

Prognozowanie *Value-at-Risk* dla portfeli inwestycyjnych

Jakub Tomczak

Uniwersytet Warszawski

e-mail: jp.tomczak@student.uw.edu.pl

ORCID: [0009-0001-8315-3381](https://orcid.org/0009-0001-8315-3381)

Natalia Nehrebecka

Uniwersytet Warszawski

e-mail: nnehrebecka@wne.uw.edu.pl

ORCID: [0000-0003-3870-7231](https://orcid.org/0000-0003-3870-7231)

© Jakub Tomczak, Natalia Nehrebecka

Praca opublikowana na licencji Creative Commons Uznanie autorstwa-Na tych samych warunkach 4.0 Międzynarodowe (CC BY-SA 4.0). Skrócona treść licencji dostępna jest online na <https://creativecommons.org/licenses/by-sa/4.0/deed.pl>

Cytuj jako: Tomczak, J. i Nehrebecka, N. (2025). Prognozowanie *Value-at-Risk* dla portfeli inwestycyjnych. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 69(1), 118–137.

DOI: [10.15611/pn.2025.1.10](https://doi.org/10.15611/pn.2025.1.10)

JEL: G17, G32

Streszczenie

Cel: Celem artykułu jest wskazanie metody szacowania *Value-at-Risk*, która generuje najdokładniejsze prognozy w przypadku portfeli inwestycyjnych. W pracy przyjmuje się, że dokładność prognoz *Value-at-Risk* mierzona jest za pomocą testów Kupca i Christoffersena. Dodatkowo zbadano również, jaki jest wpływ konstrukcji danego portfela na jego podatność na ryzyko.

Metodyka: Badanie przeprowadzono, wykorzystując dane dotyczące indeksów giełdowych pobranych ze strony Stooq.com. Do oszacowania *Value-at-Risk* wykorzystana została metoda symulacji historycznej oraz dwa modele GARCH(1,1) z różnymi rozkładami teoretycznymi. Dodatkowo wyniki zostały poddane walidacji i weryfikacji z uwzględnieniem zdarzeń, takich jak brexit, pandemia COVID-19 czy atak Rosji na Ukrainę.

Wyniki: Najdokładniejsze prognozy *Value-at-Risk* otrzymano za pomocą modelu GARCH(1,1) z rozkładem skośnym t-Studenta, ponadto w praktycznie każdym przypadku modele GARCH generowały dokładniejsze prognozy niż symulacja historyczna.

Implikacje i rekomendacje: Aktywa oraz portfele inwestycyjne często narażone są na straty spowodowane wzrostem ryzyka rynkowego w okresach dużych zmian na rynkach finansowych. Prognostowanie

VaR jest jednym z kluczowych narzędzi używanych w analizie ryzyka finansowego, szczególnie w kontekście portfeli inwestycyjnych. Inwestorzy powinni łączyć VaR z innymi narzędziami analitycznymi oraz regularnie testować i aktualizować modele w zależności od zmieniających się warunków rynkowych.

Oryginalność/wartość: Zagadnienie to jest istotne z perspektywy problemu zarządzania ryzykiem przez banki i inne instytucje finansowe. Nie ma obecnie zalecanej metody szacowania *Value-at-Risk*, która byłaby standardem w branży, dlatego została podjęta próba opracowywania i testowania podejścia w obrębie sposobów wyznaczania wartości zagrożonej.

Słowa kluczowe: *Value-at-Risk*, model GARCH, portfele inwestycyjne, ryzyko rynkowe

1. Wstęp

Jednym z największych problemów, z jakimi obecnie muszą zmagać się instytucje finansowe oraz inwestorzy, jest występowanie zwiększonego ryzyka rynkowego spowodowanego dynamicznie zmieniającą się sytuacją ekonomiczną, która w znaczący sposób wpływa na rynki finansowe. Ryzyko rynkowe zostało zdefiniowane przez Komisję Nadzoru Finansowego (2024) jako: „ryzyko utraty wartości aktywów, wzrostu poziomu zobowiązań lub negatywnej zmiany wyniku finansowego w rezultacie wrażliwości na zmienność parametrów rynkowych (cen na rynku)”. Można wyróżnić trzy główne rodzaje ryzyka rynkowego: ryzyko stóp procentowych, polegające na utracie wartości aktywów spowodowanej ich wrażliwością na zmiany stóp procentowych, ryzyko kursowe (walutowe), jakim jest niebezpieczeństwo spadku wartości aktywów w konsekwencji zmian kursów walut, oraz ryzyko zmian cen towarów i akcji. Każde z powyżej wymienionych typów ryzyka stanowi poważne zagrożenie dla funkcjonowania banków oraz innych instytucji finansowych. Niewłaściwe zarządzanie ryzykiem finansowym lub jego brak może prowadzić do poważnych konsekwencji nie tylko dla danej instytucji, ale również dla całego sektora bankowego. Dobrym przykładem jest światowy kryzys finansowy w latach 2007-2009, spowodowany przez krach na rynku kredytów hipotecznych *subprime*. Przy obecnej, mało stabilnej i dynamicznie zmieniającej się sytuacji ekonomicznej na świecie prawidłowe zarządzanie ryzykiem stało się jednym z kluczowych wyzwań dla wszystkich instytucji finansowych oraz inwestorów.

Przez lata powstało wiele metod oraz instrumentów mających na celu minimalizację ryzyka rynkowego, wciąż opracowywane są również nowe podejścia oraz instrumenty do zarządzania wskazanym problemem. Do najczęściej stosowanych należą ***Value-at-Risk (VaR, wartość zagrożona)*** oraz ***Expected Shortfall (ES, oczekiwana strata warunkowa)***. Pierwszy opisuje maksymalną oczekiwaną wartość straty na danym portfelu/aktywie, która może wystąpić w warunkach rynkowych w określonym czasie i z określonym prawdopodobieństwem. *Expected Shortfall* z kolei określa oczekiwany zwrot z danego portfela lub aktywa dla ustalonego procenta najgorszych przypadków. Miary te są ze sobą zatem ściśle powiązane, gdyż ES wskazuje na to, jak duża strata wygenerowana zostanie w przypadku przekroczenia VaR o określonym poziomie dla danego aktywa. *Value-at-Risk* jest jednak miarą stosowaną powszechnie przez instytucje finansowe, dlatego to wokół wskazanego pojęcia skupiać się będzie ten artykuł. Mimo iż VaR jest miarą obligatoryjną w zakresie zarządzania ryzykiem przez banki, to nie istnieje jedna obiektywnie najlepsza metoda jej wyznaczania.

Opracowano wiele różnych metod szacowania *Value-at-Risk*, od najprostszych, takich jak metoda symulacji historycznej, którą obliczyć można za pomocą arkusza kalkulacyjnego, aż po znacznie bardziej skomplikowane metody wykorzystujące zaawansowane modele ekonometryczne (zob. Fiszedler i in., 2023; Kakade i in., 2022). Obecnie wciąż tworzone są nowe podejścia do wyznaczania wartości zagrożonej w jak najdokładniejszy sposób, np. za pomocą modeli uczenia maszynowego. Można wyróżnić jednak pewne metody wyznaczania *Value-at-Risk*, które cieszą się znacznie większą popularnością niż pozostałe, należą do nich np. modele klasy GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*). Modele tej klasy szczególnie dobrze radzą sobie z modelowaniem finansowych szeregów czasowych oraz skutecznie wykrywają występowanie faktów stylizowanych w szeregach czasowych, takich jak grupowanie wariancji, które są charakterystyczne dla szeregów zwrotów z instrumentów

finansowych. Wskazany model jest zatem często stosowany jako jedna z podstawowych metod wyznaczania VaR, ma on również różne rozszerzenia, ponadto można go stosować w połączeniu z licznymi rozkładami teoretycznymi (nie tylko z rozkładem normalnym).

Celem pracy jest wskazanie metody szacowania *Value-at-Risk*, która generuje najdokładniejsze prognozy w przypadku portfeli inwestycyjnych, uwzględniając zdarzenia, takie jak brexit, pandemia COVID-19 czy atak Rosji na Ukrainę. W pracy przyjmuje się, że dokładność prognoz *Value-at-Risk* mierzona jest za pomocą testów Kupca i Christoffersena. Wyniki zostały poddane walidacji i weryfikacji. Dodatkowo zbadano również, jaki jest wpływ konstrukcji danego portfela na jego podatność na ryzyko.

Zagadnienie to jest istotne z perspektywy problemu zarządzania ryzykiem przez banki i inne instytucje finansowe. Nie istnieje obecnie jedna najlepsza metoda szacowania *Value-at-Risk* będąca standardem w branży, wciąż opracowywane i testowane są nowe podejścia w obrębie sposobów wyznaczania wartości zagrożonej. Ponadto aktywa oraz portfele inwestycyjne często narażone są na straty spowodowane wzrostem ryzyka rynkowego w okresach dużych zmian na rynkach finansowych. Istotne jest zatem znalezienie modelu, który jest w stanie jak najdokładniej oszacować przyszłe zmiany wartości danego aktywa i zniwelować wystąpienie ryzyka rynkowego. Rozważona zostanie metoda wyznaczania VaR – symulacja historyczna oraz dwa modele GARCH z różnymi rozkładami. Analiza przeprowadzona zostanie na dwóch różnych portfelach inwestycyjnych w okresie od początku 2016 roku do kwietnia 2023 roku. Dane dotyczące indeksów giełdowych pobrano ze strony Stooq.com.

W artykule zaprezentowano metodykę prognozowania VaR, która uwzględnia zmiany w zmienności rynków finansowych, zwłaszcza w okresach kryzysów lub dużych fluktuacji. Oryginalnym wkładem jest rozważenie podejść, które dostosowują prognozy do aktualnej sytuacji rynkowej, zamiast polegać na stałych, historycznych danych.

2. Przegląd literatury i sformułowanie hipotez

Berkowitz i O'Brien (2002) analizowali modele stosowane przez banki komercyjne służące do szacowania *Value-at-Risk* dla szeregów dziennych zysków i strat z działalności bankowej sześciu banków, następnie porównali ich skuteczność z modelem GARCH(1,1). Doszli oni do wniosku, że średni wskaźnik naruszenia dla VaR szeregu czasowego wyznaczonej metodą GARCH jest niższy niż w przypadku wartości zagrożonej wyliczanej przez banki. Maksymalne naruszenia są podobne dla obu metod, jednak banki stosują bardziej konserwatywne podejście w wyznaczaniu VaR, co potwierdza skuteczność modelu GARCH, ponieważ za jego pomocą uzyskano dokładniejsze oszacowania wartości zagrożonej niż przy użyciu skomplikowanych modeli opracowanych przez banki. Model GARCH dobrze sprawdza się również w przypadku wyznaczania *Value-at-Risk* w niestacjonarnych szeregach czasowych. Jego zaletą jest skuteczność w modelowaniu szeregów czasowych o wariancji zmiennej w czasie (szeregi heteroskedastyczne) oraz takich, w których występuje zjawisko grupowania wariancji (Karanđikar i in., 2009). Brooks i Persaud (2003) testowali, które metody najlepiej nadają się do szacowania zmienności, a jednym z kryteriów był poziom dokładności wyznaczania *Value-at-Risk*. W artykule zaprezentowano modele o różnym poziomie skomplikowania, w tym pięć różnych modeli klasy GARCH, które wyróżniają się skutecznością pod względem szacowania wartości zagrożonej. Na poziomie istotności 5% szczególnie pozytywnie wyróżnia się standardowy GARCH(1,1), którego rzeczywisty wskaźnik pokrycia (*likelihood ratio*, LR) jest bliski wskaźnikowi nominalnemu.

