

Marek Lusztyn

Unicredit SpA

## WERYFIKACJA HISTORYCZNA MODELI WARTOŚCI ZAGROŻONEJ – ZASTOSOWANIE WYBRANYCH METOD DLA RYNKU POLSKIEGO W OKRESIE KRYZYSU FINANSOWEGO

**Streszczenie:** Wiele banków regulacyjnego na pokrycie tego zagrożonej. Analiza historyczna od niedokładnego. Celem wartości zagrożonej w okresie w wartości zagrożonej w okresie kładzie polskich rynków ak nie dają podstaw do jednoznacznie. To z kolei nakazuje daleko narzędzia ilościowego zarządzania ryzykiem rynkowym. W celu ostrożnego i stabilnego zarządzania instytucją finansową miara ta powinna być uzupełniona metodami alternatywnymi.

Weryfikacja historyczna modeli wartości zagrożonej – zastosowanie wybranych metod dla rynku

kowego i wyznaczania kapitału regulacyjnego, oparte na wartości zagrożonej na odróżnienie modelu dokładnego i nieprecyzyjnego. Weryfikacja historyczna wybranych modeli wartości zagrożonej i w jego trakcie na przykładzie polskiego rynku finansowego. Przeprowadzone wyniki testów weryfikacji żadnego z analizowanych modeli. Weryfikacja historyczna wartości zagrożonej jako jedynej metody zarządzania ryzykiem rynkowym.

**Słowa kluczowe:** wartość zagrożona, VaR, backtesting.

### 1. Wstęp

*Value at Risk* (VaR, *wartość zagrożona*) jest podstawową metodą stosowaną w bankach w zarządzaniu ryzykiem rynkowym, zarówno do codziennej kontroli tego ryzyka, jak i do wyznaczania wysokości kapitału regulacyjnego koniecznego do pokrycia ponoszonego ryzyka rynkowego. Standardy nadzorcze wprowadziły w roku 1996 koncepcję *Value at Risk* jako narzędzie wyznaczania kapitału na pokrycie ryzyka rynkowego na podstawie modelu wewnętrznego [BCBS 1996], po spełnieniu przez bank wielu wymagań ilościowych i jakościowych co do procesu zarządzania ryzykiem rynkowym, i podtrzymały jego zastosowanie w ramach Filara I Nowej Umowy Kapitałowej (NUK) [BCBS 2006], pomimo szerokiej krytyki tego modelu [Danielsson i in. 2001]. Kryzys finansowy *subprime* w 2008 r. pokazał nieadekwatność modeli wartości zagrożonej, które to w znaczny sposób niedoszacowały ponoszone przez banki ryzyko rynkowe [Lazaregue-Bazard 2010]. Problem ten niósł ze sobą skutki nie tylko dla wewnętrznych systemów zarządzania ryzykiem rynkowym

w bankach, ale zwłaszcza dla instytucji nadzorczych, z perspektywy stabilności finansowej. Ponieważ straty w portfelach handlowych banków podczas kryzysu finansowego były znacznie wyższe od minimalnych wymogów kapitałowych, obliczonych zgodnie z zasadami określonymi przez instytucje nadzorcze, dla modeli wewnętrznych ryzyka rynkowego było to jednym z czynników istotnie zwiększających ryzyko systemowe. W znacznej części wynikało to z procykliczności obowiązujących minimalnych wymogów kapitałowych opartych na *Value at Risk* [Youngman 2009]. Zmiany minimalnych wymogów kapitałowych dokonane przez nadzór bankowy w rezultacie kryzysu nie tylko nie odeszły od stosowania wartości zagrożonej, ale wręcz zwiększyły stopień zależności od tej miary przez wprowadzenie wartości zagrożonej w warunkach skrajnych (*Stressed Value at Risk*) [BCBS 2009].

W tym kontekście dla instytucji, które przyjęły podejście statystyczne do pomiaru ryzyka rynkowego, sprawdzanie jakości i dokładności systemów pomiaru ryzyka jest kluczowe dla ostrożnego i stabilnego zarządzania ryzykiem rynkowym w banku. Podstawowym procesem w tym zakresie jest tzw. analiza historyczna (*backtesting*), czyli porównywanie codziennych zysków i strat z działalności handlowej z miarami VaR generowanymi przez model ryzyka w celu dokładniejszego rozróżnienia pomiędzy dokładnymi i niedokładnymi modelami ryzyka.

Celem niniejszej pracy jest weryfikacja historyczna wybranych modeli wartości zagrożonej w okresie poprzedzającym kryzys finansowy oraz w jego trakcie na przykładzie polskich rynków akcji, obligacji i rynku walutowego. Aspekt teoretyczny modeli wartości zagrożonej (*Value at Risk*) [Grabowska 2000; Iskra 2012], jak również ich weryfikacja historyczna [Jędrusik i in. 2007], jest dość szeroko opisany w polskiej literaturze przedmiotu. Tematyka weryfikacji historycznej wybranych modeli wartości zagrożonej, uwzględniająca perspektywę kryzysu finansowego, była już obszarem badań m.in. [Rossignolo i in. 2012]; w przypadku rynku polskiego głównie w aspekcie rynku walutowego [Lusztyn 2011; Ludwiczak 2012]. Niniejsza praca rozszerza obszar badawczy pod względem zarówno badanych modeli, jak i badanych segmentów polskiego rynku finansowego (rynek akcji i stóp procentowych).

