



Studium Licencjackie

Kierunek: Metody ilościowe w ekonomii i systemy informacyjne

Specjalność: Metody ilościowe w finansach

Imię i nazwisko autora: Jakub Mikołajczak

Nr albumu: 121732

Efektywność modeli GARCH i Markov-switching GARCH w ocenie ryzyka portfeli funduszy inwestycyjnych

Praca licencjacka pod kierunkiem naukowym

dr hab. Dobromiła Serwy

Instytut Ekonometrii

Warszawa 2025

Spis treści

1	Przegląd literatury	5
1.1	Teorie i modele analizy zwrotów z funduszy inwestycyjnych	5
1.2	Matematyczne wprowadzenie do modelowania szeregów czasowych	6
1.3	Modele autoregresyjne i średniej ruchomej	8
1.4	Modele zmienności warunkowej	12
1.5	Markov-Switching GARCH	15
1.6	Miary ryzyka	18
2	Metodologia	23
2.1	Backtesting i klasyczne testy dla VaR	23
2.2	Backtesting dla ES	24
3	Empiryczna analiza danych	25
3.1	Wykresy cen i zwrotów z funduszy	28
3.2	Charakterystyka danych	31
3.3	Testy statystyczne	32
3.4	Estymacja i diagnostyka modeli typu GARCH dla portfela konserwatywnego . . .	37
3.5	Estymacja modelu MS-EGARCH dla portfela konserwatywnego	40
3.6	Estymacja modeli typu GARCH dla portfela agresywnego	43
3.7	Estymacja modelu MS-EGARCH dla portfela agresywnego	44
3.8	Backtesting	46

Wstęp

Po globalnym kryzysie finansowym z 2008 roku wzrosło zainteresowanie modelami przełącznikowymi (Markov-Switching), jednak ich zastosowanie koncentrowało się głównie na analizie indeksów giełdowych oraz funduszy typu hedge. W literaturze przedmiotu wciąż brakuje badań empirycznych wykorzystujących modele MS-GARCH do opisu zmienności stóp zwrotu z jednostek uczestnictwa otwartych funduszy inwestycyjnych, zwłaszcza w kontekście rynku polskiego. Analizy portfeli funduszy inwestycyjnych otwartych (FIO) z zastosowaniem tego typu modeli są niemal nieobecne, a ponadto rzadko uwzględniają one wpływ kryzysu pandemicznego. Niniejsza praca stanowi próbę wypełnienia tej luki badawczej.

Celem pracy jest porównanie efektywności klasycznych modeli GARCH i ich rozszerzeń z modelami Markov-Switching GARCH w kontekście modelowania zmienności otwartych funduszy inwestycyjnych. Ocena jakości modeli oparta jest na analizie dwóch miar ryzyka: Value at Risk (VaR) oraz Expected Shortfall (ES), z wykorzystaniem testów backtestingu przeprowadzonych na danych empirycznych dla dwóch portfeli – konserwatywnego i agresywnego. Główne pytanie badawcze dotyczy tego, czy modele przełącznikowe, dzięki zdolności do identyfikacji zmian reżimu zmienności, zapewniają wyższą trafność estymacji ryzyka niż klasyczne podejścia jednoreżimowe. **Postawiona hipoteza badawcza zakłada, że modele Markov-Switching GARCH charakteryzują się wyższą skutecznością prognostyczną niż standardowe modele GARCH w odniesieniu do miar ryzyka analizowanych dla portfeli funduszy inwestycyjnych o odpowiednio agresywnym i konserwatywnym profilu.**

Praca składa się z trzech zasadniczych części. Rozdział pierwszy zawiera przegląd literatury dotyczącej modelowania funduszy inwestycyjnych z wykorzystaniem modeli szeregów czasowych, ze szczególnym uwzględnieniem podejść przełącznikowych. Omówiono także definicje oraz właściwości miar ryzyka stosowanych w dalszej części pracy. W rozdziale drugim zaprezentowano metodologię badania, obejmującą klasyfikację funduszy, konstrukcję portfeli inwestycyjnych oraz procedury backtestingu. Trzeci rozdział ma charakter empiryczny: zawiera opis zbioru danych, wyniki estymacji modeli oraz porównanie ich skuteczności w odwzorowywaniu ryzyka. Pracę zamykają wnioski, omówienie ograniczeń oraz propozycje dalszych kierunków badań.

Rozdział I

1 Przegląd literatury

1.1 Teorie i modele analizy zwrotów z funduszy inwestycyjnych

Znaczenie doboru strategii inwestycyjnej w kontekście zarządzania funduszami inwestycyjnymi zostało szeroko omówione w literaturze z zakresu finansów oraz teorii portfela, m.in. w pracach Sharpe'a, Eltona i Grubera (1999), Fabozziego (2002) oraz Bodie, Kane'a i Marcusa (2014), co potwierdza istotne znaczenie tego zagadnienia w badaniach ekonomicznych.

Złożoność problematyki doboru strategii inwestycyjnej sprawiła, że stanowi ona przedmiot licznych opracowań teoretycznych i empirycznych. W literaturze przedmiotu dominują dwa podejścia badawcze: klasyczne modele oparte na teorii optymalizacji portfela (Markowitz, 1952; Sharpe, 1964) oraz podejścia statystyczne, w szczególności modele szeregów czasowych, takie jak ARMA-GARCH (Engle, 1982; Bollerslev, 1986), stosowane do modelowania dynamiki stóp zwrotu oraz zmienności.

Pierwsza grupa metod zakłada jednorazową optymalizację portfela inwestycyjnego. W jej ramach wyróżnia się m.in. model Markowitza (1952), który zakłada maksymalizację oczekiwanej stopy zwrotu przy jednoczesnym ograniczaniu ryzyka mierzonego wariancją stóp zwrotu portfela. Do klasycznych rozwiązań należy również teoria wyceny aktywów kapitałowych (CAPM), rozwinięta przez Sharpe'a (1964), która wprowadza pojęcie ryzyka systematycznego, mierzonego współczynnikiem beta jako miarą wrażliwości aktywa na zmiany rynkowe. Mimo że podejścia te mają fundamentalne znaczenie teoretyczne, ich zastosowanie w warunkach dynamicznie zmieniającego się rynku jest ograniczone ze względu na statyczny charakter i brak uwzględnienia zmienności parametrów w czasie.

Drugie podejście koncentruje się na modelowaniu szeregów czasowych, w tym zmienności stóp zwrotu i ich nieliniowej struktury. W przypadku stacjonarnych szeregów, takich jak logarytmiczne stopy zwrotu, stosowane są najczęściej modele ARMA. Natomiast do modelowania zmienności wykorzystuje się modele GARCH oraz ich rozszerzenia, takie jak EGARCH czy GJR-GARCH. Modele te pozwalają uchwycić zjawiska charakterystyczne dla rynków finansowych, takie jak efekt grupowania się zmienności (volatility clustering) czy asymetrie w odpowiedzi zmienności na ujemne i dodatnie szoki.

W kolejnych podrozdziałach przedstawiono podstawy teoretyczne wspomnianych podejść do modelowania stóp zwrotu z funduszy inwestycyjnych oraz ich praktyczne zastosowania w analizie danych finansowych.

1.2 Matematyczne wprowadzenie do modelowania szeregów czasowych

W niniejszym podrozdziale przedstawione zostaną podstawy matematyczne niezbędne do zrozumienia głównego przesłania pracy oraz najnowszych wyników badań dotyczących modelowania zwrotów z funduszy inwestycyjnych.

Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzenią probabilistyczną. **Proces stochastyczny** to rodzina zmiennych losowych $\{X_t\}_{t \in T}$ określona na tej przestrzeni, indeksowana względem zbioru T , przy czym najczęściej $T = \mathbb{N}$ (czas dyskretny) lub $T = \mathbb{R}^+$ (czas ciągły). Dla każdego $t \in T$ zmienna losowa $X_t : \Omega \rightarrow \mathbb{R}$ (ewentualnie inna przestrzeń stanów) jest mierzalna względem $\mathcal{F}$.

Inaczej mówiąc, proces stochastyczny opisuje ewolucję losowego zjawiska w czasie – reprezentuje zbiór możliwych trajektorii, z których każda odpowiada jednej realizacji procesu. Dla każdej chwili t , wartość procesu X_t ma charakter zmiennej losowej o określonym rozkładzie prawdopodobieństwa.

Proces stochastyczny $\{X_t\}_{t \in T}$ na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$ nazywamy **procesem Markowa pierwszego rzędu**, jeśli spełniona jest **własność Markowa**, tj. przyszłość procesu zależy wyłącznie od jego stanu obecnego, a nie od przeszłości. Formalna definicja własności Markowa i konstrukcja procesów stochastycznych zostały zaczerpnięte z opracowania Podgórskiej i in. (2002):

$$\mathbb{P}(X_{t+1} | X_t, X_{t-1}, X_{t-2}, \dots, X_0) = \mathbb{P}(X_{t+1} | X_t). \quad (1)$$

Alternatywnie, własność Markowa można wyrazić w języku dystrybuant warunkowych:

$$\mathbb{P}(X_{t+1} \leq y | X_0, X_1, X_2, \dots, X_t) = \mathbb{P}(X_{t+1} \leq y | X_t). \quad (2)$$

Łańcuchy Markowa stanowią szczególny przypadek procesów Markowa, w których zmiana stanów zachodzi w dyskretnych chwilach czasu. W niniejszej pracy analizowane będą wyłącznie jednorodne łańcuchy Markowa z czasem dyskretnym.

Łańcuch Markowa nazywamy jednorodnym w czasie (lub stacjonarnym), jeśli prawdopodobieństwo przejścia pomiędzy stanami nie zależy od momentu obserwacji t . Oznacza to, że dla dowolnych stanów $i, j \in E$ (gdzie E to przestrzeń stanów) oraz każdej chwili $t \in \mathbb{N}$ zachodzi:

$$\mathbb{P}(X_{t+1} = j | X_t = i) = p_{ij}, \quad (3)$$

gdzie $p_{i,j}$ jest wartością niezależną od czasu, zwaną prawdopodobieństwem przejścia ze stanu i do stanu j .

W ogólności zachodzi $p_{i,j} \neq p_{j,i}$, co oznacza, że łańcuch nie musi być symetryczny. Równość ta występuje jedynie w przypadku szczególnych łańcuchów, np. symetrycznych lub takich,

które spełniają warunek odwracalności przy założeniu równego rozkładu stacjonarnego. W ogólnym przypadku łańcuch odwracalny spełnia tzw. warunek równowagi szczegółowej:

$$\pi_i p_{i,j} = \pi_j p_{j,i},$$

gdzie π jest rozkładem stacjonarnym. Formalne omówienie własności odwracalności i równowagi szczegółowej można znaleźć m.in. w pracy Norrisa (1997).

Strukturę przejść w jednorodnym łańcuchu Markowa można przedstawić w postaci macierzy przejścia:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1,N} \\ p_{21} & p_{22} & \dots & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \dots & p_{N,N} \end{bmatrix}, \quad (4)$$

gdzie $p_{i,j} = \mathbb{P}(X_{t+1} = j | X_t = i)$. Macierz $\mathbf{P}$ jest macierzą stochastyczną, czyli jej wiersze sumują się do jedności:

$$\sum_{j=1}^N p_{i,j} = 1. \quad (5)$$

Prawdopodobieństwo przejścia ze stanu i do stanu j w dokładnie t krokach definiujemy jako:

$$p_{i,j}^{(t)} = \mathbb{P}(X_{m+t} = j | X_m = i). \quad (6)$$

Oznacza to, że jeśli proces Markowa znajduje się w stanie i w chwili m , to po t krokach znajdzie się w stanie j z prawdopodobieństwem $p_{i,j}^{(t)}$.

Równania Chapmana-Kołmogorowa umożliwiają wyrażenie przejścia wielokrokowego poprzez sumę przejść pośrednich przez dowolny stan k :

$$p_{i,j}^{(t+m)} = \sum_{k \in E} p_{i,k}^{(t)} p_{k,j}^{(m)}. \quad (7)$$

W ujęciu macierzowym, równania te przyjmują postać:

$$\mathbf{P}^{m+t} = \mathbf{P}^m \mathbf{P}^t, \quad (8)$$

co pozwala na rekurencyjne wyznaczanie macierzy przejścia w kolejnych krokach czasowych. W praktyce umożliwia to szybkie obliczenia związane z prognozowaniem długookresowego za-

chowania się łańcucha.

1.3 Modele autoregresyjne i średniej ruchomej

Model autoregresyjny rzędu p , oznaczany jako $AR(p)$, opisuje on wartość zmiennej czasowej jako liniową kombinację jej p wcześniejszych realizacji oraz składnika losowego.

Ogólna postać modelu $AR(p)$, wyrażona jest wzorem:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t, \quad (9)$$

gdzie:

- $a_t \sim i.i.d(0, \sigma^2)$ - składnik losowy o charakterze białego szumu,
- $\phi_1, \phi_2, \dots, \phi_p$ - parametry modelu,
- y_t - wartość procesu w chwili t .

Model ten zakłada, że obecna wartość zmiennej zależy od jej przeszłych realizacji, co określa się jako efekt inercji (Tsay, 2005).

W modelu $AR(p)$ zakłada się, że warunkowa wartość oczekiwana procesu względem informacji dostępnej do chwili $t - 1$, oznaczonej jako $\mathcal{F}_{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$ jest równa:

$$\mathbb{E}(y_t | \mathcal{F}_{t-1}) = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i}. \quad (10)$$

Z kolei warunkowa wariancja procesu, ze względu na przyjęte założenie o stałej wariancji składnika losowego, wynosi:

$$\text{Var}(y_t | \mathcal{F}_{t-1}) = \text{Var}(a_t) = \sigma^2. \quad (11)$$

Oznacza to, że chociaż model uwzględnia zależność wartości oczekiwanej od przeszłości, wariancja pozostaje stała w czasie, co ogranicza jego zdolność do modelowania efektów heteroskedastyczności.

Model $AR(p)$ jest słabo stacjonarny, jeśli wszystkie pierwiastki wielomianu charakterystycznego leżą poza jednostkowym kołem na płaszczyźnie zespolonej. W przypadku modelu $AR(1)$, warunkiem wystarczającym dla słabej stacjonarności jest:

$$|\phi_1| < 1. \quad (12)$$

Wtedy można wyznaczyć bezwarunkowe momenty procesu, które opisują jego długookresowe własności, niezależne od konkretnej realizacji przeszłości.

Dla modelu AR(1), bezwarunkowa wartość oczekiwana procesu wyraża się wzorem:

$$\mu = \frac{\phi_0}{1 - \phi_1}, \quad (13)$$

natomiast bezwarunkowa wariancja przyjmuje postać:

$$\text{Var}(y_t) = \frac{\sigma^2}{1 - \phi_1^2} = \gamma_0, \quad (14)$$

gdzie γ_t oznacza autowariancję zerowego rzędu. Wyniki te odgrywają istotną rolę w analizie długoterminowego zachowania szeregów czasowych oraz ich stabilności w czasie (Tsay, 2005).

Analiza autokorelacji stanowi jedno z kluczowych narzędzi służących ocenie struktury zależności czasowych w modelach AR(p). Autokorelacja opisuje siłę i kierunek zależności pomiędzy wartościami procesu w różnych chwilach czasu, co ma istotne znaczenie zarówno na etapie diagnostyki modelu, jak i przy jego estymacji. Szczególne znaczenie mają dwie funkcje: funkcja autokorelacji (ACF) oraz funkcja częściowej autokorelacji (PACF), których zachowanie stanowi podstawę identyfikacji rzędu modelu – zwłaszcza w klasycznym podejściu Boxa i Jenkinsa (Tsay, 2005).

- Autokorelacja (**ACF**) - Autokorelacja (ACF) mierzy korelację pomiędzy y_t a y_{t-k} . Dla modelu AR(1), autokorelacja opóźnienia k wynosi $\rho_k = \phi_1^k$, co oznacza wykładnicze wygasanie w miarę wzrostu opóźnienia.
- Częściowa autokorelacja (**PACF**) mierzy korelację między y_t a y_{t-k} po usunięciu wpływu wszystkich pośrednich wartości. Dla modelu AR(p) funkcja PACF ucina się po opóźnieniu p , tj. $\phi_{k,k} \approx 0$ dla $k > p$.

Model AR(p) koncentruje się wyłącznie na modelowaniu warunkowej wartości oczekiwanej. Zakładając stałą wariancję składnika losowego, nie uwzględnia warunkowej heteroskedastyczności, czyli zjawiska często obserwowanego w szeregach finansowych. W szczególności nie potrafi uchwycić efektu grupowania się zmienności (volatility clustering), typowego dla logarytmicznych stóp zwrotu instrumentów finansowych. W rezultacie model AR(p) może okazać się niewystarczający w analizie danych finansowych, dla których zmienność jest zmienna w czasie.