Powyższe artykuły wskazały sytuacje, w których standardowy model GARCH ma przewagę w szacowaniu *Value-at-Risk*. Obecnie jednak coraz częściej do wyznaczania wartości zagrożonej stosuje się inne, bardziej skomplikowane modele klasy GARCH, lub inne bardziej zaawansowane modele ekonometryczne. McMillan i Speight (2007) zaproponowali wykorzystanie modeli GARCH II i III generacji w celu wyznaczenia VaR. Przewagą modeli II generacji nad standardowym GARCH jest uwzględnienie przez nie możliwości asymetrycznego wpływu pozytywnych i negatywnych szoków na wariancję warunkową. Są one również skuteczne w wykrywaniu efektu dźwigni w danym szeregu czasowym. Modele III generacji uwzględniają nie tylko asymetryczny wpływ szoków, ale również efekt długiej

pamięci, który polega na uwzględnieniu wpływu wysoce trwałych, lecz przejściowych szoków w prognozowaniu zmienności. Oznacza to, że łączna zmienność obejmuje składniki zarówno długo-okresowe, jak i bardziej przejściowe. Wong i in. (2016) pokazali, że wśród wybranych modeli klasy GARCH i modeli opartych na zrealizowanej zmienności (*realized volatility*) najbardziej skuteczny w prognozowaniu *Value-at-Risk* na wybranych poziomach istotności jest model E-GARCH, którego przewagą jest uwzględnienie asymetrycznego zachowania zmienności oraz efektu dźwigni. Elenjical i in. (2016) przeprowadzili natomiast analizę międzyreżimową (*cross-regime*), która odnosi się do pytania o odporność danego modelu w różnych okresach, w których warunki rynkowe ulegają zmianie, w celu oceny dokładności prognozowania każdego modelu VaR w stosunku do jego alternatyw w różnych warunkach rynkowych. Najlepsze oszacowanie VaR w przypadku całego badanego okresu, obejmującego różne reżimy rynkowe, uzyskane zostało za pomocą modelu z długą pamięcią (FIEGARCH). Problemem rozpatrywanego modelu jest jego tendencja do nadmiernego kompensowania rzeczywistych strat w okresach przedkryzysowych, a niedoszacowania w okresach kryzysów. W przypadku rozpatrywania podokresów osobno znacznie lepiej wypadają inne modele, a szczególnie E-GARCH.

Modele klasy GARCH mogą być rozpatrywane nie tylko z rozkładem normalnym, ale również z innymi rozkładami statystycznymi. Szczególnie dobrze w szacowaniu *Value-at-Risk* radzą sobie rozkłady leptokurtyczne, a zwłaszcza *t*-Studenta, który w połączeniu z prostym modelem E-GARCH generuje bardzo dokładne prognozy wartości zagrożonej dla większości rynków (Angelidis i in., 2004). Abounoori i in. (2016) również zaobserwowali przewagę rozkładów *t*-Studenta oraz GED nad rozkładem normalnym. Zastosowali oni również model GARCH w połączeniu z przełącznikowym modelem Markowa (*Markov switching model*) oraz modelem autoregresyjnym (AR) w celu dokładniejszego oszacowania wartości zagrożonej. Jego przewagą jest możliwość przełączania niektórych lub wszystkich parametrów modelu pomiędzy różnymi reżimami zgodnie z procesem Markowa. Kolejny skuteczny model do szacowania *Value-at-Risk* został zaproponowany przez Chen i Watanabe'ego (2018). Progowy zrealizowany GARCH (*threshold realized GARCH*) modeluje wspólnie dzienne zwroty oraz zrealizowaną zmienność, uwzględniając tendencyjność i asymetrię zrealizowanej zmienności. Stosowanie wskazanego modelu w połączeniu z rozkładem *t*-Studenta prowadzi do znacznej poprawy prognozowania zmienności i szacowania VaR w porównaniu z bardziej powszechnymi modelami GARCH. Wśród modeli klasy GARCH możemy wyróżnić modele jedno- oraz wielowymiarowe, przy czym te pierwsze są znacznie szerzej stosowane, a modele wielowymiarowe stosuje się wyłącznie do szacowania VaR dla portfeli inwestycyjnych. W większości przypadków nie ma preferencji pomiędzy tymi modelami w dokładności szacowania wartości zagrożonej, jednak gdy taka preferencja istnieje, to na korzyść modeli wielowymiarowych (Santos i in., 2013). Szczególnie widoczna jest w przypadku obliczania VaR dla dużych i zdywersyfikowanych portfeli. Wybór modelu wielowymiarowego najlepiej sprawdza się, gdy celem jest pomiar i prognoza wartości zagrożonej związanej z liniową kombinacją poszczególnych aktywów w portfelu, warto wtedy modelować wspólną dynamikę tych aktywów za pomocą wskazanego modelu. Fiszeder i in. (2023) zaobserwowali, że istnieją modele wielowymiarowe bardziej skuteczne niż DCC- oraz CCC-GARCH (wielowymiarowe GARCH). DCC-OHLC (*Dynamic Conditional Correlation – Opening, High, Low and Closing*) jest modelem wielowymiarowym, w którym dzienne ceny otwarcia, wyżu, niżu i zamknięcia wykorzystywane są w równaniach zarówno wariancji, jak i korelacji.

Wśród metod stosowanych do obliczania *Value-at-Risk* można wyróżnić metodę symulacji historycznej lub wariancji-kowariancji, nazywaną również parametryczną, która stosowana jest w połączeniu z różnymi modelami do analizy szeregów czasowych, np. GARCH. Istnieje również metoda filtrowanej symulacji historycznej, łącząca dwa powyższe podejścia i przewagi z nich wynikające, zapewniająca również lepsze wyniki niż te metody (Kakade i in., 2022). Sadeghi i Shavvalpour (2006) opracowali metodę symulacji historycznej z prognozowaniem ARMA (*autoregressive moving-average model*): HSAF (*historical simulation ARMA forecasting*), która różni się od standardowego podejścia tym, że nie wykorzystuje bezpośrednio rozkładu przeszłych stóp zwrotu, a rozkład błędów prognoz, uzyskany z oszacowanego modelu ARMA. Metoda ta w niektórych przypadkach potrafi generować prognozy VaR lepsze niż modele GARCH.

Literatura dotycząca szacowania *Value-at-Risk* za pomocą różnych metod skupia się przede wszystkim na opracowywaniu i porównywaniu coraz to nowszych i bardziej skomplikowanych modeli w celu osiągnię-

cia jak najdokładniejszych prognoz wartości zagrożonej. W niniejszym artykule duży nacisk położony został na samą konstrukcję modelu do szacowania VaR, z uwzględnieniem zmian w zmienności rynków finansowych, zwłaszcza w okresach kryzysów lub dużych fluktuacji.

Na podstawie literatury przedmiotu w artykule sformułowano hipotezy badawcze, które niniejsza praca ma na celu zweryfikować. Angelidis i in. (2004) zaobserwowali, że modele GARCH z rozkładem *t*-Studenta dokładniej szacują *Value-at-Risk* (VaR) niż modele oparte na rozkładzie normalnym, szczególnie w przypadku leptokurtycznych rozkładów zwrotów. Z kolei Guermat i Harris (2002) zauważyli, że modele GARCH z rozkładem *t*-Studenta przewyższają inne modele tej klasy w estymacji VaR dla portfeli inwestycyjnych. W związku z tym **pierwsza hipoteza badawcza** postawiona w artykule zakłada, że modele GARCH z rozkładem *t*-Studenta dokładniej szacują *Value-at-Risk* niż modele z rozkładem normalnym, zwłaszcza w warunkach występowania leptokurtycznych rozkładów zwrotów, z uwzględnieniem okresów takich jak pandemia COVID-19 oraz agresja Rosji na Ukrainę.

Brooks i Persaud (2003) stwierdzili, że portfele inwestycyjne są mniej podatne na ryzyko rynkowe niż poszczególne aktywa. Nie zbadali jednak, w jaki sposób liczba składników portfela wpływa na jego podatność na ryzyko. Dlatego warto poddać weryfikacji również **drugą hipotezę badawczą**, zgodnie z którą portfele składające się z większej liczby instrumentów są bardziej odporne na ryzyko rynkowe niż te zawierające mniejszą liczbę aktywów. W niniejszej pracy hipoteza ta zostanie zweryfikowana na podstawie dokładności oszacowań *Value-at-Risk* – im dokładniejsze są prognozy VaR dla danego portfela, tym większą odporność na ryzyko wykazuje ten portfel.

3. Metodyka

3.1. Metody użyte do szacowania *Value-at-Risk*

Jedną z metod wybranych do wyznaczania wartości zagrożonej jest symulacja historyczna. Jest to najmniej skomplikowana z wybranych metod szacowania *Value-at-Risk*, przez co wyniki osiągnięte za jej pomocą nie zawsze są bardzo dokładne. Jej precyzja wzrasta natomiast w przypadku dużej próby obserwacji. Oszacowania tą metodą wymagają wyboru odpowiedniego przedziału danych historycznych, który służy za rozkład empiryczny. Jednodniowe prognozy VaR są szacowane przy użyciu okna kroczącego dla całego szacowanego okresu *out-of-sample*.

Kolejną z metod wyznaczania *Value-at-Risk* jest zastosowanie modelu GARCH(p, q). Do badania przeprowadzonego w niniejszym artykule użyty zostanie model GARCH z parametrami $p = 1$, $q = 1$, którego warunkowa wariancja zależy od jej własnych opóźnień i kwadratów przeszłych wartości błędów. Model GARCH(1,1) został wybrany jako kluczowa metoda prognozowania *Value-at-Risk* w tym artykule między innymi ze względu na popularność – jest to jedna z najczęściej używanych metod szacowania VaR. Jest on również bardzo skuteczny w modelowaniu skupień wariancji, których występowanie stwierdzone zostało w obu analizowanych portfelach. W badaniu zastosowano do szacowania VaR modele GARCH z następującymi rozkładami: normalnym oraz skośnym *t*-Studenta. Wybór różnych rozkładów wynika z tego, że zwroty z danych instrumentów nie zawsze muszą mieć rozkład normalny, zatem wypróbowane zostaną również inne rozkłady, które mogą lepiej pasować. Zdecydowano się na wybór jednego rozkładu ze skośnością, gdyż w obu analizowanych portfelach występuje asymetria lewo- albo prawostronna, można oczekiwać zatem, że model GARCH z takim rozkładem będzie lepiej dopasowany. Angelidis i in. (2004) zauważyli, że model GARCH z rozkładem *t*-Studenta generuje znacznie dokładniejsze prognozy VaR w przypadku występowania leptokurtycznych rozkładów zwrotów. Za jego pomocą w większości przypadków można osiągnąć również lepsze oszacowania *Value-at-Risk* dla zwrotów z portfeli inwestycyjnych niż używając modelu GARCH z rozkładem normalnym (Guermat i Harris, 2002).

Value-at-Risk z zastosowaniem modelu GARCH można obliczyć za pomocą następującego wzoru:

$$VaR_{\alpha} = \mu + \sigma_{t|t-1} F_{\alpha}, \quad (1)$$

gdzie: μ – średnia wartość zwrotów z portfela, $\sigma_{t|t-1}$ – warunkowa wariancja z uwzględnieniem zwrotów z okresu $t-1$, F_α – wartość krytyczna określona przez poziom ufności α i wybrany rozkład teoretyczny.

Wzór (1) można zastosować w przypadku każdego z omawianych modeli GARCH niezależnie od ich rozkładu.

