PRACE NAUKOWE
Uniwersytetu Ekonomicznego we Wrocławiu
RESEARCH PAPERS
of Wrocław University of Economics

323

Inwestycje finansowe i ubezpieczenia – tendencje światowe a rynek polski

Redaktorzy naukowi
Krzysztof Jajuga
Wanda Ronka-Chmielowiec

Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2013

Redaktor Wydawnictwa: Agnieszka Flasińska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:
www.ibuk.pl, www.ebscohost.com,
w Dolnośląskiej Bibliotece Cyfrowej www.dbc.wroc.pl,
The Central and Eastern European Online Library www.ceeol.com,
a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon
http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się
na stronie internetowej Wydawnictwa
www.wydawnictwo.ue.wroc.pl

Kopiowanie i powielanie w jakiejkolwiek formie
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2013

ISSN 1899-3192
ISBN 978-83-7695-351-9

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp 11

Adam Adamczyk: Poziom wewnętrznych źródeł finansowania jako determinanta inwestycji w działalność B + R przedsiębiorstw 13
Roman Asyngier: Ekonomiczne i prawne aspekty nieprawidłowości funkcjonowania rynku NewConnect. Ocena i propozycje zmian 23
Jacek Białek: Zastosowanie autorskiego indeksu wydajności pracy do analizy dynamiki cen jednostek rozrachunkowych OFE 34
Magdalena Chmielowiec-Lewczuk: Zrównoważona Karta Wyników w zakładzie ubezpieczeń 43
Dawid Dawidowicz: Ocena efektywności nowych i pozostałych funduszy inwestycyjnych akcji polskich w latach 2000–2012 53
Ewa Dziwok: Weryfikacja modeli krzywej dochodowości na podstawie metod dynamicznych 66
Krzysztof Echaust: Zwroty dzienne a zwroty nocne – porównanie wybranych własności na przykładzie kontraktów *futures* notowanych na GPW w Warszawie 75
Urszula Gierałtowska: Inwestowanie w metale szlachetne jako alternatywna forma lokowania kapitału 88
Paweł Kliber: Spread WIBOR-OIS jako miara ryzyka kredytowego i premii płynnościowej 101
Karol Marek Klimczak: Struktura autoregresyjna zysku rezydualnego spółek z Polski, Niemiec i Francji 112
Anna Korzeniowska: Wybrane problemy rynku finansowego wynikające z sytuacji na rynku oszczędności gospodarstw domowych 120
Mieczysław Kowerski: Cateringowa teoria dywidend 128
Marzena Krawczyk: Adekwatność oferty instytucji rynku finansowego do potrzeb kapitałowych MŚP 142
Paweł Kufel, Magdalena Mosionek-Schweda: Wpływ doświadczenia giełdowego na koszt pozyskiwania kapitału na rynku Catalyst 151
Robert Kurek: Ewolucja konwergencji regulacji i sposobów nadzorowania na rynku ubezpieczeniowym UE 161
Sebastian Majewski, Mariusz Doszyń: Efekty wpływu czynników behawioralnych na stopy zwrotu z akcji spółek sektora budowlanego notowanych na GPW w Warszawie 170

Sebastian Majewski: Behawioralny portfel według Maslowa – analiza symulacyjna ... 180
Marta Małecka: Metody oceny jakości prognoz ryzyka rynkowego – analiza porównawcza ... 192
Aleksander R. Mercik: Wykorzystanie rozkładu *t*-Studenta do szacowania wartości zagrożonej ... 202
Artur Mikulec: Znormalizowany względem czasu τ wskaźnik Calmara i jego zastosowanie w analizie efektywności inwestycji portfelowych ... 212
Wojciech Misterek: Bariery w zakresie pozyskania zewnętrznych źródeł finansowania na realizacje projektów innowacyjnych przedsiębiorstw ... 223
Paweł Niszczota: Wpływ języka raportowania na płynność spółek zagranicznych notowanych na GPW ... 232
Dorota Pekasiewicz: Wyznaczanie współczynnika bezpieczeństwa na podstawie kwantyla rozkładu sumy roszczeń w portfelu ubezpieczeń komunikacyjnych ... 241
Agnieszka Perepeczo: Reakcja akcjonariuszy na decyzje o wypłacie dywidendy w spółkach publicznych – wyniki badań empirycznych ... 253
Tomasz Pisula: Metodyczne aspekty zastosowania modeli skoringowych do oceny zdolności kredytowej z wykorzystaniem metod ilościowych ... 265
Paweł Porcenaluk: Analiza wybranych miar ryzyka płynności dla akcji notowanych na GPW w Warszawie w latach 2001–2011 ... 289
Marcin Salamaga: Zastosowanie metody średniej kroczącej do badania zyskowności inwestycji na polskim rynku kapitałowym ... 298
Rafał Siedlecki: Prognozowanie trudności finansowych przedsiębiorstw z wykorzystaniem miary rozwoju Hellwiga ... 308
Anna Sroczyńska-Baron: Możliwości aplikacyjne gier mniejszościowych na Giełdzie Papierów Wartościowych ... 319
Michał Stachura, Barbara Wodecka: Asymetria w ujęciu Boshnakova – propozycja metody szacowania miar asymetrii z próby ... 328
Piotr Staszkiewicz: Verification of the disclosure lemma applied to the model for reputation risk for subsidiaries of non-public group with reciprocal shareholding on the Polish broker-dealers market ... 337
Anna Szymańska: Bayesowskie szacowanie stawek składki w ubezpieczeniach komunikacyjnych z wybranymi funkcjami straty ... 347
Jacek Welc: Prognozowana dynamika zysków spółek a obciążenie błędów prognoz – doświadczenia polskie ... 357
Jerzy Węcławski: Pożyczki hybrydowe jako alternatywna forma finansowania przedsiębiorstw ... 366
Ryszard Węgrzyn: Analiza wrażliwości zmienności implikowanej względem instrumentu podstawowego opcji – podejście dynamiczne ... 375
Stanisław Wieteska: Obciążenia obiektów budowlanych śniegiem jako element ryzyka w ubezpieczeniach majątkowo-osobowych w Polskim obszarze klimatycznym ... 385