Modele średnich ruchomych (ang. Moving Average, MA) stanowią kolejne z podstawowych narzędzi w modelowaniu stacjonarnych szeregów czasowych. Model $MA(q)$ opisuje bieżącą wartość procesu jako liniową kombinację składnika losowego oraz q jego opóźnionych realizacji. Formalnie model ten można zapisać jako:

$$y_t = c_0 + a_t + \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (15)$$

gdzie:

- $a_t \sim i.i.d(0, \sigma^2)$ - biały szum,
- $\theta_1, \theta_2, \dots, \theta_q$ - parametry modelu,
- c_0 - wyraz wolny,
- L operator opóźnienia, taki że $L^k a_t = a_{t-k}$, co pozwala zapisać model w postaci zwartej:

$$y_t = c_0 + a_t(1 - \theta_1 L^1 - \theta_2 L^2 - \dots - \theta_q L^q). \quad (16)$$

Modele $MA(q)$ są zawsze słabo stacjonarne, o ile składnik losowy jest niezależny i ma stałą wariancję. Dla tego modelu bezwarunkowa wartość oczekiwana wynosi:

$$\mathbb{E}(y_t) = c_0. \quad (17)$$

Z kolei bezwarunkowa wariancja wyraża się wzorem:

$$\text{Var}(y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2. \quad (18)$$

W przeciwieństwie do modeli $AR(p)$, proces $MA(q)$ nie wymaga warunku pierwiastków wielomianu charakterystycznego.

Charakterystyczną cechą modelu $MA(q)$ jest ograniczona pamięć autokorelacyjna. Oznacza to, że funkcja autokorelacji (ACF) jest niezerowa tylko do rzędu q , a następnie ulega obcięciu:

$$\rho = \begin{cases} \neq 0, & k = 0, 1, 2, \dots, q \\ 0, & k > q \end{cases} \quad (19)$$

Tym samym funkcja ACF dostarcza narzędzia diagnostycznego do identyfikacji rzędu modelu MA, podobnie jak PACF w przypadku modeli AR. Z kolei funkcja częściowej autokorelacji (PACF) dla modelu $MA(q)$ wykazuje wygasanie wykładnicze lub oscylacyjne w miarę wzrostu rzędu opóźnienia k , tj.:

$$\phi_{k,k} \rightarrow 0, \text{ gdy } k \rightarrow \infty.$$

Model ARMA (p,q) łączy w sobie elementy modelu autoregresyjnego AR (p) oraz modelu średnich ruchomych MA (q) , co czyni go elastycznym narzędziem w opisie stacjonarnych szeregów czasowych. W ogólnej postaci model ARMA (p,q) zapisuje się jako:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (20)$$

gdzie:

- $a_t \sim i.i.d(0, \sigma^2)$ - biały szum,
- $\phi_1, \phi_2, \dots, \phi_p$ - współczynniki części AR,
- $\theta_1, \theta_2, \dots, \theta_q$ - współczynniki części MA,
- ϕ_0 - wyraz wolny,

Z wykorzystaniem operatora opóźnień L , model można zapisać w zwartej postaci operatorowej:

$$\Phi(L)y_t = \Theta(L)a_t + \phi_0, \quad (21)$$

gdzie:

- $\Phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$
- $\Theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$

Proces $\{y_t\}$ jest słabo stacjonarny, jeżeli wszystkie pierwiastki wielomianu $\Phi(L)$ leżą poza kołem jednostkowym na płaszczyźnie zespolonej (Tsay, 2005).

Warunkowa wartość oczekiwana, czyli prognoza oparta na informacji dostępnej do chwili $t-1$, wynosi:

$$\mathbb{E}(y_t | \mathcal{F}_{t-1}) = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}. \quad (22)$$

Warunkowa wariancja pozostaje stała w czasie i równa:

$$\text{Var}(y_t | \mathcal{F}_{t-1}) = \sigma^2. \quad (23)$$

Bezwarunkowa wartość oczekiwana (o ile model jest stacjonarny) wynosi:

$$\mu = \frac{\phi_0}{1 - \sum_{i=1}^p \phi_i}, \quad (24)$$

przy założeniu, że $|\sum \phi_i| < 1$, co gwarantuje istnienie średniej.

Z kolei bezwarunkowa wariancja nie posiada ogólnej postaci analitycznej, lecz może zostać wyrażona jako nieskończona suma ważona kwadratów współczynników odpowiedzi impulsowej ψ_j :

$$\text{Var}(y_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2, \quad (25)$$

gdzie ψ_j to współczynniki rozwinięcia procesu w szereg nieskończony.

W modelu ARMA(p, q) obie funkcje korelacyjne – ACF oraz PACF – wykazują wygasanie wykładnicze lub oscylacyjne, w przeciwieństwie do „obciętych” funkcji w modelach czysto AR lub MA. Oznacza to brak charakterystycznych punktów ucięcia, które mogłyby posłużyć do jednoznacznej identyfikacji rzędu modelu. W praktyce, identyfikacja modelu ARMA(p, q) wymaga zazwyczaj zastosowania kombinacji narzędzi graficznych i kryteriów informacyjnych (Box i in., 2015).

1.4 Modele zmienności warunkowej

Model ARCH (ang. Autoregressive Conditional Heteroskedasticity) został zaproponowany przez Roberta F. Engle’a w przełomowej pracy opublikowanej w *Econometrica* (Engle, 1982). Głównym założeniem modelu jest uwarunkowanie zmienności składnika losowego od przeszłych szoków, co pozwala opisać zjawiska charakterystyczne dla szeregów finansowych, takie jak skupianie zmienności, czyli występowanie naprzemiennych okresów wysokiej i niskiej zmienności. W uznaniu za wkład w rozwój teorii ekonometrii, Engle otrzymał Nagrodę Nobla w dziedzinie ekonomii w 2003 roku.

W przeciwieństwie do klasycznych modeli ARMA, które zakładają stałą wariancję składnika losowego, model ARCH umożliwia dynamiczne modelowanie zmienności. Modele takie mają strukturę złożoną: jedna część opisuje średnią procesu (wartość oczekiwaną), a druga – jego warunkową wariancję (Alexander, 2008).

Model ARCH(q) przyjmuje następującą postać:

$$y_t = \mu + \varepsilon_t, \quad (26)$$

gdzie:

- μ – wartość oczekiwana procesu,
- $\varepsilon_t = \sigma_t z_t$ – składnik losowy, przy czym, $z_t \sim iid.N(0, 1)$
- σ^2 – warunkowa wariancja składnika losowego.

Zmienność warunkowa opisana jest równaniem:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2, \quad (27)$$

gdzie:

$\alpha_0 > 0$ – zapewnia dodatniość wariancji nawet w przypadku braku szoków w przeszłości,
 $\alpha_i \geq 0$ dla $i = 1, \dots, q$ – określają wpływ przeszłych wartości błędów na bieżącą zmienność.

Model ARCH umożliwia dynamiczne dostosowanie się wariancji do występujących w czasie zmian intensywności zakłóceń. Aby proces miał skończoną wariancję bezwarunkową, wymagane jest spełnienie warunku:

$$\sum_{i=1}^q \alpha_i < 1$$

Dzięki tej konstrukcji, modele ARCH są zdolne do odwzorowania leptokurtycznych rozkładów stóp zwrotu, czyli rozkładów o grubych ogonach i występowaniu silnej koncentracji wokół wartości średniej w porównaniu do rozkładu normalnego. Co więcej, uchwytują one empirycznie obserwowaną cechę rynków finansowych, jaką jest autokorelacja w kwadratach stóp zwrotu, świadcząca o występowaniu heteroskedastyczności (Tsay, 2005).

Choć model ARCH(q) wprowadzony przez Engle'a (1982) umożliwił uwzględnienie zależnej od przeszłości zmienności warunkowej, w praktyce okazało się, że dla wielu szeregów czasowych finansowych konieczne jest stosowanie wysokich rzędów q , aby właściwie uchwycić długotrwałe skupianie się zmienności. Prowadziło to do estymacji dużej liczby parametrów i utrudniało interpretację wyników.

W odpowiedzi na te ograniczenia, Tim Bollerslev (1986) zaproponował model GARCH (ang. Generalized Autoregressive Conditional Heteroskedasticity), będący uogólnieniem modelu ARCH. Nowa konstrukcja pozwala nie tylko na uwzględnienie wpływu przeszłych wartości kwadratów błędów, ale także wprowadza mechanizm autoregresji w samej wariancji, co znacząco zwiększa elastyczność modelowania zmienności finansowej przy stosunkowo niewielkiej liczbie parametrów.

Model GARCH(p, q) składa się z dwóch równań:

Procesu średniej:

$$y_t = \mu + \varepsilon_t, \quad (28)$$

Struktury zmienności składnika losowego:

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim \text{i.i.d. } \mathcal{N}(0, 1), \quad (29)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (30)$$

gdzie:

$\alpha_0 > 0$ – wartość bazowa wariancji (gwarantuje dodatniość σ^2 niezależnie od przeszłości),

$\alpha_i \geq 0$ - współczynniki typu ARCH, odpowiadające za wpływ przeszłych szoków,

$\beta_j \geq 0$ – współczynniki typu GARCH, odpowiadające za wpływ przeszłej zmienności.

Model GARCH w naturalny sposób rozszerza klasę modeli ARCH o mechanizm pamięci w wariancji, umożliwiając odwzorowanie długotrwałych efektów skupiania się zmienności. W przypadku najczęściej stosowanego modelu GARCH(1,1), warunkiem istnienia skończonej bezwarunkowej wariancji procesu jest: $\alpha_1 + \beta_1 < 1$.

Im bliższa jedności jest suma $(\alpha_1 + \beta_1)$, tym dłużej utrzymuje się podwyższona zmienność po wystąpieniu szoku. Ta cecha modelu GARCH czyni go szczególnie przydatnym w modelowaniu rynków finansowych, gdzie efekt „pamięci” zmienności jest dobrze udokumentowany empirycznie (Alexander, 2008).

Chociaż klasyczny model GARCH(1,1) skutecznie odwzorowuje skupianie zmienności, to nie jest w stanie uchwycić asymetrycznej reakcji zmienności na szoki rynkowe. W praktyce bowiem obserwuje się, że spadki cen instrumentów finansowych (ujemne stopy zwrotu) mają zazwyczaj silniejszy wpływ na wzrost zmienności niż wzrosty o tej samej skali – zjawisko to znane jest jako efekt dźwigni (leverage effect). W odpowiedzi na te ograniczenia zaproponowano szereg rozszerzeń klasycznego modelu GARCH. W niniejszej pracy zostaną rozważone dwa z nich: GJR-GARCH oraz EGARCH, ze względu na ich ugruntowane miejsce w literaturze i zastosowaniach empirycznych.

Model GJR-GARCH, zaproponowany przez Glostena, Jagannathana i Runkle’a (1993), wprowadza do klasycznego modelu GARCH dodatkowy składnik, który uwzględnia reakcję zmienności na znaki przeszłych błędów. Konstrukcja modelu umożliwia różnicowanie wpływu pozytywnych i negatywnych szoków, co pozwala uchwycić efekt dźwigni.

Formalna postać modelu GJR-GARCH(1,1) jest następująca:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 I_{\{\varepsilon_{t-1} < 0\}} + \beta_1 \sigma_{t-1}^2. \quad (31)$$

gdzie:

$\alpha_0 \geq 0, \alpha_1 \geq 0, \beta_1 \geq 0$ – standardowe parametry GARCH,

$\gamma_1 \geq 0$ – współczynnik mierzący siłę asymetrii,

$I_{\{\varepsilon_{t-1} < 0\}}$ – wskaźnik aktywujący dodatkowy efekt tylko dla ujemnych szoków.

W sytuacji, gdy $\gamma_1 > 0$, negatywne zakłócenia zwiększają zmienność bardziej niż pozytywne, co odzwierciedla empirycznie obserwowaną strukturę ryzyka na rynkach finansowych (Glosten, Jagannathan & Runkle, 1993).

Alternatywnym podejściem do modelowania asymetrii jest model EGARCH (Exponential GARCH), zaproponowany przez Nelsona (1991). Główna innowacja EGARCH polega na logarytmicznej postaci wariancji warunkowej, co eliminuje konieczność narzucania nieujemności parametrów. Dodatkowo model ten w naturalny sposób umożliwia asymetryczne reakcje zmienności na szoki, zarówno pod względem kierunku, jak i siły.

Postać ogólna modelu EGARCH(p, q) jest następująca:

$$\log(\sigma_t^2) = \omega + \sum_{i=1}^q \alpha_i \left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right) + \gamma_i \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{j=1}^p \beta_j \log(\sigma_{t-j}^2) \quad (32)$$

Współczynnik γ_i odpowiada za efekt asymetrii, natomiast składnik logarytmiczny pozwala na odwzorowanie nieliniowej dynamiki zmienności bez konieczności jawnego ograniczania parametrów (Nelson, 1991).

1.5 Markov-Switching GARCH

Kluczowym narzędziem analitycznym wykorzystywanym w niniejszej pracy jest model Markov-Switching GARCH (MS-GARCH), który łączy w sobie dwa istotne podejścia do modelowania zmienności finansowej:

- modele klasy GARCH, umożliwiające uchwycenie pamięci w procesie oraz zjawiska grupowania się zmienności,
- modele przełącznikowe (regime-switching), oparte na ukrytych procesach Markowa, pozwalające na zmianę parametrów modelu w czasie.

Zastosowanie podejścia MS-GARCH pozwala na bardziej realistyczne odwzorowanie nieliniowości, przejść między reżimami rynku (np. między okresem spadków a hossą) oraz nagłych zmian strukturalnych charakterystycznych dla szeregów finansowych (Gray, 1996; Haas, Mittnik & Paolella, 2004a; Ardia i in., 2019). W ogólnym ujęciu zakłada się, że w każdej chwili t zmienność warunkowa zależy nie tylko od przeszłych błędów oraz własnej przeszłości, jak w modelach GARCH, ale także od ukrytego stanu s_t , który ewoluuje zgodnie z jednorodnym łańcuchem Markowa z macierzą przejścia $\mathbf{P}$. Stan ten nie jest obserwowalny bezpośrednio, lecz rekonstruowany na podstawie danych empirycznych. W niniejszej analizie zakłada się istnienie dwóch reżimów zmienności – niskiej oraz wysokiej – zgodnie z podejściem zaproponowanym przez Haas et al. (2004a), a zaimplementowanym w pakiecie MSGARCH dla języka R.

Formalnie, idąc za Ardią i in., model można zapisać:

$$y_t \mid (s_t = k, \mathcal{I}_{t-1}) \sim D(0, h_{k,t}, \xi_k), \quad (33)$$

gdzie:

$s_t \in \{1, \dots, K\}$ – nieobserwowalny stan (reżim) w chwili t , podlegający jednorodnemu łańcuchowi Markowa,

$h_{k,t}$ – warunkowa wariancja w reżimie k ,

D – rozkład o zerowej średniej i wariancji $h_{k,t}$ z parametrami kształtu ξ_k .

Każdy reżim posiada własny model GARCH-typu:

$$h_{k,t} = h(y_{t-1}, h_{k,t-1}, \theta_k), \quad (34)$$

gdzie θ_k to wektor parametrów wariancji w reżimie k .

W porównaniu z klasycznymi modelami GARCH, estymacja MS-GARCH jest znacząco bardziej złożona. Wynika to z konieczności jednoczesnego modelowania:

- nieliniowej dynamiki zmienności,
- ukrytego procesu przejść między reżimami.

Podstawową metodą estymacyjną jest metoda największej wiarygodności (MLE). Ze względu na fakt, że zmienne stanu są nieobserwowalne, stosuje się zmodyfikowaną wersję filtru Hamiltona, który pozwala oszacować warunkowe prawdopodobieństwa przebywania w poszczególnych reżimach w danym momencie czasu na podstawie historii danych (Hamilton, 1989, 1994).

Filtr Hamiltona (Hamilton, 1989) to algorytm rekurencyjny. Umożliwia on wyznaczenie warunkowych prawdopodobieństw przebywania w danym reżimie na podstawie dotychczasowych

danych. Pomimo że zmienne stanu s_t są nieobserwowalne, filtr pozwala traktować je jak losowe zmienne z rozkładem warunkowym, który może być rekurencyjnie aktualizowany. W ten sposób możliwa jest konstrukcja funkcji log-wiarygodności oraz prognozowanie reżimów w czasie rzeczywistym.

Logarytmiczna funkcja wiarygodności stosowana w estymacji metodą MLE ma postać:

$$\log L(\Psi | \mathcal{I}_T) = \sum_{t=1}^T \log f(y_t | \Psi, \mathcal{I}_{t-1}), \quad (35)$$

gdzie:

- Ψ – to wektor wszystkich parametrów modelu (parametry skedastyczne, rozkładowe i przejść między stanami),
- $f(y_t | \Psi, \mathcal{I}_{t-1})$ – to warunkowa gęstość obserwacji y_t wyliczana na podstawie pełnej informacji do chwili $t - 1$.

Z uwagi na ukryty charakter reżimów, warunkowa gęstość $f(y_t | \cdot)$ jest sumą ważoną gęstości stanowych:

$$f(y_t | \Psi, \mathcal{I}_{t-1}) = \sum_{i=1}^K \sum_{j=1}^K p_{ij} \cdot z_{i,t-1} \cdot f_D(y_t | s_t = j, \Psi, \mathcal{I}_{t-1}) \quad (36)$$

- p_{ij} – prawdopodobieństwo przejścia z reżimu i do j ,
- $z_{i,t-1}$ – prawdopodobieństwo przebywania w stanie i w czasie $t - 1$,
- $f_D(\cdot)$ – gęstość obserwacji warunkowa względem reżimu j , przy zadanych parametrach.