3.2. Back-testing

Back-testing polega na ocenie prognozy otrzymanej za pomocą danej metody wyznaczania *Value-at-Risk* w testowanym okresie (*out-of-sample*), która generowana jest na podstawie danych z okresu estymacji modelu (*in-sample*). Otrzymane wartości prognozy VaR_{t+1} porównywane są z rzeczywistymi wartościami z okresu r_{t+1} , gdzie $t = T^*, \dots, T-1$ oraz T^* jest ostatnią obserwacją w okresie *in-sample*, a T ostatnią obserwacją w całym badanym okresie. Definiuje się również funkcję wskaźnika o postaci:

$$I_{t+1} = \begin{cases} 1, & \text{jeżeli } r_{t+1} \leq VaR_{t+1} \\ 0, & \text{jeżeli } r_{t+1} > VaR_{t+1} \end{cases} \quad (2)$$

Wartości tej funkcji składają się na szereg o długości $n = T - T^*$, czyli takiej samej jak szeregi prognoz i rzeczywistych wartości danej zmiennej. Liczbę przekroczeń można zatem oznaczyć jako n_1 obliczane za pomocą wzoru:

$$n_1 = \sum_{t=T^*}^{T-1} I_{t+1}. \quad (3)$$

Natomiast liczbę braku przekroczeń można zapisać jako $n_0 = n - n_1$, z kolei symbolem π oznacza się udział przekroczeń w badanej próbie liczony za pomocą wzoru:

$$\pi = \frac{n_1}{n}. \quad (4)$$

W celu oceny danej metody szacowania *Value-at-Risk* porównuje się wartość udziału przekroczeń z wcześniej ustalonym poziomem istotności α . Im bardziej wartości te zbliżone są do siebie, tym dana metoda wyznaczania wartości zagrożonej jest dokładniejsza.

Poza opisanymi powyżej intuicyjnymi metodami oceny modeli VaR, można wyróżnić również bardziej formalne sposoby ich weryfikacji, określane jako **back-testing**, do których zaliczają się testy pokrycia oraz testy niezależności. Pierwszym z nich jest **test Kupca** (test bezwarunkowego pokrycia, *unconditional coverage*), który bada, czy odsetek przekroczeń prognozowanej wartości VaR w próbie odpowiada założonemu poziomowi istotności α . Hipoteza zerowa testu zakłada, że prawdopodobieństwo przekroczenia VaR wynosi α .

$$H_0: P(I_t = 1) = \alpha. \quad (5)$$

Hipotezę tę można zweryfikować za pomocą bezwarunkowego testu pokrycia, znanego również jako test ilorazu wiarygodności (*likelihood ratio test*, LR). Test ten odrzuca hipotezę zerową o poprawności metody szacowania VaR, jeśli udział przekroczeń prognozowanej wartości VaR w badanej próbie jest statystycznie różny od przyjętego poziomu istotności α . Odrzucenie hipotezy H_0 jest w tym przypadku niepożądane, ponieważ świadczy o niedostosowaniu modelu do rzeczywistego ryzyka.

Kolejną metodą weryfikacji modeli VaR jest **test niezależności Christoffersena**. Podczas gdy test Kupca koncentruje się wyłącznie na zgodności liczby przekroczeń z założonym poziomem istotności, test Christoffersena sprawdza dodatkowo, czy przekroczenia występują w sposób losowy, czy też tworzą skupienia (klastry). Weryfikacja tego drugiego warunku jest istotna dla oceny adekwatności modelu VaR, ponieważ występowanie skupisk przekroczeń może świadczyć o niedoszacowaniu ryzyka w pew-

nich okresach, co z kolei zwiększa prawdopodobieństwo poniesienia skumulowanych strat przez portfel inwestycyjny.

Hipoteza zerowa zakłada, że przekroczenie prognozowanej wartości VaR w okresie t jest niezależne od przekroczenia w okresie $t + 1$ i że prawdopodobieństwo przekroczenia jest równe przyjętemu poziomowi istotności α :

$$H_0: P(I_t = 1 | I_t = 0) = P(I_t = 1 | I_t = 1) = \alpha. \quad (6)$$

Metodę wyznaczania *Value-at-Risk* uznaje się za poprawną, jeżeli przejdzie ona oba wyżej wymienione testy, tj. hipoteza zerowa nie zostanie w nich odrzucona. Testy te pozwalają zatem na identyfikację i eliminację nieadekwatnych modeli, jednak nie służą bezpośrednio do porównywania poszczególnych metod w celu wyłonienia tej najdokładniejszej.

Aby porównać skuteczność różnych modeli szacowania wartości zagrożonej, można zestawzić liczbę zaobserwowanych przekroczeń VaR z ich oczekiwaną liczbą. Modele, dla których liczba przekroczeń jest najbardziej zbliżona do wartości oczekiwanej, można uznać za najbardziej trafne i efektywne w estymacji ryzyka.

3.3. Opis danych

W celu przeprowadzenia badań skonstruowano dwa portfele inwestycyjne. Pierwszy z nich składa się z pięciu wybranych par walutowych: USD/PLN, EUR/PLN, GBP/PLN, CHF/PLN oraz JPY/PLN. Drugi portfel obejmuje dwa wybrane fundusze inwestycyjne typu ETF notowane na warszawskiej giełdzie: ETFSP500 (odzwierciedlający amerykański indeks S&P 500) oraz ETFDAX (odzwierciedlający niemiecki indeks DAX).

Oba portfele inwestycyjne skonstruowano poprzez przypisanie wszystkim aktywom wchodzącym w ich skład równej wagi, co w przypadku portfela ETF oznacza wagę $1/2$, a w przypadku portfela walutowego – wagę $1/5$. Dzięki temu żadne z aktywów nie będzie miało nieproporcjonalnie większego wpływu na zmienność portfela niż pozostałe. Portfel walutowy składa się z pięciu różnych aktywów, co sprawia, że powinien być bardziej odporny na ryzyko rynkowe niż mniej zdywersyfikowany portfel funduszy inwestycyjnych.

Do skonstruowania portfeli użyte zostałyienne ceny zamknięcia pobrane ze strony internetowej Stooq.com. Dane obejmują okres ponad 7 lat, od 04.01.2016 do 13.04.2023, co daje łącznie 1883 i 1813 obserwacji odpowiednio dla portfela walutowego i portfela funduszy inwestycyjnych. Okres obejmuje wydarzenia mające istotny wpływ na zachowania rynków finansowych, takie jak brexit, pandemia COVID-19 i kryzys z nią związany oraz wojna w Ukrainie, która miała szczególnie duży wpływ na państwa regionu Europy Środkowo-Wschodniej.

Badany okres podzielony został na dwie próbki: *in-sample*, na podstawie której skonstruowany zostanie model, oraz *out-of-sample*, która posłuży do weryfikacji poprawności otrzymanych wyników. Pierwsza próbka obejmuje okres od 04.01.2016 do 24.04.2022 (1633 obserwacje) dla portfela walutowego i od 04.01.2016 do 13.04.2022 (1563 obserwacje) dla portfela ETF, druga zaś po 250 obserwacji dla każdego z portfeli – odpowiednio okres od 28.04.2022 do 13.04.2023 i od 14.04.2022 do 13.04.2023.

4. Rezultaty badań

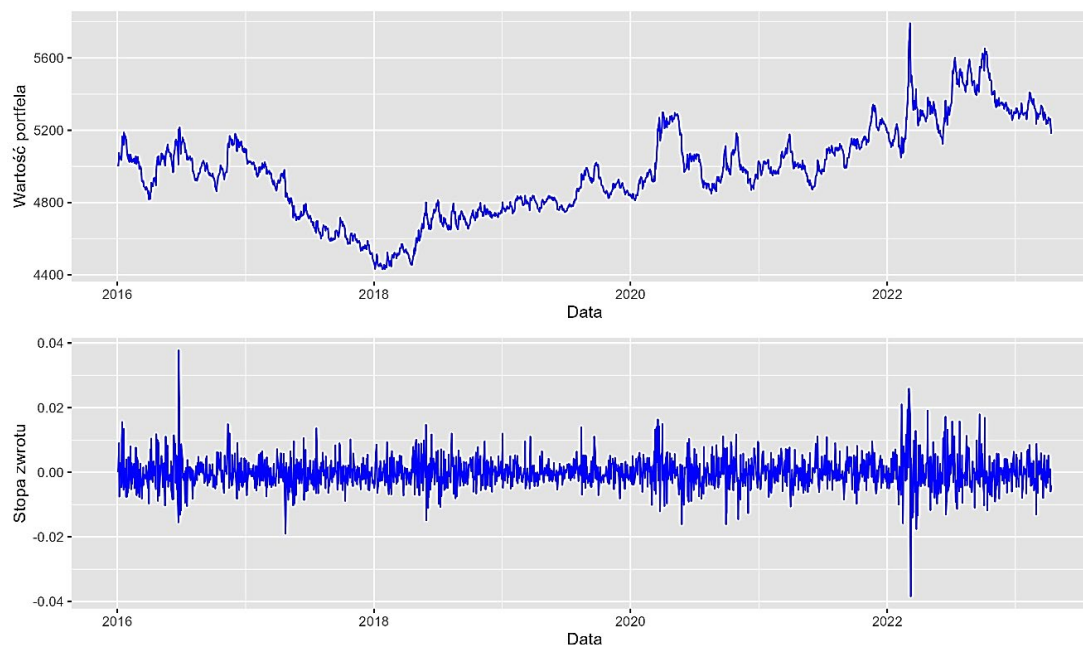
4.1. Analiza zmiennych

Stopy zwrotów z portfela przedstawione są w formie zlogarytmowanej i obliczane za pomocą następującego wzoru:

$$return_t = (\ln(p_t) - \ln(p_{t-1})), \quad (7)$$

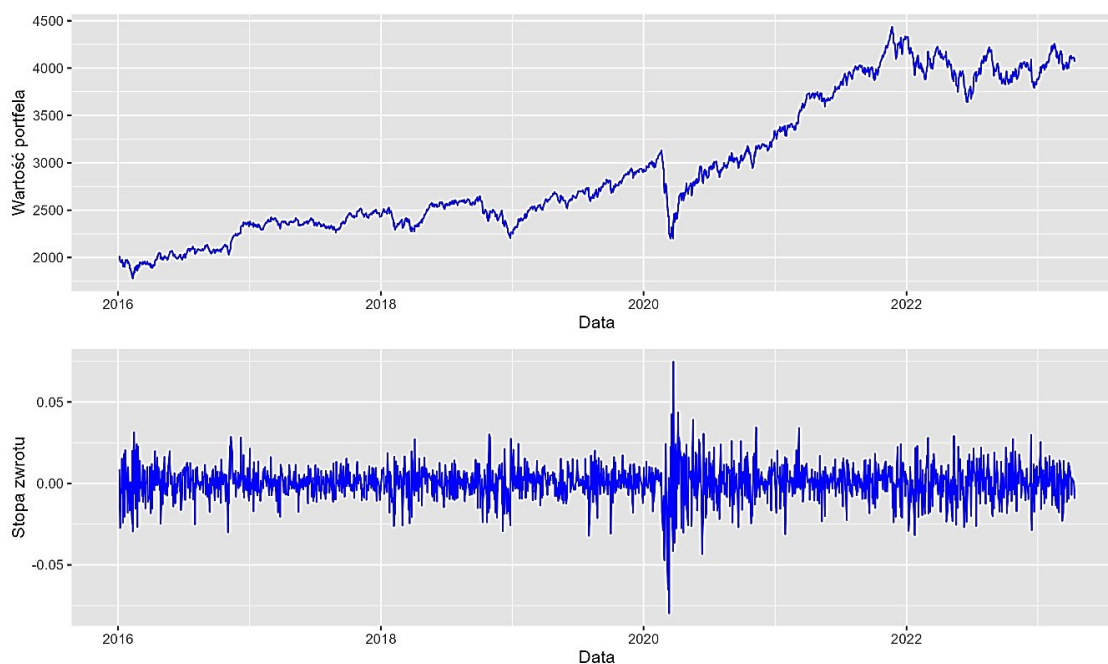
gdzie: p_t – wartość danego portfela w okresie t .