Zuzanna Wośko: Odporność sektora bankowego w Polsce na szoki zewnętrzne w kontekście ryzyka kredytowego. Badanie zależności między zmiennymi makroekonomicznymi ... 397
Anna Zamojska: Wskaźnik Sharpe'a w teorii i w praktyce ... 406
Aneta Zglińska-Pietrzak: Bootstrapowe prognozy zmienności stóp zwrotu na podstawie modelu GARCH ... 415
Monika Zielińska-Sitkiewicz: Ocena kondycji rynku nieruchomości mieszkaniowych na podstawie badania danych z raportów finansowych firm deweloperskich ... 423

Summaries

Adam Adamczyk: The level of internal sources of finance as a determinant of investment in R & D of enterprises ... 22
Roman Asyngier: Economic and legal aspects of irregularities in the functioning of the NewConnect market. Assessment and suggestions for changes ... 33
Jacek Białek: Application of the original index of labour productivity in the analysis of open pension funds' units dynamics ... 42
Magdalena Chmielowiec-Lewczuk: Balanced Scorecard in insurance company ... 52
Dawid Dawidowicz: Evaluation of efficiency of new Polish equity investment funds in comparison to the other investment funds in the period 2000–2012 ... 65
Ewa Dziwok: Yield curve verification based on the correlation surface method ... 74
Krzysztof Echaust: Traded period returns and non-traded period returns – comparison of selected properties on the basis of futures contracts quoted on Warsaw Stock Exchange ... 87
Urszula Gierałtowska: Investing in precious metals as an alternative form of capital investment ... 100
Paweł Kliber: WIBOR-OIS spread as a measure of liquidity and default risk 111
Karol Marek Klimczak: Autoregressive structure of residual income of Polish, French and German firms ... 119
Anna Korzeniowska: Selected problems of financial market resulting from the situation on household savings market ... 127
Mieczysław Kowerski: Catering theory of dividends ... 141
Marzena Krawczyk: Adequacy of the offer given by financial market institution to capital needs of SMEs ... 150
Paweł Kufel, Magdalena Mosionek-Schweda: The impact of the stock-market experience on the cost of capital gained on the Catalyst market ... 160

Robert Kurek: The evolution in convergence of supervision regulations and methods on the European Union insurance market 169
Sebastian Majewski, Mariusz Doszyń: The effects of impact of behavioural factors on the rate of return of construction companies stocks listed on the Warsaw Stock Exchange..... 179
Sebastian Majewski: Behavioural portfolio according to Maslov – simulation analysis 191
Marta Małecka: Methods for evaluating Value-at-Risk forecasts – comparative analysis 201
Aleksander R. Mercik: Using the Student's *t* distribution in Value-at-Risk estimation..... 211
Artur Mikulec: Tau-normalized-Calmar ratio and its application in the analysis of portfolio investment efficiency 222
Wojciech Misterek: Barriers in obtaining external funding to the realization of innovative projects in companies 231
Paweł Niszczota: The language used in filings and the trading activity of foreign companies listed on the Warsaw Stock Exchange 240
Dorota Pekasiewicz: Determination of the safety factor based on quantile of the sum of claims distribution in the portfolio of automobile insurance 252
Agnieszka Perepeczo: Market reactions to dividend announcements in public companies – empirical evidence 264
Tomasz Pisula: Methodological aspects of the application of credit scoring models to assess the creditworthiness with the use of quantitative methods 288
Paweł Porcenaluk: The analysis of the selected liquidity risk measures for stocks listed on the Warsaw Stock Exchange in 2001–2011 period 297
Marcin Salamaga: An application of moving average rules for testing the profitability of Polish stock market..... 307
Rafał Siedlecki: Forecasting financial problems of companies based on Hellwig measurement of development 318
Anna Sroczyńska-Baron: The application of the minority games and gambling on the stock exchange..... 327
Michał Stachura, Barbara Wodecka: Boshnakov's approach to asymmetry – proposal of estimation of sample asymmetry measures 336
Piotr Staszkiewicz: Weryfikacja lematu ujawnienia dla modelu ryzyka reputacji niepublicznych grup kapitałowych z powiązaniami wzajemnymi na polskim rynku firm inwestycyjnych 346
Anna Szymańska: Bayesian estimation of premium rates in motor insurance with selected loss functions 356
Jacek Welc: Forecasted earnings growth of companies and earnings forecast bias – Polish experience..... 365
Jerzy Węcławski: Hybrid loans as an alternative form of corporate finance .. 374