Dla prawidłowego rozpoczęcia algorytmu filtracji konieczne jest przyjęcie rozkładu stacjonarnego π jako początkowego wektora prawdopodobieństw stanów:

$$\pi P = \pi, \quad \text{oraz} \quad \sum_{k=1}^K \pi_k = 1 \quad (37)$$

Rozkład ten pełni rolę prawdopodobieństw początkowych w filtrze Hamiltona.

Po zbudowaniu funkcji logarytmu wiarygodności następuje jej numeryczna maksymalizacja względem wektora parametrów Ψ , obejmującego zarówno parametry modelu warunkowej wariancji, jak i macierzy przejścia między reżimami.

Ze względu na brak analitycznej postaci estymatora, w implementacji wykorzystuje się metody numerycznej optymalizacji. W pakiecie MSGARCH dla języka R, wybrany przez autora parametr `solver = "hybrid"` wskazuje – zgodnie z dokumentacją – na zastosowanie dwustopniowego podejścia, łączącego metody Neldera–Meada oraz BFGS.

- Etap 1: algorytm Neldera–Meada (NM) – metoda ta nie wymaga pochodnych funkcji celu i opiera się na manipulacji prostokąta (sympleksu) w przestrzeni parametrów. Jest odporna na nierówności i lokalne nieregularności funkcji wiarygodności, przez co stanowi dobrą strategię do znajdowania punktu początkowego.
- Etap 2: quasi-Newtonowska metoda BFGS (Broyden–Fletcher–Goldfarb–Shanno) – po uzyskaniu wstępnych estymat z etapu 1, dalsza maksymalizacja przeprowadzana jest z użyciem metody BFGS, która przybliża macierz Hessego i wykorzystuje gradienty do szybkiej konwergencji do lokalnego maksimum.

Możliwe jest również zastosowanie alternatywnych metod optymalizacji poprzez zmianę parametru `solver` w funkcji estymacyjnej pakietu MSGARCH. Przykładowo, ustawienie `solver = "solnp"` umożliwia wykorzystanie algorytmu sekwencyjnego programowania kwadratowego (Sequential Quadratic Programming), zaimplementowanego w pakiecie Rsolnp.

1.6 Miary ryzyka

W matematyce finansowej miara ryzyka to odwzorowanie (funkcjonał), który każdej zmiennej losowej (np. zyskom lub stratom portfela) przypisuje jedną liczbę — reprezentującą ryzyko.

Formalnie rzecz biorąc:

$$\rho : \mathcal{X} \rightarrow \mathbb{R},$$

gdzie:

- $\mathcal{X}$ - zbiór zmiennych losowych
- $\rho(X)$ - miara ryzyka dla instrumentu/portfela o profilu zysków i strat opisanym przez zmienną losową X .

Miara Value at Risk (VaR) dla danego poziomu istotności $\alpha \in (0, 1)$ określa kwantyl rozkładu strat portfela:

$$\text{VaR}_\alpha(X) = -\inf \{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq \alpha\} \quad (38)$$

Miara Expected Shortfall (ES), znana również jako Conditional Value at Risk (CVaR), określa średnią wartość strat przekraczających VaR, czyli oczekiwaną stratę w najgorszych $\alpha \cdot 100\%$ przypadkach.

Expected Shortfall można zapisać matematycznie jako:

$$\text{ES}_\alpha(X) = -\frac{1}{\alpha} \int_0^\alpha \text{VaR}_u(X) du, \quad (39)$$

$$\text{ES}_\alpha(X) = -\mathbb{E}[X \mid X \leq q_\alpha(X)]. \quad (40)$$

Powyższe definicje są szeroko stosowane w praktyce zarządzania ryzykiem, a ich własności formalne opisano m.in. w pracach McNeila, Freya i Embrechtsa (2015) oraz Joriona (2007). Dodatkowo, Expected Shortfall jest uznawany za spójną miarę ryzyka, co zostało wykazane przez Artznera i in. (1999).

Wprowadzenie pojęcia spójnych miar ryzyka miało na celu nadanie rygorystycznych podstaw matematycznych intuicji towarzyszącej zarządzaniu ryzykiem. Warunki koherencji, takie jak monotoniczność, podaddytywność, jednorodność i translacyjna niezmienniczość, są formalnym zapisem oczekiwań, które powinny spełniać miary ryzyka, jeśli mają właściwie odzwierciedlać zagrożenia związane z pozycjami portfelowymi.

Intuicyjnie oznacza to, że:

- większe straty powinny wiązać się z większym ryzykiem (**monotoniczność**):

$$\rho(X) \geq \rho(Y)$$

- dywersyfikacja powinna obniżać ryzyko, a nie je zwiększać (**subaddytywność**):

$$\rho(X + Y) \leq \rho(X) + \rho(Y)$$

- zmiana skali inwestycji proporcjonalnie zmienia skalę ryzyka (**jednorodność**):

$$\rho(\lambda X) = \lambda \rho(X), \quad \lambda > 0$$

- dodanie pewnej kwoty gotówki do portfela powinno zmniejszyć ryzyko (**translacyjna niezmienniczość**):

$$\rho(X + c) = \rho(X) - c, \quad c \in \mathbb{R}.$$

Choć Value at Risk jest jedną z najbardziej rozpowszechnionych miar ryzyka, nie spełnia wszystkich warunków koherencji – w szczególności nie jest subaddytywna.

W odpowiedzi na te ograniczenia zaproponowano Expected Shortfall, który spełnia wszystkie warunki koherencji, w tym subaddytywność.

W literaturze oraz praktyce finansowej wyróżnia się trzy podstawowe metody szacowania miar ryzyka, takich jak Value at Risk oraz Expected Shortfall: **metodę historyczną**, **metodę parametryczną** oraz **metodę numeryczną** opartą na symulacjach Monte Carlo (Jorion, 2007; Dowd, 2005; Hull, 2015).

W niniejszej pracy do procedur backtestingu zastosowano metodę rolling window w połączeniu z numeryczną estymacją miar ryzyka przy użyciu symulacji Monte Carlo. Podejście to umożliwia elastyczne odwzorowanie struktury zmienności i uwzględnienie pełnego rozkładu prognozowanych stóp zwrotu w danym horyzoncie czasowym.

Ze względu na problemy numeryczne, które pojawiły się podczas szacowania miary Expected Shortfall dla modeli przełącznikowych Markowa, konieczne było zastosowanie w jednym przypadku alternatywnej procedury — w szczególności, parametrycznej metody obliczania ES przy założeniu określonego rozkładu składnika losowego. Szczegóły teoretyczne dotyczące wykorzystanych metod (Monte Carlo oraz estymacji parametrycznej) przedstawiono w kolejnych fragmentach pracy.

Schemat zastosowanego algorytmu symulacyjnego opiera się na podejściu przedstawionym przez Danielssona (2011) w monografii Financial Risk Forecasting. Algorytm ten został zaadaptowany do prognozowania miar ryzyka w oparciu o szacowane modele zmienności i jest szczegółowo opisany w kolejnych podrozdziałach.

1. W pierwszym kroku dopasowano do serii logarytmicznych stóp zwrotu r_t odpowiednie modele heteroskedastyczności warunkowej GARCH oraz MS-GARCH.
2. Modele oszacowano metodą największej wiarygodności. Dla każdego modelu uzyskano:
 - Parametry średniej (np. ARMA),
 - Parametry wariancji (np. parametry log-skali i asymetrii),
 - Parametry rozkładu reszt (np. stopnie swobody, parametr skośności),
 - dla modeli przełącznikowych uzyskano także:
 - zestawy parametrów oddzielnie dla każdego stanu
 - macierz przejścia
 - prawdopodobieństwa filtrowane i stabilne
3. Generujemy losowe ścieżki z rozkładu reszt N razy:

(a) Dla modeli typu GARCH - zakładamy, że:

$$z^{(i)} \sim sstd(\hat{\xi}, \hat{\nu}), i = 1, \dots, N.$$

gdzie $z^{(i)}$ to i-ta realizacja z rozkładu reszt (skewed t-Studenta). Parametry $\hat{\xi}, \hat{\nu}$ są wspólne dla każdej generacji.

(b) Dla modeli typu MS-GARCH -dla każdej ścieżki i przeprowadzono:

- i. Losowanie reżimu $s_{t+1}^{(i)}$ zgodnie z macierzą przejścia $\mathbf{P}$ i filtrowanym prawdopodobieństwem stanu w czasie t ,
- ii. Pobranie parametrów odpowiadających danemu stanowi:

$$z^{(i)} \sim sstd(\hat{\xi}^{(s)}, \hat{\nu}^{(s)}), i = 1, \dots, N.$$

Oznacza to, że każdy punkt ścieżki jest generowany z rozkładu specyficznego dla wylosowanego reżimu.

4. Generujemy warunkowe ścieżki stóp zwrotu:

(a) Dla modeli typu GARCH - dla każdej symulacji stosujemy wzór:

$$r^{(i)} = \hat{\mu}_t + \hat{\sigma}_t \cdot z^{(i)}, \quad i = 1, \dots, N.$$

(b) Dla modeli typu MS-GARCH - Dla każdej ścieżki i , po wylosowaniu stanu $s_{t+1}^{(i)}$:

$$r^{(i)} = \hat{\mu}_t^{(s)} + \hat{\sigma}_t^{(s)} \cdot z^{(i)}, \quad i = 1, \dots, N.$$

5. Obliczenie miar ryzyka VaR i ES - dla każdego poziomu istotności $\alpha \in \{0.01, . - 0.05\}$:

- Posortowano wartości $\{r^{(i)}\}_{i=1}^N$ rosnąco (automatycznie dzieje się to w funkcji quantile() w R),
- Obliczono:

$$\text{VaR}_\alpha = -\text{quantile}(r^{(i)}, \alpha),$$

$$\text{ES}_\alpha = -\mathbb{E} \left[r^{(i)} \mid r^{(i)} \leq -\text{VaR}_\alpha \right].$$

6. Zastosowanie rolling window:

zastosowano metodę rolling window (80% danych treningowych, 20% testowych). Dla każdej daty testowej:

- Estymowano model na ostatnim oknie danych,
 - Wygenerowano N ścieżek Monte Carlo dla dnia,
 - Obliczono VaR i ES według powyższej procedury,
 - Zapisano wartości do późniejszego backtestingu.

Rozdział ten zakończony zostanie opisem drugiej metody wykorzystanej do wyznaczenia miary Expected Shortfall, tj. podejścia parametrycznego. Zastosowanie tej metody było podyktowane problemami numerycznymi, które pojawiły się w trakcie szacowania ES przy użyciu symulacji Monte Carlo w modelu MSGARCH — prowadziły one do nienaturalnie zaniżonych wartości oczekiwanego niedoboru. Alternatywna procedura, oparta na założeniu określonego rozkładu składnika losowego, została zaimplementowana zgodnie ze schematem przedstawionym przez Danielssona (2011).

1. Estymacja MS-GARCH

Na danych dziennych logarytmicznych stóp zwrotu dopasowano model MS-EGARCH z dwoma reżimami i rozkładem reszt typu skośny t-Studenta (skewed-t). Estymację przeprowadzono metodą największej wiarygodności.

W wyniku dopasowania modelu uzyskano:

- szereg warunkowych odchyłeń standardowych,
- szereg wartości oczekiwanych otrzymanych z oddzielnie estymowanego modelu AR(1),
- parametry rozkładu reszt, takie jak skośność i liczba stopni swobody.

Dla uproszczenia przyjęto uśrednione wartości:

$$\hat{v} = \frac{\hat{v}^{(1)} + \hat{v}^{(2)}}{2} \quad \hat{\xi} = \frac{\hat{\xi}^{(1)} + \hat{\xi}^{(2)}}{2}$$

2. Wyznaczenie parametrycznego Expected Shortfall

Zakładając, że reszty mają znany rozkład $sstd(\hat{\xi}, \hat{v})$, dla każdego poziomu istotności $\alpha \in \{0.01, . . . 0.05\}$ oraz dla każdego dnia t w próbie testowej, wyznaczono wartość parametrycznego Expected Shortfall według formuły:

$$ES_t(\alpha) = \hat{\mu}_t + \hat{\sigma}_t \cdot ES_{sstd}(\alpha; \hat{\xi}, \hat{v})$$

3. Estymację przeprowadzono dla ostatnich $n = 248$ obserwacji szeregu czasowego (20% zbioru danych), tworząc wektor prognozowanych wartości $ES_t(\alpha)$, które następnie poddano analizie jakościowej.

Rozdział II

2 Metodologia

2.1 Backtesting i klasyczne testy dla VaR

Backtesting to empiryczna procedura oceny jakości prognoz modelu ryzyka, polegająca na porównaniu wyznaczonych ex-ante miar ryzyka z rzeczywistymi ex-post obserwacjami stóp zwrotu. Celem backtestingu jest sprawdzenie, czy model dostarcza wiarygodnych i zgodnych z rzeczywistością prognoz.

W kontekście niniejszej pracy, zastosowano dwa testy statystyczne służące do backtestingu: **test Kupca i Christoffersena.**

Test Kupca (1995) sprawdza, czy częstość przekroczeń prognozowanego VaR odpowiada zadanej wartości poziomu istotności α . Hipotezy testowe są następujące:

$$\begin{cases} \mathcal{H}_0 : \mathbb{P}(\text{przekroczenia}) \neq \alpha \\ \mathcal{H}_1 : \mathbb{P}(\text{przekroczenia}) = \alpha \end{cases}$$

Test Christoffersena (1998) rozszerza test Kupca, badając również, czy przekroczenia pojawiają się losowo w czasie (czyli są niezależne). Ocenia on tzw. warunkowe pokrycie (conditional coverage), łącząc dwa składniki:

- Unconditional coverage — czy liczba przekroczeń zgadza się z oczekiwaniami:

$$\begin{cases} \mathcal{H}_0 : \pi \neq \alpha \\ \mathcal{H}_1 : \pi = \alpha \end{cases}$$

- Independence — czy przekroczenia są niezależne w czasie (czy np. nie występują w klastrach):

$$\begin{cases} \mathcal{H}_0 : \text{przekroczenia niezależne} \\ \mathcal{H}_1 : \text{przekroczenia zależne} \end{cases}$$

Statystyka testu ma rozkład χ^2 .

2.2 Backtesting dla ES

Expected Shortfall, jako spójna, koherentna miara ryzyka, jest szeroko rekomendowana zarówno przez środowisko akademickie, jak i instytucje regulacyjne (np. Komitet Bazylejski w ramach Bazylei III), jako alternatywa dla Value at Risk – szczególnie ze względu na swoją zdolność do uwzględniania pełnego ryzyka ogonowego.

Jednak przez długi czas ograniczeniem dla praktycznego zastosowania ES była trudność związana z jej empiryczną weryfikacją (backtestingiem). W szczególności, fakt że ES nie jest **uzyskiwalne**, czyli z ang. **elicitable** w sensie klasycznym – tj. nie istnieje funkcja scoringowa, której wartość oczekiwana byłaby minimalizowana wyłącznie przez poprawną wartość ES – prowadził do przekonania, że nie można jej skutecznie testować.

Nieporozumienie dotyczące ograniczeń wynikających z braku uzyskiwalności miary ryzyka Expected Shortfall zostało zakwestionowane w pracy *Backtesting Expected Shortfall* autorstwa Acerbiego i Székely’ego (2014). Autorzy ci pokazali, że choć ES nie jest samodzielnie uzyskiwalny, nie wyklucza to możliwości jego efektywnego backtestowania. Głównym wnioskiem płynącym z ich pracy jest stwierdzenie, że uzyskiwalność miary ryzyka nie ma istotnego znaczenia w kontekście bezwzględnej walidacji modelu, a jej brak nie uniemożliwia konstrukcji testów oceniających jakość prognoz.

Co więcej, Acerbi i Székely przedstawili intuicyjne wyjaśnienie, dlaczego ES nie spełnia warunku uzyskiwalności jako pojedyncza miara. W odpowiedzi na to ograniczenie zaproponowali konstrukcję wspólnej funkcji scoringowej dla pary miar ryzyka: Value at Risk oraz Expected Shortfall. Na podstawie tej funkcji opracowali nieparametryczne testy backtestowe, które nie wymagają założeń dotyczących rozkładu danych i koncentrują się na penalizacji przypadków niedoszacowania ryzyka.

Formalny dowód wspólnej uzyskiwalności pary (VaR, ES) został przedstawiony dopiero dwa lata później przez Fisslera i Ziegela (2016), którzy wykazali matematycznie, że para ta jest uzyskiwalna. W przeciwieństwie do Acerbiego i Székely’ego, którzy oparli swoje podejście na przesłankach intuicyjnych, Fissler i Ziegel ugruntowali tę koncepcję w ścisłych ramach teoretycznych.