Na rysunkach 1 i 2 pokazano zmiany wartości poszczególnych zmiennych oraz ich zwrotów, odpowiednio dla portfela walutowego i portfela ETF. Oba portfele zyskały na wartości w omawianym okresie, choć osiągnięty zysk różni się między nimi na korzyść portfela funduszy inwestycyjnych. Portfel ETF wykazuje trend wzrostowy, widoczny aż do początku 2022 roku, czyli prawie przez cały badany okres. Z kolei w przypadku portfela walutowego brak jest jednoznacznego trendu wzrostowego. Można jednak podzielić badany okres na dwie części: w pierwszej, do połowy 2018 roku, widoczny jest trend spadkowy, a w drugiej, aż do końca okresu, trend wzrostowy.



Rys. 1. Wartość portfela walutowego oraz zlogarytmowane zwroty z portfela

Źródło: opracowanie własne.



Rys. 2. Wartość portfela ETF oraz zlogarytmowane zwroty z portfela

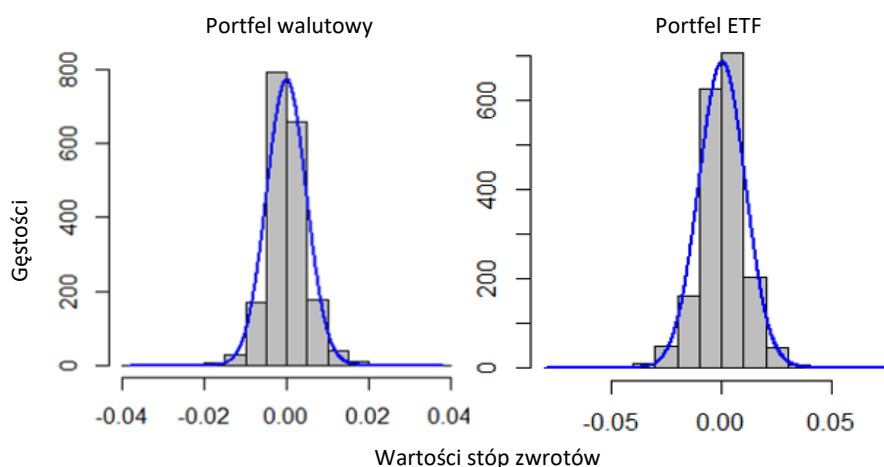
Źródło: opracowanie własne.

Warto zwrócić uwagę na obserwacje odstające, czyli nagłe zmiany wartości, które znacznie odbiegają od średniej i wpływają na dynamikę szeregu czasowego. Zjawiska te najczęściej są wynikiem istotnych wydarzeń o globalnym zasięgu, które miały wyraźny wpływ na rynki finansowe. Największy wpływ na zmiany wartości portfeli miały następujące wydarzenia: referendum w sprawie opuszczenia Unii Europejskiej przez Wielką Brytanię w czerwcu 2016 roku, wybuch pandemii COVID-19 w Europie w marcu 2020 roku oraz początek wojny w Ukrainie w lutym 2022 roku.

Na dynamikę portfela ETF szczególnie wpłynęła pandemia COVID-19, natomiast referendum w sprawie brexitu oraz atak Rosji na Ukrainę miały istotny wpływ na zmiany zwrotów w portfelu walutowym. Wpływ tych wydarzeń widać również na wykresach zlogarytmowanych zwrotów. W pierwszym kwartale 2020 roku w portfelu ETF pojawiły się zwroty (zarówno ujemne, jak i dodatnie) wyraźnie odbiegające od średniej, podobną sytuację można zaobserwować w połowie 2016 roku oraz na początku 2022 roku w przypadku portfela walutowego.

Portfel funduszy inwestycyjnych cechuje się większą zmiennością niż portfel walutowy. Wskazuje na to między innymi ponad dwukrotnie większa różnica pomiędzy najmniejszą a największą wartością zwrotu z portfela ETF w porównaniu z portfelem walutowym. Ponadto zwroty w portfelu są średnio ponad dwukrotnie większe co do wartości bezwzględnej. W portfelu walutowym występuje jednak więcej obserwacji odstających, wśród których najbardziej widoczne są te związane z brexitem oraz wybuchem wojny w Ukrainie. To drugie wydarzenie w szczególności doprowadziło do znacznych wahań w zwrotach z portfela. Portfel funduszy inwestycyjnych jest bardziej odporny na wymienione powyżej istotne dla gospodarki światowej wydarzenia, gdyż nie wpłynęły one znacząco na wielkość zwrotów w portfelu. Wyjątek stanowi początek pandemii COVID-19 w 2020 roku, który znacząco obniżył wartość portfela, wpływając istotnie na zmiany w wielkościach zwrotów, które widoczne były przez cały rok 2020.

Zjawisko grupowania wariancji (*volatility clustering*) polega na tym, że okresy o wysokiej zmienności są następstwami innych okresów o wysokiej zmienności, a okresy o niskiej zmienności występują po innych okresach o niskiej zmienności. Analizując wykresy zwrotów z obu portfeli (rys. 1 i 2), można zaobserwować skupiska, w których zmiany są znacznie wyższe od średniej. Zjawisko to najczęściej wynika z obecności obserwacji odstających, które zwiększają wariancję zwrotów w kilkudziesięciu kolejnych obserwacjach. W szczególności, jest ono najbardziej widoczne w pierwszym kwartale 2020 roku dla portfela ETF oraz w drugim kwartale 2016 roku i pierwszym kwartale 2022 roku dla portfela walutowego. Dla obu portfeli przeprowadzono również test ARCH-LM, w którym uzyskano *p-value* bliskie zeru, co oznacza, że hipoteza o braku efektu ARCH została odrzucona. Test ten stanowi formalne potwierdzenie występowania zjawiska grupowania wariancji w obu szeregach.



Rys. 3. Histogram zlogarytmowanych zwrotów z portfela walutowego i portfela ETF

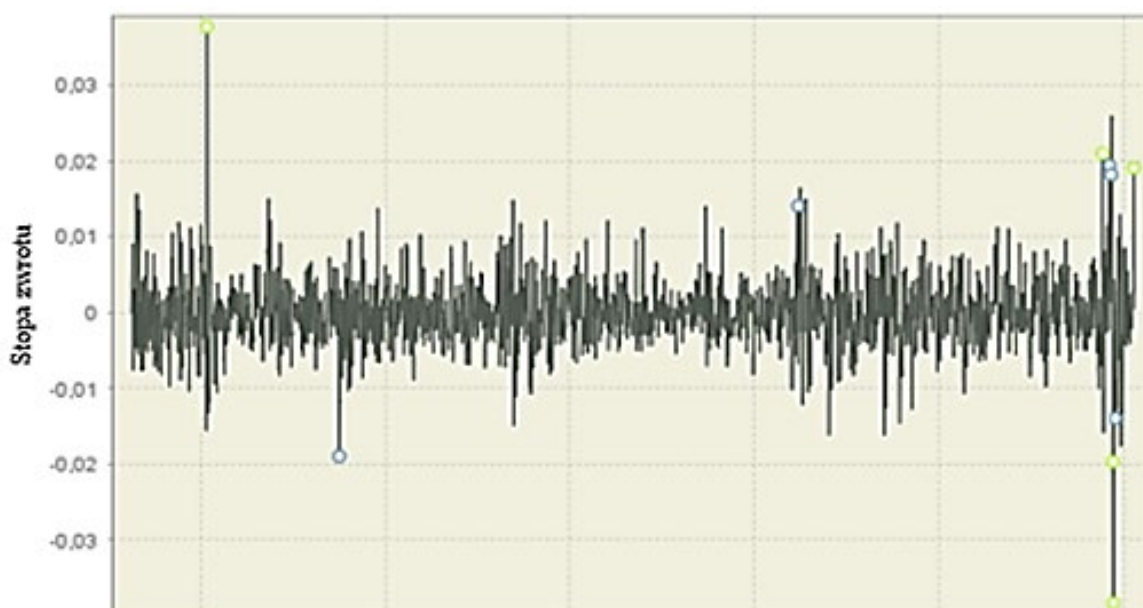
Źródło: opracowanie własne.

Portfel funduszy inwestycyjnych cechuje się większą zmiennością niż portfel walutowy, zależność ta wygląda podobnie również pomiędzy poszczególnymi składowymi tych portfeli (rys. 3). Wskazuje na to znacznie większe odchylenie standardowe portfela ETF oraz większy wpływ obserwacji odstających na zmienność zwrotów portfela. Zwroty żadnego z portfeli nie mają rozkładu normalnego, o czym świadczy bliskie zera *p-value* otrzymane w teście Jarque’a-Bery. W obu portfelach widoczna jest wysoka dodatnia kurtოza, wskazująca na dużą liczbę obserwacji odstających oraz na występowanie grubych ogonów. Ponadto zwroty z portfela funduszy inwestycyjnych cechują się delikatną lewostronną skośnością, a z portfela walutowego prawostronną, co wskazuje na występowanie asymetrii rozkładu zwrotów. Analizując część tabeli dotyczącą danych wielowymiarowych, można zauważyć, że poszczególne składowe portfela walutowego są bardziej zróżnicowane niż w przypadku portfela ETF. Wskazują na to znaczące różnice pomiędzy maksymalnymi a minimalnymi wartościami poszczególnych statystyk.

4.2. Identyfikacja obserwacji odstających

Istotnym etapem w konstrukcji poprawnego modelu GARCH jest identyfikacja obserwacji odstających. Posłużono się programem JDemetra+, który wykrył po 10 takich obserwacji dla każdego z omawianych szeregów. Dla portfela walutowego zidentyfikowano pięć obserwacji o charakterze jednorazowym (*additive outlier*, AO), które powodują zmianę wartości tylko w jednym okresie, i pięć o charakterze przejściowym (*temporary change*, TC), które powodują tymczasowe przesunięcie poziomu zmiennej zależnej, która wraca do poziomu pierwotnego po pewnym czasie zgodnie z funkcją wykładniczą. Na rysunku 4 przedstawione zostały wszystkie obserwacje odstające dla szeregu portfela walutowego.

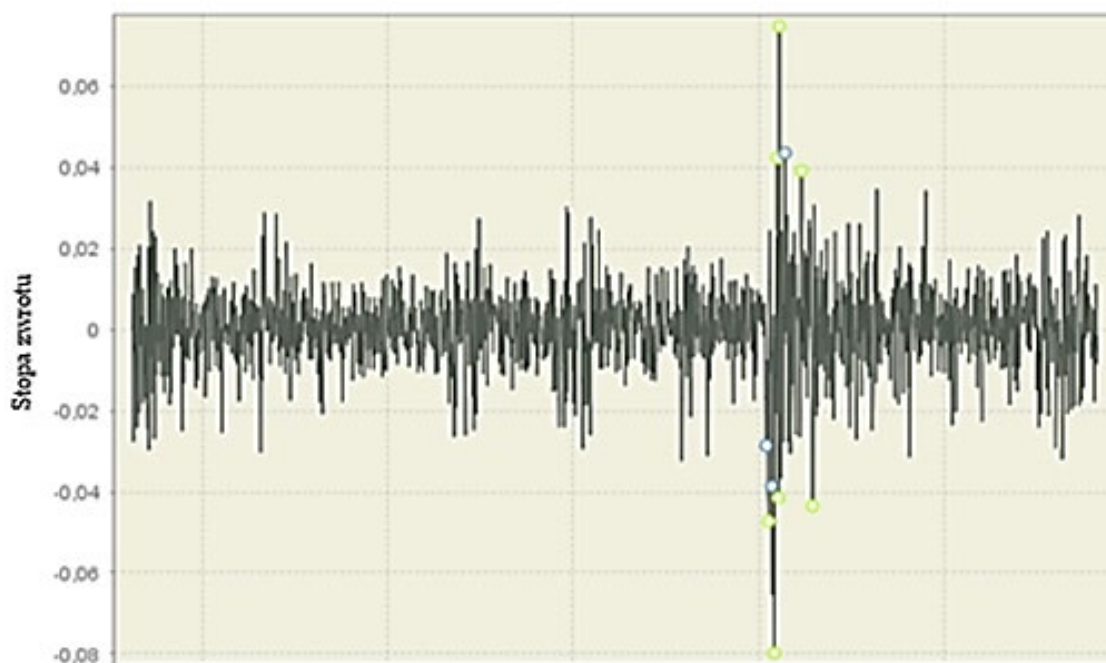
Największe ich skupisko przypada na końcówkę badanego okresu (rok 2022), co prawdopodobnie spowodowane było atakiem Rosji na Ukrainę, ponadto w 2016 roku można było zauważyć obserwację znacząco odbiegającą od średniej spowodowaną wyjściem Wielkiej Brytanii z Unii Europejskiej. W szeregu portfela ETF wykryto natomiast siedem obserwacji o charakterze jednorazowym i trzy o charakterze przejściowym – przedstawiono je na rysunku 5. Wszystkie skupione były w 2020 roku i związane bezpośrednio z wybuchem pandemii COVID-19 oraz z kryzysem przez nią spowodowanym.



