

# SGH

**Szkoła Główna  
Handlowa  
w Warszawie**

## **Metody ekonometryczne - analiza zwrotów akcji PKN Orlen S.A**

Opracowanie:

Marek Polit 121564

Dawid Dmitroca 129270

22 maja 2025

# **Spis treści**

<b>1</b>	<b>Wprowadzenie</b>	<b>3</b>
<b>2</b>	<b>Pozyskiwanie danych i konstrukcja zmiennych</b>	<b>3</b>
<b>3</b>	<b>Przegląd zmiennych</b>	<b>5</b>
<b>4</b>	<b>Testy i Transformacje zmiennych</b>	<b>8</b>
<b>5</b>	<b>Budowanie modeli</b>	<b>11</b>
<b>6</b>	<b>Testy na ostatecznym modelu</b>	<b>14</b>
<b>7</b>	<b>Interpretacja zmiennych ostatecznego modelu</b>	<b>17</b>

# 1 Wprowadzenie

Analiza zwrotów z akcji jest jednym z podstawowych narzędzi wykorzystywanych w ocenie opłacalności inwestycji na rynku kapitałowym. Obserwując, jak zmienia się cena akcji w czasie, inwestorzy oraz analitycy finansowi mogą lepiej zrozumieć, jakie czynniki wpływają na ich wartość i jakie ryzyko wiąże się z daną spółką. Szczególnie interesującym przypadkiem jest PKN Orlen, jedna z największych firm w Polsce i kluczowy podmiot w sektorze paliwowym. Ze względu na swoją skalę działalności oraz powiązania z globalnymi rynkami ropy naftowej i gazu, akcje tej spółki są podatne na wpływ wielu zmiennych ekonomicznych.

Celem niniejszej analizy jest zbadanie, w jaki sposób różne czynniki — takie jak zmiany kursów walut, ceny surowców energetycznych czy sytuacja na rynku krajowym — wpływają na miesięczne stopy zwrotu z akcji PKN Orlen. W tym celu wykorzystane zostaną narzędzia ekonometrii, w szczególności modele liniowe oraz testy diagnostyczne. Analiza pozwoli ocenić, które zmienne mają istotny wpływ na zwroty z akcji, a także sprawdzić jakość dopasowania modelu do danych.

Dzięki temu możliwe będzie lepsze zrozumienie mechanizmów kształtujących notowania PKN Orlen oraz ocena, czy i w jakim stopniu można przewidywać ich zmiany na podstawie dostępnych danych ekonomicznych.

## 2 Pozyskiwanie danych i konstrukcja zmiennych

W analizie wykorzystano dane miesięczne od 2000 roku do 2024 roku. Zmiennymi objaśniającymi były zarówno wskaźniki makroekonomiczne, jak i rynkowe oraz sektorowe. Wszystkie dane zostały przekształcone do formatu miesięcznego, połączone w jedną ramkę danych oraz przygotowane do dalszej obróbki ekonometrycznej. Poniżej przedstawiono zestaw zmiennych użytych w modelu:

- *PKN\_WA\_Close*<sup>1</sup>

Zmienna objaśniana: cena zamknięcia akcji PKN Orlen na koniec każdego miesiąca.

- *Brent\_USD*<sup>2</sup>

Średniomiesięczna cena surowca Brent w USD za baryłkę.

Uzasadnienie: Cena Brent może obrazować generalne zmiany cen ropy naftowej, przy założeniu konkurencyjności cen

---

<sup>1</sup><https://finance.yahoo.com/quote/PKN.WA/history?p=PKN.WA>

<sup>2</sup><https://fred.stlouisfed.org/series/POILBREUSDM>

- *Gas\_USD*<sup>3</sup>

Średniomiesięczna cena gazu ziemnego (Henry Hub) w USD.

Uzasadnienie: Cena gazu ziemnego (podobnie) może obrazować generalne zmiany cen na rynku surowców gazowniczych (stanowi istotny komponent kosztów energetycznych i petrochemicznych)

- *EURPLN*<sup>4</sup>

Średniomiesięczny kurs wymiany euro względem złotego.

Uzasadnienie: W przypadku eksportu paliwa na rynki europejskie warto uwzględnić zmiany kursów EUR/PLN jako czynnika wpływającego na zmiany cen akcji

- *USDPLN*<sup>5</sup>

Średniomiesięczny kurs wymiany dolara amerykańskiego względem złotego.

Uzasadnienie: z drugiej strony, pozyskanie surowców z różnorodnych rynków często nominowane jest w dolarach amerykańskich, dlatego uwzględniono ten czynnik w analizie

- *WIG\_Close*<sup>6</sup>

Wartość indeksu WIG na koniec miesiąca.

Uzasadnienie: indeks giełdowy odzwierciedla ogólny sentyment inwestorów na polskim rynku oraz generalną kondycję rynku

- *ref, lom, red*<sup>7</sup>

Podstawowe stopy procentowe Banku Centralnego (referencyjna, lombardowa, redyskontowa)

Uzasadnienie: stopy procentowa NBP odzwierciedlają ogólny kierunek i intensywność polityki monetarnej, która wpływa na wszystkie kosztowo-finansowe aspekty funkcjonowania przedsiębiorstw oraz na zachowania inwestorów.

- *IP\_manuf*<sup>8</sup>

Indeks produkcji przemysłowej w przetwórstwie (Indeks 2015 = 100).

Uzasadnienie: poziom produkcji w przetwórstwie wpływa na popyt na produkty rafinerijne i petrochemiczne – jest istotnym wskaźnikiem koniunktury sektorowej.

---

<sup>3</sup><https://fred.stlouisfed.org/series/MHHNGSP>

<sup>4</sup><https://stoq.pl/q/?s=eurpln>

<sup>5</sup><https://stoq.pl/q/?s=usdpln>

<sup>6</sup><https://stoq.pl/q/?s=wig>

<sup>7</sup>[https://www.nbp.pl/home.aspx?f=/dzienne/stopy\\_archiwum.htm](https://www.nbp.pl/home.aspx?f=/dzienne/stopy_archiwum.htm),

<sup>8</sup>[https://ec.europa.eu/eurostat/databrowser/view/sts\\_inpr\\_m/default/table](https://ec.europa.eu/eurostat/databrowser/view/sts_inpr_m/default/table)

- *Unemp*<sup>9</sup>

Miesięczna stopa bezrobocia (bez korekty sezonowej).