W rezultacie, w pracy Acerbiego i Székely’ego (2014) zaproponowano klasę nieparametrycznych testów backtestowych, które wykorzystują wspólną funkcję scoringową dla pary (VaR, ES), skonstruowaną w taki sposób, by penalizować niedoszacowanie ryzyka. Dla pojedynczej obserwacji r_t , funkcja scoringowa może być zdefiniowana w postaci:

$$S_t = \frac{1}{\alpha} [(V_t - r_t) \cdot \mathbb{I}(r_t < V_t) + (E_t - V_t)], \quad (41)$$

gdzie:

- V_t – prognozowana wartość Value at Risk na poziomie istotności
- E_t – prognozowana wartość Expected Shortfall na poziomie istotności
- r_t – rzeczywista obserwacja (stopa zwrotu / strata) w chwili
- $\mathbb{I}(\cdot)$ –funkcja wskaźnikowa (przyjmuje wartość 1, gdy warunek wewnątrz jest spełniony, inaczej 0).

Warto uściślić, że zastosowany w pracy zapis funkcji scoringowej jest szczególnym, uproszczonym przypadkiem ogólniejszej konstrukcji zaproponowanej przez Fisslera i Ziegela (2016), służącej do oceny uzyskiwalności pary miar ryzyka VaR i ES.

Zestaw hipotez do testu jednostronnego można zapisać tak:

$$\begin{cases} \mathcal{H}_0 : & \text{Model poprawnie estymuje ES (średni score } \bar{S} = 0) \\ \mathcal{H}_1 : & \text{Model niepoprawnie estymuje ES (średni score } \bar{S} > 0) \end{cases}$$

Należy jednak podkreślić, że test Acerbiego–Székely’ego został skonstruowany z perspektywy ochrony przed niedoszacowaniem ryzyka (tzn. sytuacjami, w których model przewiduje zbyt niskie wartości ES). W konsekwencji, przypadki przeszacowania ryzyka – w których prognozowane wartości ES są nadmiernie konserwatywne – skutkują ujemnymi wartościami funkcji scoringowej, co nie prowadzi do istotności testu. Test nie jest więc narzędziem do wykrywania nadmiernego oszacowania ryzyka, lecz wyłącznie do identyfikacji sytuacji, w których ryzyko zostało zaniżone w sposób statystycznie istotny.

W niniejszej pracy autor zaimplementował test backtestowy dla miary Expected Shortfall w języku R, opierając się na istniejącym repozytorium GitHub autorstwa Lluca Puiga Codiny, które zawiera implementację testu zaproponowanego przez Acerbiego i Székely’ego (2014).

Rozdział III

3 Empiryczna analiza danych

W celu zaklasyfikowania 365 funduszy inwestycyjnych (na podstawie dziennych danych) pobranych ze strony BossaFund do dwóch kategorii — konserwatywnej oraz agresywnej — zastosowano podejście oparte na analizie językowej opisów funduszy oraz deklarowanych polityk inwestycyjnych. W przeciwieństwie do metod klasyfikacji bazujących na wskaźnikach ilościowych, przyjęta

metoda opiera się na założeniu, że strategia inwestycyjna funduszu, zakotwiczona w jego deklaracji polityki inwestycyjnej, odzwierciedla trwałą orientację ryzyka. Innymi słowy, zakłada się, że fundusze określone jako konserwatywne faktycznie stosują konserwatywną strategię inwestycyjną, a nie agresywną, i analogicznie w przypadku funduszy agresywnych. Przyjmuje się również, że zarządzający funduszem dążą do realizacji obranej strategii w sposób spójny, co ma bezpośrednie znaczenie przy doborze modeli statystycznych służących do szacowania miar ryzyka.

Dane tekstowe pozyskano z materiałów informacyjnych publikowanych przez Towarzystwa Funduszy Inwestycyjnych (TFI), a następnie scalono w jeden jednolity korpus służący jako podstawa do dalszej analizy semantycznej. Na potrzeby klasyfikacji przeprowadzono ekspercką selekcję 40 funduszy, których opisy wykazywały wyraźne i jednoznaczne cechy stylistyczne odpowiadające profilom ryzyka. Część z nich została następnie wykluczona z powodu braków w danych ilościowych, pozostawiając próbę gotową do dalszej obróbki.

W ramach klasyfikatora stworzono dwie listy słów kluczowych: jedną odpowiadającą profilowi agresywnemu, drugą — konserwatywnemu. Uwzględniono zarówno rdzenie wyrazów (np. „akcj”, „obligacj”), jak i wyrażenia odnoszące się do poziomu ryzyka, klas aktywów czy stylu zarządzania. Dla słów zawierających rdzenie o dużym ładunku semantycznym, takie jak „agres” i „konserwaty”, przypisano wyższą wagę (2 zamiast 1). Listy słów kluczowych są dostępne na żądanie.

Następnie, dla każdego funduszu, zliczono wystąpienia słów z obu grup z uwzględnieniem wag, a wynik wykorzystano do przypisania jednej z trzech etykiet:

- agresywny — jeśli suma ważona słów agresywnych przewyższała konserwatywne,
- konserwatywny — w przypadku zależności odwrotnej,
- niepewny — gdy klasyfikacja nie była jednoznaczna.

Całą procedurę przeprowadzono w środowisku R z wykorzystaniem pakietów `dplyr`, `stringr` oraz `readr`. Opisy funduszy zapisano w pliku CSV, który następnie wczytano i przetworzono zgodnie z założeniami klasyfikacyjnymi.

Na podstawie uzyskanych wyników klasyfikacji wybrano:

- 5 funduszy o najwyższych wartościach wskaźnika `score_agres`, reprezentujących profil agresywny,
- 5 funduszy o najwyższych wartościach `score_kons`, odpowiadających profilowi konserwatywnemu.

Dobór składników do każdego z dwóch portfeli — konserwatywnego i agresywnego — stanowił kompromis pomiędzy reprezentatywnością a przejrzystością konstrukcji. Wybrane fundusze

cechowały się najbardziej jednoznacznym profilem językowym, co pozwoliło uniknąć rozmycia charakterystyki portfela oraz zapewniło semantyczną spójność analizowanych zbiorów.

Takie podejście umożliwiło stworzenie dwóch wyraźnie zróżnicowanych grup funduszy. W dalszej części analizy przyjęto podejście agregacyjne, polegające na utworzeniu dwóch skalarnych szeregów czasowych: jednego dla portfela konserwatywnego, drugiego — dla portfela agresywnego. W obu przypadkach szereg ten zbudowano jako średnią ważoną logarytmicznych stóp zwrotu funduszy należących do danego portfela, przy założeniu równych wag (portfel equally weighted).

Wybór tego podejścia wynikał z chęci uproszczenia estymacji modeli GARCH i MS-GARCH, ale znajduje również oparcie w literaturze. McAleer i da Veiga (2008) wykazali, że w przypadku portfeli o niewielkiej liczbie składników, prognozy miar ryzyka (np. Value at Risk) uzyskiwane z modeli opartych na zagregowanym szeregu czasowym nie odbiegają istotnie od wyników modeli wielowymiarowych. Dodatkowo, podejście to pozwala uniknąć problemów związanych z tzw. klątwą wymiarowości (curse of dimensionality), czyli nadmiernym wzrostem liczby parametrów w modelach obejmujących wiele zmiennych jednocześnie.

Wybrane fundusze zaklasyfikowane do analizy zostały przedstawione w Tabeli 1. Szczegółowe informacje dotyczące profilu poszczególnych funduszy oraz ich deklarowanej polityki inwestycyjnej dostępne są w odrębnym pliku (format CSV/XLSX), który może zostać udostępniony na żądanie.

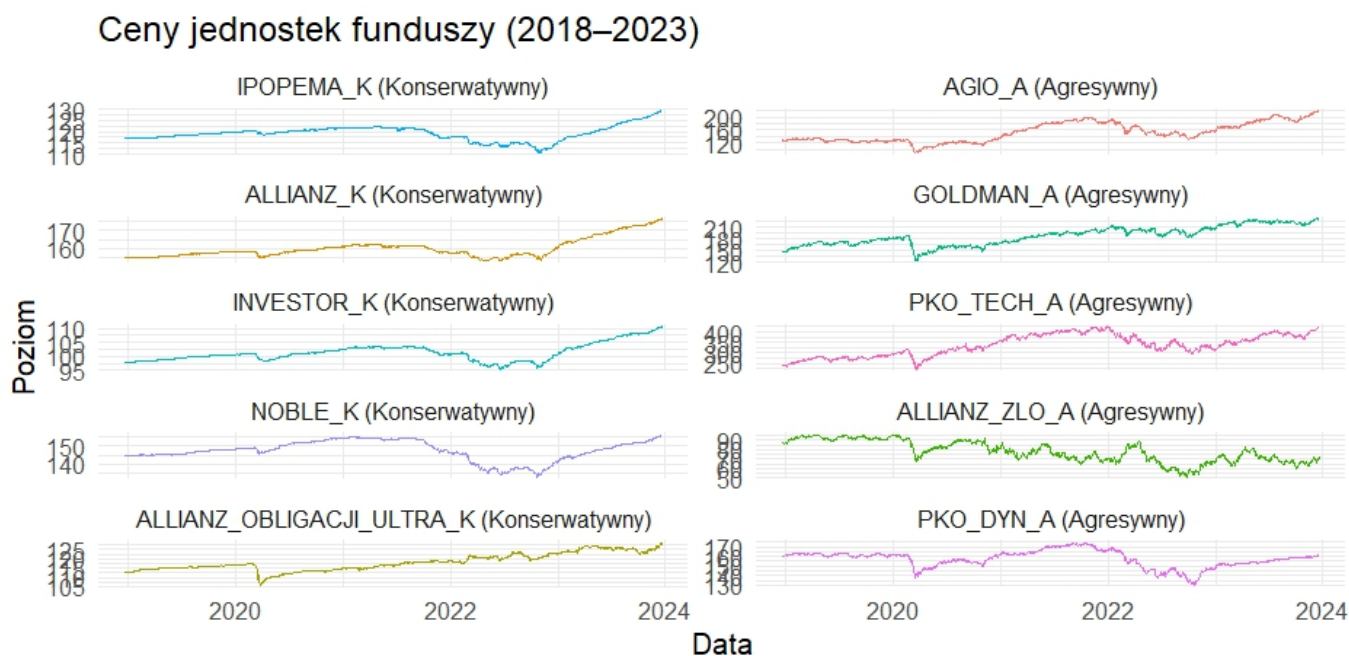
Tabela 1: Najbardziej agresywne i konserwatywne fundusze w próbie badawczej

Nazwa funduszu (agresywne)	Kod	Scoring	Nazwa funduszu (konserwatywne)	Kod	Scoring
PKO Technologii i Innowacji Globalny	PKO027	13	INVESTOR Konserwatywny	DWS037	10
GOLDMAN SACHS Europejski Spółek Dywidendowych	ING027	13	IPOPEMA Konserwatywny Uniwersalny A	IPO082	8
ALLIANZ Akcji Rynku Złota	ALL033	9	ALLIANZ Obligacji Ultra Długoterminowych	ALL038	7
AGIO Akcji PLUS	AGI044	8	ALLIANZ Konserwatywny	ALL005	6
PKO Dynamicznej Akcji	PKO002	8	NOBLE FUND Konserwatywny	NOB001	6

Kolumna *Scoring* zawiera przypisane punkty odzwierciedlające stopień „agresywności” lub „konserwatywności” funduszy, przyznane na podstawie częstości występowania charakterystycznych słów w ich opisach.

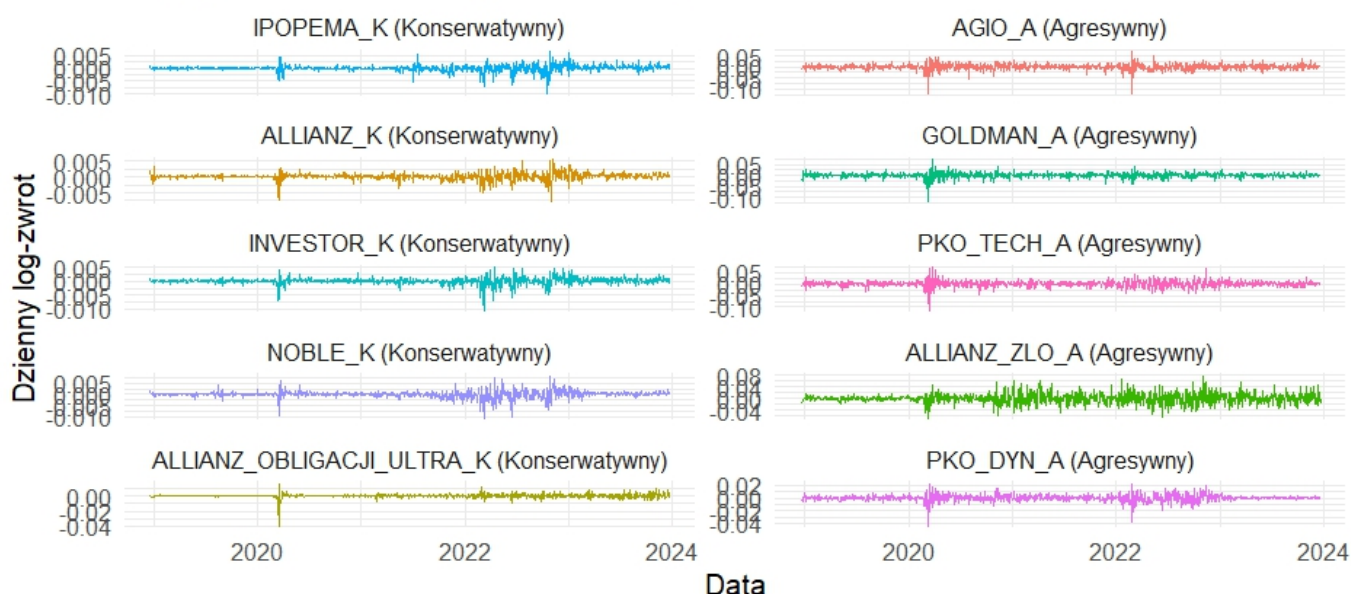
3.1 Wykresy cen i zwrotów z funduszy

Empiryczną część analizy rozpoczyna się od graficznej prezentacji szeregów cen oraz odpowiadających im logarytmicznych stóp zwrotu dla każdego z wybranych funduszy. Zakres czasowy badania obejmuje okres od 20 grudnia 2018 roku do 20 grudnia 2023 roku, co odpowiada pięcioletniemu horyzontowi obserwacji. Długość tego okna została dobrana zgodnie z powszechnie przyjętymi praktykami w literaturze empirycznej oraz zaleceniami metodologicznymi dotyczącymi analizy ryzyka (Khanniche, 2009). Analiza oparta jest na dziennych danych.



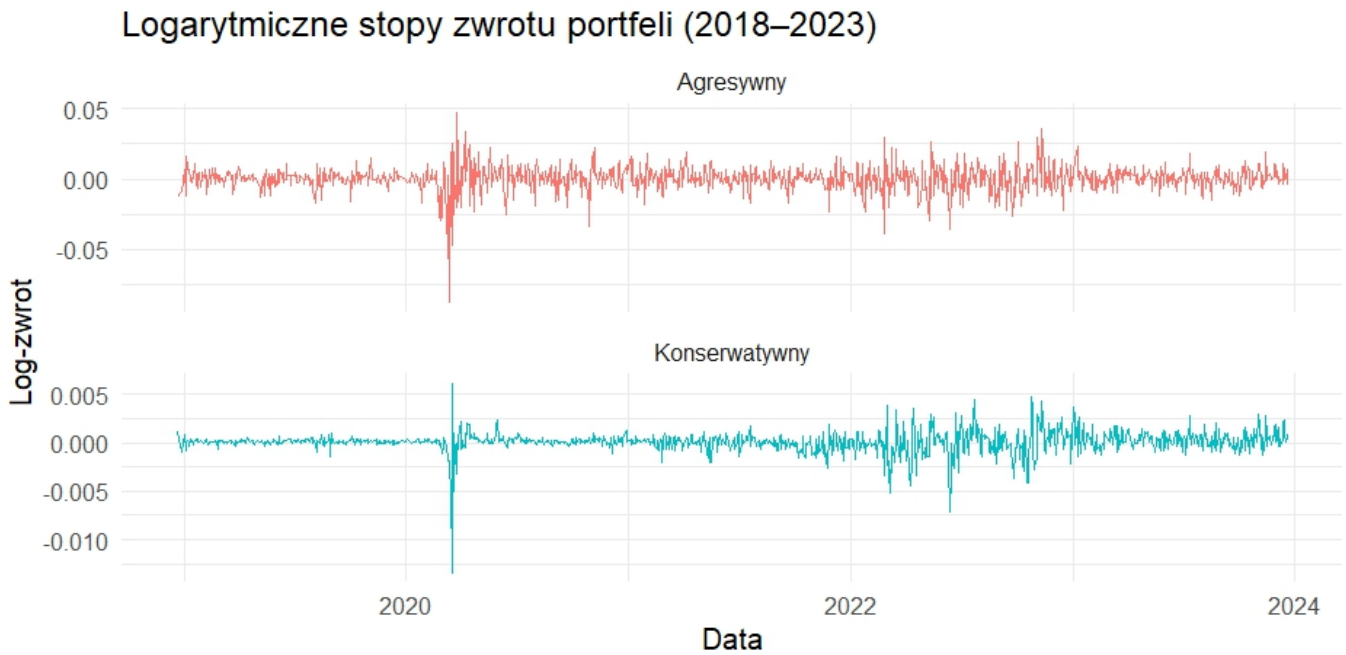
Rysunek 1: Ceny jednostek funduszy 2018-2023