Rys. 4. Obserwacje odstające dla szeregu portfela walutowego

Oznaczenia: niebieski znacznik – *temporary change*; zielony znacznik – *additive outlier*.

Źródło: opracowanie własne.



Rys. 5. Obserwacje odstające dla szeregu portfela ETF

Oznaczenia: niebieski znacznik – *temporary change*; zielony znacznik – *additive outlier*.

Źródło: opracowanie własne.

W celu wymodelowania tych obserwacji w okresie *out-of-sample* do każdego z modeli dodane zostały dwa rodzaje zmiennych odpowiednio dla obserwacji o charakterze jednorazowym (AO) i o charakterze przejściowym (TC):

$$AO_t^{(t_0)} = \begin{cases} 1 & \text{dla } t = t_0 \\ 0 & \text{dla } t \neq t_0 \end{cases}, TC_t^{(t_0)} = \begin{cases} 0 & \text{dla } t < t_0 \\ 0,8^{t-t_0} & \text{dla } t \geq t_0 \end{cases} \quad (8)$$

gdzie: t_0 – okres, w którym wystąpiła dana obserwacja nietypowa.

4.3. Estymacja modeli klasy GARCH

W badaniu skonstruowane zostaną modele *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) z różnymi rozkładami teoretycznymi dla każdego z omawianych portfeli. Pierwszym krokiem w identyfikacji procesu *AutoRegressive Integrated Moving Average* (ARIMA(p, d, q)) jest wyznaczenie wartości parametru d , który odpowiada stopniowi zintegrowania analizowanej zmiennej. Posłużono się odpowiednimi testami na stacjonarność: testem Dickeya-Fullera (ADF) oraz testem Kwiatkowskiego-Phillipsa-Schmidta-Shina (KPSS), których wyniki dla obu szeregów przedstawione zostały w tabeli 1. Wyniki testów ADF oraz KPSS dla obu portfeli wskazują, że szeregi te są stacjonarne.

Kolejnym krokiem identyfikacji procesu ARIMA(p, d, q) jest wyznaczenie pozostałych parametrów p oraz q (z badania stacjonarności wiadomo, że parametr $d = 0$ dla obu szeregów). Posłużono się w tym celu wykresami autokorelacji (*autocorrelation function*, ACF) oraz autokorelacji częściowej (*partial autocorrelation function*, PACF). Przed przystąpieniem do analizy wykresów ACF i PACF przeprowadzono również test Ljunga-Boxa (LB) dla 40. opóźnień, w którym uzyskano p -value poniżej 0,05 dla każdego z szeregów. Oznacza to, że szeregi te nie są białym szumem i można je modelować za pomocą modelu ARIMA, uzasadnione jest zatem przejście do dalszej analizy.

Tabela 1. Wyniki testów na stacjonarność dla portfela walutowego oraz portfela ETF

Portfel walutowy							
Test	Statystyka testowa	Statystyka krytyczna	Test Breuscha-Godfrey				
			1. rząd	2. rząd	3. rząd	4. rząd	5. rząd
Dickeya-Fullera	-41,732	-1,95	0,384	0,01	0,016	0,018	0,003
ADF z lag = 1	-31,83	-1,95	0,044	0,1	0,202	0,265	0,042
ADF z lag = 2	-24,783	-1,95	0,772	0,463	0,392	0,037	0,036
ADF z lag = 3	-20,871	-1,95	0,051	0,146	0,001	0,002	0,004
ADF z lag = 4	-17,773	-1,95	0,008	0,005	0,008	0,014	0,007
ADF z lag = 5	-17,473	-1,95	0,053	0,127	0,247	0,013	0,025
ADF z lag = 6	-16,949	-1,95	0,508	0,802	0,743	0,407	0,459
KPSS	0,073	0,463					
Portfel ETF							
Dickeya-Fullera	-44,885	-1,95	0,897	0,071	0,003	0,007	0,015
ADF z lag = 1	-31,653	-1,95	0,571	0,391	0,051	0,1	0,169
KPSS	0,027	0,463					

Źródło: opracowanie własne.

Na podstawie analizy wykresów ACF i PACF oraz przeprowadzonej następnie procedury „od ogółu do szczegółu” udało się wybrać parę najlepszych modeli ARMA dla każdego z omawianych szeregów. W celu wyłonienia najlepiej dopasowanego z nich posłużono się dwoma kryteriami informacyjnymi: kryterium informacyjnym Akaikego (*Akaike information criterion*, AIC) oraz bayesowskim kryterium informacyjnym Schwarza (*bayesian information criterion*, BIC). Dla portfela walutowego model ARMA(3,2) uzyskał najniższe wartości dla obu kryteriów, potwierdzając wnioski uzyskane z przeprowadzonej wcześniej procedury „od ogółu do szczegółu”. Podobnie sytuacja wygląda z modelem ARMA(2,1) dla portfela ETF, który również uzyskał najniższe wartości obu kryteriów. Po zastosowaniu wszystkich trzech metod identyfikacji procesu ARIMA ustalono więc, że dla portfela walutowego najlepiej dopasowanym modelem jest ARMA(3,2), a dla portfela funduszy inwestycyjnych ARMA(2,1).

W tabeli 2 przedstawiono wyniki oszacowań modeli ARMA(2,3) dla portfela walutowego i ARMA(2,1) dla portfela funduszy inwestycyjnych, przeprowadzono również odpowiednie testy diagnostyczne dla tych modeli. W teście Ljunga-Boxa dla obu modeli uzyskano wartość $p\text{-value} > 0,05$, co oznacza, że reszty z tych modeli są białym szumem. Potwierdzają to również wykresy ACF i PACF dla każdego z tych modeli, na których nie uwidoczniło żadnych istotnych opóźnień. Oba modele są poprawne i mogą zostać użyte do konstrukcji modeli GARCH.

Tabela 2. Parametry modelu ARMA(2,3) i ARMA(2,1) kolejno dla portfela walutowego i ETF oraz wyniki testu Ljunga-Boxa i odpowiadające mu $p\text{-value}$

Parametr	ARMA(2,3)		ARMA(2,1)	
	wartość parametru	$p\text{-value}$	wartość parametru	$p\text{-value}$
AR L1	0,76	0,00	-0,93	0,00
AR L2	-0,83	0,00	-0,08	0,00
AR L3	0,09	0,00		
MA L1	-0,72	0,00	0,88	0,00
MA	0,76	0,00		0,00
Test	statystyka testowa	$p\text{-value}$	statystyka testowa	$p\text{-value}$
Ljung-Box	42,85	0,35	54,59	0,06

Źródło: opracowanie własne.

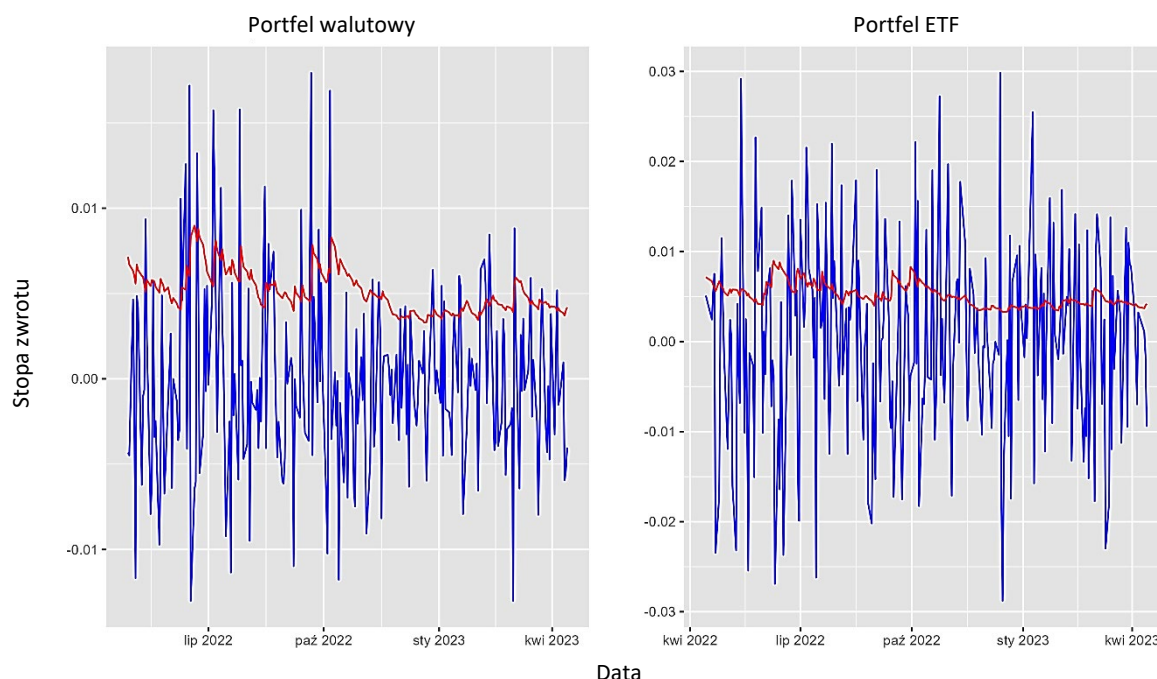
Przeprowadzono estymację modelu GARCH(1,1) z rozkładem normalnym dla każdego z szeregów, każdy z modeli oparty jest na odpowiednim modelu ARIMA oraz zawiera zmienne wykrywające obserwacje odstające. Wybrano GARCH(p,q) z oboma parametrami równymi 1, gdyż taki jest najczęściej używany w literaturze dotyczącej modelowania szeregów czasowych za pomocą rozpatrywanego modelu (Wong i in., 2016). W tabeli 3 przedstawione zostały wyniki estymacji parametrów modelu wraz z odpowiadającymi im statystykami p -value dla portfela walutowego oraz portfela ETF.

Tabela 3. Parametry modelu GARCH(1,1) z rozkładem normalnym, wyniki testów Ljunga-Boxa i Jarque'a-Bery oraz odpowiadające im p -value

Portfel walutowy						
Parametr	ω	α	β	test	Ljunga-Boxa	Jarque'a-Bery
Wyniki estymacji	0,000	0,127	0,804	statystyka testowa	33,88	50,56
p -value	0,000	0,000	0,000	p -value	0,741	0,000
Portfel ETF						
Parametr	ω	α	β	test	Ljunga-Boxa	Jarque'a-Bery
Wyniki estymacji	0,000	0,071	0,904	statystyka testowa	35,14	66,24
p -value	0,000	0,008	0,011	p -value	0,689	0,000

Źródło: opracowanie własne.