Uzasadnienie: ogólny miernik warunków makroekonomicznych i siły nabywczej konsumentów

- *CPI\_index*<sup>10</sup>

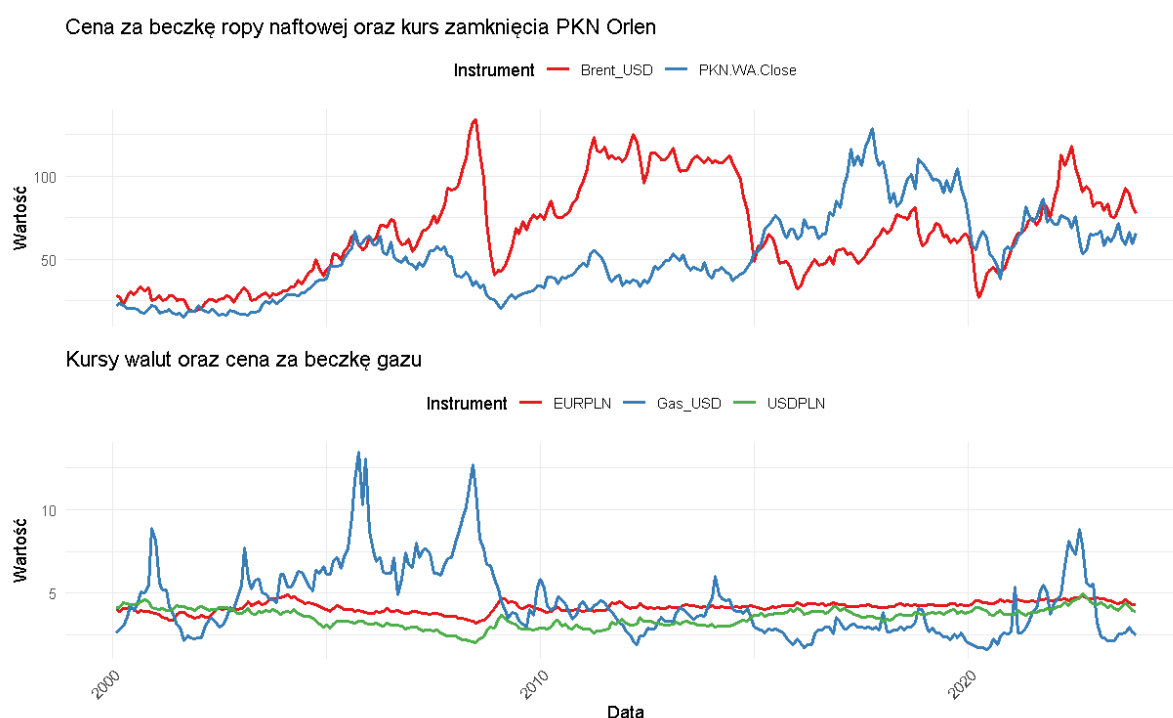
Indeks cen towarów i usług konsumpcyjnych (Indeks 2015 = 100).

Uzasadnienie: zmiany inflacji wpływają na koszty operacyjne i realną siłę nabywczą; uwzględnienie krótkookresowych odchyleń CPI pozwala wychwycić wpływ nagłych skoków cen.

### 3 Przegląd zmiennych

#### Wizualizacje zmiennych

Przegląd zmiennych warto rozpocząć od zwrócenia uwagi na zmienną objaśnianą - cenę akcji PKN Orlen S.A



Rysunek 1: Wykres zmian poszczególnych zmiennych

Źródło: Opracowanie własne

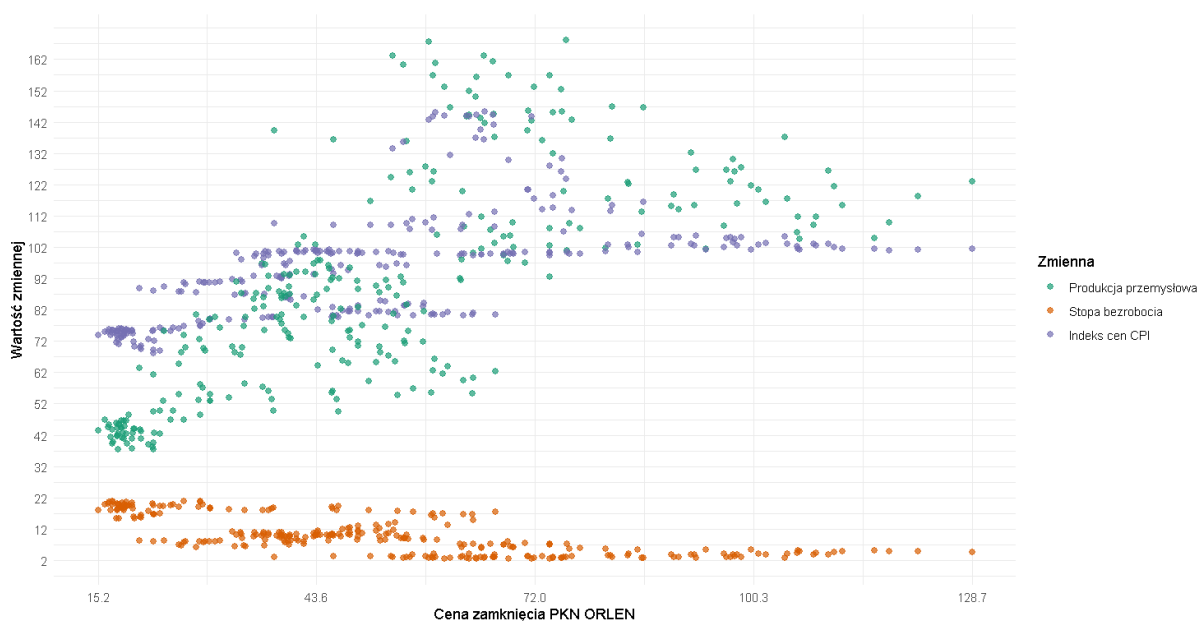
<sup>9</sup>[https://ec.europa.eu/eurostat/databrowser/view/une\\_rt\\_m/default/table](https://ec.europa.eu/eurostat/databrowser/view/une_rt_m/default/table)

<sup>10</sup>[https://ec.europa.eu/eurostat/databrowser/view/prc\\_hicp\\_midx/default/table](https://ec.europa.eu/eurostat/databrowser/view/prc_hicp_midx/default/table)

Na Rysunku 1 w panelu górnym pokazano miesięczną cenę ropy Brent (linia czerwona) oraz kurs zamknięcia akcji PKN Orlen (linia niebieska). W panelu dolnym uwzględniono kursy EUR/PLN (czerwona), USD/PLN (zielona) oraz cenę gazu Henry Hub (niebieska). Obserwowalne są korelacje zmian pomiędzy zmiennymi objaśniającymi w modelu, dlatego dalsza analiza uwzględni również korelacje pomiędzy zmiennymi. Ponadto, Wykres wyraźnie wskazuje, że cena gazu ziemnego cechuje się dużą dynamiką zmian cen, co sugeruje dodanie nowych zmiennych egzogenicznych, które "obsłużą" dużą wariancję cen w przypadku tych okresów. Spośród tych właśnie okresów, wyróżniono cztery główne epizody gwałtownych wahań:

- *D\_Crisis2008* (styczeń 2008–grudzień 2009),
- *D\_Shock2014\_16* (lipiec 2014–czerwiec 2016),
- *D\_COVID2020* (marzec 2020–grudzień 2020),
- *D\_War2022* (luty 2022–grudzień 2022).