Stopy zwrotu funduszy (2018–2023)

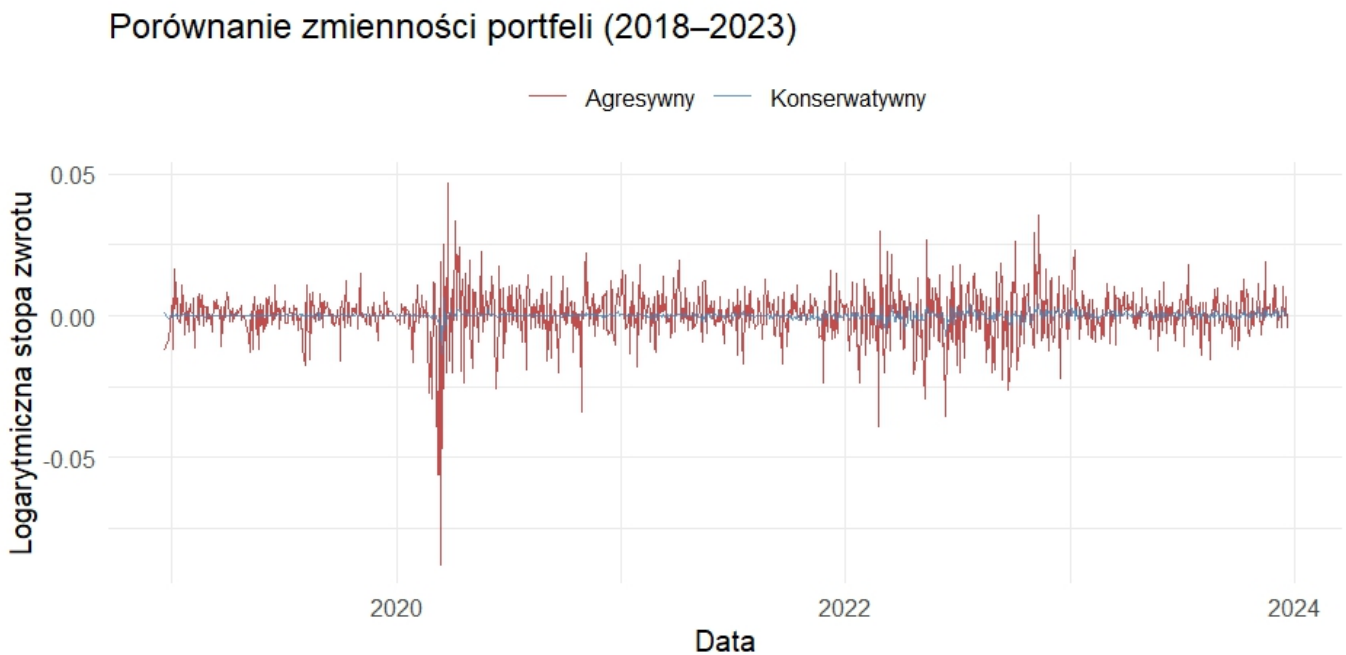


Rysunek 2: Logarytmiczne stopy zwrotu funduszy 2018-2023

Zagregowane skalarne szeregi stóp zwrotów przedstawiono na rysunkach 3. i 4. Już na podstawie analizy wizualnej można zaobserwować wyraźne zróżnicowanie poziomu zmienności między oboma portfelami. Portfel konserwatywny cechuje się zdecydowanie niższą zmiennością, natomiast portfel agresywny wykazuje liczne epizody podwyższonego ryzyka. Wstępne szacunki wskazują, że wariancja portfela agresywnego może być nawet około dziesięciokrotnie wyższa niż wariancja portfela konserwatywnego.



Rysunek 3: Porównanie logarytmicznych stóp zwrotu portfeli



Rysunek 4: Porównanie zmienności portfeli

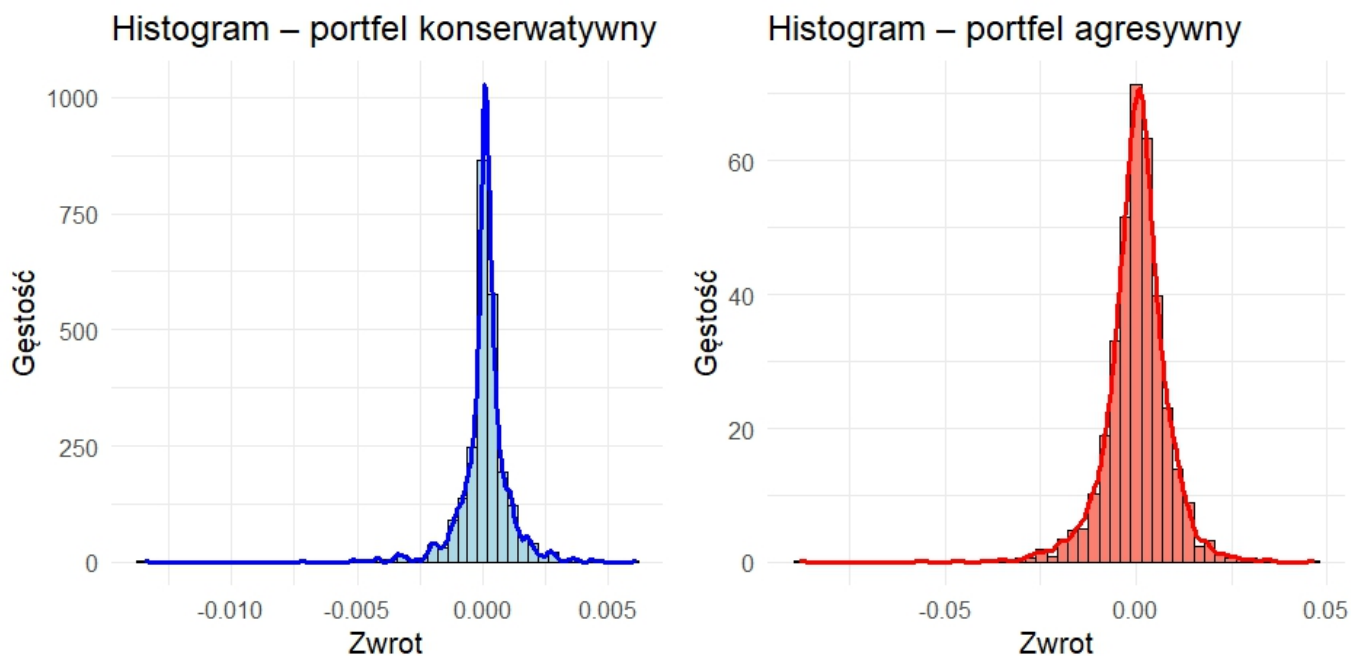
3.2 Charakterystyka danych

Statystyki opisowe potwierdzają znaczącą różnicę w zmienności między portfelami — odchylenie standardowe portfela agresywnego jest ponad **osiem razy większe** niż konserwatywnego. Oba rozkłady charakteryzują się **silną lewostronną asymetrią** oraz **bardzo wysoką kurtozą**, wskazującą na obecność ekstremalnych obserwacji i odchyłeń od normalności, co znajduje potwierdzenie w wartościach **statystyki Jarque-Bera**.

Tabela 2: Statystyki opisowe dla portfela konserwatywnego i agresywnego

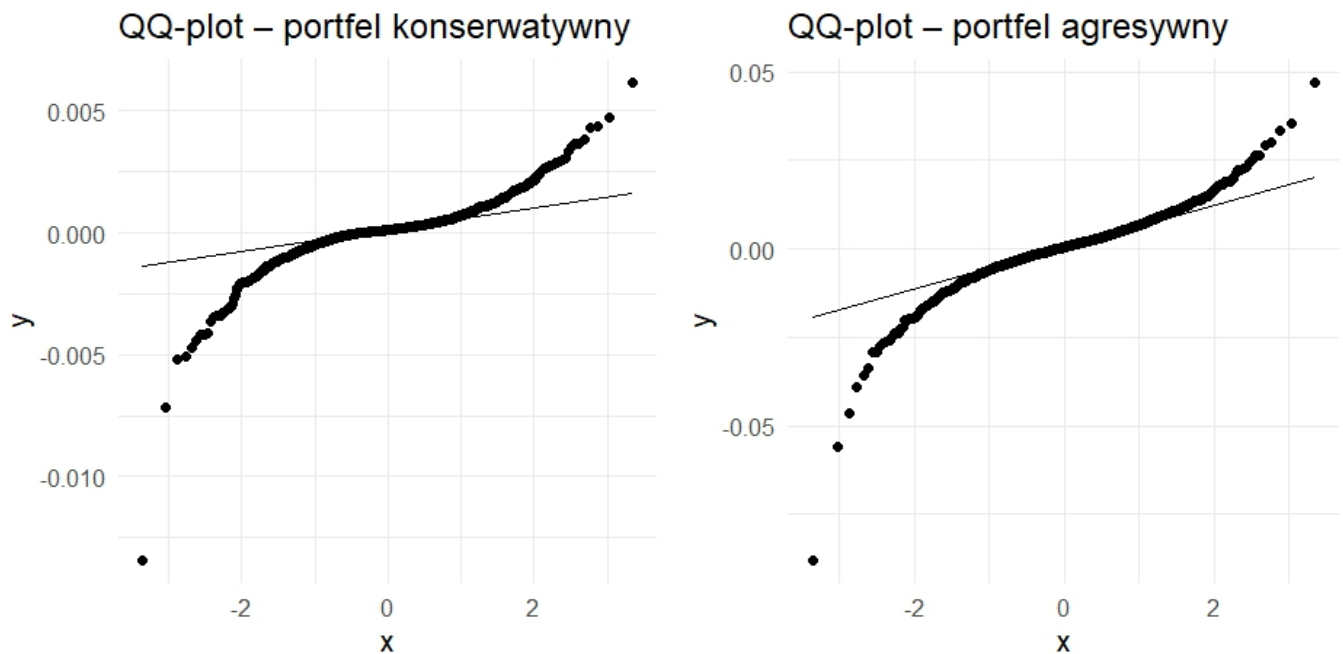
Portfel	Średnia	Mediana	Min	Max	SD	Skośność	Kurtoza
Konserwatywny	0.000090	0.000113	-0.013435	0.006155	0.001059	-2.156	27.01
Agresywny	0.000179	0.000517	-0.087985	0.046968	0.008607	-1.300	12.90

Wysokie wartości statystyki testowej Jarque’a-Bery — jak w przypadku obu analizowanych portfeli (38 632 dla konserwatywnego i 8 890 dla agresywnego) — świadczą o zdecydowanym odrzuceniu hipotezy zerowej na rzecz alternatywnej. Oznacza to, że rozkłady stóp zwrotu znacząco odbiegają od normalności, co potwierdzają także obserwowane wartości skośności i kurtozy. Taki wynik uzasadnia zastosowanie modeli i metod uwzględniających niestandardowe cechy danych, takie jak **leptokurtyczność** i **asymetria**.



Rysunek 5: Histogramy stóp zwrotów

Histogramy potwierdzają nienormalność rozkładów stóp zwrotu – obie serie charakteryzują się leptokurtycznością oraz asymetrią, szczególnie widoczną w przypadku portfela konserwatywnego. Dodatkowo, estymacja gęstości wykonana metodą jądrową uwidacznia wąski szczyt i wydłużone ogony, co również wskazuje na istotne odstępstwa od rozkładu normalnego.



Rysunek 6: Wykresy kwantyl-kwantyl

Wykresy kwantyl-kwantyl porównują kwantyle empiryczne (oś OY) z teoretycznymi kwantylami rozkładu normalnego (oś OX). Widoczne odstępstwa punktów od linii prostej, szczególnie w ogonach, wskazują na nienormalność rozkładów. Oba portfele wykazują leptokurtyczność i obecność obserwacji ekstremalnych, które odbiegają od oczekiwań przy rozkładzie normalnym.

3.3 Testy statystyczne

Wyniki testów Augmented Dickey–Fuller (ADF) przedstawione w tabeli 3. wskazują na stacjonarność logarytmicznych stóp zwrotu zarówno dla portfela konserwatywnego, jak i agresywnego. W obu przypadkach wartość statystyki testowej znacząco przekracza wartość krytyczną, co pozwala na odrzucenie hipotezy zerowej o obecności pierwiastka jednostkowego na wysokim poziomie istotności.

Tabela 3: Wyniki testów stacjonarności dla logarytmicznych stóp zwrotu

Portfel	Statystyka ADF	Wartości krytyczne ADF	Statystyka KPSS	<i>p</i> -wartość KPSS
Konserwatywny	-19,296	-3,43 / -2,86 / -2,57	0,557	0,029
Agresywny	-22,428	-3,43 / -2,86 / -2,57	0,052	0,100

Odmienne wyniki przynosi test KPSS (Kwiatkowski–Phillips–Schmidt–Shin), który zakłada odwrotny zestaw hipotez — hipotezą zerową jest stacjonarność szeregu. W tym przypadku szereg stóp zwrotu dla portfela konserwatywnego został oceniony jako niestacjonarny, natomiast dla portfela agresywnego hipotezy zerowej nie odrzucono, co potwierdza jego stacjonarność.

Należy zauważyć, że wpisana w wynikach testu KPSS wartość $p = 0.1$ nie odzwierciedla dokładnego poziomu istotności, lecz stanowi przyjętą konwencję — wartości p większe niż 0.1 są sprowadzane do tej granicy. W przypadku portfela agresywnego nieodrzućenie hipotezy zerowej o stacjonarności można zatem uznać za jednoznaczne.

W odniesieniu do portfela konserwatywnego wyniki testu KPSS mogą być częściowo zafałszowane przez obecność efektu ARCH lub autokorelacji, które nie są w pełni uwzględniane przez tę procedurę testową. Niska zmienność oraz silne skupienie obserwacji wokół średniej mogą być błędnie interpretowane przez test KPSS jako oznaki niestacjonarności. Tego rodzaju fałszywe sygnały są znane w literaturze jako ograniczenia testu KPSS w kontekście heteroskedastyczności, w szczególności przy obecności efektów typu ARCH czy GARCH (Lamoureux & Lastrapes, 1990). Mając na uwadze możliwe ograniczenia testu KPSS oraz jednoznaczne wyniki testu ADF, przyjęto założenie o stacjonarności szeregu czasowego portfela konserwatywnego na gruncie definicyjnym.

Tabela 4: Wyniki testów diagnostycznych dla logarytmicznych stóp zwrotu

Portfel	ARCH statystyka	<i>p</i> -wartość ARCH	McLeod–Li statystyka	<i>p</i> -wartość McLeod–Li	Ljung–Box statystyka (dla serii)
Konserwatywny	156,988	0,000	251,686	0,000	191,670
Agresywny	222,984	0,000	423,067	0,000	43,315

Wyniki testów diagnostycznych przedstawionych w tabeli 4. jednoznacznie wskazują na obecność autokorelacji oraz efektów warunkowej heteroskedastyczności w szeregach logarytmicznych stóp zwrotu obu portfeli.

Wysokie wartości statystyki testu ARCH oraz bardzo niskie wartości *p*-value sugerują, że dla obu portfeli występuje efekt ARCH, czyli zmienność wykazuje zależność od wartości z poprzednich okresów. Zjawisko to jest typowe dla danych finansowych i objawia się w postaci skupisk zmienności (tzw. volatility clustering).

Podobnie, test McLeoda–Li, będący rozszerzoną wersją testu Ljunga–Boxa stosowaną do kwadratów reszt, potwierdza obecność autokorelacji w wariancji, co dodatkowo wskazuje na dyna-

miczne zmiany zmienności w czasie.

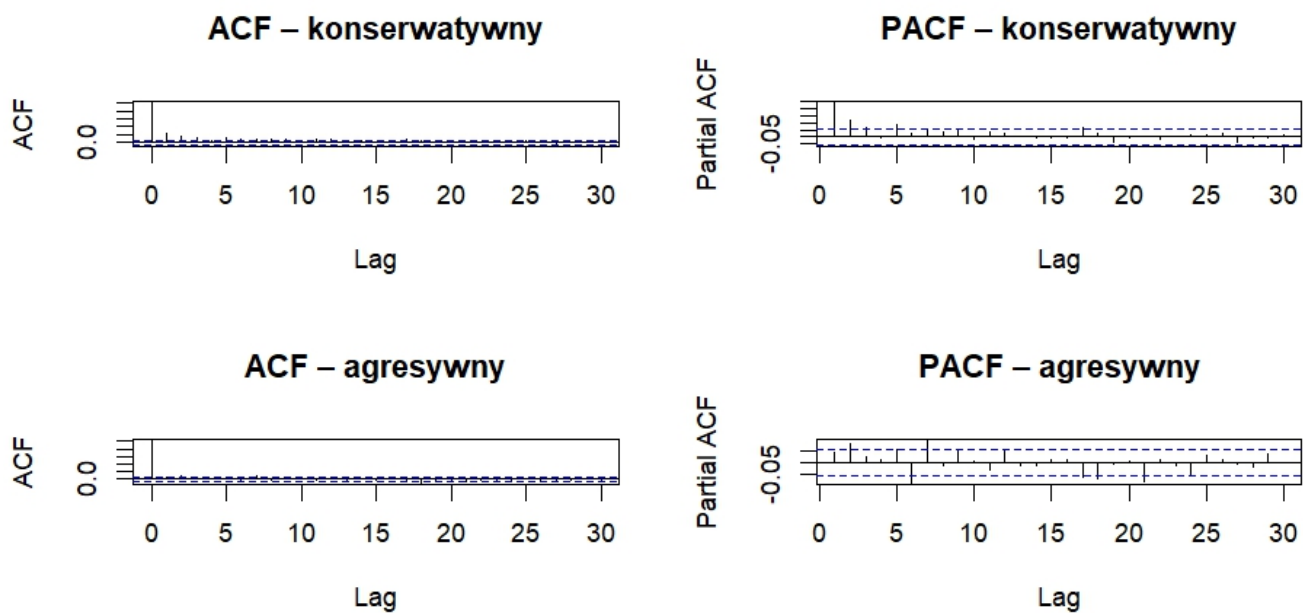
Test Ljunga–Boxa dla samej serii stóp zwrotu wykazuje znacznie silniejsze efekty autokorelacji w portfelu konserwatywnym niż w agresywnym (191,7 vs. 43,3), co może sugerować wyższą regularność krótkookresowych wahań w konserwatywnym portfelu. Jest to zgodne z charakterystyką instrumentów konserwatywnych – stabilnych i mniej podatnych na gwałtowne szoki rynkowe.

