201	-12,862	-12,056
1999	-7,011	-7,019	-6,407	-6,179	-6,665	-7,263	-5,743	-6,213
2000	-7,237	-7,085	-7,303	-5,532	-7,091	-6,872	-6,784	-7,236
2001	-5,382	-5,005	-5,472	-6,458	-5,517	-7,065	-5,928	-5,878
2002	-4,954	-6,796	-6,699	-7,349	-5,568	-5,886	-5,914	-5,658
2003	-4,790	-5,948	-6,950	-4,952	-5,994	-5,278	-5,481	-5,150
2004	-6,375	-6,772	-6,798	-7,583	-5,821	-7,624	-6,463	-7,005
2005	-5,988	-6,504	-6,103	-7,461	-6,774	-5,306	-5,480	-5,416
2006	-5,673	-6,649	-5,585	-5,721	-5,334	-5,534	-6,054	-5,818
2007	-9,197	-5,790	-6,997	-5,757	-6,653	-7,685	-5,996	-6,476

Źródło: opracowanie własne na podstawie danych z GPW w Warszawie.

Kolejnym etapem było zbadanie współczynnika korelacji. Otrzymane współczynniki były istotne statystycznie na poziomie 0,01.

Najsilniej skorelowany z indeksem WIG w całym badanym okresie, jak również w poszczególnych latach, był bank PEKAO. Najsłabiej skorelowaną spółką była spółka CERSANIT.

Tabela 2. Współczynniki korelacji logarytmicznych stóp zwrotu indeksu WIG z innymi spółkami w okresie od 1.07.1998 do 20.02.2008

Okres	BRE	BUDIMEX	BZWBK	CERSANIT	KGHM	PEKAO	PROKOM
1998-2008	0,45	0,35	0,51	0,36	0,59	0,60	0,52
1999	0,36	0,42	0,40	0,44	0,42	0,51	0,35
2000	0,35	0,40	0,37	0,39	0,41	0,47	0,38
2001	0,37	0,40	0,40	0,35	0,47	0,49	0,46
2002	0,40	0,38	0,43	0,35	0,51	0,52	0,50
2003	0,42	0,39	0,45	0,34	0,54	0,54	0,52
2004	0,42	0,38	0,46	0,34	0,55	0,55	0,53
2005	0,42	0,36	0,47	0,33	0,55	0,56	0,52
2006	0,43	0,35	0,49	0,34	0,57	0,58	0,52
2007	0,44	0,34	0,50	0,35	0,58	0,59	0,52

Źródło: opracowanie własne na podstawie danych z GPW w Warszawie.

Z tabeli 2 można odczytać, że współczynniki korelacji pomiędzy stopami zwrotu indeksu WIG a stopami zwrotu spółek nie przekraczały wielkości 0,60. Ponadto korelacje w badanym okresie utrzymywały się na podobnym poziomie w stosunku do indeksu WIG.

5. Badanie stabilności współczynników beta

Zgodnie z początkowymi założeniami pracy badaniu poddano szczególny przypadek modelu zależności liniowej, mianowicie jednowskaźnikowy model Sharpe'a dany następującym wzorem:

$$r_i = \alpha_i + \beta_i r_M + \varepsilon,$$

gdzie r_i to stopa zwrotu dla:

$i = 1$ BRE Bank, $i = 2$ Budimex, $i = 3$ BZWBK, $i = 4$ Cersanit,

$i = 5$ KGHM, $i = 6$ PEKAO, $i = 7$ Prokom,

oraz r_M to stopa zwrotu indeksu WIG.

W pracy porównywano współczynniki beta otrzymane następującymi metodami:

β^{MNK} – klasyczną metodą najmniejszych kwadratów (MNK),

β^{LMS} – metodą najmniejszej mediany kwadratów (LMS),

β^{LTS} – metodą najmniejszych uciętych kwadratów (LTS),

β^{hub} – M-estymatorem z funkcją Hubera,

β^{pk} – M-estymatorem z funkcją podwójnie kwadratową.