## 2. Miary używane w weryfikacji historycznej

*Value at Risk* opisuje poziom strat, jaki może zostać osiągnięty na danym portfelu w określonym czasie i z zadaniem poziomem ufności [Jorion 2006]. Mając dany poziom ufności  $\alpha \in (0,1)$ , wartość VaR dla danego portfela na poziomie ufności  $\alpha$  jest dana wartością  $l$ , taką że prawdopodobieństwo straty  $L$  przekraczającej  $l$  jest nie większe niż  $(1 - \alpha)$ :

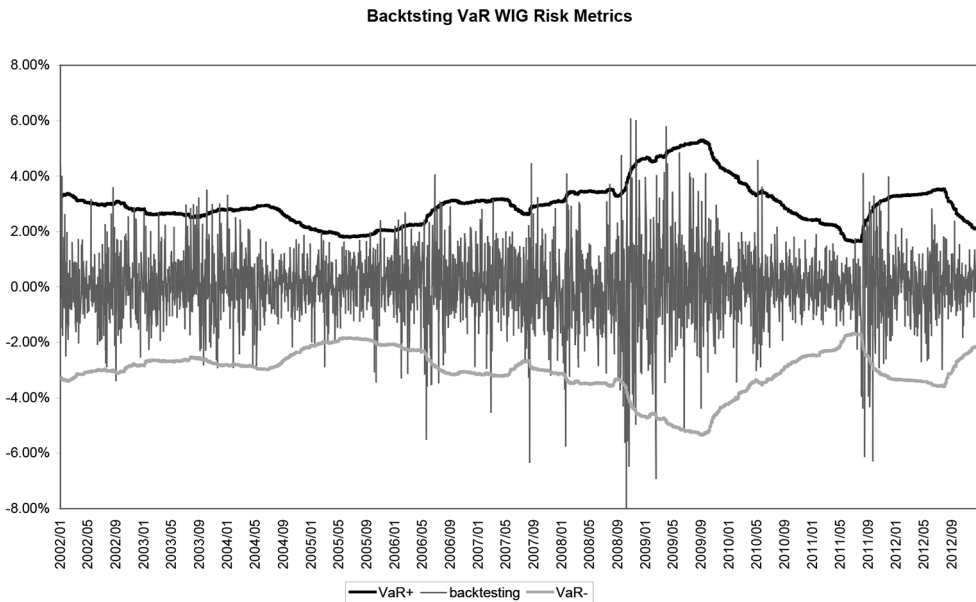
$$VaR_\alpha = \inf \{l \in \mathfrak{R}: P(L > l) \leq 1 - \alpha\}. \quad (1)$$

Definicja wyrażona wzorem (1) daje możliwość zastosowania różnorodnych modeli, które jednak w każdym przypadku powinny zostać poddane weryfikacji. Istotę wszystkich zastosowań analizy historycznej stanowi ocena jakości modelu przez porównanie rzeczywistych wyników handlowych z miarami ryzyka genero-

wanymi przez dany model wartości zagrożonej. Najczęściej pojedyncza metoda testowania nie daje jednak jednoznacznych rozstrzygnięć – konieczna jest ocena z użyciem różnych testów. Celem niniejszego rozdziału jest przegląd omawianych w literaturze metod pomiaru efektywności modeli wartości zagrożonej, użytych następnie w niniejszym badaniu.

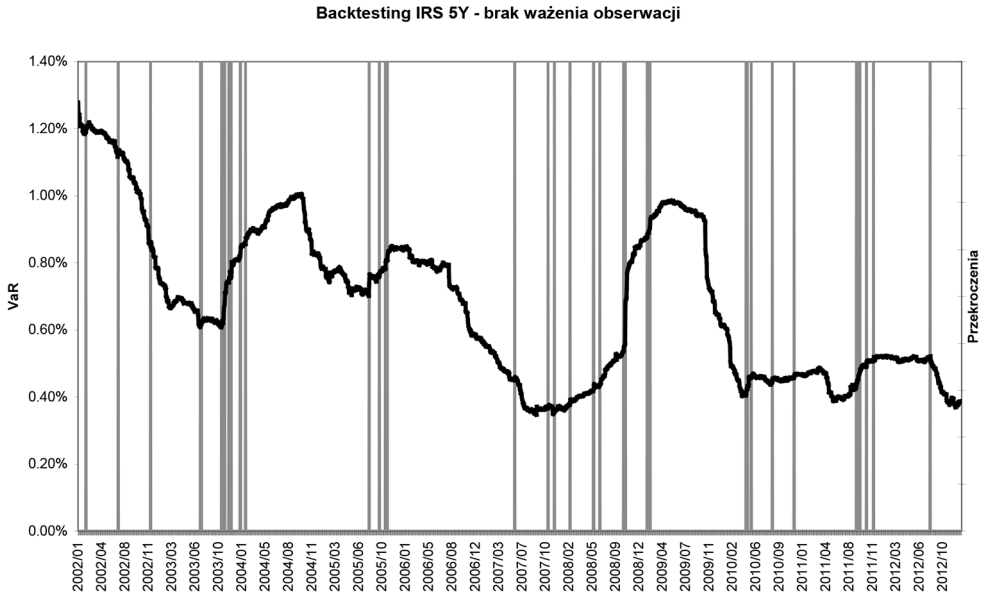
## 2.1. Stopień pokrycia

Najbardziej popularnym sposobem testowania poprawności modeli wartości zagrożonej jest analiza liczby przekroczeń backtestingu, tj. liczby dni (lub wybranych okresów o innej długości), w których wartość zagrożona jest przekroczona przez poziom strat na analizowanym portfelu. Jeśli liczba przekroczeń jest większa niż określony poziom ufności, wskazuje to na model wartości zagrożonej niedoszacowujący ryzyko. W przeciwnym przypadku, gdy liczba przekroczeń jest niższa niż zdefiniowany poziom ufności, jest to wskazówka, iż model może przeszacowywać poziom ryzyka. W rzeczywistości rzadko kiedy można zaobserwować liczbę przekroczeń dokładnie odpowiadającą założonemu poziomowi ufności. W tym celu potrzebna jest więc analiza statystyczna, pozwalająca określić, czy analizowany model wartości zagrożonej powinien być zaakceptowany bądź odrzucony. Przykłady graficznej prezentacji backtestingu zostały przedstawione na rys. 1 i 2.