Łącznie, wyniki te potwierdzają zasadność modelowania zmienności przy użyciu warunkowych modeli heteroskedastycznych typu GARCH oraz ich wariantów, a także ewentualnej potrzeby uwzględnienia struktury autokorelacyjnej.

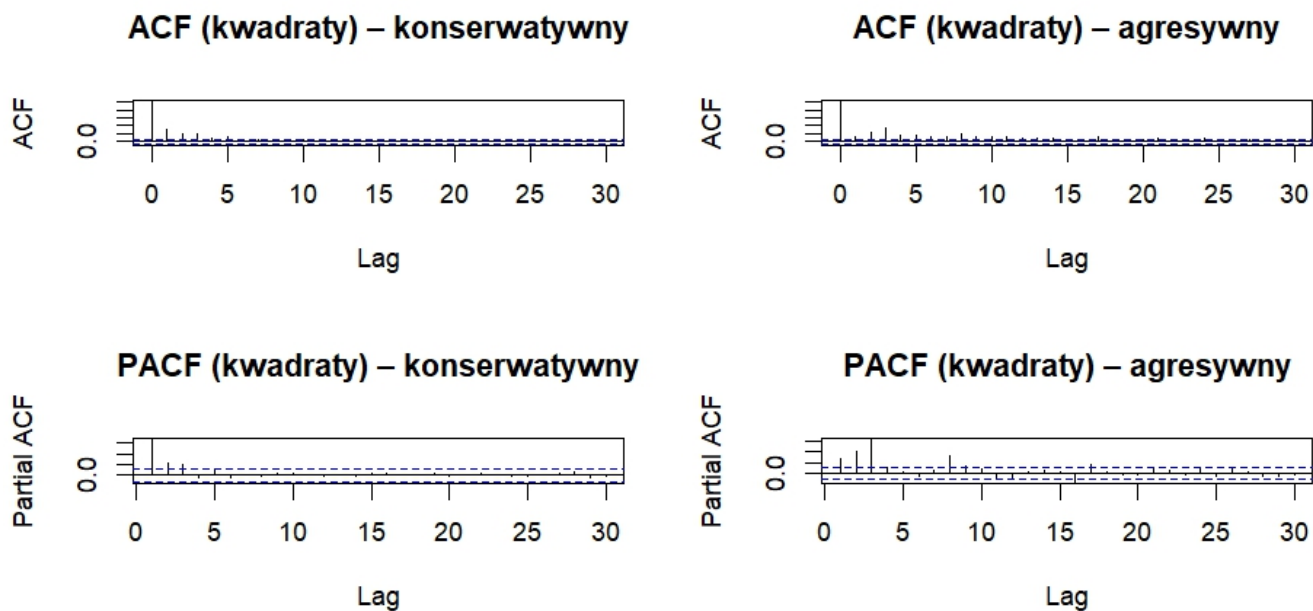
Rysunek 7 przedstawia wykresy funkcji autokorelacji (ACF) oraz cząstkowej funkcji autokorelacji (PACF) dla logarytmicznych stóp zwrotu z obu portfeli. Na wykresach zaznaczono granice istotności statystycznej przy poziomie ufności 95%, wyznaczone według wzoru: $\pm \frac{1.96}{\sqrt{T}}$, gdzie T to liczba obserwacji, 1.96 to wartość krytyczna z rozkładu normalnego dla poziomu istotności 5%. Przekroczenie tych granic przez słupki świadczy o istotnej autokorelacji w danym opóźnieniu.

Dla portfela konserwatywnego zauważalne są istotne autokorelacje w pierwszych kilku opóźnieniach na wykresie ACF, co sugeruje obecność liniowej struktury czasowej i większą przewidywalność zmian. Z kolei w przypadku portfela agresywnego wykres ACF nie wskazuje na istotne autokorelacje, więc można wnioskować o większej przypadkowości i mniejszej regularności procesu generującego dane.

Warto zwrócić uwagę na wykresy PACF, które ukazują cząstkową autokorelację – czyli wpływ danego opóźnienia po odjęciu efektów wcześniejszych lagów. Dla portfela agresywnego pojawiają się pojedyncze istotne wartości, co może wskazywać na obecność zależności wyższego rzędu, które nie są uchwytne w klasycznym ACF.



Rysunek 7: Panel ACF i PACF dla stóp zwrotu



Rysunek 8: Panel ACF i PACF dla kwadratów stóp zwrotu

Na podstawie analizy wstępnej zgodnej z podejściem Boxa–Jenkinsa stwierdzono zasadność uwzględnienia składników autoregresyjnych (AR) oraz średniej ruchomej (MA) w modelowaniu struktury średniej, co przemawia za zastosowaniem modelu typu ARMA-GARCH. Istotność opóźnień w funkcjach ACF i PACF wskazuje na obecność systematycznych zależności liniowych w szeregach stóp zwrotu, których nieuwzględnienie mogłoby prowadzić do błędnej specyfikacji wariancji warunkowej.

Dodatkowo, analiza ACF i PACF dla kwadratów stóp zwrotu, przedstawiona na Rysunku 8., potwierdza występowanie autokorelacji w drugich momentach, co jest charakterystycznym sygnałem obecności heteroskedastyczności warunkowej. Obserwacje te są zgodne z wynikami testów diagnostycznych (np. ARCH-LM), które również wskazują na obecność efektu ARCH.

Tym samym, zarówno analiza graficzna, jak i testy formalne, dostarczają spójnych podstaw do estymacji modeli GARCH.

Wstępnego doboru komponentów ARMA dokonano z wykorzystaniem funkcji `auto.arima()` z pakietu `forecast` w języku R. Algorytm ten automatycznie wybiera optymalną specyfikację modelu ARIMA na podstawie maksymalizacji funkcji wiarygodności estymowanej metodą największej wiarygodności, przy jednoczesnym minimalizowaniu wybranego kryterium informacyjnego.

Zgodnie z wcześniejszym założeniem o stacjonarności logarytmicznych stóp zwrotu, nałożono dodatkowy warunek ograniczający parametr różnicowania $d = 0$, co zapewnia wybór wyłącznie modeli ARMA(p,q), bez komponentu różnicującego.

Tabela 5. prezentuje wyniki wstępnego doboru komponentów ARMA, które następnie zostały wykorzystane jako część struktury modeli GARCH w dalszej analizie empirycznej.

Tabela 5: Porównanie modeli ARMA dla portfela konserwatywnego i agresywnego

Parametr	Portfel konserwatywny	Portfel agresywny
Model ARMA	(3,0,2)	(5,0,0)
AR(1)	0.3394 (0.0917)	−0.0308 (0.0318)
AR(2)	−0.4213 (0.0973)	0.0965 (0.0376)
AR(3)	0.3015 (0.0382)	0.1937 (0.0327)
AR(4)	–	0.0681 (0.0326)
AR(5)	–	−0.1161 (0.0368)
MA(1)	−0.1837 (0.0983)	–
MA(2)	0.5709 (0.1071)	–
Stała (mean)	0.0001 (0.0001)	–
σ^2	$9.7 \cdot 10^{-7}$	$6.997 \cdot 10^{-5}$
Log-likelihood	6787.58	4131.37
AIC	−13561.15	−8250.74
AICc	−13561.09	−8250.70
BIC	−13522.58	−8217.68

3.4 Estymacja i diagnostyka modeli typu GARCH dla portfela konserwatywnego

Tabela 6 przedstawia wyniki estymacji wybranych specyfikacji modeli GARCH dla szeregu czasowego portfela konserwatywnego. Wstępnie przyjęto model GARCH(1,1) ze składnikiem średniej ARMA(3,2) oraz rozkładem normalnym dla składnika losowego. W kolejnych krokach autor przeprowadził iteracyjny proces optymalizacji, obejmujący modyfikację zarówno typu rozkładu, jak i struktury modelu wariancji.

Pierwsze porównanie modelu opartego na rozkładzie normalnym z modelem wykorzystującym rozkład t-Studenta wykazało, że ten drugi zapewnia wyraźnie lepsze dopasowanie do danych. Świadczy o tym wyższa wartość logarytmu funkcji wiarygodności oraz niższe wartości kryteriów informacyjnych AIC i BIC, co potwierdza lepsze właściwości statystyczne modelu.

W dalszej analizie oszacowano również modele typu EGARCH i GJR-GARCH, zakładając rozkład t-Studenta dla składnika losowego. Wstępna diagnostyka wskazała jednak na niższą stabilność i gorsze dopasowanie modeli innych niż EGARCH — zarówno pod względem wartości funkcji wiarygodności, jak i kryteriów informacyjnych. Dodatkowo, jak pokazuje Tabela 7, nie wszystkie parametry tych modeli uzyskały istotność statystyczną na konwencjonalnych poziomach istotności.

Na podstawie tych obserwacji wybrano do dalszej analizy model EGARCH(1,2) z rozkładem skośnym t-Studenta (skewed-t), który wykazał najlepsze dopasowanie do danych. Model ten osiągał wyższą wartość funkcji wiarygodności oraz niższe wartości AIC i BIC w porównaniu do EGARCH z rozkładem t-studenta. Wybrana specyfikacja została ostatecznie przedstawiona w Tabeli 6.

Tabela 6: Estymacja parametrów modeli GARCH dla portfela konserwatywnego

Parametr	sGARCH(1,1), norm	sGARCH(1,1), std	eGARCH(1,2), sstd	gjrGARCH(1,1), std
μ	0.000114	0.000121	0.000094	0.000129
AR(1)	0.254	0.258	0.284	0.210
AR(2)	0.771	0.693	0.697	0.771
AR(3)	-0.086	-0.053	-0.085	-0.066
MA(1)	-0.081	-0.128	-0.106	-0.068
MA(2)	-0.741	-0.665	-0.688	-0.754
ω	0.000	0.000	-0.360	0.000
α_1	0.125	0.095	-0.093	0.050
β_1	0.859	0.887	0.831	0.896
β_2	–	–	0.144	–
γ_1	–	–	0.483	0.060
Skew (λ)	–	–	0.938	–
Shape (ν)	–	4.684	4.257	4.044
LogLik	7373.423	7451.621	7501.022	7437.490
AIC	-11.888	-12.012	-12.087	-11.988
BIC	-11.850	-11.971	-12.033	-11.942

Tabela 7. prezentuje istotność estymowanych parametrów modeli zgodnie z następującą konwencją oznaczeń poziomów istotności statystycznej:

- *** – istotność na poziomie 1% ($p < 0,01$),
- ** – istotność na poziomie 5% ($p < 0,05$),
- – istotność na poziomie 10% ($p < 0,10$),
- brak symbolu – brak istotności ($p \geq 0,10$),
- „–” – parametr nie występuje w danej specyfikacji modelu.

Tabela 7: Istotność parametrów estymowanych w modelach GARCH (na podstawie testu t-studenta)

Parametr	sGARCH(1,1), norm	sGARCH(1,1), std	eGARCH(1,2), sstd	gjrGARCH(1,1)
μ	***	***	***	***
AR(1)			***	
AR(2)	**		***	
AR(3)	*		***	
MA(1)			**	
MA(2)	**		***	*
ω			***	
α_1	***	***	**	***
β_1	***	***	***	***
β_2	—	—	***	—
γ_1	—	—	***	***
Skew (λ)	—	—	***	—
Shape (ν)	—	***	***	***

Tabela 8 prezentuje wyniki testów diagnostycznych, w tym testu stabilności parametrów opartego na statystyce Nybloma (1989). Test ten służy do oceny hipotezy zerowej o stabilności parametrów w czasie. Statystyka testowa porównywana jest z wartościami krytycznymi z rozkładu asymptotycznego (1% - 1.93; 5% - 1.52; 10% - 1.34).

Jeśli wartość statystyki przekracza wartość krytyczną, hipoteza o stabilności parametrów zostaje odrzucona.

W badanym przypadku model EGARCH(1,2) z rozkładem skośnym t-Studenta osiągnął wartość statystyki stabilności równą 4.46, co przekracza próg istotności 1%. Choć formalnie oznacza to brak stabilności parametrów, model ten wciąż wykazuje znaczną przewagę empiryczną nad pozostałymi specyfikacjami, które uzyskały jeszcze wyższe statystyki testowe (przekraczające 400), jednoznacznie dyskwalifikujące je jako niestabilne.

Z tego względu, pomimo niespełnienia formalnego warunku stabilności, model EGARCH(1,2) został wybrany jako optymalna struktura opisująca zmienność portfela konserwatywnego — zarówno pod względem dopasowania, jak i przejścia testów diagnostycznych związanych z resztami oraz strukturą warunkowej wariancji.

Zastanowienie może budzić obecność autokorelacji w modelu, sugerowana wstępnie przez wyniki testu Ljunga–Boxa na standaryzowanych resztach, domyślnie implementowanego w pakiecie rugarch jako tzw. ważony test Ljunga–Boxa (Weighted Ljung–Box Test). Ze względu na ograniczoną przejrzystość interpretacyjną tej wersji testu, autor zdecydował się przeprowadzić klasyczną, nieważoną wersję testu Ljunga–Boxa na standaryzowanych resztach, wykorzystując 14 opóźnień.

Wyniki przeprowadzonego testu wskazują, że statystyka testowa wyniosła $\chi^2 = 17,8432$ przy 14 stopniach swobody, a odpowiadająca jej wartość p była równa 0,214. Ponieważ wartość ta znacząco przekracza konwencjonalne poziomy istotności, nie ma podstaw do odrzucenia hipotezy

zerowej o braku autokorelacji w resztach.

W związku z tym autor uznaje, że model EGARCH(1,2) z rozkładem skośnym t-Studenta spełnia wymogi poprawnej specyfikacji również pod względem struktury reszt, a zastosowana specyfikacja nie wymaga dalszych korekt w tym zakresie.

Tabela 8: Testy diagnostyczne modeli GARCH dla portfela konserwatywnego

Statystyka	sGARCH-norm	sGARCH-std	eGARCH-sstd	gjrGARCH-std
ARCH LM (p-val)	0.1610	0.0046	0.3197	0.0392
Ljung-Box reszty [14]	0.0292	0.0000	5.3e-08	0.0000
Ljung-Box kw. reszt [5]	0.0139	2.7e-05	0.8296	0.0007
Sign Bias (p-val)	0.8537	0.7521	0.5850	0.9406
Neg. Sign Bias (p-val)	0.0035	7.1e-05	0.3403	0.0013
Pos. Sign Bias (p-val)	0.7374	0.9783	0.6980	0.6931
Joint Bias Test (p-val)	0.0140	0.0004	0.7866	0.0074
Stabilność (joint stat.)	431.72	439.85	4.46	436.37
Pearson GoF (p-val, g=30)	1.7e-06	0.0045	0.7180	2.0e-08

3.5 Estymacja modelu MS-EGARCH dla portfela konserwatywnego

Jak wykazano w poprzednim podrozdziale, spośród analizowanych modeli najlepiej do danych portfela konserwatywnego dopasował się model EGARCH(1,2) z rozkładem skośnym t-Studenta. Zasadniczym celem dalszej części badania jest jednak ocena, czy wprowadzenie mechanizmu przełączania reżimów, charakterystycznego dla modeli Markov-Switching, może prowadzić do poprawy jakości prognozowanych miar ryzyka. W tym celu niezbędne jest oszacowanie modelu MS-EGARCH, który pozwala na uwzględnienie zmian reżimu w dynamice zmienności. Autor zdecydował się także na zmianę rozkładu sstd na zwykły rozkład t-studenta, jako że model z tym rozkładem osiągnął niższe wartości kryteriów informacyjnych.

W praktyce implementacyjnej napotkano jednak istotne ograniczenia techniczne. Popularny pakiet MSGARCH w języku R, wykorzystywany do estymacji modeli GARCH z przełączaniem Markowa, umożliwia zmianę specyfikacji modelu wariancji (np. na EGARCH, GJR-GARCH lub APARCH), jednak ogranicza strukturę modelu GARCH wyłącznie do postaci (1,1). Ponadto pakiet nie oferuje możliwości jednoczesnej estymacji składników ARMA w równaniu średniej oraz składników wariancji warunkowej. Oznacza to, że mechanizm przełączania może być zaimplementowany wyłącznie w procesie wariancji, natomiast wartość oczekiwana musi zostać oszacowana niezależnie.