Wartości p -value dla wszystkich parametrów modelu GARCH(1,1) są większe od 0,05, co oznacza, że wszystkie zmienne w modelu są istotne statystycznie. Ponadto sumy parametrów α i β dla obu modeli są mniejsze niż 1, zatem ich wariancja warunkowa jest stacjonarna. W modelach GARCH dla portfela walutowego otrzymano p -value > 0,05 w teście Ljunga-Boxa, co wskazuje na brak podstaw do odrzucenia hipotezy o braku autokorelacji reszt, a to oznacza, że proces jest białym szumem. W teście Jarque'a-Bery otrzymano natomiast p -value < 0,05 i odpowiadającą mu statystykę testową równą 50,56, co wskazuje na konieczność odrzucenia hipotezy o normalności rozkładu reszt w modelu. Zastosowanie modelu GARCH(1,1) z uwzględnieniem zmiennych opisujących obserwacje odstające sprawiło, że statystyka testowa w teście Jarque'a-Bery zmalała względem początkowych danych prawie czterokrotnie, nie udało się jednak uzyskać rozkładu normalnego reszt w modelu. Podjęto próbę zastosowania wszystkich możliwych kroków, by uzyskać rozkład normalny reszt.



Rys. 6. Oszacowania wariancji warunkowej za pomocą modelu GARCH(1,1) z rozkładem normalnym odpowiednio dla portfela walutowego i portfela ETF

Źródło: opracowanie własne.

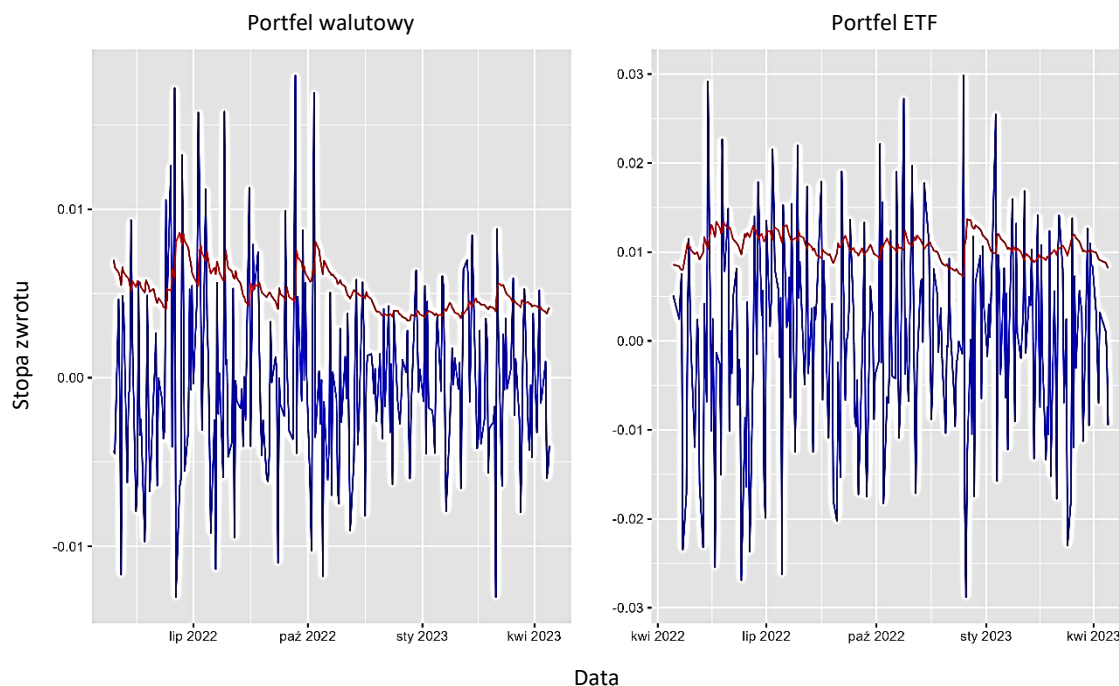
Dla modelu GARCH wyestymowanego dla portfela ETF sytuacja wygląda analogicznie, również nie udało się uzyskać rozkładu normalnego mimo trzykrotnego zmniejszenia statystyki testowej w teście Jarque'a-Bery względem początkowych danych. Na rysunku 6 przedstawiono oszacowania wartości wariancji warunkowej za pomocą modelu GARCH(1,1) z rozkładem normalnym (czerwona linia). Oszacowania wariancji wydają się odpowiednio dopasowane szczególnie w przypadku zwrotów z portfela walutowego, model dobrze radzi sobie z wykrywaniem skupień wariancji zwrotów. Dla portfela ETF prognozy wariancji wydają się mniej dokładne, szczególnie w przypadku obserwacji odstających.

Zdecydowano się na estymację modelu GARCH(1,1) z rozkładem skośnym *t*-Studenta w celu porównania uzyskanych w nim oszacowań z tymi z modelu GARCH z rozkładem normalnym. Oszacowania jego parametrów oraz wyniki wybranych testów statystycznych przedstawiono w tabeli 4. Ponownie dla obu modeli wszystkie parametry modelu są istotne ($p\text{-value} < 0,05$). Ponadto oszacowania parametrów są zbliżone do tych z modelu GARCH z rozkładem normalnym. W obu przypadkach suma parametrów α i β również jest mniejsza niż jeden, co wskazuje na stacjonarność wariancji warunkowej każdego

Tabela 4. Parametry modelu GARCH(1,1) z rozkładem skośnym *t*-Studenta, wyniki testów Ljunga-Boxa i Jarque'a-Bery oraz odpowiadające im $p\text{-value}$

Portfel walutowy							
Parametr	ω	α	β	kształt	test	Ljunga-Boxa	Jarque'a-Bery
Wyniki estymacji	0,000	0,123	0,816	6,903	statystyka testowa	28,337	118,03
$p\text{-value}$	0,000	0,000	0,000	0,000	$p\text{-value}$	0,916	0,000
Portfel ETF							
Parametr	ω	α	β	kształt	test	Ljunga-Boxa	Jarque'a-Bery
Wyniki estymacji	0,0000	0,077	0,892	6,344	statystyka testowa	4,344	63,6
$p\text{-value}$	0,002	0,000	0,000	0,000	$p\text{-value}$	0,0371	0,000

Źródło: opracowanie własne.



Rys. 7. Oszacowania wariancji warunkowej za pomocą modelu GARCH(1,1) z rozkładem skośnym *t*-Studenta odpowiednio dla portfela walutowego i portfela ETF

Źródło: opracowanie własne.

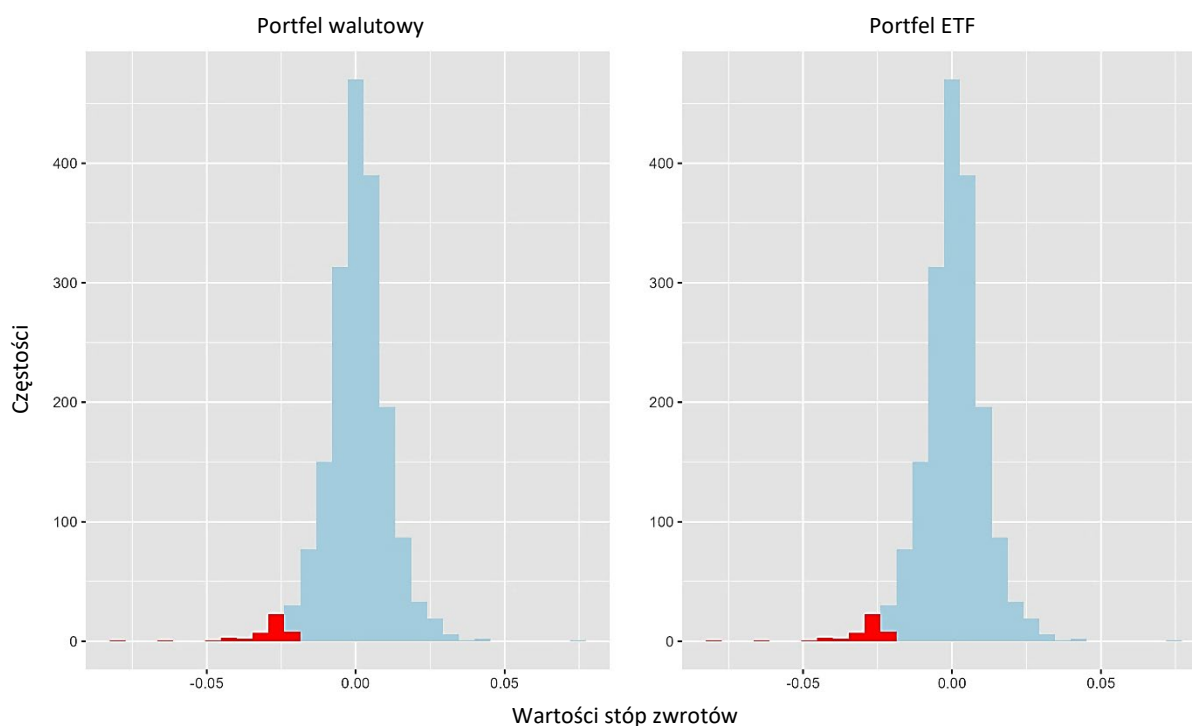
z modeli. Wyniki testu Ljunga-Boxa dla obu portfeli nie różnią się od tych dla modelu GARCH z rozkładem normalnym. Podobnie sytuacja wygląda w teście Jarque'a-Bery, w którym otrzymano statystykę testową znacznie przewyższającą wartość krytyczną na poziomie 5%, co wskazuje na konieczność odrzucenia hipotezy zerowej o normalności rozkładu reszt z modeli. W modelu dla portfela walutowego otrzymano jednak statystykę testową w teście Jarque'a-Bery ponad dwukrotnie większą niż w modelu GARCH z rozkładem normalnym. Na rysunku 7 przedstawiono oszacowania wariancji warunkowej dla każdego z omawianych modeli GARCH(1,1) z rozkładem skośnym t -Studenta. Stwierdzono, iż oszacowana postać poprawnie modeluje zarówno zjawisko skupienia wariancji, jak i występowanie obserwacji odstających.

W celu uzyskania dokładnej informacji dotyczącej tego, który model generuje najlepsze prognozy VaR, należy posłużyć się odpowiednimi testami, które przeprowadzone zostaną w kolejnym podpunkcie.

4.3.1. Szacowanie *Value-at-Risk* oraz porównanie wyników

Przeprowadzona zostanie analiza porównawcza oraz ocena wyników oszacowań *Value-at-Risk* uzyskanych za pomocą wcześniej wymienionych metod dla portfela walutowego oraz portfela ETF. Na podstawie wyników zostanie wyłoniona najbardziej dokładna metoda wyznaczania VaR, a także wyciągnięte zostaną odpowiednie wnioski.