**Rys. 1.** Graficzna reprezentacja backtestingu na przykładzie modelu *Risk Metrics*. Dane dzienne w okresie 2002-2012

Źródło: opracowanie własne.



**Rys. 2.** Graficzna reprezentacja backtestingu na przykładzie modelu parametrycznego. Dane dzienne w okresie 2002-2012

Źródło: opracowanie własne.

Oznaczając liczbę przekroczeń jako  $x$ , a całkowitą liczbę obserwacji jako  $T$ , można zdefiniować odsetek niepowodzeń jako  $x/T$ , która to wielkość, zakładając poprawność funkcjonowania modelu, będzie odpowiadała estymatorowi nieobciążonemu założonego poziomu ufności  $\alpha$ . Każdy dzień podlegający analizie backtestingu może się zakończyć jedynie dwoma możliwymi wynikami – wystąpieniem przekroczenia bądź też jego brakiem, co jest szczególnym przypadkiem próby Bernoullego. W związku z tym liczba przekroczeń może zostać opisana rozkładem dwumianowym:

$$f(x) = \binom{T}{x} p^x (1-p)^{T-x}, \quad (2)$$

co wraz ze wzrostem liczby obserwacji można przybliżyć rozkładem normalnym:

$$z = \frac{x - pT}{\sqrt{p(1-p)T}} \approx N(0,1), \quad (3)$$

gdzie  $pT$  jest wartością oczekiwaną liczby przekroczeń, zaś  $p(1-p)T$  – wariancją liczby przekroczeń backtestingu. Używając rozkładu dwumianowego, możemy następnie zbadać dokładność modelu wartości zagrożonej. Jednak weryfikując hipotezę

zę zerową o prawidłowości testowanego modelu VaR, musimy znaleźć kompromis pomiędzy możliwością popełnienia błędu pierwszego i drugiego rodzaju.

Rozważając model wartości zagrożonej obliczonej zgodnie z opisanymi niżej wytycznymi Komitetu Bazylejskiego, na poziomie ufności 99% dla 250 dni roboczych, oczekując poprawności modelu, należy stwierdzić, że oczekiwana liczba dni, dla których faktyczne straty przekraczają poziom VaR, wynosi 2,5 dnia. W przykładowej sytuacji, w której bank zdecydowałby się przyjąć poziom 5 przekroczeń jako punkt odcięcia, prawdopodobieństwo popełnienia błędu pierwszego rodzaju (odrzućenia poprawnego modelu VaR) wynosiłoby 10,8%, zaś prawdopodobieństwo popełnienia błędu drugiego rodzaju (przyjęcia błędnego modelu VaR) wynosi 12,8%.

## 2.2. Test Kupca

W ocenie jakości oszacowanego VaR-u często wykorzystywany jest test Kupca [Kupiec 1995]. Test ten weryfikuje, czy udział przekroczeń jest zgodny z zadaniem poziomem istotności (*unconditional coverage test*). Hipoteza zerowa tego testu jest odrzucana zarówno w przypadku, gdy model niedoszacowuje, jak i wtedy, gdy model przeszacowuje ryzyko. Przyjmuje się następujący układ hipotez:

$$\begin{aligned} H_0 : f &= \alpha \\ H_1 : f &\neq \alpha \end{aligned} \quad (4)$$

gdzie  $f$  oznacza liczbę przekroczeń backtestingu, zaś  $\alpha$  oznacza poziom ufności, z którym wyznaczany jest VaR. Oznaczmy jako  $I$  ciąg wartości 0 i 1 przyjmujący podane wartości w zależności od, odpowiednio, braku przekroczenia backtestingu bądź jego wystąpienia. Dla pozycji długiej elementy tego ciągu przyjmą następujące wartości [Lopez 1998]:

$$I_t^L = \begin{cases} 0 & \text{dla } r_t \geq VaR_t^\alpha \\ 1 & \text{dla } r_t < VaR_t^\alpha \end{cases} \quad (5)$$

Statystyka testu Kupca określona jest następująco:

$$LR_{POF} = 2 \left( \ln \left( \left( \frac{T_1}{T_0 + T_1} \right)^{T_1} \left( 1 - \frac{T_1}{T_0 + T_1} \right)^{T_0} \right) - \ln \left( \alpha^{T_1} (1 - \alpha)^{T_0} \right) \right), \quad (6)$$

gdzie  $T_1$  jest liczbą przekroczeń, a  $T_0 + T_1$  liczbą obserwacji. Przy prawdziwej hipotezie zerowej  $LR_{POF}$  ma rozkład asymptotycznie zmierny do  $\chi^2(1)$ .