Tabela 3. Wyliczone kwantyle różnic dla estymatorów MNK, LMS, LTS, M-estymatorów z funkcją Hubera i podwójnie kwadratową

Miesiące		BRE			BUDIMEX			BZWBK			CERSANIT			KGHM			PEKAO			PROKOM		
		$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$
$k = 3$	MNK	0,375	0,709	0,767	0,491	0,830	1,117	0,280	0,537	0,883	0,413	0,709	0,875	0,445	0,717	0,802	0,308	0,494	0,695	0,452	0,704	0,761
	LMS	0,739	1,118	1,193	0,733	1,466	2,014	0,831	1,211	1,587	0,580	1,019	1,528	0,699	1,083	1,269	0,562	1,345	1,591	0,641	1,245	1,541
	LTS	0,436	0,694	0,817	0,478	0,739	0,962	0,341	0,613	0,722	0,454	0,752	0,856	0,427	0,666	0,861	0,299	0,591	0,826	0,484	0,731	0,862
	hub	0,373	0,558	0,753	0,467	0,721	0,995	0,325	0,519	0,785	0,387	0,622	0,752	0,430	0,625	0,805	0,311	0,535	0,699	0,414	0,645	0,741
	pk	0,406	0,570	0,748	0,467	0,768	1,055	0,349	0,565	0,773	0,375	0,639	0,846	0,421	0,651	0,868	0,304	0,533	0,696	0,431	0,660	0,804
$k = 6$	MNK	0,317	0,649	0,678	0,359	0,514	0,616	0,239	0,373	0,425	0,322	0,552	0,583	0,410	0,612	0,711	0,269	0,451	0,497	0,348	0,586	0,753
	LMS	0,575	0,872	0,948	0,522	0,979	1,553	0,498	0,950	1,146	0,435	0,686	1,032	0,832	1,170	1,335	0,443	0,962	1,144	0,615	1,008	1,218
	LTS	0,247	0,419	0,597	0,419	0,597	0,683	0,341	0,470	0,500	0,317	0,482	0,568	0,424	0,682	0,773	0,338	0,459	0,586	0,348	0,559	0,731
	hub	0,298	0,481	0,556	0,358	0,516	0,694	0,265	0,381	0,425	0,259	0,415	0,526	0,404	0,624	0,667	0,301	0,460	0,533	0,319	0,562	0,690
	pk	0,255	0,421	0,649	0,369	0,544	0,704	0,273	0,406	0,445	0,282	0,449	0,535	0,373	0,588	0,656	0,299	0,455	0,545	0,324	0,571	0,718
$k = 12$	MNK	0,338	0,474	0,535	0,326	0,442	0,484	0,283	0,390	0,419	0,247	0,348	0,390	0,391	0,904	1,001	0,331	0,533	0,576	0,432	0,722	0,769
	LMS	0,399	0,666	0,776	0,455	0,728	0,823	0,377	0,536	0,711	0,330	0,609	0,677	0,583	1,075	1,167	0,360	0,654	0,746	0,446	1,118	1,320
	LTS	0,327	0,476	0,522	0,342	0,433	0,512	0,291	0,475	0,526	0,226	0,345	0,376	0,434	0,962	0,993	0,390	0,527	0,573	0,449	0,868	0,986
	hub	0,294	0,378	0,460	0,307	0,410	0,473	0,284	0,420	0,436	0,185	0,303	0,324	0,421	0,904	0,974	0,345	0,510	0,549	0,366	0,824	0,876
	pk	0,292	0,388	0,477	0,316	0,420	0,476	0,296	0,429	0,444	0,172	0,289	0,334	0,398	0,874	0,945	0,350	0,477	0,554	0,373	0,875	0,944
$k = 18$	MNK	0,376	0,506	0,551	0,320	0,444	0,503	0,280	0,529	0,561	0,234	0,334	0,384	0,405	0,957	0,996	0,341	0,600	0,622	0,427	0,892	1,006
	LMS	0,401	0,623	0,683	0,340	0,580	0,657	0,367	0,629	0,686	0,273	0,502	0,619	0,640	1,104	1,669	0,426	0,606	0,674	0,607	1,386	1,635
	LTS	0,332	0,445	0,498	0,381	0,493	0,556	0,313	0,594	0,623	0,210	0,308	0,342	0,480	0,960	1,005	0,381	0,607	0,633	0,437	1,080	1,221
	hub	0,310	0,406	0,441	0,311	0,438	0,516	0,295	0,532	0,554	0,183	0,293	0,342	0,439	0,958	0,987	0,354	0,581	0,618	0,389	0,996	1,141
	pk	0,293	0,397	0,429	0,317	0,455	0,540	0,296	0,549	0,573	0,190	0,296	0,360	0,422	0,950	0,972	0,385	0,584	0,627	0,385	1,056	1,208
$k = 24$	MNK	0,459	0,554	0,584	0,256	0,285	0,324	0,401	0,547	0,569	0,230	0,321	0,375	0,611	0,940	0,960	0,422	0,618	0,631	0,425	0,981	0,994
	LMS	0,553	0,726	0,759	0,405	0,535	0,688	0,325	0,458	0,575	0,217	0,523	0,624	0,703	1,470	1,544	0,502	0,593	0,684	0,564	1,204	1,412
	LTS	0,318	0,458	0,530	0,294	0,400	0,445	0,442	0,528	0,559	0,198	0,249	0,310	0,601	0,940	0,963	0,510	0,651	0,664	0,438	0,994	1,074
	hub	0,343	0,449	0,461	0,248	0,307	0,356	0,425	0,538	0,549	0,196	0,268	0,318	0,598	0,940	0,960	0,457	0,608	0,616	0,430	1,028	1,045
	pk	0,296	0,407	0,426	0,269	0,343	0,400	0,444	0,551	0,561	0,209	0,292	0,299	0,620	0,947	0,975	0,494	0,614	0,624	0,424	1,054	1,078