W następnej kolejności sprawdzono zmienne makroekonomiczne takie jak wskaźnik *produkcji przemysłowej*, *stopa bezrobocia* oraz *Consumer Price Index*



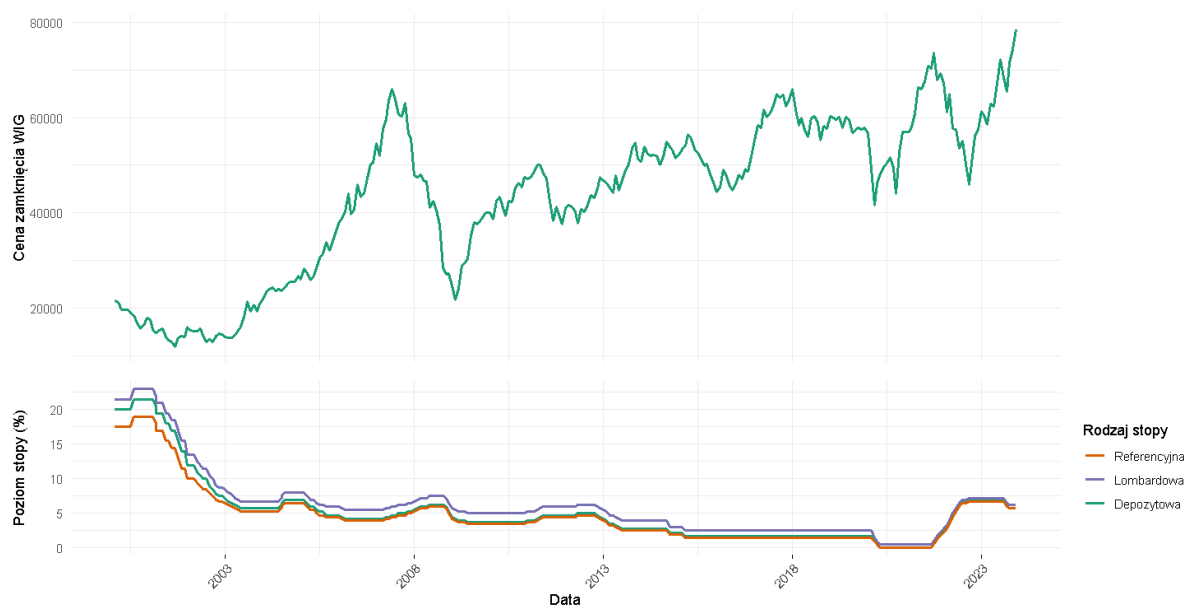
Rysunek 2: Wykres zmian poszczególnych zmiennych

Źródło: Opracowanie własne

Wykres jasno wskazuje na to, że stopa bezrobocia nie jest czynnikiem różnicującym cenę zamknięcia akcji PKN Orlen S.A. Z drugiej strony CPI oraz produkcja przemysłowa okazały się istotnymi zmiennymi, które mogą różnicować cenę akcji. Ponadto, warto również zwrócić uwagę na rozrzut danych wraz ze zwiększaniem się ceny akcji zamknięcia, który wyraźnie wskazuje heteroskedastyczność danych w przypadku obu pozostałych zmiennych. Implikuje to

konieczne sprawdzenie tej przesłanki po zbudowaniu modelu ekonometrycznego by zobaczyć, czy składnik losowy również będzie heteroskedastyczny.

Ostatecznie stworzono panel ostatnich zmiennych, które określają ogólną kondycję rynkową (Rysunek 3). Cena zamknięcia WIG, która również jest bardzo zróżnicowana, wykazuje ewidentną tendencję wzrostową. Ponadto, podobne okresy już dodane jako zmienne binarne wcześniej, cechują się dużym zróżnicowaniem wartości. W przypadku stóp procentowych, postanowiono o pozostawieniu tylko stopy referencyjnej, gdyż jest to naistotniejsza stopa w kontekście rozwiązywanego problemu, a reszta stóp zachowuje się proporcjonalnie do niej. Ewidencja widoczny jest znaczny spadek wysokości stóp procentowych, które są do siebie proporcjonalne, w okresie pomiędzy latami 2000-2003. W kolejnych latach, stopy utrzymywały podobny poziom, aż do okresu Wojny na Ukrainie (rok 2022), kiedy to stopy procentowe wróciły do poziomu sprzed 20 lat.