W związku z powyższym zdecydowano się na zastosowanie procedury estymacji dwustopniowej. W pierwszym kroku wyestymowano model ARMA(3,2) dla składnika średniej, zgodnie z wynikami przedstawionymi we wcześniejszych częściach analizy. W drugim kroku, na podstawie uzyskanych reszt, oszacowano model MS-EGARCH(1,1), co wiązało się również z uproszczeniem

liczby parametrów beta do jednego.

Tabela 9 przedstawia wyniki estymacji modelu MS-EGARCH(1,1) z rozkładem t-Studenta dla portfela konserwatywnego. Wszystkie oszacowane parametry są istotne statystycznie, co potwierdzono w Tabeli 10, prezentującej poziomy istotności przyjęte zgodnie z konwencją opisaną wcześniej.

Model zakłada istnienie dwóch odmiennych reżimów zmienności, z których każdy charakteryzuje się odrębną dynamiką warunkowej wariancji. W pierwszym reżimie obserwuje się umiarkowaną reaktywność zmienności na nowe informacje. Parametr asymetrii $\alpha_1 = 0,0085$ przyjmuje niską, dodatnią wartość, natomiast efekt szoku (parametr $\alpha_2 = 0,0268$) również pozostaje na umiarkowanym poziomie. Wysoka wartość parametru $\beta = 0,9608$ wskazuje na dużą trwałość procesów zmienności, zaś parametr rozkładu t-Studenta $\nu = 5,8237$ sugeruje stosunkowo cieńsze ogony rozkładu i mniejszą częstość obserwacji ekstremalnych.

W drugim reżimie dynamika zmienności ulega istotnej zmianie. Wartość parametru asymetrii wzrasta do $\alpha_1 = 0,4252$, a efekt szoku przyjmuje ujemną wartość $\alpha_2 = -0,1322$. Taka struktura może sugerować asymetryczną reakcję rynku, polegającą na silniejszym tłumieniu pozytywnych wstrząsów lub silniejszym wzroście zmienności po negatywnych informacjach. Trwałość procesu pozostaje wysoka ($\beta = 0,9559$), natomiast wartość parametru $\nu = 5,2126$ wskazuje na grubsze ogony, czyli większe prawdopodobieństwo wystąpienia ekstremalnych obserwacji w tym reżimie.

Tabela 9: Estymacja parametrów modelu MS-eGARCH(1,1) dla portfela konserwatywnego

Parametr	Reżim 1	Reżim 2
α_0	-0.6853	-0.6200
α_1	0.0085	0.4252
α_2	0.0268	-0.1322
β	0.9608	0.9559
ν	5.8237	5.2126

Tabela 10: Istotność estymowanych parametrów w modelu MS-eGARCH(1,1) dla portfela konserwatywnego

Parametr	Reżim 1	Reżim 2
α_0	***	***
α_1	*	***
α_2	***	***
β	***	***
ν	***	***

Tabela 11: Macierz przejścia dla modelu MS-eGARCH(1,1) dla portfela konserwatywnego

	Przejście do stanu 1	Przejście do stanu 2
Z reżimu 1	0.9847	0.0153
Z reżimu 2	0.0025	0.9975

Macierz przejścia widoczna w tabeli 11. wskazuje na wysoką trwałość obu reżimów – prawdopodobieństwo pozostania w reżimie 1 wynosi 98,47%, a w reżimie 2 aż 99,75%. Niskie prawdopodobieństwa przejść między reżimami ($P_{1 \rightarrow 2} = 0,0153$ i $P_{2 \rightarrow 1} = 0,0025$) potwierdzają, że model skutecznie wyodrębnia względnie stabilne fazy rynku o odmiennych charakterystykach zmienności.

Tabela 12: Prawdopodobieństwa stacjonarne dla modelu MS-eGARCH(1,1)

Stan	Prawdopodobieństwo
1	0.1398
2	0.8602

Obliczone prawdopodobieństwa stacjonarne (13,98% dla reżimu 1 oraz 86,02% dla reżimu 2) informują, że model oczekuje, iż w długim okresie portfel konserwatywny będzie znajdować się głównie w reżimie drugim. W kontekście wcześniejszej analizy parametrów, reżim ten charakteryzował się silną asymetrią i bardziej leptokurtycznym rozkładem (niższym parametrem ν), co może odpowiadać okresom większego ryzyka lub napięcia rynkowego.

Tabela 13: Kryteria informacyjne dla modelu MS-eGARCH(1,1)

Wskaźnik	Wartość
Log-likelihood	7504.0794
AIC	-14984.1588
BIC	-14922.6940

Wysoka wartość funkcji wiarygodności (Log-likelihood = 7504,08) oraz bardzo niskie wartości kryteriów informacyjnych (AIC = -14984,16, BIC = -14922,69) wskazują na dobrą jakość dopasowania modelu do danych. Należy jednak podkreślić, że wartości AIC i BIC należy interpretować względnie. W literaturze wskazuje się, że bezpośrednie porównywanie wartości AIC/BIC między modelami o odmiennych strukturach (np. między standardowymi modelami GARCH a modelami Markov-switching GARCH) może prowadzić do mylących wniosków, jeśli nie uwzględni się różnic w liczbie parametrów, sposobie estymacji oraz liczbie stanów ukrytych (Haas et al., 2004 oraz Burnham i Anderson, 2002). Dlatego porównań takich należy dokonywać ostrożnie, a najlepiej ograniczyć je do modeli o tej samej klasie i analogicznej konstrukcji.

3.6 Estymacja modeli typu GARCH dla portfela agresywnego

Wzorując się na dobrych wynikach uzyskanych dla modelu eGARCH z rozkładem skośnym t-Studenta dla portfela konserwatywnego, wszystkie specyfikacje prezentowane w tabeli 14. oszacowano z założeniem rozkładu skośnego t-Studenta dla składnika losowego.

Tabela 14: Estymacja parametrów modeli GARCH dla portfela agresywnego

Parametr	sGARCH(1,1)	eGARCH(1,1)	gjrGARCH(1,1)	eGARCH(1,2)
μ	0.000311	-0.000035	0.000134	0.000031
AR(1)	0.072	0.085	0.085	0.081
AR(2)	-0.034	-0.015	-0.019	—
AR(3)	-0.042	-0.025	-0.040	—
AR(4)	-0.010	0.010	0.002	—
AR(5)	0.040	0.069	0.056	—
ω	0.000001	-0.122	0.000001	-0.163
α_1	0.0859	-0.143	0.0000	-0.184
α_2	—	—	—	—
β_1	0.908	0.987	0.913	0.611
β_2	—	—	—	0.372
γ_1	—	0.083	0.152	0.121
Skew (λ)	0.901	0.881	0.895	0.889
Shape (ν)	4.596	5.245	4.971	5.431
LogLik	4379.641	4409.869	4398.470	4405.877
AIC	-7.0919	-7.1394	-7.1209	-7.1378
BIC	-7.0463	-7.0896	-7.0711	-7.1004

Tabela 15: Istotność parametrów estymowanych w modelach GARCH dla portfela agresywnego (na podstawie testu t-studenta)

Parametr	sGARCH(1,1)	eGARCH(1,1)	gjrGARCH(1,1)	eGARCH(1,2)
μ	*			
AR(1)	**	*	**	**
AR(2)				—
AR(3)				—
AR(4)				—
AR(5)			*	—
ω		***		***
α_1	***	***		***
α_2	—	—	—	—
β_1	***	***	***	***
β_2	—	—	—	***
γ_1	—	***	***	***
Skew (λ)	***	***	***	***
Shape (ν)	***	***	***	***

Najlepsze dopasowanie według kryteriów AIC i BIC osiągnięto dla modeli eGARCH(1,1) oraz

eGARCH(1,2), które charakteryzują się także wysoką wartością log-likelihood. Modele różnią się pod względem obecności efektu asymetrii (γ_1) oraz zastosowanego rzędu w części średniej (AR). Parametry rozkładu pozostają istotne we wszystkich przypadkach.

Tabela 15 wskazuje, że większość parametrów modeli – w tym parametry wariancji, skośności i kurtozy – jest statystycznie istotna na poziomie 1% lub 5%. Wyjątki stanowią wybrane parametry AR oraz dodatkowe składniki wariancji w modelu eGARCH(1,2), które nie uzyskały istotności.

Tabela 16: Testy diagnostyczne modeli GARCH dla portfela agresywnego

Statystyka	sGARCH(1,1)	eGARCH(1,1)	gjrGARCH(1,1)	eGARCH(1,2)
ARCH LM (p-val)	0.030	0.0468	0.4627	0.4128
Ljung-Box reszty [14]	1.18e-11	0.8594	0.9345	0.1547
Ljung-Box kw. reszt [5]	0.2697	0.3977	0.7761	0.4897
Sign Bias (p-val)	0.2133	0.6518	0.0780	0.5164
Neg. Sign Bias (p-val)	0.1176	0.4498	0.8221	0.8804
Pos. Sign Bias (p-val)	0.1409	0.1060	0.1777	0.1743
Joint Bias Test (p-val)	0.0005	0.0469	0.0074	0.1344
Stabilność (joint stat.)	141.11	3.768	104.54	3.595
Pearson GoF (p-val, g=30)	0.3831	0.3831	0.4190	0.7905

Tabela 16. przedstawia wyniki testów diagnostycznych oceniających poprawność dopasowania modeli GARCH dla portfela agresywnego. Spośród analizowanych modeli jedynie gjrGARCH oraz eGARCH(1,2) nie wykazują istotnego efektu ARCH. Model eGARCH(1,2) przechodzi pozostałe testy diagnostyczne, nie wskazując na istotne problemy ze specyfikacją. Mimo że model eGARCH(1,1) osiągnął wyższą wartość funkcji log-likelihood, to wynik testu ARCH-LM ($p = 0,0468$) świadczy o niewystarczającym uchwyceniu zmienności, co czyni go mniej wiarygodnym. Model jest też mniej stabilny i nie uchwycił poprawnie reakcji zmienności na pozytywne i negatywne szoki (Join Bias test). Z tego względu jako model najlepiej opisujący zmienność portfela agresywnego autor przyjmuje eGARCH(1,2).

3.7 Estymacja modelu MS-EGARCH dla portfela agresywnego

Dla portfela agresywnego przeprowadzono estymację modelu MS-EGARCH(1,1) z założeniem rozkładu skośnego t-Studenta dla składnika losowego. Ze względu na ograniczenia implementacyjne pakietu MSGARCH, zastosowano procedurę estymacji dwustopniowej: w pierwszym kroku oszacowano wartość oczekiwaną przy użyciu modelu AR(1), natomiast w drugim kroku modelowano reszty z wykorzystaniem mechanizmu przełącznikowego w wariancji — zgodnie ze specyfikacją modelu MS-EGARCH.

Tabela 17: Estymacja parametrów modelu MS-eGARCH(1,1) z rozkładem sstd dla portfela agresywnego

Parametr	Reżim 1	Reżim 2
α_0	-0.5524	-0.7636
α_1	0.0069	0.4494
α_2	0.0231	-0.1134
β	0.9685	0.9450
ν	5.2718	5.4139
ξ	0.8839	0.9320

Tabela 17. przedstawia wyniki estymacji modelu MS-eGARCH(1,1) z rozkładem skośnym t-Studenta (sstd) dla portfela agresywnego. Parametry wskazują na istotne różnice między reżimami: w reżimie 2 obserwuje się wyraźnie silniejszy efekt asymetrii ($\alpha_1 = 0,4494$) oraz ujemny efekt szoku ($\alpha_2 = -0,1134$). W obu reżimach utrzymuje się wysoka pamięć zmienności ($\beta > 0,94$) oraz leptokurtyczność rozkładu ($\nu \approx 5,2 - 5,4$). Parametr skośności ξ przyjmuje wartości poniżej 1, co wskazuje na przewagę lewego ogona, a więc większe prawdopodobieństwo ekstremalnych ujemnych stóp zwrotu.

Tabela 18: Istotność estymowanych parametrów w modelu MS-eGARCH(1,1) dla portfela agresywnego

Parametr	Reżim 1	Reżim 2
α_0	***	***
α_1	*	***
α_2	***	***
β	***	***
ν	***	***
ξ	***	***

Tabela 19. prezentuje macierz przejścia dla modelu MS-eGARCH(1,1) dopasowanego do portfela agresywnego. Wysokie wartości diagonalne ($P_{11} = 0,9870, P_{22} = 0,9964$) wskazują na dużą trwałość obu reżimów oraz rzadkie zmiany stanu, co sugeruje wyraźne rozdzielenie faz rynku.

Tabela 19: Macierz przejścia dla modelu MS-eGARCH(1,1) – portfel agresywny

	Przejście do stanu 1	Przejście do stanu 2
Z reżimu 1	0.9870	0.0130
Z reżimu 2	0.0036	0.9964

Tabela 20. przedstawia prawdopodobieństwa stacjonarne dla każdego z reżimów. Zgodnie z nimi, portfel agresywny przebywa długookresowo głównie w reżimie 2 (78,45%), który – zgodnie z wcześniejszą analizą parametrów – może odpowiadać okresom podwyższonego ryzyka i zmienności.

Tabela 20: Prawdopodobieństwa stacjonarne dla modelu MS-eGARCH(1,1)

Stan	Prawdopodobieństwo
1	0.2155
2	0.7845

Tabela 21. zawiera wartości funkcji wiarygodności oraz kryteriów informacyjnych AIC i BIC. Wysoka wartość log-likelihood oraz bardzo niskie wartości AIC i BIC wskazują na dobre dopasowanie modelu do danych.

Tabela 21: Kryteria informacyjne dla modelu MS-eGARCH(1,1) – portfel agresywny

Wskaźnik	Wartość
Log-likelihood	7498.3542
AIC	-14968.7083
BIC	-14896.9995

3.8 Backtesting

W kolejnej części analizy przeprowadzono procedurę backtestingu w celu porównania skuteczności modeli GARCH oraz ich rozszerzeń w postaci modeli z przełączaniem reżimów Markowa w prognozowaniu ryzyka. Analiza została przeprowadzona niezależnie dla obu portfeli — konserwatywnego i agresywnego. Kluczowym celem było sprawdzenie, czy uwzględnienie zmienności reżimów w modelu przynosi poprawę jakości prognozowanych miar ryzyka, takich jak Value at Risk i Expected Shortfall.

Do oszacowania miar ryzyka wykorzystano symulację Monte Carlo, zgodnie ze schematem opisanym w części teoretycznej. Dane podzielono w proporcji 80:20, gdzie 80% zbioru zostało wykorzystane do estymacji modeli (zbiór treningowy), natomiast pozostałe 20% posłużyło do oceny jakości prognoz (zbiór testowy). Zastosowanie metody Monte Carlo umożliwiło symulację wielu możliwych trajektorii cenowych, co pozwoliło lepiej uchwycić strukturalne zmiany zmienności oraz potencjalne skrajne scenariusze rynkowe, szczególnie w kontekście modeli MS-GARCH, które uwzględniają losowe przejścia między reżimami.

W analizie przyjęto dwa poziomy istotności dla miar ryzyka: 1% oraz 5%. Czas obliczeń był znacząco zróżnicowany w zależności od modelu. Dla portfela konserwatywnego estymacja modelu EGARCH(1,2) trwała 491,3 sekundy, natomiast wersja przełącznikowa MS-EGARCH(1,1) wymagała 4166,25 sekundy. Jeszcze większe różnice zaobserwowano dla portfela agresywnego, gdzie model EGARCH(1,1) został oszacowany w 232 sekundy, natomiast jego wersja z przełączaniem reżimów wymagała aż 7317,33 sekundy obliczeń.

Warto zaznaczyć, że model MS-EGARCH(1,1) dla portfela agresywnego sprawił istotne trudności numeryczne. W szczególności, estymowane wartości miary Expected Shortfall były znacząco

zaniżone, co prowadziło do wyników backtestingu pozostających w wyraźnej sprzeczności z intuicją ekonomiczną i realiami rynkowymi. Z tego względu autor zdecydował się na zastosowanie alternatywnego podejścia: dla modelu MS-EGARCH(1,1) oszacowano miarę ES przy użyciu metody parametrycznej, co umożliwiło uzyskanie stabilnych i wiarygodnych wyników dla dalszego porównania modeli pod względem jakości prognoz ryzyka.

Tabela 22. pokazuje wyniki testów Kupca i Christoffersena przy podziale danych treningowych i testowych w proporcji 80/20. Wyniki wskazują, że tylko model eGARCH(1,1) dla portfela konserwatywnego przechodzi oba testy dla poziomu istotności 1% oraz mniej przekonująco dla poziomu 5%, co świadczy o zgodności modelowanej częstości przekroczeń z rzeczywistością. Modele MS-eGARCH nie mogą być formalnie przetestowane, gdy liczba przekroczeń wynosi zero, przez co testy nie są obliczane. Wyjątkiem jest MS-eGARCH z jednym przekroczeniem.