## 2.3. Testy Christoffersena

Jeśli przekroczenia backtestingu nie występują w próbie równomiernie, lecz mają miejsce w krótkich odstępach czasu, oznacza to, że model przeszacowuje wartość

zagrożoną w normalnych warunkach rynkowych i niedoszacowuje jej w sytuacjach kryzysowych, gdy następuje istotne zwiększenie zmienności na rynku. W związku z tym istotne jest nie tylko przeprowadzenie testów opisujących liczbę przekroczeń wartości zagrożonej, ale także weryfikacja, czy przekroczenia backtestingu są niezależne w czasie (*conditional coverage test*).

Najczęściej wykorzystywany w tym celu jest test Christoffersena [1998] z następującym układem hipotez i statystyką testową wyrażoną za pomocą następującego wzoru:

$H_0$ : przekroczenia niezależne w czasie.

$H_1$ : przekroczenia zależne w czasie.

$$LR_{IND} = 2 \ln \left( \left( 1 - \frac{T_{01}}{T_{00} + T_{01}} \right)^{T_{00}} \left( \frac{T_{01}}{T_{00} + T_{01}} \right)^{T_{01}} \right) - 2 \ln \left( \left( \frac{T_{10}}{T_{00} + T_{01} + T_{10}} \right)^{T_{10}} \left( 1 - \frac{T_{10}}{T_{00} + T_{01} + T_{10}} \right)^{T_{00} + T_{01}} \right), \quad (7)$$

gdzie:  $T_{00}$  – liczba braków przekroczeń VaR następujących bezpośrednio po brakach przekroczeń,

$T_{01}$  – liczba przekroczeń VaR następujących bezpośrednio po brakach przekroczeń,

$T_{10}$  – liczba braków przekroczeń VaR następujących bezpośrednio po przekroczeniach,

$T_{11}$  – liczba przekroczeń VaR następujących bezpośrednio po przekroczeniach.

Przy prawdziwej hipotezie zerowej  $LR_{IND}$  ma rozkład asymptotycznie zmierny do  $\chi^2(1)$ . Biorąc pod uwagę niezależność statystyk liczby przekroczeń oraz ich niezależności w czasie, Christoffersen [2004] proponuje też test mieszany  $LR_{MIX}$ , uwzględniający zarówno liczbę przekroczeń, jak i ich niezależność w czasie:

$$LR_{MIX} = LR_{POF} + LR_{IND} \sim \chi^2_2. \quad (8)$$

### 3. Ramy nadzorcze

Ramy opracowane przez Bazylejski Komitet Nadzoru Bankowego określiły sposób włączenia analizy historycznej (*backtesting*) do podejścia statystycznego wykorzystującego modele wewnętrzne do wyznaczania wymogów kapitałowych z tytułu ryzyka rynkowego. Porównując miary ryzyka z wynikami handlowymi, bank liczy liczbę przypadków, kiedy miary ryzyka były większe niż wynik handlowy (tzw. przekroczenia backtestingu), porównując następnie odsetek rzeczywiście pokrytych wyników z założonym poziomem pokrycia. Ze statystycznego punktu widzenia

stosowanie liczby wyjątków jako podstawy do oceny modelu banku wymaga stosunkowo niewielu silnych założeń. W szczególności podstawowym założeniem jest to, że codzienny test (wyjątek/brak wyjątku) jest niezależny od wyniku uzyskanego w jakimkolwiek innym dniu. Komitet Bazylejski [BCBS 1996] podkreślił, że analizy tego typu są ograniczone, jeśli chodzi o ich możliwości rozróżnienia modelu dokładnego od niedokładnego. Dla statystyka oznacza to, że nie można zaprojektować analizy tak, by poprawnie sygnalizowała wszystkie problematyczne modele, bez dawania fałszywych sygnałów o trudnościach w przypadku wielu innych. W celu zrównoważenia dwóch typów błędów statystycznych Komitet Bazylejski nakreślił, omówione niżej, trzy strefy i wybrał ich granice, mając na względzie: (1) możliwość, że dokładny model ryzyka byłby sklasyfikowany jako niedokładny na podstawie jego wyników analizy historycznej, oraz (2) możliwość, że niedokładny model zostałby sklasyfikowany jako dokładny na podstawie jego wyników analizy historycznej. Pamiętając o statystycznych ograniczeniach analizy historycznej, Komitet wprowadził ramy dla interpretacji nadzorczej wyników analizy historycznej, które obejmują wiele możliwych reakcji, w zależności od siły sygnału generowanego przez analizę historyczną. Reakcje te są podzielone na trzy strefy, rozróżnione za pomocą kolorów, w hierarchię reakcji. Strefa zielona odpowiada wynikom analizy historycznej, które same nie sugerują problemu z jakością lub dokładnością modelu VaR. Strefa żółta obejmuje wyniki, które nie rodzą pytań w tym względzie, ale gdzie taki wniosek nie jest jednoznaczny. Strefa czerwona wskazuje na wynik analizy historycznej, który prawie na pewno wskazuje na problem z modelem ryzyka banku.

Banki używające modelu wewnętrznego do wyznaczenia poziomu kapitału regulacyjnego na pokrycie ryzyka rynkowego  $c$  czynią to zgodnie z następującą formułą [BCBS 1996]:

$$c = \text{Max} \left( \text{VaR}_{t-1}, (K + k) \frac{1}{60} \sum_{i=1}^{60} \text{VaR}_{t-i} \right), \quad (9)$$

gdzie  $K$  jest mnożnikiem ustalonym przez Komitet Bazylejski [1996] na 3,  $k$  w zależności od jakości prognoz modelu (*backtesting*) przyjmuje wartości z przedziału  $[0,1]$ , zgodnie z tab. 1. Te szczególne wartości odzwierciedlają ogólną koncepcję, że zwiększenie mnożnika powinno być wystarczające, by sprowadzić model do normy 99 percentyla. VaR obliczony jest z horyzontem utrzymywania pozycji 10 dni. W praktyce wyznaczenie VaR odbywa się z horyzontem utrzymywania pozycji 1 dzień, skalowanym następnie do 10-dniowego okresu utrzymywania pozycji na potrzeby wyznaczenia kapitału regulacyjnego, pomimo znanych ograniczeń tej metody, wskazanych m.in. przez [Alexander 2009].