Ryszard Węgrzyn: Analysis of the sensitivity of implied volatility to the underlying instrument of option − a dynamic approach 384
Stanisław Wieteska: Overload of roofs of buildings with snow as an element of risk in property insurance in the Polish climate area..... 396
Zuzanna Wośko: Resilience of the Polish banking sector to external shocks in the context of credit risk. Analysis of the relationship between macroeconomic variables 405
Anna Zamojska: Sharpe ratio – theory and practice..... 414
Aneta Zglińska-Pietrzak: Bootstrap predictions of returns for GARCH processes 422
Monika Zielińska-Sitkiewicz: Assessment of the condition of the Polish real estate market based on the data analysis from the financial statements of developers 437

PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU
RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS nr 323 • 2013
Inwestycje finansowe i ubezpieczenia – tendencje światowe a rynek polski ISSN 1899-3192

Marta Małecka
Uniwersytet Łódzki

METODY OCENY JAKOŚCI PROGNOZ RYZYKA RYNKOWEGO – ANALIZA PORÓWNAWCZA

Streszczenie: Ze względu na brak obserwowalności realizacji VaR oraz innych, stosowanych popularnie na rynku kapitałowym, miar ryzyka, nie istnieje bezpośrednia możliwość oceny średniego błędu estymatora. Stosowanie analitycznych wzorów pozwalających na wyznaczenie wariancji estymatora jest sprzeczne z założeniem o braku stałości rozkładów zmiennych obserwowanych na rynku finansowym. W pracy przedstawiono analizę porównawczą metod oceny jakości prognoz ryzyka. Omówione zostały własności testów służących ocenie jakości prognoz VaR. Analizie poddano test ilości przekroczeń oraz testy niezależności przekroczeń w czasie. Pokazano, że powszechnie stosowane testy zaproponowane w latach 90. ubiegłego stulecia charakteryzują się gorszymi własnościami statystycznymi w porównaniu z testem Engla i Manganellego z 2004 r. W części empirycznej do oceny ryzyka wykorzystano cztery modele VaR, które oceniono z wykorzystaniem rozważanych procedur testowych.

Słowa kluczowe: VaR, test Kupca, test Markowa, dynamiczny test kwantylowy.

1. Wstęp

Począwszy od lat 90. ubiegłego wieku następuje dynamiczny rozwój metod statystycznych związanych z oceną ryzyka rynkowego. Punktem przełomowym było wprowadzenie modelu VaR (*Value-at-Risk*) przez J.P. Morgan w 1993 r. W świetle praktyki rynkowej, w ciągu dwóch ostatnich dekad VaR stała się najpopularniejszą miarą ryzyka rynkowego. Wraz z rozwojem metod estymacji VaR w literaturze przedmiotu pojawiła się dyskusja dotycząca oceny jakości prognoz generowanych przez alternatywne modele. W nurcie badań nad skutecznością prognoz VaR wyłoniły się jakościowo różne grupy. Pierwsza z nich wywodzi się z propozycji P. Kupca [1995], który wykorzystał wnioskowanie statystyczne oparte na empirycznej realizacji procesu stochastycznego przekroczeń VaR. W ramach drugiej grupy metod wykorzystuje się statystyki opisowe mierzące wielkość przekroczeń VaR [Doman, Doman 2004].

Procedury testowe oceniające jakość prognoz VaR w różny sposób wykorzystują ciąg realizacji procesu przekroczeń VaR. W teście Kupca z 1995 r. wnioskowanie oparte zostało na porównaniu empirycznego udziału przekroczeń z prawdopodo-

bieństwem sukcesu w teoretycznym modelu niezależnych prób Bernoulliego. Oddzielną grupę testów stanowią procedury sprawdzające autokorelację przekroczeń VaR, gdzie naturalnym kandydatem jest test Ljunga-Boxa z 1978 r. W 1998 r. zaproponowano test autokorelacji, nazywany testem Markowa, oparty na idei ilorazu wiarygodności, na podstawie założenia, że ciąg przekroczeń VaR stanowi fragment procesu Markowa pierwszego rzędu [Christoffersen 1998]. W odpowiedzi na problem niskiej mocy powyższych testów w 2004 r. zaproponowano test oparty na regresji liniowej wykorzystującej realizacje procesu przekroczeń [Engle, Manganelli 2004].