Źródło: opracowanie własne.

Dla metody LTS za stałą regulującą przyjęto $h = [0{,}95n] + 1$, gdzie $[x]$ oznacza część całkowitą z x. Natomiast dla M-estymatora z funkcją Hubera wybrano $k = 1{,}345$, zaś dla M-estymatora z funkcją podwójnie kwadratową $k = 4{,}685$. Przy tak dobranych stałych otrzymane estymatory są bardzo efektywne i wciąż pozostają odporne.

Pierwszym etapem badań było zbadanie zasadności stosowania klasycznej metody najmniejszych kwadratów w jednowskaźnikowym modelu Sharpe'a. Zbadano stopień zależności liniowej między zmienną objaśnianą a objaśniającą. Średnie współczynniki determinacji w badanym okresie dla modelów liczonych w trzymiesięcznych podokresach były niższe niż 0,435, co świadczy o słabym dopasowaniu modelu. Najwyższe średnie współczynniki determinacji posiadały spółki KGHM, PEKAO i PROKOM (odpowiednio 0,435, 0,426 i 0,359). Najniższe współczynniki miały spółki BUDIMEX i CERSANIT (0,114 i 0,137 odpowiednio). Analizie poddano również reszty dla modelu regresji, w której zbadano normalność reszt w rozpatrywanym okresie. Do badania użyto testu Shapiro-Wilka. Dla prób badanych za 12 miesięcy w większości przypadków odrzucono hipotezę o normalności rozkładu reszt na poziomie istotności 0,05. Zatem z powyższych badań wynika, że stosowanie metody KMNK nie jest pożądane na polskim rynku akcji.

Kolejnym etapem badań było porównanie stabilności otrzymanych współczynników w czasie. Stabilność jest rozumiana tu jako bezwzględna wartość różnicy $\beta_{t_k} - \beta_{t_{k-1}}$, gdzie β_{t_k} – współczynnik beta wyliczony za okres t, w momencie k. Współczynniki beta były obliczane za 3miesiące, 6, 12, 18 miesięcy oraz za 24 miesiące, a różnice były brane na następny okres.

Badano, jak zmieniają się poszczególne współczynniki beta obliczone za okres n miesięcy w ciągu następnych n miesięcy. Zmiany rozpatrywano w odstępie comiesięcznym.

Dane w tabeli 3 przedstawiono w postaci kwantyli. Wartości zostały wyliczone w sposób następujący: dla danego okresu n miesięcy wyliczono współczynnik beta, a następnie odjęto wartość współczynnika beta wyliczonego za okres n miesięcy w ciągu następnych n miesięcy. Badane różnice były wyliczane w odstępach comiesięcznych, zatem im mniejsza bezwzględna wartość różnic, tym większa stabilność danego estymatora. Następnie dla danego rozkładu różnic wyliczono kwantyle (75%, 95%, 99% oznaczone odpowiednio jako $q_{0.75}$, $q_{0.95}$, $q_{0.99}$) i umieszczono w tab. 3. Obliczenia były wykonywane dla estymatorów MNK, LMS, LTS, M-estymatorów z funkcjami Hubera i podwójnie kwadratową.

Jak wynika z tab. 3, najgorsze wyniki osiągały estymatory LMS, natomiast najbardziej stabilnymi estymatorami były M-estymatory.