Rysunek 3: Wykres zmian poszczególnych zmiennych

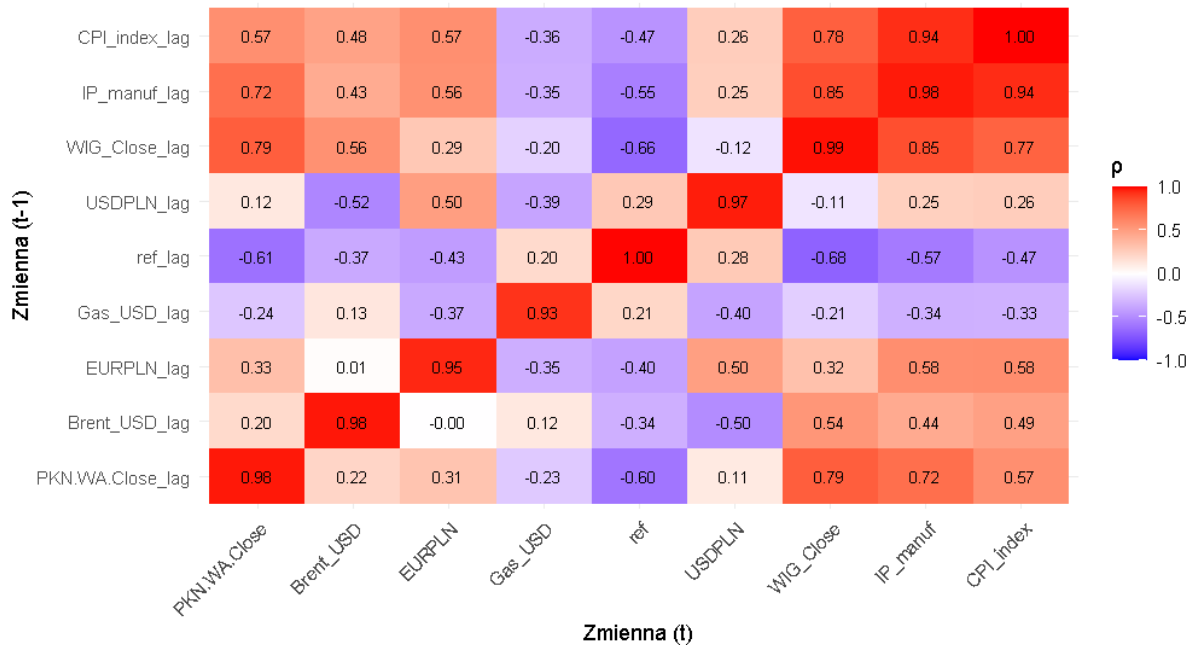
Źródło: Opracowanie własne

## Macierz autokorelacji

W celu wstępnego zbadania potencjalnych zależności czasowych pomiędzy zmiennymi, przeanalizowano macierz korelacji Pearsona, w której zestawiono wartości każdej zmiennej z okresu  $t$  z odpowiadającą jej wartością w okresie  $t - 1$ . Wyniki przedstawiono na Rysunku 4.

Na podstawie analizy można zauważyć, że zmienne są dość silnie skorelowane ze sobą, co dodatkowo sugeruje wykonanie odpowiednich testów diagnostycznych w późniejszych etapach analizy. Przykładowo zmienna objaśniana *PKN.WA.Close* wykazuje bardzo wysoką korelację z własnymi wartościami opóźnionymi ( $\rho = 0,98$ ), co wskazuje na dużą bezwładność notowań giełdowych spółki w czasie. Podobne zależności obserwuje się w przypadku indeksu

$WIG\_Close$  ( $\rho = 0,99$ ) oraz zmiennych makroekonomicznych, takich jak  $ref$  ( $\rho = 1,00$ ) i  $IP\_manuf$  ( $\rho = 0,98$ ). Tego rodzaju analiza dostarcza przesłanek do dalszego testowania obecności autokorelacji oraz ewentualnego doboru odpowiednich transformacji zmiennych w modelach regresyjnych, np. przez uwzględnienie opóźnień lub różnicowanie.



Rysunek 4: Macierz korelacji zmiennych

Źródło: Opracowanie własne

## 4 Testy i Transformacje zmiennych

### Transformacja

W celu poprawy spełnienia założeń klasycznego modelu regresji liniowej, takich jak liniowość relacji, stałość wariancji oraz normalność rozkładu reszt, zastosowano transformację Boxa-Coxa dla wszystkich zmiennych ciągłych. Metoda ta pozwala dobrać odpowiednią funkcję potęgową (zależną od parametru  $\lambda$ ), która stabilizuje wariancję i może przybliżyć rozkład zmiennej do normalnego:

Transformacja Boxa-Coxa dla zmiennej  $Y$  definiowana jest następująco:

$$Y^{(\lambda)} = \begin{cases} \frac{Y^\lambda - 1}{\lambda}, & \text{dla } \lambda \neq 0 \\ \log(Y), & \text{dla } \lambda = 0 \end{cases}$$

Parametr  $\lambda$  został oszacowany osobno dla każdej zmiennej, a jego optymalna wartość została przedstawiona w Tabeli 1. Zmiennym, dla których  $\lambda$  było bliskie zeru (tj.  $|\lambda| < 0,30$ ), przypisano transformację logarytmiczną – odpowiadającą klasycznemu przypadkowi Boxa-



Coxa przy  $\lambda = 0$ . Ze względu na przejrzystość analizy, nie zastosowano innego rodzaju transformacji, tak aby móc jasno interpretować wpływ poszczególnych zmiennych w dalszej części analizy. Ostatecznie transformacji poddano 4 zmienne, które są zaprezentowane w tabeli:

Tabela 1: Optymalne wartości parametru  $\lambda$  w transformacji Boxa-Coxa oraz zastosowanie transformacji

Zmienna	$\lambda$ optymalne	Zastosowana transformacja	Opis ekonomiczny
Brent_USD	-0,18	$\log(\text{Brent\_USD})$	Cena ropy Brent
Gas_USD	-0,26	$\log(\text{Gas\_USD})$	Cena gazu ziemnego
PKN.WA.Close	0,08	$\log(\text{PKN.WA.Close})$	Kurs akcji PKN Orlen
WIG_Close	0,46	brak	Indeks giełdowy WIG
EURPLN	1,62	brak	Kurs EUR/PLN
USDPLN	1,64	brak	Kurs USD/PLN
IP_manuf	-0,02	$\log(\text{IP\_manuf})$	Produkcja przemysłowa
CPI_index	-1,00	brak	Indeks cen konsumpcyjnych
ref	0,66	brak	Stopa referencyjna