Celem niniejszej pracy było przeprowadzenie analizy porównawczej własności testów statystycznych oceniających jakość prognoz VaR. Porównanie rozmiaru oraz mocy badanych testów przeprowadzono z wykorzystaniem eksperymentu symulacyjnego. Dodatkowo zaprezentowano badanie empiryczne wykorzystujące badane testy statystyczne do oceny modeli VaR dla indeksu DJIA oraz dokonano analizy wyników w kontekście wniosków płynących z przeprowadzonego eksperymentu symulacyjnego. Wyniki testów rozważano w połączeniu z wartościami statystyk opisowych mierzących wielkość przekroczeń VaR.

W drugim punkcie opracowania przedstawione zostały testy statystyczne stosowane do oceny jakości prognoz VaR oraz statystyki opisowe proponowane w literaturze przedmiotu do pomiaru wielkości przekroczeń. Trzeci i czwarty punkt zawierają kolejno: opis wniosków z eksperymentu symulacyjnego badającego rozmiar i moc rozważanych testów oraz wyniki badania empirycznego. Piąty punkt podsumowuje pracę.

2. Metody oceny jakości prognoz VaR

Punktem wyjścia przy konstrukcji testów statystycznych oceniających jakość modeli VaR jest założenie, że prognozy wartości zagrożonej powinny spełniać dwa postulaty. Pierwszy z nich mówi, że udział przekroczeń VaR (wyjątków) powinien być równy założonemu poziomowi tolerancji (*unconditional coverage property*). Drugi postulat nakazuje testowanie autokorelacji przekroczeń (*independence property*). Wystąpienie przekroczenia VaR w okresie t nie powinno wpływać na prawdopodobieństwo wystąpienia przekroczenia w okresie $t + 1$ [Campbell 2005].

Wnioskowanie statystyczne przy ocenie modeli VaR opiera się na procesie przekroczeń:

$$I_{t+1} = \begin{cases} 1, & r_{t+1} < VaR_t(p) \\ 0, & r_{t+1} \geq VaR_t(p) \end{cases}, \tag{1}$$

gdzie: p – ustalony poziom tolerancji,
r_t – wartość stopy zwrotu w okresie t,
$VaR_t(p)$ – wartość prognozy VaR z okresu t.

W teście Kupca, sprawdzającym pierwszy wymieniony postulat, zaproponowano, aby ciąg realizacji procesu traktować jako ciąg niezależnych prób Bernoulliego o tym samym prawdopodobieństwie sukcesu [Kupiec 1995]. Kryterium oceny *ex post* prognoz wartości zagrożonej sformułowane zostało w postaci testu ilorazu wiarygodności restrykcji parametru p w ramach powyższego modelu. Wnioskowanie oparte zostało zatem na porównaniu empirycznego udziału przekroczeń, stanowiącego oszacowanie bezwarunkowego prawdopodobieństwa sukcesu, z parametrem p [Pipień 2006]. Hipoteza zerowa postaci $H_0 : \pi_1 = p$ sprawdzana jest za pomocą statystyki

$$LR_{uc} = -2\log\frac{p^{t_1}(1-p)^{t_0}}{\hat{\pi}_1^{t_1}(1-\hat{\pi}_1)^{t_0}} \sim_{as} \chi^2_{(1)}, \tag{2}$$

gdzie: π_1 – bezwarunkowe prawdopodobieństwo przekroczenia prognozy VaR,
t_0 – liczba realizacji procesu, dla których nie wystąpiło przekroczenie prognozy VaR,
t_1 – liczba realizacji procesu, dla których wystąpiło przekroczenie prognozy VaR,
$\hat{\pi}_1 = \dfrac{t_1}{t_0 + t_1}$.

W kontekście testowania autokorelacji możliwe jest wykorzystanie testu Ljunga-Boxa z 1978 r., opartego na badaniu współczynników autokorelacji między przekroczeniami. Hipoteza zerowa tego testu przybiera postać $H_0 : \gamma_k = 0$, $k = 1, \ldots, K$, natomiast statystyka testowa

$$LB = t(t+2)\sum_{k=1}^{K}\frac{\hat{\gamma}_k^2}{t-k} \sim_{as} \chi^2_{(K)}, \tag{3}$$

gdzie: t – liczba obserwacji,
γ_k – współczynnik autokorelacji między przekroczeniami rzędu k [Berkowitz i in. 2011].

W 1998 r., w celu testowania autokorelacji przekroczeń, zaproponowany został test Markowa [Christoffersen 1998], w którym wykorzystuje się założenie, że proces stanowi fragment jednorodnego łańcucha Markowa z macierzą przejść:

$$\begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix}, \tag{4}$$

gdzie: π_{ij} – prawdopodobieństwo wystąpienia stanu j w okresie $t + 1$, pod warunkiem że w okresie t wystąpił stan i.