W tabeli 4 znajdują się zbiorcze wyniki z tab. 3. Każdy kwantyl porównano z pozostałymi w danej grupie, a następnie została im przyporządkowana ranga od 1 (najwyższe wartości) do 5 (najniższe wartości). Następnie wyliczono średnią z otrzymanych rang, co przedstawiono w tab. 4.

Tabela 4. Zbiorcze wyniki z tab. 2. Kwantylom zostały przyporządkowane rangi od 1 (najgorszy) do 5 (najlepszy). Następnie wyliczono średnie wartości rang dla każdego z estymatorów, *k* – liczba miesięcy

Wyszczególnienie	$k = 3$			$k = 6$			$k = 12$			$k = 18$			$k = 24$		
	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$	$q_{0.75}$	$q_{0.95}$	$q_{0.99}$
MNK	2,75	2,63	2,88	3,00	3,13	2,88	3,13	2,88	2,63	3,13	3,13	3,13	2,88	2,75	2,88
LMS	0,88	0,88	0,88	0,88	0,88	0,88	1,13	0,88	0,88	1,00	1,00	0,88	1,63	1,88	0,88
LTS	2,63	2,25	2,75	2,25	2,75	2,50	1,88	2,25	2,25	2,00	1,88	2,13	2,38	2,75	2,63
hub	3,38	4,13	3,50	3,63	3,13	3,88	3,63	3,63	4,00	3,63	3,75	3,88	3,50	3,38	3,75
pk	3,50	3,25	3,13	3,38	3,25	3,00	3,38	3,50	3,38	3,38	3,38	3,13	2,75	2,38	3,00

Źródło: opracowanie własne.

Z danych zawartych w tab. 3 i 4, wynika, że najbardziej stabilnym estymatorem dla badanego okresu był M-estymator z funkcją Hubera. Wysoką stabilnością charakteryzował się również drugi z M-estymatorów.

Z przeprowadzonych badań można również wnioskować, że lepszymi estymatorami współczynnika beta (w sensie stabilności) są estymatory odporne, jednak stosując inne podejście niż klasyczna metoda najmniejszych kwadratów, należy mieć na uwadze, że rozpatrujemy inne miary ryzyka.

Rozpatrując również badane estymatory w podziale na spółki, można zauważyć wyraźnie większą stabilność M-estymatora z funkcją Hubera niż w przypadku pozostałych estymatorów.

W tabeli 5 znajdują się wyliczone współczynniki beta za okres od 20.08.2006 do 20.05.2007 dla wcześniej rozpatrywanych spółek. Wyliczono oceny estymatorów KMNK oraz M-estymatora z funkcją Hubera. Dodatkowo, aby porównać wyliczone współczynniki beta, wyliczono dla każdego z indeksów wartość średnią w następnym okresie (od 20.05.2007 do 20.02.2008), a następnie wyliczono iloraz wartości średniej danej spółki i wartości średniej indeksu WIG. Tak wyliczony iloraz pokazuje, jak średnio zachowywały się stopy zwrotu danej spółki w porównaniu z indeksem WIG.

Ponieważ współczynnik beta wskazuje, o ile jednostek w przybliżeniu wzrośnie stopa zwrotu akcji, gdy stopa zwrotu wskaźnika rynku wzrośnie o jednostkę, wyliczony iloraz posłuży do oceny trafności wybranych metod.

Tabela 5. Współczynniki beta wyliczone za okres 20.08.2006 do 20.05.2007

Wyszczególnienie	BRE	BUDIMEX	BZWBK	CERSANIT	KGHM	PEKAO	PROKOM
MNK	1,033	0,835	1,322	0,929	1,630	1,254	1,059
hub	0,969	0,574	1,316	0,672	1,617	1,288	0,958
$\bar{x}_i / \bar{x}_{WIG}$	**0,604**	1,618	2,341	1,872	-0,229	1,149	**0,749**

Źródło: opracowanie własne.

Z danych przedstawionych w tab. 5 wynika, że błędne wyliczenie współczynnika beta może spowodować niewłaściwą strategię inwestycyjną. W przypadku, gdy inwestor rozważałby stworzenie portfela z akcji mniej ryzykownych niż indeks rynku, opierając się na KMNK, wybrałby akcje spółek CERSANIT i BUDIMEX, które (jak się okazało w następnym okresie) były bardziej ryzykowne niż indeks WIG. Natomiast w skład portfela inwestora opierającego się na M-estymatorach weszłyby dodatkowo spółki BRE i PROKOM, które były mniej ryzykowne niż indeks rynku.