Tabela 22: Wyniki testów Kupca i Christoffersena dla VaR - (trening/test = 0,8)

Model	Poziom	Przekroczenia	Oczekiwane	Kupiec (p)	Christ. (p)
EGARCH (kons)	1%	1	2.00	0.2831	0.5598
EGARCH (kons)	5%	19	12.00	0.0735	0.1834
MS-eGARCH (kons)	1%	0	2.48	–	–
MS-eGARCH (kons)	5%	1	12.00	0.000019	0.0001
EGARCH (agres)	1%	0	2.47	–	–
EGARCH (agres)	5%	3	12.00	0.0011	0.0049
MS-eGARCH (agres)	1%	0	2.47	–	–
MS-eGARCH (agres)	5%	0	12.35	–	–

Z racji, że modele MS-eGARCH poradziły sobie nie najlepiej, zdecydowanie przeszacowując ryzyko oraz że były bardzo kosztowne obliczeniowo, autor postanowił wydłużyć okno do backtestingu i przeprowadzić ponownie estymacje dla modeli eGARCH, aby zobaczyć czy poprawi to ich moc prognostyczną.

Tabela 23., oparta na podziale 60/40, wskazuje na istotne niedoszacowanie ryzyka przez model eGARCH dla portfela konserwatywnego — zarówno test Kupca, jak i Christoffersena jednoznacznie odrzucają poprawność oszacowań ($p < 0.01$). Dla portfela agresywnego sytuacja wygląda lepiej — p-values przekraczają 0.05, co oznacza brak podstaw do odrzucenia hipotezy o poprawności prognozowanego VaR.

Tabela 23: Wyniki testów Kupca i Christoffersena dla VaR - (proporcja trening/test = 0.6)

Model	Poziom	Przekroczenia	Oczekiwane	Kupiec (p)	Christ. (p)
EGARCH (kons)	1%	46	4.00	0.0000	0.0000
EGARCH (kons)	5%	77	24.00	0.0000	0.0000
EGARCH (agres)	1%	7	4.00	0.3781	0.6130
EGARCH (agres)	5%	29	24.00	0.3813	0.1832

Tabela 24: Wyniki testu Acerbiego–Szekelyego dla Expected Shortfall - (trening/test = 0,8)

Model	Poziom	Liczba obs.	Średni Score	p-value
EGARCH (kons)	1%	248	-0.0620	1.000
EGARCH (kons)	5%	248	-0.0088	1.000
MS-eGARCH (kons)	1%	248	-0.1937	1.000
MS-eGARCH (kons)	5%	248	-0.0289	1.000
EGARCH (agres)	1%	247	-0.6932	1.000
EGARCH (agres)	5%	247	-0.1129	1.000

Tabela 25: Wyniki testu Acerbiego–Szekelyego dla Expected Shortfall - (trening/test = 0,6)

Model	Poziom	Liczba obs.	Średni Score	p-value
EGARCH (kons)	1%	496	-0.0508	1.000
EGARCH (kons)	5%	496	-0.0055	1.000
EGARCH (agres)	1%	493	-0.6831	1.000
EGARCH (agres)	5%	493	-0.1050	1.000

Wyniki zaprezentowane w tabelach 24 i 25 wskazują, że żaden z estymowanych modeli nie został odrzucony w teście Acerbiego–Szekelyego, niezależnie od przyjętej proporcji zbioru treninowego do testowego (0.8 lub 0.6). Wszystkie wartości p-value są równe 1, a uzyskane średnie wartości funkcji scoringowej mają znak ujemny.

Należy jednak zaznaczyć, że taka interpretacja wynika ze specyfiki samego testu, którego konstrukcja została opracowana z myślą o wykrywaniu niedoszacowania ryzyka (czyli sytuacji, w których rzeczywiste straty przekraczają prognozowany Expected Shortfall). Ujemne wartości funkcji scoringowej oznaczają natomiast, że prognozowany ES przewyższał faktycznie zrealizowane straty – a więc ryzyko zostało przeszacowane. W takiej sytuacji test nie ma siły, by odrzucić hipotezę zerową, ponieważ jego konstrukcja penalizuje jedynie zbyt konserwatywne prognozy w ograniczonym zakresie.

Zatem, choć formalnie wszystkie modele „przechodzą” test Acerbiego–Szekelyego, uzyskane wyniki należy interpretować ostrożnie. Wysoka zgodność z hipotezą zerową może nie oznaczać trafności predykcji, lecz wynikać z faktu, że modele były nadmiernie ostrożne w estymacji ryzyka.

Wnioski i zakończenie

Celem niniejszej pracy była empiryczna ocena skuteczności modeli typu GARCH oraz Markov-Switching GARCH w prognozowaniu ryzyka dla dwóch portfeli inwestycyjnych o odmiennych charakterystykach: konserwatywnego i agresywnego. W tym celu estymowano odpowiednie modele dla skalarne go szeregu czasowego każdego portfela, a następnie przeprowadzono procedury backtestingu miar ryzyka Value at Risk oraz Expected Shortfall.

Modele zostały najpierw poddane standardowej diagnostyce ekonometrycznej, która wykazała poprawność ich dopasowania do danych w sensie teoretycznym. Celem testów backtestingu było zweryfikowanie, czy deklarowana jakość modeli znajduje odzwierciedlenie w praktyce.

Wyniki empiryczne wskazują jednoznacznie, że modele EGARCH znacząco przewyższają swoje przełącznikowe odpowiedniki pod względem skuteczności prognozowania miar ryzyka. Dla portfela konserwatywnego model EGARCH poprawnie prognozował wartość VaR przy poziomie istotności 1%, przechodząc zarówno test Kupca, jak i Christoffersena. Przy wyższym poziomie istotności (5%) trafność predykcji uległa osłabieniu, jednak model nadal przechodził formalne testy prognostyczne. Można zatem stwierdzić, że model EGARCH jest odpowiedni do prognozowania ryzyka dla portfeli o niskiej zmienności w typowym układzie backtestingu (z podziałem danych 80:20).

W przypadku portfela agresywnego model EGARCH wykazywał jednak tendencję do przeszacowywania ryzyka, co skutkowało zbyt konserwatywnymi prognozami VaR. Zjawisko to można interpretować jako „nadmierną ostrożność” modelu wobec wysokiej zmienności danych — model niejako „przestraszony” dynamiką szeregu generuje zbyt wysokie wartości prognozowanych strat. Wprowadzenie mechanizmu przełączania reżimów w modelach MS-EGARCH dodatkowo spotęgowało to zjawisko. Modele te często nie generowały żadnych przekroczeń progowego poziomu VaR, co uniemożliwiło przeprowadzenie testów backtestingu i wskazuje na skrajną konserwatywność tych konstrukcji. W praktyce modele MS-GARCH w analizowanej postaci nie nadają się do prognozowania VaR, szczególnie w krótszych horyzontach estymacyjnych.

Wydłużenie okna testowego (do 40%, czyli podziału 60:40) zmieniło wyniki: modele EGARCH przestały radzić sobie dobrze w przypadku portfela konserwatywnego, natomiast zaczęły skutecznie przechodzić testy dla portfela agresywnego. Wynik ten sugeruje, że dobór długości okna testowego wpływa na trafność prognoz, przy czym krótsze okna lepiej sprawdzają się w warunkach niskiej zmienności, a dłuższe – w przypadku danych o wyższej dynamice i większej częstotliwości ekstremalnych obserwacji.

Autor przeprowadził również rzadziej spotykaną procedurę backtestingu dla Expected Shortfall, wykorzystując test Acerbiego–Székely’ego (2014). Wyniki tych testów okazały się jednak mało informacyjne. Wszystkie modele wykazywały przeszacowanie ryzyka, co skutkowało ujemnymi wartościami funkcji scoringowej i brakiem podstaw do odrzucenia hipotezy zerowej. Wynika to z konstrukcji testu, który penalizuje wyłącznie niedoszacowanie ryzyka. W efekcie, w obecnym kształcie nie umożliwia on porównywania modeli w przypadku systematycznego przeszacowywania ES. W przyszłych badaniach warto rozważyć zastosowanie testów umożliwiających analizę przeszacowania — przykładem mogą być prace Fisslera i Ziegela (2016, 2019), oparte na metodzie ridge regression i uwzględniające asymetrię w funkcji scoringowej.

Należy również wskazać istotne ograniczenia badania. Po pierwsze, estymacja modeli MS-

GARCH miała charakter dwukrokowy, tj. wartość oczekiwana i wariancja były estymowane niezależnie. Pełna estymacja łączna, uwzględniająca ARMA oraz GARCH w jednym modelu z przełączaniem, mogłaby przynieść inne wyniki, jednak nie jest obecnie możliwa przy użyciu pakietu MSGARCH. Dodatkowo, pakiet ten nie umożliwia estymacji modeli GARCH o rządach innych niż (1,1), co ogranicza elastyczność analizy. Istotnym uproszczeniem było również sprowadzenie danych do postaci skalarne szeregów czasowych, co wykluczyło analizę współzależności pomiędzy funduszami. Rozwinięciem badania mogłoby być zatem wykorzystanie wielowymiarowych modeli GARCH, a także ich odpowiedników z przełączaniem reżimów.

Mimo wspomnianych ograniczeń, uzyskane wyniki są spójne i pozwalają sformułować jednoznaczny wniosek: **hipoteza badawcza o wyższej skuteczności modeli MS-GARCH względem klasycznych modeli GARCH w prognozowaniu ryzyka dla portfeli funduszy zostaje odrzucona**. Modele Markov-Switching GARCH, w obecnej postaci implementacyjnej, nie wykazują przewagi nad klasycznymi modelami GARCH w zakresie szacowania miar ryzyka, takich jak Value at Risk oraz Expected Shortfall, dla portfeli funduszy inwestycyjnych o charakterystyce konserwatywnej i agresywnej.

Literatura

- [1] Acerbi, C., & Székely, B. (2014). Backtesting Expected Shortfall. *Risk*, 27(11), 76–81.
- [2] Alexander, C. (2008). *Market risk analysis. Volume II: Practical financial econometrics*. John Wiley & Sons.
- [3] Ardia, D., Bluteau, K., & Borms, S. (2019). MSGARCH: Markov-switching GARCH models in R. *Journal of Statistical Software*, 91(4), 1–38. <https://doi.org/10.18637/jss.v091.i04>
- [4] Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228. <https://doi.org/10.1111/1467-9965.00068>
- [5] Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862. <https://doi.org/10.2307/2527341>
- [6] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- [7] Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments* (10th ed.). McGraw-Hill Education.
- [8] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.
- [9] Danielsson, J. (2011). *Financial risk forecasting: The theory and practice of forecasting market risk with implementation in R and Matlab*. John Wiley & Sons.
- [10] Dowd, K. (2005). *Measuring market risk* (2nd ed.). John Wiley & Sons.
- [11] Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2014). *Modern portfolio theory and investment analysis* (9th ed.). John Wiley & Sons.
- [12] Enders, W. (2014). *Applied econometric time series* (4th ed.). Wiley.
- [13] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007. <https://doi.org/10.2307/1912773>

- [14] Fabozzi, F. J. (2002). *The theory and practice of investment management*. John Wiley & Sons.
- [15] Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5), 1779–1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>
- [16] Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1), 27–62. [https://doi.org/10.1016/0304-405X\(96\)00875-6](https://doi.org/10.1016/0304-405X(96)00875-6)
- [17] Haas, M., Mittnik, S., & Paoletta, M. S. (2004a). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, 2(4), 493–530. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=821731
- [18] Hamilton, J. D. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica*, 57(2), 357–384. <https://doi.org/10.2307/1912559>
- [19] Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.
- [20] Hull, J. C. (2015). *Risk management and financial institutions* (4th ed.). Wiley Finance.
- [21] Jorion, P. (2007). *Value at Risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.
- [22] Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2), 73–84. <https://doi.org/10.3905/jod.1995.407942>
- [23] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.2307/2975974>
- [24] McAleer, M., & da Veiga, B. (2008). Single-index and portfolio models for forecasting value-at-risk thresholds. *Journal of Forecasting*, 27(3), 217–235.
- [25] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools* (Revised ed.). Princeton University Press.
- [26] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370. <https://doi.org/10.2307/2938260>

- [27] Norris, J. R. (1997). Markov chains. Cambridge University Press.
- [28] Podgórska, M. et al. (2002). Łańcuchy Markowa w teorii i zastosowaniach. Warszawa: Szkoła Główna Handlowa – Oficyna Wydawnicza.
- [29] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance, 19(3), 425–442. <https://doi.org/10.2307/2977928>
- [30] Tsay, R. S. (2005). Analysis of financial time series (2nd ed.). John Wiley & Sons.

Spis tabel

1	Najbardziej agresywne i konserwatywne fundusze w próbie badawczej	27
2	Statystyki opisowe dla portfela konserwatywnego i agresywnego	31
3	Wyniki testów stacjonarności dla logarytmicznych stóp zwrotu	33
4	Wyniki testów diagnostycznych dla logarytmicznych stóp zwrotu	33
5	Porównanie modeli ARMA dla portfela konserwatywnego i agresywnego	36
6	Estymacja parametrów modeli GARCH dla portfela konserwatywnego	38
7	Istotność parametrów estymowanych w modelach GARCH (na podstawie testu t-studenta)	39
8	Testy diagnostyczne modeli GARCH dla portfela konserwatywnego	40
9	Estymacja parametrów modelu MS-eGARCH(1,1) dla portfela konserwatywnego .	41
10	Istotność estymowanych parametrów w modelu MS-eGARCH(1,1) dla portfela konserwatywnego	41
11	Macierz przejścia dla modelu MS-eGARCH(1,1) dla portfela konserwatywnego . .	42
12	Prawdopodobieństwa stacjonarne dla modelu MS-eGARCH(1,1)	42
13	Kryteria informacyjne dla modelu MS-eGARCH(1,1)	42
14	Estymacja parametrów modeli GARCH dla portfela agresywnego	43
15	Istotność parametrów estymowanych w modelach GARCH dla portfela agresywnego (na podstawie testu t-studenta)	43
16	Testy diagnostyczne modeli GARCH dla portfela agresywnego	44
17	Estymacja parametrów modelu MS-eGARCH(1,1) z rozkładem sstd dla portfela agresywnego	45
18	Istotność estymowanych parametrów w modelu MS-eGARCH(1,1) dla portfela agresywnego	45
19	Macierz przejścia dla modelu MS-eGARCH(1,1) – portfel agresywny	45
20	Prawdopodobieństwa stacjonarne dla modelu MS-eGARCH(1,1)	46
21	Kryteria informacyjne dla modelu MS-eGARCH(1,1) – portfel agresywny	46
22	Wyniki testów Kupca i Christoffersena dla VaR - (trening/test = 0,8)	47
23	Wyniki testów Kupca i Christoffersena dla VaR - (proporcja trening/test = 0.6) . . .	47
24	Wyniki testu Acerbiego–Szekelyego dla Expected Shortfall - (trening/test = 0,8) . .	48
25	Wyniki testu Acerbiego–Szekelyego dla Expected Shortfall - (trening/test = 0,6) . .	48

Spis rysunków

1	Ceny jednostek funduszy 2018-2023	28
2	Logarytmiczne stopy zwrotu funduszy 2018-2023	29
3	Porównanie logarytmicznych stóp zwrotu portfeli	30
4	Porównanie zmienności portfeli	30
5	Histogramy stóp zwrotów	31
6	Wykresy kwantyl-kwantyl	32
7	Panel ACF i PACF dla stóp zwrotu	35
8	Panel ACF i PACF dla kwadratów stóp zwrotu	35

Streszczenie

Celem niniejszej pracy była empiryczna ocena efektywności modeli klasy GARCH oraz ich rozszerzeń w porównaniu do modeli Markov-Switching GARCH w prognozowaniu miar ryzyka, takich jak Value at Risk oraz Expected Shortfall, dla dwóch zagregowanych portfeli funduszy inwestycyjnych: konserwatywnego i agresywnego. Analiza została przeprowadzona na danych z polskiego rynku finansowego, z wykorzystaniem procedury rolling window oraz symulacji Monte Carlo.

Pomimo teoretycznych przesłanek sugerujących wyższość modeli MS-eGARCH wynikającą z możliwości uwzględniania zmian reżimów zmienności, uzyskane wyniki wskazują, że to klasyczne modele eGARCH – w szczególności eGARCH(1,2) – zapewniły lepsze dopasowanie do danych empirycznych oraz trafniejsze prognozy ryzyka. Testy backtestingu potwierdziły zgodność prognozowanych wartości VaR z rzeczywistymi realizacjami.

Z kolei zastosowanie testu Acerbiego–Székely’ego nie pozwoliło na jednoznaczną ocenę skuteczności prognozowania miary Expected Shortfall, co wynika z konstrukcji testu skoncentrowanej na niedoszacowaniu ryzyka, podczas gdy analizowane modele wykazywały raczej tendencję do jego przeszacowywania. Uzyskane wyniki sugerują, że zarówno modele GARCH, jak i ich przełącznikowe odpowiedniki mają wrodzoną skłonność do konserwatywnej oceny ryzyka.

Przeprowadzona analiza podkreśla znaczenie empirycznej weryfikacji stosowanych modeli oraz konieczność ostrożnego wdrażania zaawansowanych narzędzi prognostycznych w praktyce zarządzania ryzykiem, zwłaszcza w kontekście funduszy inwestycyjnych funkcjonujących w warunkach zmiennej koniunktury rynkowej.