Hipoteza zerowa, formułowana w języku warunkowych prawdopodobieństw przejścia, H0 : $\pi_{01} = \pi_{11}$, sprawdzana jest za pomocą statystyki postaci

$$LR_{ind} = -2\log\frac{\hat{\pi}_1^{t_1}(1-\hat{\pi}_1)^{t_0}}{\hat{\pi}_{01}^{t_{01}}(1-\hat{\pi}_{01})^{t_{00}}\hat{\pi}_{11}^{t_{11}}(1-\hat{\pi}_{11})^{t_{10}}} \sim_{as} \chi^2_{(1)}, \tag{5}$$

gdzie: t_{ij} – liczba obserwacji dla których po stanie i nastąpił stan j, $\hat{\pi}_{01} = \frac{t_{01}}{t_0}$, $\hat{\pi}_{11} = \frac{t_{11}}{t_1}$.

W dynamicznym teście kwantylowym (teście DQ), stanowiącym alternatywę dla powyższych testów autokorelacji, zaproponowano wykorzystanie procesu postaci $Hit_t = I_t - p$ [Engle, Manganelli 2004]. Wówczas, jeżeli zachodzi $E(I_{t+1}) = p$, to $E(Hit_{t+1}) = 0$ oraz jeżeli przekroczenia VaR nie są skorelowane, to $E(Hit_{t+1}\omega_t) = 0$, $\forall\omega_t \in \Omega_t$. Łączny test dwóch powyższych własności oparty został na regresji $Hit_t = X\beta + \varepsilon$, gdzie X – macierz m zmiennych objaśniających, w szczególności stała, opóźnienia zmiennej I_t oraz prognozy VaR na okres t[1]. Procedura testowa polega na badaniu istotności powyższej regresji, tj. hipoteza $H_0 : \beta = 0$ sprawdzana jest za pomocą statystyki

$$DQ = \frac{\hat{\beta}X'X\hat{\beta}}{p(1-p)} \sim_{as} \chi^2_{(m)}, \tag{6}$$

gdzie: $\hat{\beta}$ – oszacowanie MNK wektora parametrów w regresji względem zmiennej Hit_t.

Poza testami statystycznymi do oceny jakości modelu ryzyka stosuje się statystyki opisowe, będące oszacowaniami tzw. oczekiwanego niedoboru, czyli oczekiwanej wielkości straty, pod warunkiem że będzie ona większa niż VaR. Pierwsza z nich, ESF1 definiowana jest jako średnia arytmetyczna przekroczeń VaR dla tych stóp zwrotu, dla których nastąpiło przekroczenie, natomiast druga, ESF2 jest średnią stosunków wartości przekroczeń do odpowiednich wartości VaR [Doman, Doman 2004].

3. Wyniki eksperymentu symulacyjnego

W celu zbadania rozmiaru czterech rozważanych testów wygenerowano T wyników niezależnych prób Bernoulliego z prawdopodobieństwem sukcesu p, równym założonemu poziomowi tolerancji przy szacowaniu VaR. W ten sposób otrzymano ciąg realizacji procesu $\{I_t\}$, cechującego się wartością oczekiwaną p oraz brakiem autokorelacji. Otrzymany ciąg był zatem zgodny z hipotezą zerową testu Kupca,

[1] W niniejszym opracowaniu uwzględniono pięć opóźnień zmiennej I_t.

mówiącą o bezwarunkowym prawdopodobieństwie przekroczenia VaR na poziomie p oraz z hipotezami zerowymi pozostałych testów, mówiącymi o braku autokorelacji przekroczeń. W teście DQ oszacowania VaR wygenerowano z procesu GARCH(1,1)-normal ze składnikiem losowym niezależnym od zmiennej I_t^2. Eksperyment powtórzono dla długości prób $T = 250, 500, \ldots, 1500$, wykonując każdorazowo 5000 replikacji.

W tabeli 1 pokazano oszacowania rozmiaru rozważanych testów wyznaczone w oparciu o opisany wyżej eksperyment symulacyjny. Największą rozbieżność między otrzymanym symulacyjnie rozmiarem testu a nominalnym poziomem istotności zaobserwowano dla testu Markowa. Rozmiar pozostałych testów, dla prób powyżej 500 obserwacji, był podobny w różnych procedurach oraz zbliżony do nominalnego. Powyżej 1000 obserwacji, dla wszystkich testów zaobserwowano ustabilizowanie się rozmiaru testu między poziomami 5 a 6%.

Tabela 1. Rozmiar testów jakości prognoz VaR

	Wielkość próby					
	250	500	750	1000	1250	1500
Test Kupca	0,0637	0,0597	0,0487	0,0573	0,0573	0,0573
Test Markowa	0,0896	0,0950	0,0788	0,0913	0,0913	0,0913
Test LB	0,0627	0,0593	0,0493	0,0513	0,0513	0,0513
Test DQ	0,0700	0,0563	0,0550	0,0570	0,0570	0,0570

Źródło: obliczenia własne.