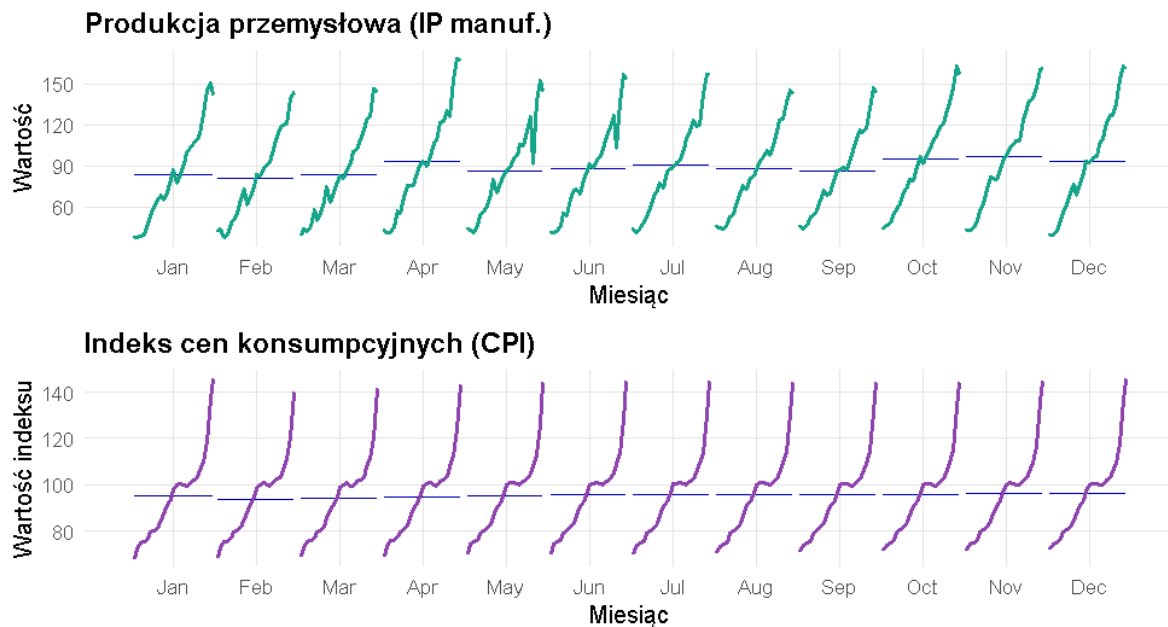
## Stacjonarność

Tabela 2: Liczba wymaganych różnic do osiągnięcia stacjonarności

Zmienna	Różnicowanie $d$	Różnicowanie sezonowe $D$
PKN.WA.Close	1	0
Brent_USD	1	0
EURPLN	1	0
Gas_USD	1	0
ref	2	0
USDPLN	1	0
WIG_Close	1	0
IP_manuf	1	1
CPI_index	2	0

Przed przystąpieniem do modelowania, zgodnie z klasycznymi założeniami analizy szeregów czasowych oraz ekonometrii czasowej, zweryfikowano stacjonarność wszystkich zmiennych za pomocą dwóch komplementarnych testów: testu Dickeya-Fullera (ADF) oraz testu

KPSS. Wyniki testów stacjonarności wskazały, że wszystkie analizowane zmienne w poziomie są niestacjonarne. W związku z tym określono liczbę wymaganych różnic (zwykłych i sezonowych), niezbędnych do uzyskania stacjonarności każdej ze zmiennych. Podsumowanie przedstawiono w Tabeli 2.



Rysunek 5: Wykres sezonowości zmiennych

Źródło: Opracowanie własne

Dla większości zmiennych wystarczyło pierwsze różnicowanie ( $d = 1$ ), natomiast w przypadku zmiennej *ref* oraz *CPI\_index* konieczne było drugie różnicowanie ( $d = 2$ ). Zmienna *IP\_manuf* została dodatkowo poddana sezonowemu różnicowaniu z okresem 12 miesięcy ( $D = 1$ ), co wynikało z występowania wyraźnej sezonowości w danych. Ponieważ, pewne testy wskazywały również na niestacjonarność zmiennej *Consumer Price Index*, postanowiono o zweryfikowaniu tego wizualnie. Wyniki wskazują na błędy tamtego testu, gdyż różnice (w szczególności przy uwzględnieniu skali), nie są różnicowane znacznie ze względu na miesiąc co przedstawiono na rysunku 5.

## 5 Budowanie modeli

### Pełny model ekonometryczny

W celu wybrania odpowiedniego modelu ekonometrycznego regresji liniowej, po transformacji i różnicowaniu zmiennych, rozpoczęto proces oszacowywania logarytmicznych zwrotów akcji PKN ORLEN S.A. Pełna struktura modelu prezentuje się w sposób następujący:

$$\begin{aligned}\Delta \ln(\text{PKN}_t) = & \beta_0 + \beta_1 \Delta \text{EURPLN}_t + \beta_2 \Delta \text{USDPLN}_t + \beta_3 \Delta \text{WIG}_t \\ & + \beta_4 \Delta \ln(\text{Brent}_t) + \beta_5 \Delta \ln(\text{Gas}_t) + \beta_6 \Delta \text{ref}_t \\ & + \beta_7 \Delta \ln(\text{IP}_t) + \beta_8 \Delta \text{CPI}_t + \gamma_1 D_{\text{Crisis2008},t} \\ & + \gamma_2 D_{\text{Shock2014\_16},t} + \gamma_3 D_{\text{COVID2020},t} + \gamma_4 D_{\text{War2022},t} + \varepsilon_t,\end{aligned}\tag{1}$$

Tabela 3: Estymacja pełnego modelu

Zmienna	Estymator	Błąd std.	<i>t</i>	<i>p</i> -value
(Intercept)	-0,003465	0,004443	-0,780	0,4362
$\Delta \text{ EUR/PLN}$	-0,1154	0,05753	-2,005	0,0460 *
$\Delta \text{ USD/PLN}$	0,1058	0,04500	2,352	0,0194 *
$\Delta \text{ WIG}$	0,000026	0,000002	14,336	<2e-16 ***
$\Delta \ln(\text{Brent})$	-0,02985	0,04584	-0,651	0,5155
$\Delta \ln(\text{Gas})$	0,04681	0,02693	1,738	0,0833 ·
$\Delta \text{ ref}$	-0,02579	0,01279	-2,017	0,0447 *
$\Delta \ln(\text{IP})$	0,01600	0,06122	0,261	0,7940
$\Delta \text{ CPI}$	-0,01246	0,008543	-1,459	0,1458
$D_{\text{Crisis2008}}$	0,003418	0,01369	0,250	0,8030
$D_{\text{Shock2014\_16}}$	0,03012	0,01379	2,184	0,0298 *
$D_{\text{COVID2020}}$	-0,01567	0,02055	-0,763	0,4462
$D_{\text{War2022}}$	0,01217	0,01965	0,619	0,5364
Residual Std. Error: 0,06298 on 263 DF				
Multiple R <sup>2</sup> : 0,5298, Adjusted R <sup>2</sup> : 0,5084				
F-statistic: 24,7 on 12 and 263 DF (p < 2,2e-16)				
AIC: -728,29				

Uwagi: \*\*\* $p < 0,001$ , \*\* $p < 0,01$ , \* $p < 0,05$ , · $p < 0,1$ .

Oszacowania pełnego modelu ekonometrycznego wskazują na istotną statystyczność 5 zmiennych na poziomie 0,05 oraz 6 zmiennych, jeżeli za próg istotności statystycznej przyjmujemy 0,1. W następnej kolejności zdecydowano o usunięciu zmiennych nieistotnych, tak aby móc sprawdzić, czy redukcja modelu może wpłynąć na poprawę oszacowań oraz dopasowania.

## Zredukowany model ekonometryczny

Na podstawie wyników pełnego modelu usunięto zmienne nieistotne statystycznie (na poziomie 0,1), pozostawiając te o istotnym wpływie na logarytmiczne stopy zwrotu z akcji PKN Orlen. Ostateczny, zredukowany model ma postać:

$$\Delta \ln(\text{PKN}_t) = \beta_0 + \beta_1 \Delta \text{USDPLN}_t + \beta_2 \Delta \text{EURPLN}_t + \beta_3 \Delta \text{WIG}_t + \beta_4 \Delta \text{ref}_t + \gamma D_{\text{Shock2014\_16},t} + \delta \Delta \ln(\text{Gas}_t) + \varepsilon_t, \quad (2)$$

gdzie oznaczenia są takie same jak we wzorze (1).