Komitet Bazylejski [BCBS 1996] zwraca uwagę, iż w praktyce miary wartości zagrożonej nie mogą być bezpośrednio porównywane z wynikami handlowymi, ponieważ wyniki rzeczywiste w nieunikniony sposób są „skażone” zmianami struktury portfela w czasie okresu utrzymywania – włączenie dochodów z tytułu

**Tabela 1.** Mnożniki regulacyjne dla modeli wewnętrznych

Strefa	Liczba przekroczeń backtestingu	Zwiększenie wskaźnika regulacyjnego $k$	Prawdopodobieństwo, iż model jest dokładny
'Zielona'	0	0,00	100,0%
	1	0,00	91,9%
	2	0,00	71,4%
	3	0,00	45,7%
	4	0,00	24,2%
'Żółta'	5	0,40	10,8%
	6	0,50	4,1%
	7	0,65	1,4%
	8	0,75	0,4%
	9	0,85	0,1%
'Czerwona'	10 i więcej	1,00	0,0%

Źródło: [BCBS 1996].

provizji i opłat do zysków i strat handlowych wynikających ze zmian składu portfela nie powinno być uwzględniane w definicji wyników handlowych, ponieważ nie są one związane z ryzykiem statycznego portfela, który był założony przy konstruowaniu miary wartości zagrożonej. Uznając analizę historyczną za statystyczny test rzetelności obliczeń miary wartości zagrożonej, należy stwierdzić, że zastosowanie definicji dziennego wyniku handlowego, pozwalającej na test „nieskażony”, jest najbardziej właściwe (np. analizy historyczne oparte na hipotetycznych zmianach wartości portfela, które wystąpiłyby, gdyby pozycje na koniec dnia miały pozostać niezmiennione).

#### 4. Przegląd wybranych modeli wartości zagrożonej

W przypadku modeli parametrycznych problem estymacji VaR sprowadza się do wyznaczenia nieznanymi parametrów rozkładu. W przypadku najczęściej używanego w praktyce w tej klasie modeli rozkładu normalnego: średniej  $\mu$  i odchylenia standardowego  $\sigma$ . Mając dane te parametry rozkładu, można obliczyć dowolny kwantyl, co prowadzi do określenia *Value at Risk* jako  $VaR(\alpha) = \mu + \sigma\Phi(\alpha)$ , gdzie  $\Phi$  jest poziomem ufności, a  $\alpha$  oznacza dystrybuantę rozkładu normalnego  $N[0,1]$ . W przypadku pierwszej z wymienionych wyżej metod (oznaczonej dalej jako *VaR OS*) odchylenie standardowe stóp zwrotu z portfela dla szeregu danych historycznych długości  $k$  dane jest wzorem:

$$\sigma_t = \sqrt{\frac{1}{k-1} \sum_{i=t-k}^{t-1} (x_i - \mu_k)^2}, \quad (10)$$

gdzie  $\sigma_t$  oznacza estymator odchylenia standardowego stóp zwrotu portfela na początku okresu  $t$ ,  $x_i$  oznacza zmianę wartości portfela w okresie  $i$ ,  $\mu_k$  zaś oznacza średnią zmianę wartości portfela w okresie  $k$ .



Powyższy model bazuje na założeniu, iż stopy zwrotu z portfela podlegają takiemu samemu rozkładowi (normalnemu) i są od siebie niezależne. Badania empiryczne różnych finansowych szeregów czasowych [Tsay 2002] wykazały jednak występowanie w szeregach stóp zwrotu efektu skupiania zmienności (*volatility clustering*), oznaczającego, iż zarówno małe, jak i duże zmiany kursu następują seriami, a tym samym powodują niestałość wariancji stóp zwrotu w czasie, co jest sprzeczne z założeniem przedstawionego powyżej modelu. Jednym z rozwiązań tego problemu jest drugi najpopularniejszy wariant metody parametrycznej, tj. wyznaczenie zmienności z użyciem wykładniczo ważonej średniej ruchomej. W odróżnieniu od metody opisanego powyżej podejście to przypisuje różne wagi poszczególnym obserwacjom w szeregu historycznym. Ponieważ wagi zmniejszają się wykładniczo, najnowsze obserwacje mają dużo większe znaczenie dla wyznaczenia zmienności, wpływ zaś obserwacji znajdujących się daleko w przeszłości jest relatywnie mniejszy.