Symulacyjne porównanie mocy testów przeprowadzono na podstawie eksperymentów, gdzie zastosowano dwustopniową procedurę, gwarantującą otrzymanie szeregu przekroczeń charakteryzującego się liczbą przekroczeń VaR niezgodną z hipotezą zerową testu Kupca oraz różną siłą autokorelacji. Pierwszy krok polegał na wygenerowaniu procesu stopy zwrotu, natomiast w drugim kroku otrzymano oszacowania VaR. Do generowania wartości stopy zwrotu wykorzystano proces GARCH(1,1), odzwierciedlający zjawisko grupowania zmienności odpowiadające za autokorelację przekroczeń. Dzięki temu zapewniono niezgodność z hipotezami zerowymi w testach autokorelacji. Szeregi przekroczeń otrzymano jako dodatnią różnicę między wartościami stopy zwrotu a oszacowaniami VaR. Oszacowania VaR generowano z dziewięciu modeli homoskedastycznych z rozkładem normalnym, w których odchylenie standardowe przyjmowano na poziomie od 50 do 90% rzeczywistego odchylenia standardowego procesu. Tak samo jak przy badaniu rozmiaru, symulacyjne badanie mocy testów przeprowadzono dla długości prób $T = 250, 500, \ldots, 1500$, wykonując każdorazowo 5000 replikacji.

[2] Przegląd modeli GARCH można znaleźć np. w pracy [Fiszeder 2009].

W tabelach 2 i 3 przedstawiono oszacowania mocy testów, odpowiednio dla długości próby 500 i 1000 obserwacji, przy różnych parametrach modelu generującego dane. W eksperymentach, w których przyjęto odchylenie standardowe na poziomie 50–60% rzeczywistej wartości tego parametru, co skutkowało wyraźną niezgodnością z hipotezą zerową testu Kupca, otrzymano oszacowanie mocy tego testu powyżej 70%. Eksperyment symulacyjny pokazał skalę spadku mocy testu Kupca wraz ze zbliżaniem się parametrów modelu VaR do parametrów procesu generującego dane (rys. 1)[3]. Przy 90-procentowym odchyleniu standardowym moc tego testu, w zależności od długości próby, spadała do ok. 60–70%.

Tabela 2. Moc testów jakości prognoz VaR, wielkość próby $T = 500$

	Odchylenie standardowe w modelu generującym oszacowania VaR				
	$\sigma = 0{,}50\sigma_{real}$	$\sigma = 0{,}55\sigma_{real}$	$\sigma = 0{,}60\sigma_{real}$	$\sigma = 0{,}65\sigma_{real}$	$\sigma = 0{,}70\sigma_{real}$
Test Kupca	0,8743	0,7970	0,7233	0,6793	0,6725
Test Markowa	0,2800	0,2916	0,2937	0,3003	0,3001
Test LB	0,3648	0,4000	0,4333	0,4475	0,4664
Test DQ	0,8868	0,8195	0,7508	0,7030	0,6755

Źródło: obliczenia własne.

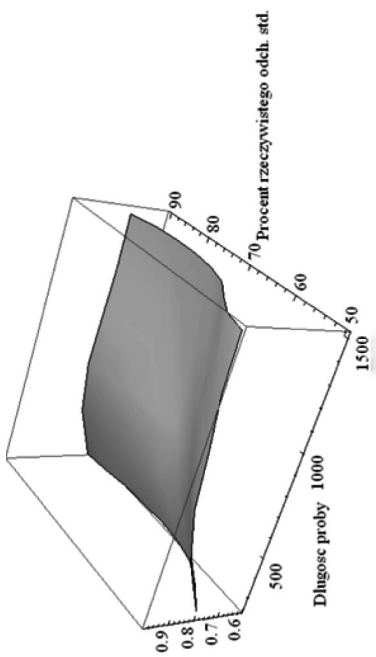
Tabela 3. Moc testów jakości prognoz VaR, wielkość próby $T = 1000$

	Odchylenie standardowe w modelu generującym oszacowania VaR				
	$\sigma = 0{,}50\sigma_{real}$	$\sigma = 0{,}55\sigma_{real}$	$\sigma = 0{,}60\sigma_{real}$	$\sigma = 0{,}65\sigma_{real}$	$\sigma = 0{,}70\sigma_{real}$
Test Kupca	0,9564	0,8966	0,8240	0,7510	0,6962
Test Markowa	0,4828	0,4966	0,5035	0,5013	0,5025
Test LB	0,7026	0,7330	0,7552	0,7728	0,7736
Test DQ	0,9714	0,9374	0,9014	0,8780	0,8612

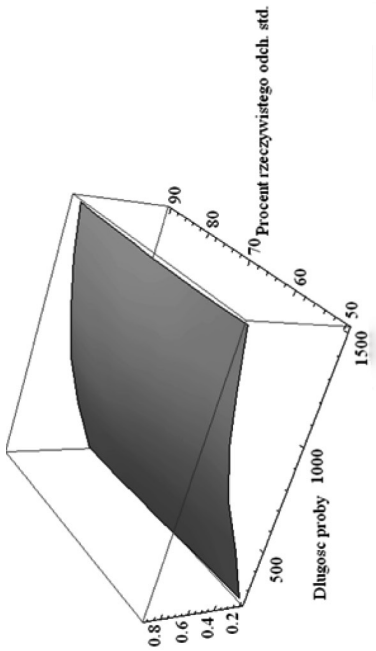
Źródło: obliczenia własne.