Tabela 4: Estymacja zredukowanego modelu

Zmienna	Estymator	Błąd std.	<i>t</i>	<i>p</i> -value
(Intercept)	-0,003248	0,003986	-0,815	0,4159
$\Delta \text{USD/PLN}$	0,1079	0,04449	2,425	0,0159 *
$\Delta \text{EUR/PLN}$	-0,1200	0,05690	-2,110	0,0358 *
$\Delta \text{WIG}$	0,000025	0,000002	14,506	<2e-16 ***
$\Delta \text{ref}$	-0,02824	0,01237	-2,282	0,0232 *
$D_{\text{Shock2014\_16}}$	0,03060	0,01351	2,266	0,0243 *
$\Delta \ln(\text{Gas})$	0,04503	0,02609	1,726	0,0855 .
Residual Std. Error: 0,06276 on 269 DF				
Multiple R <sup>2</sup> : 0,5225, Adjusted R <sup>2</sup> : 0,5119				
F-statistic: 49,06 on 6 and 269 DF (p < 2,2e-16)				
AIC: -736,02				

Uwagi: \*\*\* $p < 0,001$ , \*\* $p < 0,01$ , \* $p < 0,05$ , . $p < 0,1$ .

Model oszacowuje istotność statystyczną wszystkich zmiennych z wyjątkiem wyrazu wolnego. Model cechuje się lepszym dopasowaniem *Adjusted R<sup>2</sup>*, co sugeruje, że redukcja zmiennych przyczyniła się do lepszej interpretowalności parametrów. Poprawiło się również kryterium *AIC*, gdzie niższa wartość oznacza lepsze dopasowanie.

## Zoptymalizowany model pod względem kryterium AIC

Ostateczny model uzyskany przez procedurę krokową minimalizującą wartość *AIC* ma postać:

$$\begin{aligned} \Delta \ln(\text{PKN}_t) = & \beta_0 + \beta_1 \Delta \text{EURPLN}_t + \beta_2 \Delta \text{USDPLN}_t + \beta_3 \Delta \text{WIG}_t \\ & + \beta_4 \Delta \ln(\text{Gas}_t) + \beta_5 \Delta \text{ref}_t + \beta_6 \Delta \text{CPI}_t + \gamma D_{\text{Shock2014\_16},t} + \varepsilon_t, \end{aligned} \quad (3)$$

A tak prezentowała się tabela oszacowań parametrów:

Tabela 5: Estymacja modelu AIC-optimalnego

Zmienna	Estymator	Błąd std.	$t$	$p$ -value
(Intercept)	-0,003328	0,003974	-0,838	0,4030
$\Delta$ EUR/PLN	-0,1179	0,05673	-2,079	0,0386 *
$\Delta$ USD/PLN	0,1080	0,04435	2,434	0,0156 *
$\Delta$ WIG	0,000025	0,000002	14,642	<2e-16 ***
$\Delta \ln(\text{Gas})$	0,04228	0,02607	1,622	0,1060
$\Delta$ ref	-0,02748	0,01234	-2,226	0,0268 *
$\Delta$ CPI	-0,01358	0,00830	-1,637	0,1029
$D_{\text{Shock2014\_16}}$	0,03074	0,01347	2,283	0,0232 *

Residual Std. Error: 0,06256 (268 DF)  
Multiple  $R^2 = 0,5272$ , Adjusted  $R^2 = 0,5149$   
F-statistic: 42,70 on 7 and 268 DF ( $p < 2,2\text{e-}16$ )  
AIC: -736,76

Uwagi: \*\*\* $p < 0,001$ , \* $p < 0,05$ .

W modelu AIC-optimalnym pozostała tylko jedna zmienna binarna oznaczająca epizod rynkowy — *Shock2014\_16*. Pozostałe trzy epizody, mimo wcześniejszego uwzględnienia w modelu pełnym (*Crisis2008*, *COVID2020*, *War2022*), zostały usunięte przez procedurę stopniową. Oznacza to, że nie poprawiały one wartości kryterium AIC i nie wykazywały istotnego wpływu na zmienność zwrotów z akcji PKN Orlen w obecności pozostałych zmiennych objaśniających. Z punktu widzenia konstrukcji modelu, ich odrzucenie jest zgodne z celem minimalizacji błędu ogólnego i sugeruje, że po uwzględnieniu zmiennych ciągłych (takich jak ceny surowców, kursy walut czy indeksy), ich informacyjna wartość była marginalna.

## Wnioski

Poniższa tabela zestawia kluczowe miary jakości dopasowania trzech modeli: pełnego, zredukowanego (usunięte zmienne nieistotne przy  $p < 0,1$ ) oraz optymalnego pod względem AIC.

Tabela 6: Porównanie modeli: pełny vs. zredukowany vs. AIC-optimalny

Miara	Pełny model	Zredukowany model	AIC-optimalny model
Liczba parametrów (łącznie z interceptem)	13	7	8
Adjusted $R^2$	0,5084	0,5119	0,5149
AIC	-728,29	-736,02	-736,76

Zestawienie wskazuje, że:

- Model zredukowany (7 regresorów) już poprawia skorygowany współczynnik determinacji w porównaniu z pełnym (0,5119 vs. 0,5084) i jednocześnie obniża wartość AIC o

$\approx 7,7$  pkt.

- Dodanie zmiennej CPI do modelu AIC-optimalnego przyniosło dalszą, choć niewielką, poprawę obu wskaźników (Adjusted  $R^2 = 0,5149$ , AIC =  $-736,76$ ), kosztem jednego dodatkowego parametru.
- Najlepszym kompromisem między prostotą a jakością dopasowania jest więc model AIC-optimalny: dysponuje jednym parametrem więcej niż wersja zredukowana, ale osiąga najwyższy Adjusted  $R^2$  i najniższe AIC.

W świetle powyższych wyników rekomendujemy przyjęcie modelu (3) jako ostatecznego narzędzia do analizy krótkoterminowych zwrotów z akcji PKN Orlen.