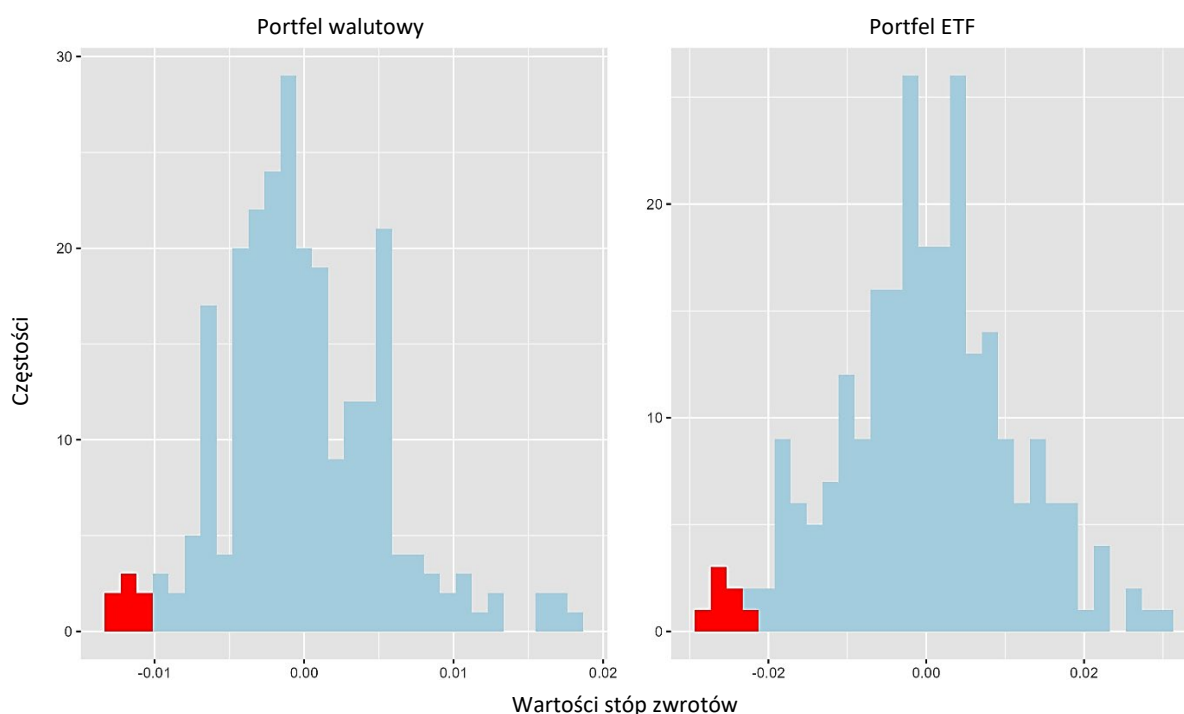
Prognozowany okres obejmuje 250 obserwacji zarówno dla portfela walutowego, jak i funduszy inwestycyjnych, co stanowi mniej niż 15% całego badanego okresu. Daje to zatem wystarczającą wielkość próby *in-sample*, by prognozy VaR wygenerowane na podstawie były dokładne. Wartość zagrożona zostanie oszacowana na poziomie ufności 97,5%, co oznacza, że z takim prawdopodobieństwem straty w portfelu nie przekroczą wartości VaR w prognozowanym okresie. Na rysunku 8 przedstawione zostały histogramy zwrotów z poszczególnych portfeli w całym badanym okresie, zaznaczono na nich czerwonym kolorem zwroty poniżej kwantyla równego 2,5%. Można zaobserwować na nich występowanie obserwacji odstających.



Rys. 8. Zwroty z portfela walutowego i portfela ETF poniżej kwantyla 2,5%

Źródło: opracowanie własne.

Na rysunku 9 przedstawiono analogiczne histogramy dla prognozowanego okresu *out-of-sample*. Nie widać na nich obserwacji odstających, ponadto w przedziale czasowym zwroty z portfela funduszy inwestycyjnych mają rozkład normalny (co potwierdzone zostało testem Jarque'a-Bery, w którym otrzymano statystykę *p-value* równą 0,94) w przeciwieństwie do zwrotów z całego badanego okresu.



Rys. 9. Zwroty z portfela walutowego i portfela ETF poniżej kwantyla 2,5% w okresie *out-of-sample*

Źródło: opracowanie własne.

Do oceny poszczególnych metod wyznaczania *Value-at-Risk* zastosowane zostaną testy statystyczne: test bezwarunkowego pokrycia (test Kupca) oraz test warunkowego pokrycia (test Christoffersena). Dodatkowo wykorzystana zostanie funkcja wskaźnika, która sumuje liczbę przypadków przekroczenia prognozowanej wartości VaR przez rzeczywiste zwroty w danym portfelu. Funkcja ta przyjmuje postać:

$$I_{t+1} = \begin{cases} 1, & \text{jeżeli } r_{t+1} \leq VaR_{t+1} \\ 0, & \text{jeżeli } r_{t+1} > VaR_{t+1} \end{cases} \quad (9)$$

Pożądana liczba przekroczeń VaR na poziomie istotności 2,5% dla prognozowanego okresu (250 obserwacji) dla danego portfela wynosi sześć przekroczeń. W tabeli 5 przedstawiono wyniki testów Kupca i Christoffersena oraz oczekiwaną i rzeczywistą liczbę przekroczeń VaR dla danej metody szacowania *Value-at-Risk* dla portfela walutowego oraz funduszy inwestycyjnych.

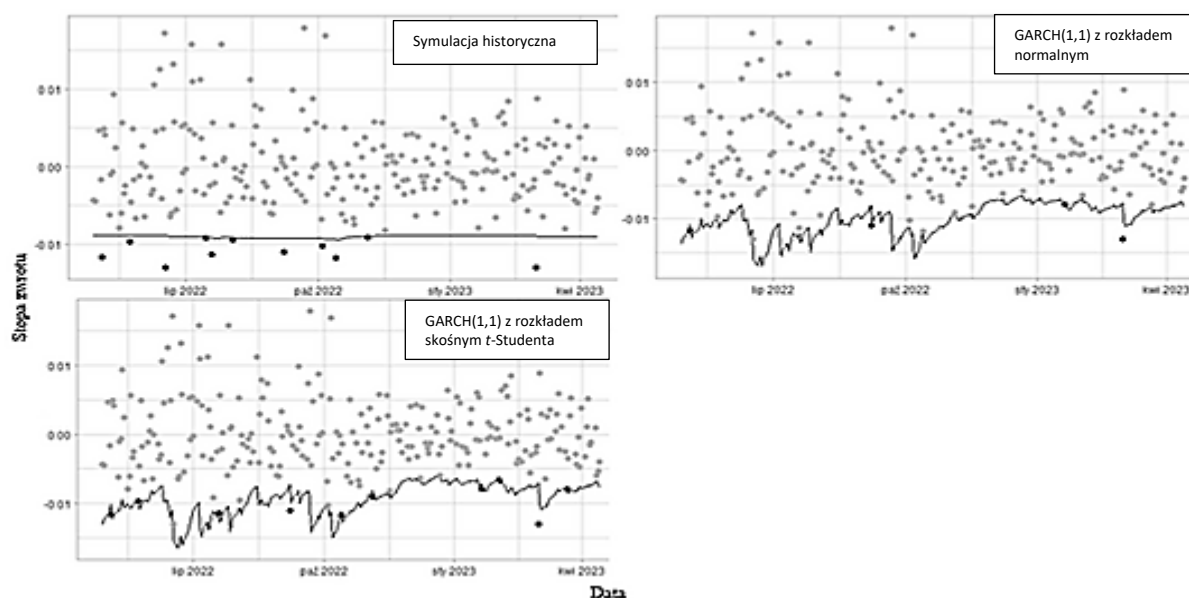
Tabela 5. Ocena metod poszczególnych szacowania VaR

Metoda	Oczekiwane przekroczenia	Uzyskane przekroczenia	Test Kupca	Test Christoffersena
Portfel walutowy				
Symulacja historyczna	6	11	brak podstaw do odrzucenia H_0	brak podstaw do odrzucenia H_0
GARCH(1,1) z rozkładem normalnym	6	4	brak podstaw do odrzucenia H_0	brak podstaw do odrzucenia H_0
GARCH(1,1) z rozkładem skośnym <i>t</i> -Studenta	6	10	brak podstaw do odrzucenia H_0	brak podstaw do odrzucenia H_0

Portfela ETF				
Symulacja historyczna	6	8	brak podstaw do odrzucenia H_0	brak podstaw do odrzucenia H_0
GARCH(1,1) z rozkładem normalnym	6	10	brak podstaw do odrzucenia H_0	brak podstaw do odrzucenia H_0
GARCH(1,1) z rozkładem skośnym t -Studenta	6	6	brak podstaw do odrzucenia H_0	brak podstaw do odrzucenia H_0

Źródło: opracowanie własne.

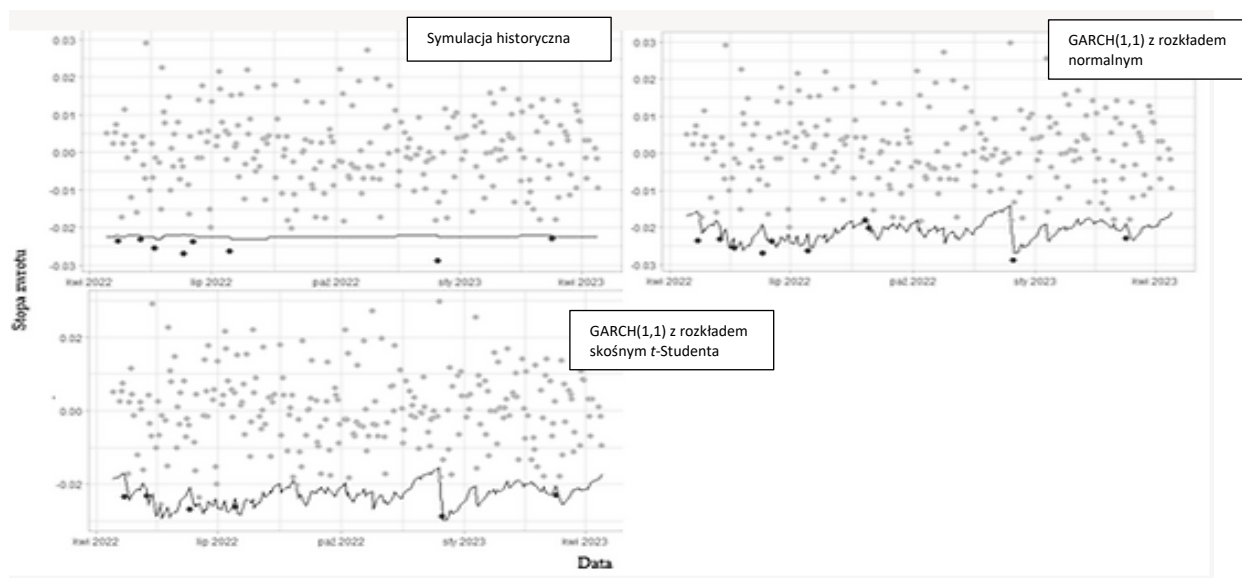
Wyniki testów Kupca i Christoffersena wskazują na brak możliwości odrzucenia hipotezy o dopuszczalnej liczbie przekroczeń VaR oraz hipotezy o niezależności przekroczeń VaR i ich dopuszczalnej liczbie. Oznacza to, że każda z wybranych w tej pracy metod szacowania *Value-at-Risk* spełnia oba testy i można jej skutecznie używać w celu wyznaczania wartości zagrożonej dla portfela walutowego. Ponadto na rysunku 10 przedstawione zostały przekroczenia VaR uzyskane za pomocą każdej z wymienionych metod dla portfela walutowego. Najdokładniejsze prognozy VaR otrzymano za pomocą modelu GARCH(1,1) z rozkładem normalnym, za jego pomocą osiągnięto jedynie cztery przekroczenia, co oznacza, że model delikatnie przeszacowuje ryzyko. Pozostałe modele mają natomiast problem z niewielkim niedoszacowaniem ryzyka (liczba przekroczeń VaR > 6). Najmniej dokładna zgodnie z przewidywaniami okazała się metoda symulacji historycznej, która miała problem z doszacowaniem ryzyka, w szczególności w pierwszej części prognozowanego okresu.



Rys. 10. Przekroczenia VaR uzyskane za pomocą kolejnych metod: symulacji historycznej, modeli GARCH(1,1) z rozkładem normalnym i skośnym t -Studenta dla portfela walutowego

Źródło: opracowanie własne.