W modelu RiskMetrics [Morgan 1995] (oznaczonym dalej jako *VaR RM*) odchylenie standardowe stóp zwrotu z portfela dla szeregu danych historycznych długości  $k$  dane jest wzorem:

$$\sigma_t = \sqrt{(1-\lambda) \sum_{i=t-k}^{t-1} \lambda^{t-i-1} (x_i - \mu_k)^2}, \quad (11)$$

gdzie waga  $\lambda$  określa tempo, w jakim obserwacje historyczne tracą na znaczeniu. Celem tego podejścia jest przypisanie większego znaczenia krótkoterminowym wahaniom zmienności, a więc w konsekwencji zapewnienie większego tempa zmian *VaR* w zależności od aktualnego zachowania rynku. Model ten jest odpowiednikiem modelu IGARCH(1,1). Zmienność wyznaczona na dzień  $t$  jest wypadkową średniej ważonej poprzednich dni, które otrzymują wagę  $\lambda$ , i obserwacji z ostatniego dnia z wagą  $(1-\lambda)$ . W konsekwencji im niższa wartość  $\lambda$ , tym większe znaczenie mają obserwacje najbliższe i jednocześnie tym szybciej tracą na znaczeniu obserwacje bardziej odległe.

W metodzie symulacji historycznej (oznaczonej dalej jako *VaR HS*), zamiast dokonywać estymacji odchylenia standardowego z użyciem szeregów czasowych, używa się ich bezpośrednio do budowy dystrybuanty empirycznej, zaś *VaR*( $\alpha$ ) wyznacza się jako  $\alpha$ -kwantyl tej dystrybuanty. Jest to najpopularniejsza w bankach metoda wyznaczania *VaR* – jak podaje opracowanie [Perignon, Smith 2006] spośród 64,9% ankietowanych banków, które ujawniły stosowaną metodę *VaR*, aż 74% stosowało metodę symulacji historycznej. Jeśli oznaczymy przez  $L_p, \dots, L_k$  uporządkowane rosnąco wartości z dystrybuanty empirycznej zysków/strat, to estymator symulacji historycznej  $VaR_{HS,\alpha}$  będzie dany wzorem [Bałamut 2002]:

$$VaR_{HS,\alpha} = L_i, \quad \text{gdzie: } \frac{i-1}{k} < \alpha < \frac{i}{k}. \quad (12)$$

## 5. Wyniki weryfikacji historycznej wybranych modeli wartości zagrożonej na przykładzie polskiego rynku finansowego

W części niniejszej przedstawiono wyniki badań, w których dokonano, z użyciem wcześniej omówionych metod, weryfikacji historycznej wybranych modeli wartości zagrożonej na przykładzie polskiego rynku walutowego, akcji oraz stóp procentowych w latach 2002-2012. Do analizy użyto ogółem 9039 obserwacji trzech szeregów czasowych: indeksu WIG, kursu USD/PLN oraz rentowności 5-letnich swapów procentowych w PLN. Szczegółowym badaniom poddano dwa najczęściej używane w praktyce modele parametryczne – model, w którym wyznaczenia zmienności dokonuje się z użyciem równo- oraz wykładniczo ważonych obserwacji stóp zwrotu z portfela, oraz model symulacji historycznej. Analizy dokonano dla wartości zagrożonej wyznaczonej z użyciem 3 poziomów ufności: 99, 97,5 oraz 95%. Łącznie pozwoliło to na zbadanie 74 574 dziennych obserwacji wartości zagrożonej i ich weryfikacji historycznej.

W tabeli 2 podsumowano wyniki przeprowadzonych testów. Z przedstawionych wyników widać, że przy najwyższym z poziomów ufności (99%) użytych do wyznaczenia wartości zagrożonej praktycznie wszystkie z użytych testów weryfikacji historycznej sugerują odrzucenie badanych modeli wartości zagrożonej. Przy współczynniku ufności 95% przedział ufności w tym przypadku wynosi  $\langle 23, 32 \rangle$  przekroczeń w badanym szeregu. Pozytywnie wypada jedynie test  $LR_{IND}$  weryfikujący hipotezę o niezależności przekroczeń backtestingu w czasie, jednak ma on jedynie znaczenie wtórne w kontekście weryfikacji negatywnej liczby przekroczeń w stosunku do poziomu ufności użytego do wyznaczenia wartości zagrożonej, potwierdzonej przez rezultaty testów *Lopez I* i  $LR_{POF}$ .

W przypadku testów przeprowadzanych dla wartości zagrożonej wyznaczonej z poziomem ufności 97,5% test *Lopez I* sugeruje odrzucenie wszystkich, z wyjątkiem jednego, z badanych modeli wartości zagrożonej. Przy współczynniku ufności 95% przedział ufności w tym przypadku wynosi  $\langle 62, 76 \rangle$  przekroczeń w badanym szeregu. Test  $LR_{POF}$  generuje nieco odmienne wyniki, sugerując akceptację 6 z 9 badanych modeli. Jednakże weryfikacja kolejnej hipotezy – o niezależności przekroczeń backtestingu w czasie, za pomocą testu  $LR_{IND}$ , ogranicza liczbę modeli, które mogłyby być zaakceptowane. Podsumowując, należy stwierdzić, że dla wartości zagrożonej wyznaczonej z poziomem ufności 97,5% łączna analiza wyników testu *Lopez I* oraz testów *LR* nie daje podstaw do akceptacji żadnego z analizowanych modeli *VaR*.