## 6 Testy na ostatecznym modelu

### Autokorelacja reszt

Aby sprawdzić, czy reszty modelu nie wykazują autokorelacji, przeprowadzono następujące testy:

- **Durbin–Watson:**

$$DW = 2.0106, \quad p = 0.5289$$

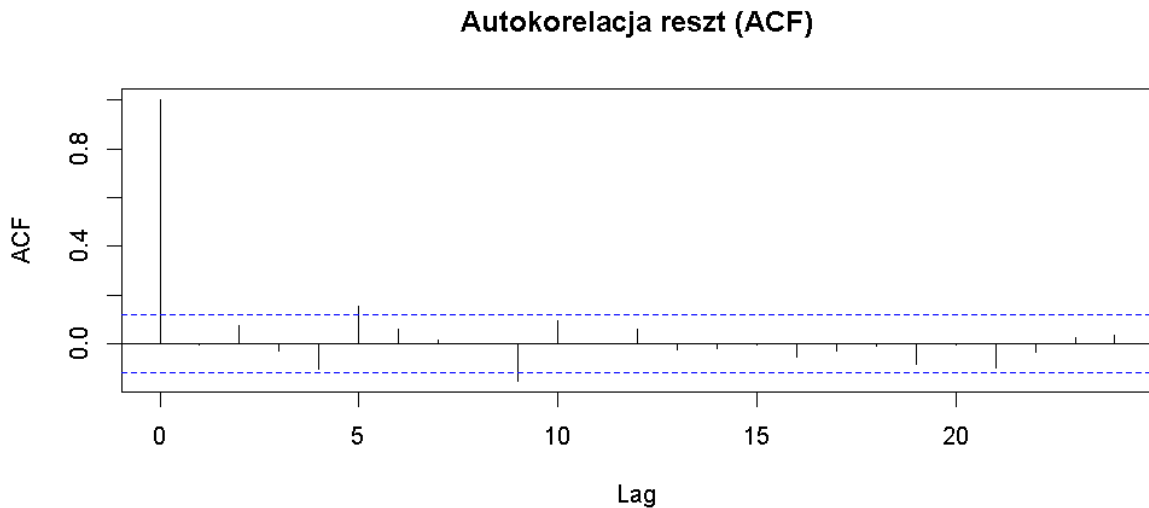
Statystyka bliska 2 oraz wysokie  $p$ -value wskazują na brak autokorelacji pierwszego rzędu.

- **Breusch–Godfrey** (do rzędu 4):

$$LM = 5.2692, \quad df = 4, \quad p = 0.2608$$

Brak istotnej autokorelacji reszt do czwartego rzędu.

- **Wykresy ACF i PACF reszt** (Rysunek 6):
  - Żadne odległe opóźnienie (lag) poza zerowym nie przekracza granicy istotności.
  - Nie obserwuje się uporczywych wzorców ani malejących „ogonów” sugerujących autokorelację.



Rysunek 6: Autokorelacja reszt

Źródło: Opracowanie własne

Podsumowując, zarówno testy statystyczne, jak i diagnostyka graficzna nie ujawniają obecności autokorelacji w resztach, co potwierdza spełnienie jednego z kluczowych założeń modelu liniowego.

## Normalność rozkładu reszt

Aby zweryfikować założenie normalności rozkładu reszt, przeprowadzono:

- **Test Jarque–Bera:**

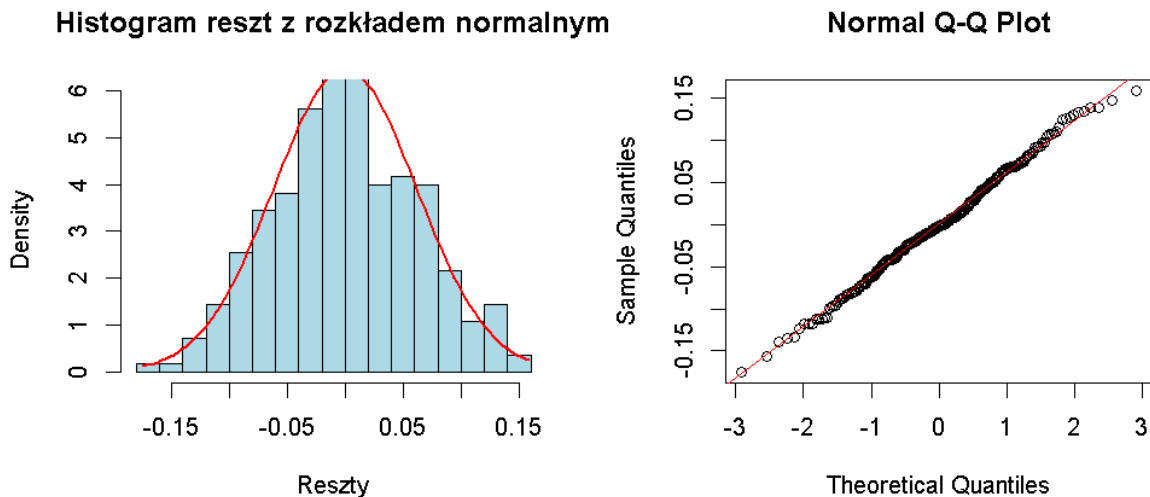
$$JB = 0.2352, \quad df = 2, \quad p = 0.889$$

Wysokie  $p$ -value oznacza brak podstaw do odrzucenia hipotezy o normalności reszt.

- **Diagnostyka graficzna:**

1. **Histogram reszt** (Rysunek 7) – rozkład zbliżony do symetrycznego dzwonu.
2. **Wykres Q–Q** (Rysunek 7) – większość punktów leży blisko linii przekątnej, co sugeruje zgodność z rozkładem normalnym.

Wnioski: zarówno test statystyczny, jak i analiza wykresów potwierdzają, że reszty modelu spełniają założenie normalności.



Rysunek 7: Histogram oraz QQ Plot

Źródło: Opracowanie własne

## Współliniowość zmiennych objaśniających

Dla oceny współliniowości obliczono wskaźniki VIF (Variance Inflation Factor) dla każdej zmiennej objaśniającej:

$$\begin{aligned} \text{VIF}_{d\_EURPLN} &= 2.147, \\ \text{VIF}_{d\_USDPLN} &= 2.310, \\ \text{VIF}_{d\_WIG\_Close} &= 1.418, \\ \text{VIF}_{d\_Gas\_USD\_log} &= 1.027, \\ \text{VIF}_{d\_ref} &= 1.008, \\ \text{VIF}_{d\_CPI\_index} &= 1.019, \\ \text{VIF}_{D\_Shock2014\_16} &= 1.015. \end{aligned}$$

Wszystkie uzyskane wartości VIF są znacznie poniżej powszechnie przyjmowanego progu (Próg krytyczny = 5), co wskazuje na brak istotnej współliniowości między zmiennymi objaśniającymi. Dzięki temu model nie wymaga poprawy specyfikacji zmiennych, ich eliminacji lub też kolejnych transformacji z powodu nadmiernej korelacji.

## Heteroskedastyczność

Zgodnie z twierdzeniem Gaussa–Markowa, przy homoskedastyczności reszt estymatory OLS są *BLUE* (Best Linear Unbiased Estimators), czyli nieobciążone i o najmniejszej wariancji w klasie estymatorów liniowych. W razie wystąpienia heteroskedastyczności estymatory pozostają nieobciążone, ale tracą efektywność.



