

**UNIWERSYTET EKONOMICZNY W POZNANIU**  
**WYDZIAŁ INFORMATYKI I GOSPODARKI ELEKTRONICZNEJ**

# **Badanie finansowych szeregów czasowych za pomocą modeli z czasem ciągłym**

---

**Piotr Pluciennik**

**Promotor: prof. UAM dr hab. Ryszard Doman**

**Poznań 2008**

# Spis treści

Wstęp .....	5
<b>1. Pojęcia wstępne .....</b>	<b>11</b>
1.1. Procesy dyfuzji .....	11
1.2. Martyngały.....	13
1.2.1. Martyngały lokalne.....	14
1.2.2. Semimartyngały.....	16
1.3. Elementy analizy stochastycznej.....	17
1.3.1. Błądzenie losowe i ruch Browna.....	18
1.3.2. Całka stochastyczna $It\hat{o}$ .....	19
1.3.3. Stochastyczne równanie różniczkowe $It\hat{o}$ .....	21
1.4. Schematy dyskretyzacji stochastycznych równań różniczkowych $It\hat{o}$ .....	24
<b>2. Analiza ekonometryczna zmienności instrumentów finansowych .....</b>	<b>29</b>
2.1. Pojęcia wstępne .....	30
2.2. Wariancja zrealizowana.....	32
2.2.1. Zależności pomiędzy wariancją zrealizowaną, a zmiennością aktualną.....	34
2.2.2. Estymatory wariancji zrealizowanej.....	36
2.3. Wariacje wielopotęgowe .....	38
2.3.1. Wariacja potęgowa .....	39
2.3.2. Wariacja dwupotęgowa .....	40
2.3.3. Uogólnienie wariacji dwupotęgowej.....	43
<b>3. Wykorzystanie wariancji zrealizowanej do estymacji procesów stochastycznych z czasem ciągłym.....</b>	<b>46</b>
3.1. Prezentacja wybranych modeli.....	47
3.1.1. Model Vašíčka.....	47
3.1.2. Model Coxa, Ingersolla i Rossa.....	48
3.1.3. Model CKLS.....	49

3.1.4.	Model Hestona.....	50
3.2.	Estymacja parametrów stochastycznych równań różniczkowych Itô .....	50
3.2.1.	Metody elementarne .....	51
3.2.2.	Metoda największej wiarygodności.....	53
3.2.3.	Dwukrokowa metoda estymacji procesu dyfuzji.....	54
3.3.	Zastosowanie metody Phillipsa i Yu do estymacji modeli zmienności stochastycznej.....	57
3.4.	Wyznaczenie jednodniowych prognoz ex-post z modeli dyfuzji.....	59
<b>4.</b>	<b>Charakterystyka procesów stochastycznych ze składową skokową .....</b>	<b>61</b>
4.1.	Testy Barndorff-Nielsen i Shepharda na występowanie skoków .....	63
4.2.	Testowanie skoków za pomocą tzw. wariacji swapowej .....	67
4.3.	Wpływ zaburzeń mikrostruktury rynku na wariacje wielopotęgowe.....	70
4.4.	Nieparametryczne testy na występowanie skoków .....	71
4.5.	Badanie mocy i rozmiaru testów na występowanie skoków .....	73
<b>5.</b>	<b>Uogólniona metoda momentów .....</b>	<b>90</b>
5.1.	Macierz wag .....	92
5.2.	Warunki momentów .....	95
5.3.	Estymacja parametrów modeli dyfuzji .....	96
<b>6.</b>	<b>Modelowanie indeksów giełdowych i kursów walutowych .....</b>	<b>98</b>
6.1.	Modelowanie danych dziennych za pomocą modeli z czasem ciągłym .....	99
6.1.1.	Prezentacja danych .....	99
6.1.2.	Omówienie wyników empirycznych.....	100
6.2.	Modelowanie procesów na podstawie danych ultra wysokiej częstotliwości.....	108
6.2.1.	Prezentacja danych .....	108
6.2.2.	Omówienie wyników empirycznych.....	109
6.3.	Modelowanie procesów cen logarytmicznych za pomocą modeli zmienności stochastycznej.....	128
6.4.	Prognozy średniej warunkowej .....	130
6.4.1.	Testy na występowanie pierwiastka jednostkowego.....	131

6.4.2. Ocena jakości prognoz.....	134
6.5. Podsumowanie.....	138
<b>7. Testowanie skoków w procesach generujących notowania indeksów giełdowych i kursów walutowych .....</b>	<b>140</b>
7.1. Omówienie wyników empirycznych .....	143
7.2. Testowanie skoków a mikrostruktura rynku .....	151
7.3. Wnioski końcowe .....	163
<b>Zakończenie.....</b>	<b>165</b>
<b>Dodatek A.....</b>	<b>168</b>
<b>Literatura .....</b>	<b>170</b>