Wyniki badania mocy rozważanych testów autokorelacji (Markowa, LB i DQ) pokazały małą moc testu Markowa w porównaniu do dwóch pozostałych procedur. Dla próby 500 obserwacji wynosiła ona ok. 30% (tab. 2). Spośród testów autokorelacji najlepszymi własnościami charakteryzował się test DQ. Dla próby 500 obserwacji, przy wszystkich wariantach doboru parametrów, oszacowanie mocy przekraczało 60%. Najmniejsza zaobserwowana moc w tym teście, otrzymana dla próby 250 obserwacji w modelu z 90-procentowym odchyleniem standardowym, wynosiła ok. 30%.

[3] W tabelach zaprezentowane zostały wyniki dla wybranych modeli oraz wybranych wielkości próby, natomiast wszystkie otrzymane wyniki symulacji przedstawione są na rys. 3–6.



Rys. 1. Moc testu Kupca



Rys. 3. Moc testu LB

Rys. 2. Moc testu Markowa

Rys. 4. Moc testu DQ

Wyraźna poprawa mocy testów autokorelacji występowała przy wydłużaniu szeregu czasowego (rys. 2, 3, 4). Szczególnie szybki wzrost częstości odrzuceń zaobserwowano dla testu DQ. Dla prób większych niż 250 obserwacji, w każdym z rozważanych modeli eksperymentu, moc testu DQ przekraczała 60%, zbliżając się do jedności przy największych próbach.

4. Wyniki badania empirycznego

Badanie empiryczne pozwoliło na zaprezentowanie analizy wybranych modeli VaR, wykorzystującej rozważane testy, na przykładzie rzeczywistego szeregu czasowego realizacji stopy zwrotu z indeksu DJIA od początku 1995 r. Zastosowano cztery popularne modele VaR (*normal*, GARCH(1,1)-*normal*, symulacji historycznej oraz *Peak over Threshold* – POT)[4]. Po przeprowadzeniu analizy dla całego szeregu od 1995 r, wszystkie testy powtórzono dla podprób długości 250 obserwacji przesuwanych w badanym okresie.

Tabela 4. Wyniki oceny jakości prognoz VaR w badaniu empirycznym dla indeksu WIG20, $T = 3445$

	Model VaR			
	Normal	GARCH-*normal*	Symulacja historyczna	POT
Test Kupca (*p*-wartość)	0,0000	0,0000	0,0000	0,2904
Test Markowa (*p*-wartość)	0,0000	0,0000	0,0000	0,0000
Test LB (*p*-wartość)	0,0000	0,0000	0,0000	0,0000
Test DQ (*p*-wartość)	0,0000	0,0000	0,0000	0,0000
Udział przekroczeń	0,0681	0,0730	0,0695	0,0461
ESF1	0,0083	0,0083	0,0083	0,0078
ESF2	1,6765	1,7163	1,6957	1,4905

Źródło: obliczenia własne.

Tabela 5. Średnie wyniki oceny jakości prognoz VaR w badaniu empirycznym dla indeksu WIG20, $T = 250$

	Model VaR			
	Normal	GARCH-*normal*	Symulacja historyczna	POT
Test Kupca (*p*-wartość)	0,1243	0,0703	0,0979	0,5870
Test Markowa (*p*-wartość)	0,4272	0,3713	0,4053	0,7231
Test LB (*p*-wartość)	0,0791	0,0850	0,0975	0,2770
Test DQ (*p*-wartość)	0,0047	0,0019	0,0096	0,0126

Źródło: obliczenia własne.

[4] Omówienie modeli VaR: *normal*, GARCH(1,1)-*normal* i symulacji historycznej można znaleźć w pracy [Jorion 2007], natomiast opis modelu POT w pracy [McNeil, Saladin 1997].

Dla całego szeregu, liczącego ponad 3000 obserwacji, wartości statystyk opisowych ESF1 i ESF2 oraz wyniki wszystkich testów przedstawione są w tabeli 4. Miara ESF1 pokazała, że udział przekroczeń w rozważanych modelach różnił się wyraźnie od założonych 5%. Z wartości ESF2 wyniknęło, że przeciętne przekroczenie VaR stanowiło od ok. 50 do ok. 70% stopy zwrotu. Wszystkie rozważane testy, poza testem Kupca dla modelu POT, nakazały odrzucenie zastosowanych modeli. Otrzymane *p*-wartości były biskie zera.