Zaprezentowane w artykule rozważania pokazują, że wybór metody służącej do wyliczenia współczynnika beta jest kluczowy.

6. Podsumowanie

Przeprowadzone badania potwierdziły, że dla polskiego rynku akcji stosowanie metod odpornych pozwoli uzyskać estymatory parametrów beta bardziej stabilne niż wyznaczane metodą KMNK. Spośród przeanalizowanych estymatorów odpornych na szczególną uwagę zasługuje M-estymator z funkcją Hubera, którego cechowała największa stabilność. Pokazano również, że stosowanie klasycznej metody najmniejszych kwadratów dla jednowskaźnikowego modelu Sharpe'a jest niepożądane na polskim rynku akcji, ponieważ dopasowanie modelu jest małe, rozkład reszt nie spełnia zaś założeń tej metody (nie jest rozkładem normalnym).

Otrzymane wyniki bazujące na polskim rynku kapitałowym powinny zachęcić inwestorów do stosowania bardziej skutecznych i efektywnych metod, takich jak metody odporne. Również przeprowadzone badania na akcjach zagranicznych spółek [Carroll, Ruppert 1980; Chan, Lakonishok 1992; Rousseeuw 1984; Shanken 1992] dawały lepsze rezultaty dla estymatorów odpornych niż dla estymatorów KMNK. Należy jednak pamiętać, że model Sharpe'a ma u podstaw założenie, że kształtowanie się stóp zwrotu akcji na rynku zależy od działania pewnego ogólnego czynnika charakteryzującego rynek akcji [Jajuga, Jajuga 2007]. W przypadku przeświadczeń inwestora, że rynek zależy od więcej niż jednego czynnika, zaleca się zatem przeprowadzenie podobnych badań.

Literatura

Carroll R.J., Ruppert D., *Trimmed Least Squares Estimation in the Linear Model*, „Journal of the American Statistical Association" 1980.

Chan L.K.C., Lakonishok J., *Robust Measurement of Beta Risk*, „The Journal of Financial and Quantitative Analysis" 1992.

Chen S., Hong X.A., *Forward Regression Algorithm Based on M-Estimators*, [w:] *2005 WSEAS Int. Conf. Dynamical Systems and Control*, 2-4 November 2005, Venice 2005.

Huber P.J., *Robust Regression: Asymptotics, Conjectures and Monte Carlo*, „Annals of Statistics" 1973, 1.

Jajuga K. (red.), *Zarządzanie ryzykiem*, PWN, Warszawa 2007.

Jajuga K., Jajuga T., *Inwestycje finansowe*, PWN, Warszawa 2007.

Kowerski M., *Ryzyko na warszawskiej Giełdzie Papierów Wartościowych*, „Zamojskie Studia i Materiały" 2003, 3.
Maronna R.A., Martin R.D., Yohai V.J. , *Robust Statistics Theory and Methods*, John Wiley & Sons Ltd, 2006.
Ostasiewicz W., *Statystyczne metody analizy danych*, AE, Wrocław 1998.
Rao C.R., *Linear Statistical Inference and Its Applications,* John Wiley, New York 1973.
Rousseeuw P.J., *Least Median of Squares Regression*, „Journal of the American Statistical Association" 1984, vol. 79, nr 388.
Shanken J., *On the Estimation of Beta-Pricing Models*, „The Review of Financial Studies", 1992, vol. 5, nr 1.
Stigler S.M., *Gauss and the Invention of Least Squares*, „The Annals of Statistics" 1981, vol. 9, nr 3.

ROBUST MEASUREMENT OF BETA RISK

Summary

Many empirical studies find that the distribution of stock returns departs from normality. In such cases, it is desirable to employ a statistical estimation procedure that may be more efficient than ordinary least squares (OLS). The main goal of this article is compare OLS estimation with various robust methods in the context of estimating beta risk.

This article studies the stability of the beta risks in single index Sharp's model, calculated using robust methods and OLS estimation. The stability of the beta risks was examined for 7 companies included in the WIG20 index, basing on Stock Exchange quotations since 1997-07-01.

Futhermore, this paper describes and compares various estimators of the parameters of linear regression models. This article describe least absolute values methods, Huber M-estimation, bisquare M-estimation, least median of squares and least trimmed mean estimation.