Aby sprawdzić, czy w zoptymalizowanym modelu nie zachodzi heteroskedastyczność, wykonano następujące testy:

- **Breusch–Pagana:**

$$BP = 6.2792, \quad df = 7, \quad p = 0.5076$$

- **Goldfeld–Quandt** (dzielenie danych na 20% obserwacji o najniższych i najwyższych wartościach dopasowanych):

$$GQ = 1.2324, \quad df_1 = 103, \quad df_2 = 102, \quad p = 0.1460$$

- **White's test** (uproszczona wersja testu Breuscha–Pagana, wykorzystująca wartość dopasowaną  $\hat{y}$  i jej kwadrat):

$$BP = 4.4839, \quad df = 2, \quad p = 0.1063$$

*Uwaga: zastosowano uproszczony wariant testu White'a, oparty jedynie na funkcji dopasowanej i jej kwadracie. Stąd liczba stopni swobody wynosi 2, a nie tyle, ile wskazywałby pełny test z uwzględnieniem wszystkich regresorów, ich kwadratów oraz interakcji.*

- **NCV Test** (Non-constant Variance Score Test):

$$\chi^2 = 0.09995, \quad df = 1, \quad p = 0.7519$$

We wszystkich testach  $p > 0.05$ , co oznacza brak podstaw do odrzucenia hipotezy o homoskedastyczności reszt. W połączeniu z brakiem wzorca „lejka” na wykresie reszt vs. wartości dopasowane, potwierdza to spełnienie założenia stałej wariancji. W konsekwencji estymatory OLS w naszym modelu zachowują efektywność i pozostają BLUE.

## 7 Interpretacja zmiennych ostatecznego modelu

Ostateczny model (równanie 3) opisuje miesięczne zmiany logarytmicznych stóp zwrotu z akcji PKN Orlen jako funkcję sześciu regresorów i jednej zmiennej zero-jedynkowej. Poniżej interpretacja poszczególnych parametrów (estymacje z Tabeli 5):

- **Stała (Intercept = -0,003328, p=0,4030)** Wartość wyrazu wolnego nie jest statystycznie istotna ( $p > 0,05$ ), co oznacza, że przy zerowych zmianach wszystkich regresorów model nie wyróżnia jednoznacznego średniego zwrotu.
- **$\Delta$  EUR/PLN (-0,1179, p=0,0386)** Wzrost średniomiesięcznego kursu euro o 1% (względem złotego) wiąże się ze spadkiem miesięcznej stopy zwrotu PKN Orlen o około 0,12

pkt proc. Interpretacja ekonomiczna: osłabienie złotego względem euro podnosi koszty importu surowców denominowanych w euro i obniża marże rafineryjne, co negatywnie przekłada się na wycenę akcji.

- **$\Delta \text{USD/PLN}$  (0,1080,  $p=0,0156$ )** Wzrost kursu dolara względem złotego o 1 % prowadzi średnio do wzrostu miesięcznego zwrotu z akcji o około 0,11 pkt proc. Z ekonomicznego punktu widzenia efekt ten może wynikać z dwóch przeciwnych mechanizmów. Z jednej strony, wyższy kurs USD podnosi koszt importowanych surowców energetycznych (np. ropy i gazu), co mogłoby negatywnie wpływać na marże spółki. Z drugiej strony, wyższy USD/PLN zwiększa wartość przychodów eksportowych przeliczanych na złote oraz może poprawiać konkurencyjność cenową na rynkach zagranicznych. Wydaje się, że w analizowanym okresie przeważał efekt przychodowy, stąd dodatni i istotny wpływ tej zmiennej.
- **$\Delta \text{WIG}$  (0,000025,  $p<2e-16$ )** Wzrost indeksu WIG o 1 pkt odpowiada wzrostowi stopy zwrotu PKN Orlen o 0,000025 pkt proc. Choć współczynnik jest niewielki, jest wysoce istotny statystycznie. Oznacza to, że PKN Orlen porusza się w zgodzie z ogólnym sentymentem na rynku: rosnące notowania całego rynku sprzyjają także akcjom Orlenu.
- **$\Delta \ln(\text{Gas})$  (0,04228,  $p=0,1060$ )** Zmiana logarytmu ceny gazu o 1% wiąże się z około 0,04 pkt proc. wzrostem zwrotu, jednak parametr nie jest istotny na poziomie 0,05. Mimo że wzrost ceny gazu może podnosić koszty produkcji paliw, częściowo przenoszony jest on na ceny końcowe, co może mieć mieszany wpływ na marże i wycenę.
- **$\Delta \text{ref}$  (-0,02748,  $p=0,0268$ )** Jednostkowy wzrost podstawowej stopy procentowej NBP o 1 pkt proc. prowadzi średnio do spadku stopy zwrotu z akcji o około 0,027 pkt proc. Wyższe stopy procentowe zwiększają koszt kapitału, obniżają skłonność do ryzyka inwestorów i mogą hamować konsumpcję paliw, co osłabia notowania Orlenu.
- **$\Delta \text{CPI}$  (-0,01358,  $p=0,1029$ )** Jednostkowy (1 pkt) wzrost miesięcznego wskaźnika inflacji wiąże się ze spadkiem zwrotu o około 0,013 pkt proc., choć parametr jest istotny jedynie na poziomie wyższym niż 10%. Wyższa inflacja może ograniczać siłę nabywczą konsumentów i zwiększać koszty operacyjne, co negatywnie wpływa na wycenę spółki.
- **$D\_Shock_{2014\_16}$  (0,03074,  $p=0,0232$ )** W okresie epizodu rynkowego oznaczonego zmienną zero-jedynkową *Shock<sub>2014\_16</sub>* (lipiec 2014–czerwiec 2016) średnia stopa zwrotu była wyższa o około 0,031 pkt proc. w porównaniu do okresów pozostałych. Może to odzwierciedlać relatywne korzyści cenowe PKN Orlen podczas tamtych wahań cen surowców czy zmian kursowych.

#### Uwagi ogólne:

- Wszystkie istotne na poziomie 5% zmienne mają znaki zgodne z oczekiwaniami ekonomicznymi.
- Regresory walutowe ( $\Delta$  EUR/PLN,  $\Delta$  USD/PLN) oraz makroekonomiczne ( $\Delta ref$ ) pełnią rolę głównych determinant krótkoterminowych zwrotów, co potwierdza rolę polityki monetarnej i wahań kursowych w kształtowaniu notowań PKN Orlen.
- Wysoka istotność  $\Delta$ WIG sygnalizuje silną zależność od sentymentu całego rynku akcyjnego.