W tabeli 5 przedstawione są uśrednione wyniki testów dla podprób długości 250 obserwacji. Porównanie wyników dla szeregu liczącego ponad 3000 obserwacji z uśrednionymi wynikami dla podprób 250-elementowych pozwoliło na zaobserwowanie wpływu długości próby na wynik testu. Dla dwóch rozważanych rozmiarów próby otrzymano znaczne różnice we wskazaniach testów. W przeciwieństwie do szeregów obejmujących historię indeksu od 1995 r., dla 250-elementowych szeregów brak było podstaw do odrzucenia większości modeli na poziomie istotności 5%. Zatem, zgodnie z wynikami eksperymentu symulacyjnego, sprzeczność w wynikach dla długich i krótkich prób w przykładzie empirycznym pokazała możliwość popełniania błędów we wnioskowaniu statystycznym przy krótkich szeregach czasowych.

5. Podsumowanie

W niniejszym opracowaniu przedstawiono eksperyment symulacyjny, którego celem było porównanie własności statystycznych testów stosowanych do oceny prognoz modeli ryzyka rynkowego. Pokazano, że powszechnie stosowane testy zaproponowane w latach 90. ubiegłego stulecia charakteryzują się gorszymi własnościami statystycznymi w porównaniu z testem Engla i Manganellego z 2004 r. Największą rozbieżność między wyznaczonym symulacyjnie rozmiarem testu a nominalnym poziomem istotności otrzymano dla testu Markowa z 1998 r. Ponadto wyniki badania symulacyjnego pokazały małą moc tego testu w porównaniu z innymi rozważanymi testami autokorelacji. Spośród testów autokorelacji najlepszymi własnościami charakteryzował się dynamiczny test kwantylowy. Wyraźna poprawa mocy testów występowała przy wydłużaniu szeregu czasowego. Dla prób większych niż 250, w każdym z rozważanych modeli, moc testu DQ przekraczała 60%, zbliżając się do jedności przy największych próbach.

Badanie empiryczne pozwoliło na zaprezentowanie wyników rozważanych testów na przykładzie rzeczywistego szeregu czasowego. Porównanie wyników pochodzących z szeregu ponad 3000 obserwacji z uśrednionymi wynikami dla podprób 250-elementowych pokazało znaczne różnice we wskazaniach testów. W pierwszym przypadku wyniki prawie wszystkich testów odrzuciły zastosowane modele, natomiast w drugim, na poziomie istotności 5%, brak było podstaw do odrzucenia większości modeli. Różnice w wynikach dla krótkich i długich prób pokazały na przykładzie empirycznym, zgodnie z wynikami badania symulacyjnego, możliwość popełniania błędów we wnioskowaniu statystycznym przy krótkich szeregach czasowych.

Literatura

Berkowitz J., Christoffersen P., Pelletier D., 2011, *Evaluating value-at-risk models with desk-level data*, Management Science, vol. 57, no. 12, s. 2213–2227.

Campbell S.D., 2005, *A Review of Backtesting and Backtesting Procedures*, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Christoffersen P., 1998, *Evaluating interval forecasts*, International Economic Review, vol. 39, no. 4, s. 841–862.

Doman M., Doman R., 2004, *Ekonometryczne modelowanie dynamiki polskiego rynku finansowego*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.

Engle R., Manganelli S., 2004, *CAViaR: Conditional Autoregressive Value at Risk by regression quantiles*, Journal of Business & Economic Statistics, no. 22, s. 367–381.

Fiszeder P., 2009, *Modele klasy GARCH w empirycznych badaniach finansowych*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.

Jorion P., 2007, *Value at Risk. The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York.

Kupiec P., 1995, *Techniques for verifying the accuracy of risk measurement models*, Journal of Derivatives, vol. 3, no. 2, s. 174–184.

McNeil A., Saladin T., 1997, *The peaks over thresholds method for estimating high quantiles of loss distributions*, [w:] *Proceedings of 28th International ASTIN Colloquium*, Institute of Actuaries Australia, Sydney.

Pipień M., 2006, *Wnioskowanie bayesowskie w ekonometrii finansowej*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

METHODS FOR EVALUATING VALUE-AT-RISK FORECASTS – COMPARATIVE ANALYSIS

Summary: Models for risk measurement in capital market rise problems with their evaluation due to the lack of possibility to observe real VaR series. Analytical formulas cannot be used in case of time-changing distributions of financial variables. In consequence, in the last two decades a lot of testing procedures have been proposed to verify quality of risk models. In the paper, we presented comparative analysis of methods for evaluating VaR forecasts. We analysed test for the number of VaR exceptions as well as tests for autocorrelation of VaR violations. We showed that widely used tests from the 1990s have lower power that the proposition of Engle and Mangianelli from 2004. In the empirical part we used four VaR models for WIG20 index which were applied with testing procedures.

Keywords: VaR, Kupiec test, Markow test, dynamic quantile test.