W przypadku portfela ETF dla każdej metody wyznaczania *Value-at-Risk* nie znaleziono podstaw do odrzucenia hipotez zerowych w testach Kupca i Christoffersena. Wszystkie omawiane modele dobrze sprawdzają się w szacowaniu wartości zagrożonej. Na rysunku 11 przedstawione zostały przekroczenia VaR osiągnięte za pomocą omawianych w tej pracy metod szacowania *Value-at-Risk*. Pod względem kryterium liczby przekroczeń najlepiej wykazuje się model GARCH(1,1) z rozkładem skośnym t -Studenta, ponieważ dzięki jego zastosowaniu otrzymano równo sześć takich przekroczeń. Pozostałe modele również są dokładne w prognozowaniu *Value-at-Risk*, jednak zdarza im się nie doszacować ryzyka. W przypadku portfela funduszy inwestycyjnych najgorzej wypadł model GARCH(1,1) z rozkładem normalnym, który miał szczególnie problem z doszacowaniem ryzyka w pierwszym półroczu okresu *out-of-sample*.



Rys. 11. Przekroczenia VaR uzyskane za pomocą kolejnych metod: symulacji historycznej, modeli GARCH(1,1) z rozkładem normalnym i skośnym t -Studenta dla portfela ETF

Źródło: opracowanie własne.

Podsumowując, należy stwierdzić, że wszystkie wybrane w tej pracy metody szacowania *Value-at-Risk* okazały się skuteczne dla obu portfeli. W przypadku portfela walutowego najlepszą metodą jest model GARCH(1,1) z rozkładem normalnym, a dla portfela ETF model GARCH(1,1) z rozkładem skośnym t -Studenta. Ponadto zauważono, że oszacowania *Value-at-Risk* są z reguły dokładniejsze dla portfela funduszy inwestycyjnych niż dla portfela walutowego, szczególnie w przypadku modeli GARCH(1,1).

5. Dyskusja i wnioski

Wybór metod wyznaczania *Value-at-Risk* (VaR) ma istotne znaczenie praktyczne, zwłaszcza w okresach wystąpienia wydarzeń wpływających na rynki finansowe, takich jak brexit, pandemia COVID-19 czy atak Rosji na Ukrainę. W niniejszym badaniu przeanalizowano, jak różne metody szacowania VaR reagują na okresy zwiększonego ryzyka rynkowego oraz jak skutecznie radzą sobie z tymi problemami w kontekście portfeli walutowego i ETF.

W artykule przetestowano hipotezę pierwszą, która dotyczyła wyższości modelu GARCH(1,1) z rozkładem skośnym t -Studenta nad modelem GARCH z rozkładem normalnym w kontekście dokładności modelowania szeregów czasowych oraz szacowania VaR, zwłaszcza w okresach związanych z pandemią COVID-19 oraz atakiem Rosji na Ukrainę. Wyniki oszacowania *Value-at-Risk* dla obu portfeli w prognozowanym okresie nie pozwalają na jednoznaczny wybór najlepszej metody szacowania VaR. Testy diagnostyczne, takie jak test Kupca i test Christoffersena nie pozwoliły na odrzucenie żadnej z metod jako nieodpowiedniej w prognozowaniu VaR.

Dzięki zastosowaniu kryterium liczby przekroczeń VaR, określono, która z metod najlepiej prognozuje wartość zagrożoną dla danego portfela. Okazało się, że dla portfela walutowego najlepszym modelem jest GARCH(1,1) z rozkładem normalnym, podczas gdy dla portfela ETF model GARCH(1,1) ze skośnym rozkładem t -Studenta. Model GARCH(1,1) z rozkładem normalnym okazał się najbardziej trafny w prognozowaniu VaR dla portfela walutowego, a rozkład t -Studenta, uwzględniający leptokurtyczność rozkładu zwrotów, najlepiej pasował do danych z portfela ETF.

W artykule wykazano również, że konstrukcja portfela ma wpływ na jego podatność na ryzyko. Została przetestowana hipoteza, że portfele składające się z większej liczby instrumentów są bardziej odporne na ryzyko rynkowe od tych, które zawierają mniejszą liczbę aktywów. Wyniki badania wskazały na

dodatnią korelację między liczbą instrumentów w portfelu a jego podatnością na ryzyko, co prowadzi do odrzucenia tej hipotezy. Przeciwnie zależności, które były spodziewane, wynikały z dodatniej korelacji pomiędzy instrumentami, co nie prowadziło do zmniejszenia ryzyka portfela. W związku z tym warto poddać dalszym badaniom portfele składające się z bardziej zróżnicowanych instrumentów, które nie są ze sobą dodatnio skorelowane.

W niniejszym artykule skoncentrowano się na portfelach inwestycyjnych, w których składniki są wzajemnie skorelowane. Kolejnym krokiem mogłoby być zastosowanie wielowymiarowych modeli GARCH, które mogłyby dostarczyć dokładniejszych oszacowań VaR niż modele jednowymiarowe (Santos i in., 2013). Badanie mogłoby także zostać rozszerzone o uwzględnienie bardziej zdywersyfikowanych portfeli inwestycyjnych, aby sprawdzić, czy są one odporniejsze na ryzyko niż portfele składające się z instrumentów jednego typu.

Aktywa oraz portfele inwestycyjne są często narażone na straty spowodowane wzrostem ryzyka rynkowego, szczególnie w okresach dużych zmian na rynkach finansowych. Prognozowanie VaR jest jednym z kluczowych narzędzi w analizie ryzyka finansowego, szczególnie w kontekście portfeli inwestycyjnych. Inwestorzy powinni łączyć VaR z innymi narzędziami analitycznymi, regularnie testować i aktualizować swoje modele, aby dostosować je do zmieniających się warunków rynkowych.

Bibliografia

- Abounoori, E., Elmi, Z. i Nademi, Y. (2016). Forecasting Tehran Stock Exchange Volatility; Markov Switching GARCH Approach. *Physica A*, 445, 264-282.
- Alexander, C., Lazar, E. i Stancu, S. (2013). Forecasting VaR Using Analytic Higher Moments for GARCH Processes. *International Review of Financial Analysis*, 30, 36-45.
- Alexander, C. i Lazar, E. (2006). Normal Mixture GARCH(1,1): Applications to Exchange Rate Modelling. *Journal of Applied Econometrics*, 21(3), 307-336.
- Angelidis, T., Benos, A. i Degiannakis, S. (2004). The Use of GARCH Models in VaR Estimation. *Statistical Methodology*, 1(1-2), 105-128.
- Berkowitz, J. i O'Brien, J. (2002). How Accurate Are Value-at-Risk Models at Commercial Banks? *The Journal of Finance*, 57(3), 1093-1111.
- Brooks, C. i Persaud, G. (2003). Volatility Forecasting for Risk Management. *Journal of Forecasting*, 22(1), 1-22.
- Chen, C. W. S. i Watanabe, T. (2018). Bayesian Modeling and Forecasting of Value-at-Risk via Threshold Realized Volatility. *Applied Stochastic Models in Business and Industry*, 35, 747-765.
- Duda, M. i Schmidt, H. (2009). *Evaluation of Various Approaches to Value at Risk*. Lund University.
- Elenjal, T., Mwangi, P., Panulo, B. i Huang, C. (2016). A Comparative Cross-Regime Analysis on the Performance of GARCH-based Value-at-Risk Models: Evidence from the Johannesburg Stock Exchange. *Risk Management*, 18(2/3), 89-110.
- Fan, J. i Gu, J. (2003). Semiparametric Estimation of Value at Risk. *Econometrics Journal*, 6(2), 261-290.
- Fiszeder, P., Fałdziński, M. i Molnár, P. (2023). Modeling and Forecasting Dynamic Conditional Correlations with Opening, High, Low, and Closing Prices. *Journal of Empirical Finance*, 70, 308-321.
- Guermat, C. i Harris, R. D. F. (2002). Forecasting Value at Risk Allowing for Time Variation in the Variance and Kurtosis of Portfolio Returns. *International Journal of Forecasting*, 18(3), 409-419.
- Jorion, P. (2007). *Value-at-Risk – The New Benchmark for Managing Financial Risk* (3. wyd., s. 274-285). McGraw-Hill.
- Kakade, K., Jain, I. i Mishra, A. K. (2022). Value-at-Risk Forecasting: A Hybrid Ensemble Learning GARCH-LSTM Based Approach. *Resources Policy*, 78, 102903.
- Karandikar, R. G., Deshpande, N. D., Khaparde, S. A. i Kulkarni, S. V. (2009). Modelling Volatility Clustering in Electricity Price Return Series for Forecasting Value at Risk. *European Transactions on Electrical Power*, 19, 15-38.
- Komisja Nadzoru Finansowego (2024). Metodyka badania i oceny nadzorczej BION powszechnych towarzystw emerytalnych. https://www.knf.gov.pl/knf/pl/komponenty/img/METODYKA_BION_PTE_wydanie%203_2024_89141.pdf
- Lazar, E. i Xue, X. (2020). Forecasting Risk Measures Using Intraday Data in a Generalized Autoregressive Score Framework. *International Journal of Forecasting*, 36(3), 1057-1072.
- McAleer, M. i Da Veiga, B. (2008). Single-Index and Portfolio Models for Forecasting Value-at-Risk Thresholds. *Journal of Forecasting*, 27(3), 217-235.
- McMillan, D. G. i Speight, A. E. H. (2007). Value-at-Risk in Emerging Equity Markets: Comparative Evidence for Symmetric, Asymmetric, and Long-Memory GARCH Models. *International Review of Finance*, 7(1-2), 1-19.

- Sadeghi, M. i Shavvalpour, S. (2006). Energy Risk Management and Value at Risk Modeling. *Energy Policy*, 34(18), 3367-3373.
- Santos, A. A. P., Nogales, F. J. i Ruiz, E. (2013). Comparing Univariate and Multivariate Models to Forecast Portfolio Value-at-Risk. *Journal of Financial Econometrics*, 11(2), 400-441.
- Sharma, M. (2012). Evaluation of Basel III Revision of Quantitative Standards for Implementation of Internal Models for Market Risk. *IIMB Management Review*, 24(4), 234-244.
- Wong, C. i So, M. K. P. (2003). On Conditional Moments of GARCH Models, with Applications to Multiple Period Value at Risk Estimation. *Statistica Sinica*, 13, 1015-1044.
- Wong, Z. Y., Chin, V. C. i Tan, S. H. (2016). Daily Value-at-Risk Modeling and Forecast Evaluation: The Realized Volatility Approach. *The Journal of Finance and Data Science*, 2(3), 171-187.

Value-at-Risk Forecasting for Investment Portfolios

Abstract

Aim: The aim of this paper is to identify the Value-at-Risk estimation method that generates the most accurate forecasts for investment portfolios. The paper assumes that the accuracy of Value-at-Risk forecasts is measured using the Kupiec and Christoffersen tests. Additionally, the impact of the portfolio construction on its vulnerability to risk was also analyzed.

Methodology: The study was conducted using data on stock market indices obtained from Stooq.com. The historical simulation method and two GARCH(1,1) models with different theoretical distributions were used to estimate Value-at-Risk. Furthermore, the results were validated and verified, taking into account events such as Brexit, the COVID-19 pandemic, and Russia's attack on Ukraine.

Results: The most accurate Value-at-Risk forecasts were obtained using the GARCH(1,1) model with a skewed *t*-student distribution. Additionally, in almost every case, the GARCH models generated more accurate forecasts than the historical simulation.

Implications and recommendations: Assets and investment portfolios are often exposed to losses caused by an increase in market risk during periods of significant market changes. VaR forecasting is one of the key tools used in financial risk analysis, particularly in the context of investment portfolios. Investors should combine VaR with other analytical tools and regularly test and update models depending on changing market conditions.

Originality/value: This issue is crucial from the perspective of risk management by banks and other financial institutions. Currently, there is no recommended Value-at-Risk estimation method that serves as an industry standard. Therefore, an attempt was made to develop and test an approach within the ways of determining the value at risk.

Keywords: Value-at-Risk, GARCH model, investment portfolios, market risk