W przypadku testów przeprowadzanych dla wartości zagrożonej wyznaczonej z poziomem ufności 95% test *Lopez I* sugeruje akceptację wszystkich z wyjątkiem dwóch z badanych modeli wartości zagrożonej. Przy współczynniku ufności 95% przedział ufności w tym przypadku wynosi  $\langle 129, 147 \rangle$  przekroczeń w badanym szeregu. Również test  $LR_{POF}$  sugeruje akceptację wszystkich, z wyjątkiem jednego,

**Tabela 2.** Wyniki przeprowadzonej weryfikacji historycznej. Dane dzienne w okresie 2002-2012

	WIG			USDPLN			IRS SY PLN		
	VaR RM	VaR OS	VaR HS	VaR RM	VaR OS	VaR HS	VaR RM	VaR OS	VaR HS
średnia		0,04%			-0,01%			0,01%	
odch std		1,32%			0,98%			0,29%	
kurtოza		3,07			3,44			5,75	
skośność		-0,36			0,09			-0,27	
<b>Ile przekroczeń (99%)</b>	51	51	46	40	47	44	45	47	39
Lopez I	1,85%	1,85%	1,67%	1,45%	1,70%	1,59%	1,63%	1,70%	1,41%
zielona	39,59%	51,50%	48,61%	69,00%	56,90%	62,48%	57,04%	55,92%	52,30%
żółta	59,72%	25,93%	43,10%	31,00%	33,47%	30,50%	39,01%	32,71%	47,70%
czerwona	0,69%	22,56%	8,29%	0,00%	9,63%	7,03%	3,95%	11,37%	0,00%
LR <sub>POF</sub>	16,01	16,01	10,31	4,93	11,36	8,33	9,29	11,36	4,21
LR <sub>IND</sub>	<b>0,94</b>	22,21	15,49	<b>0,22</b>	<b>0,05</b>	<b>0,12</b>	7,53	24,82	14,23
LR <sub>MIX</sub>	16,95	38,23	25,80	<b>5,15</b>	11,41	8,44	16,83	36,19	18,44
<b>Ile przekroczeń (97,5%)</b>	82	92	79	86	85	93	82	<b>70</b>	78
Lopez I	2,97%	3,33%	2,86%	3,11%	3,08%	3,37%	2,97%	<b>2,54%</b>	2,83%
LR <sub>POF</sub>	<b>2,36</b>	7,11	<b>1,41</b>	3,98	<b>3,54</b>	7,72	<b>2,36</b>	<b>0,01</b>	<b>1,15</b>
LR <sub>IND</sub>	<b>2,23</b>	23,50	19,43	<b>0,04</b>	5,49	<b>2,25</b>	6,19	33,28	23,76
LR <sub>MIX</sub>	<b>4,59</b>	30,62	20,84	<b>4,02</b>	9,02	9,97	8,55	33,29	24,91
<b>Ile przekroczeń (95%)</b>	<b>134</b>	<b>129</b>	<b>145</b>	<b>147</b>	<b>137</b>	<b>159</b>	<b>145</b>	<b>110</b>	<b>131</b>
Lopez I	<b>4,85%</b>	<b>4,67%</b>	<b>5,25%</b>	<b>5,32%</b>	<b>4,96%</b>	5,76%	<b>5,25%</b>	3,98%	<b>4,74%</b>
LR <sub>POF</sub>	<b>0,13</b>	<b>0,64</b>	<b>0,36</b>	<b>0,60</b>	<b>0,01</b>	<b>3,20</b>	<b>0,36</b>	6,43	<b>0,39</b>
LR <sub>IND</sub>	9,26	43,08	29,37	<b>0,19</b>	<b>3,66</b>	<b>1,63</b>	8,06	34,44	34,72
LR <sub>MIX</sub>	9,39	43,72	29,74	<b>0,79</b>	<b>3,67</b>	<b>4,82</b>	8,42	40,87	35,11

Źródło: opracowanie własne. W tabeli kolorem szarym oznaczono pozytywne wyniki testów.

badanych modeli wartości zagrożonej. Jednakże analiza wyników testów hipotezy o niezależności przekroczeń backtestingu w czasie ( $LR_{IND}$  i  $LR_{MIX}$ ) sugeruje odrzucenie wszystkich modeli wartości zagrożonej wyznaczonych na podstawie danych historycznych pochodzących z rynku akcji oraz stóp procentowych.

## 6. Podsumowanie

Celem niniejszej pracy była weryfikacja historyczna wybranych modeli wartości zagrożonej w okresie kryzysu finansowego na przykładzie polskich rynków akcji, obligacji i rynku walutowego. Przeprowadzone wyniki testów nie dają podstaw do jednoznacznej statystycznie akceptacji żadnego z analizowanych modeli. Jak pokazało niniejsze badanie, im wyższy poziom ufności, dla którego wyznaczana jest wartość zagrożona, tym niższa jakość progностyczna analizowanych modeli  $VaR$ . Dla najwyższego z badanych poziomów ufności wszystkie z analizowanych modeli niedoszacowywały wartości zagrożonej. Wyniki testów hipotez o niezależności przekroczeń backtestingu w czasie pokazały, iż w większości przypadków dla badanych szeregów przekroczenia te nie są niezależne.

Pomimo iż wyniki te mogą wydawać się nieco alarmujące, podkreślić należy, iż wartość zagrożona została wprowadzona jako miara ryzyka portfelowego w ‘normalnych’ warunkach rynkowych [RiskMetrics 1994]. W niniejszym badaniu użyto szeregów danych historycznych obejmujących kilka kryzysów finansowych, szczególnie moment po upadku Lehman Brothers (przełom 2008/2009), oraz okres gwał-

townych wahań rynkowych związanych z obawami inwestorów o potencjalną wypłacalność niektórych z państw strefy euro (wiosna 2010). W tym kontekście należy zwrócić uwagę, iż już Komitet Bazylejski [BCBS 1996], wprowadzając miarę *VaR* jako podstawę wyznaczania wymogów kapitałowych na ryzyko rynkowe z użyciem metod statystycznych, zwrócił uwagę, że w przypadku istotnych zmian reżimu rynkowego parametry rynkowe mogą ulec istotnym przesunięciom i żaden z modeli wartości zagrożonej nie będzie w stanie poprawnie ich uwzględnić w momencie ich wystąpienia, jako że korzystają one wyłącznie z danych historycznych do oceny przyszłych wydarzeń. To z kolei nakazuje daleko idącą ostrożność w stosowaniu wartości zagrożonej jako jedyne narzędzia ilościowego zarządzania ryzykiem rynkowym. W celu ostrożnego i stabilnego zarządzania instytucją finansową miara ta powinna być uzupełniona metodami alternatywnymi, takimi jak np. analiza scenariuszy i stress testy.

Opisane w niniejszej pracy podejścia koncentrowały się na ocenie prawidłowości modeli wartości zagrożonej z punktu widzenia liczby przekroczeń występujących w rzeczywistym modelu w stosunku do zakładanego teoretycznie stopnia pokrycia. Takie podejście pomija jednak zasób informacji, który może być przydatny w ocenie modeli wartości zagrożonej w praktycznym zastosowaniu jako miary determinującej wysokość minimalnego kapitału regulacyjnego na pokrycie ryzyka rynkowego, tj. informację o skali strat w przypadku wystąpienia przekroczenia. Badania takie byłyby potencjalnym rozwinięciem tematu poruszonego w niniejszej pracy.

## Literatura

- Alexander C., *Market Risk Analysis*, Volume IV, „Value at Risk Models”, John Wiley & Sons, 2009.
- Bałamut T., *Metody estymacji Value at Risk*, Materiały i Studia NBP, Zeszyt nr 147, Warszawa, sierpień 2002.
- Basel Committee on Banking Supervision, *Amendment to the capital accord to incorporate market risks*, Basel, January 1996.
- Basel Committee on Banking Supervision, *Revisions to the Basel II market risk framework – final version*, Basel, July 2009.
- Basel Committee on Banking Supervision, *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework – Comprehensive Version*, Basel, June 2006.
- Basel Committee on Banking Supervision, *Supervisory framework for the use of ‘backtesting’ in conjunction with the internal models approach to market risk capital requirements*, Basel, January 1996.
- Christoffersen P., *Evaluating interval forecasts*, “International Economic Review”, 39, 1998.
- Christoffersen P., Pelletier D., *Backtesting value-at-risk: a duration-based approach*, “Journal of Empirical Finance” 2004, 2.
- Danielsson J., Embrechts P., Goodhart C., Keating C. i in., *An academic response to Basel II*, Special Paper 130, ESRC Research Centre, Swindon, Wiltshire 2001.
- Grabowska A., *Metody kalkulacji wartości narażonej na ryzyko (VaR)*, „Bank i Kredyt” 2000, 10.
- Jędrusik S., Paliński A., Chmiel W., Kadłuczka P., *Testowanie wsteczne modeli wartości narażonej na stratę*, „Ekonomia Menedżerska”, Wydawnictwo AGH, 1/2007.

- Iskra D., *Wartość zagrożona instrumentu finansowego szacowana przedziałowo*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 254, UE, Wrocław 2012.
- Morgan J.P., *RiskMetrics Technical Document*, 3 edition, New York 1995.
- Jorion P., *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill Professional, 3 edition, 2006.
- Kupiec P., *Techniques for verifying the accuracy of risk management models*, "Journal of Derivatives" 1995, 3.
- Lazaregue-Bazard C., *Exceptions to the rule*, "Risk Magazine", January 2010.
- Lopez J.A., *Methods for evaluating VaR estimates*, "Economic Policy Review" 1998, 4.
- Lopez J.A., *Regulatory evaluation of value-at-risk models*, "Journal of Risk" 1999.
- Ludwiczak B., *Wykorzystanie metody VAR w procesie pomiaru ryzyka*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 1(271), UE, Wrocław 2012.
- Lusztyn M., *Porównanie efektywności funkcjonowania wybranych modeli wartości zagrożonej (VaR) na polskim rynku walutowym w okresie 1996-2009*, Studia i Prace Kolegium Zarządzania i Finansów SGH, 105/2011.
- Perignon C., Smith D., *The level of quality of Value at Risk disclosures by commercial banks*, Working Paper, Simon Fraser University, 2006.
- Rossignolo A., Fethi M., Shaban M., *Value-at-Risk models and Basel capital charges: evidence from emerging and frontier stock markets*, "Journal of Financial Stability", vol. 8, Issue 4, December 2012.
- Tsay R., *Analysis of Financial Time Series*, Wiley & Sons, Chicago 2002.
- Youngman P., *Procyclicality and value at risk*, "Bank of Canada, Financial System Review", June 2009.

## **BACKTESTING OF VALUE AT RISK MEASURES – ANALYSIS OF SELECTED METHODS BASED ON THE EXAMPLE OF POLISH MARKET DURING FINANCIAL CRISIS**

**Summary:** Several banks use internal Value at Risk models to measure market risk and to calculate regulatory capital necessary to cover that risk. Backtesting is a statistical tool that allows differentiating precise and imprecise risk models. The objective of this paper is to backtest selected Value at Risk models in a period preceding and during the financial crisis, based on the example of Polish currency, equity and bond markets. The obtained results do not justify unequivocal statistical acceptance of any of the analyzed models. This in turn suggest extreme caution in using Value at Risk as the only quantitative risk management tool. Stable and cautious risk management of a financial institution calls for supplementing Value at Risk with alternative risk measures.

**Keywords:** Value at Risk, VaR, backtesting.