## Wstęp

We współczesnej ekonometrii panuje przekonanie, że charakter zjawisk finansowych jest uzależniony od czynników losowych. Przedmiotem sporu pozostaje jednak kwestia przyczyn występowania tych czynników. Dlatego podstawowym jest pytanie, czy zjawiska finansowe mają charakter losowy wynikający z ich natury, czy też występowanie czynników losowych jest konsekwencją niemożności ogarnięcia pełnej informacji potrzebnej do budowy modelu oraz niemożność przewidzenia reakcji inwestorów w określonych sytuacjach? Mimo trudności z podaniem jednoznacznej odpowiedzi, wszystkie ekonometryczne modele służące do opisu finansowych szeregów czasowych, począwszy od stosowanych już przed 25 laty prostych modeli autoregresji znanych także jako modele Boxa-Jenkinsa, zawierają czynnik losowy. Wyjątkowo ważną klasą modeli służących do opisu procesów finansowych są modele z czasem ciągłym oraz ich szczególny przypadek – modele dyfuzji. Obecnie na rozwiniętych rynkach finansowych dostępne są notowania instrumentów finansowych zawierające pełną informację na temat cen transakcyjnych, tzw. dane tick-by-tick. Stanowi to naturalną motywację do stosowania modeli dyfuzji oraz modeli dyfuzji ze skokami w badaniach procesów cen instrumentów finansowych. Modele te, stosowane początkowo do wyceny krótkoterminowej stopy procentowej (Merton 1973, Vašíček 1977) zyskały na znaczeniu na początku lat siedemdziesiątych, gdy Black i Scholes (1973) wprowadzili model wyceny europejskiej opcji kupna i sprzedaży, w którym cena instrumentu podstawowego opisana była jednym z najprostszych modeli dyfuzji – geometrycznym ruchem Browna. Powstałe w następnych latach modyfikacje modelu Blacka i Scholesa – model Mertona (1974) i Hestona (1993) zakładają, odpowiednio, że wolna od ryzyka stopa procentowa i zmienność opisywane są modelami dyfuzji, Chen (1996) buduje model wyceny opcji na stopy procentowe oparty o model średniej i zmienności stochastycznej, natomiast Kou (2000, 2002) oraz Kou i Wang (2004) wykorzystują do wyceny opcji różne modele dyfuzji ze skokami.

Kolejnym obszarem zastosowań wspomnianych modeli są zagadnienia dotyczące opisu i prognozowania stóp procentowych (struktury terminowej) a także restrukturyzowanych instrumentów dłużnych. Pojawiające się obecnie na rynkach finansowych zaburzenia związane są między innymi z trudnościami dotyczącymi poprawnej wyceny bardzo złożonych instrumentów pochodnych opartych na instrumentach dłużnych. Tematyka ta została podjęta m.in. przez Detemple i Osakwe (1999) którzy zaproponowali

model wyceny pojawiających się obecnie w obrocie giełdowym opcji na zmienność, Jagannathana, Kaplina i Suna (2004) którzy zastosowali wielowymiarowe modele CIR do wyceny kontraktów cap i swapcji oraz Mannoliniego, Mariego, Renò (2008) wyceniających kontrakty cap i floor za pomocą rozszerzonego modelu CIR. W ostatnich latach pojawiają się także modele służące do wyceny instrumentów pochodnych na podstawie modeli zmienności stochastycznej, czyli dwuwymiarowych modeli dyfuzji, w których pierwsze równanie opisuje cenę instrumentu podstawowego, a drugie jego zmienność. Elliott, Siu i Chan (2007) wycenili swapcje na podstawie przełącznikowego modelu Hestona, natomiast Sepp (2008a) wykorzystał w tym celu model Hestona ze skokami. Znaczenie wyżej wspomnianych instrumentów pochodnych jest bardzo duże. Odgrywają one ważną rolę zarówno w procesach zarządzania ryzykiem, jak i w agresywnych strategiach inwestycyjnych. Pierwsze strategie zabezpieczające zostały zaproponowane już przez Blacka i Scholesa (1973). Obecnie znane są już strategie pozwalające zabezpieczyć pozycje m.in. w kontraktach swapowych (Javaheri i in. 2004, Howison i in. 2004) czy opcjach na zmienność (Psychoyios, Skiadopoulos 2006, Sepp 2008b, Broadie, Jain 2008).

Skuteczne stosowanie modeli w praktyce wymaga narzędzi pozwalających na ich zadowalającą estymację. Dlatego podjęcie w pracy tematyki modelowania i prognozowania finansowych szeregów czasowych za pomocą modeli z czasem ciągłym wydaje się szczególnie istotne.

Celem pracy jest ocena jakości oszacowań parametrów oraz prognoz otrzymanych za pomocą szybkich metod estymacji modeli z czasem ciągłym, zastosowanych do szeregów cen oraz kursów logarytmicznych. Rozstrzygnięcie tego problemu ma istotne znaczenie w zastosowaniach praktycznych, gdyż może przyczynić się do redukcji kosztów estymacji. Weryfikując występowanie skoków w badanych szeregach czasowych odpowiadamy również na pytanie, czy ceny logarytmiczne mogą być opisane jako procesy dyfuzji ze skokami. Dążymy również do rozstrzygnięcia wpływu częstotliwości danych oraz efektów mikrostruktury na wyniki testów weryfikujących występowanie skoków.

W oparciu o powyższe cele sformuowano następujące hipotezy badawcze rozprawy. Pierwsza, główna hipoteza badawcza mówi, że szybkie metody estymacji procesów z czasem ciągłym zastosowane do szeregów cen logarytmicznych pozwalają na szybsze i dokładniejsze modelowanie i prognozowanie. W hipotezie drugiej postuluje-

my, że skoki w procesach generujących dane na polskim rynku finansowym są zjawiskiem powszechnym. Trzecia hipoteza badawcza stwierdza, że otrzymanie precyzyjnych wyników testów na występowanie skoków wymaga dostatecznie dużej częstotliwości danych. Ostatnia, czwarta hipoteza mówi o istotnym wpływie efektów mikrostruktury rynku na wyniki testów na występowanie skoków.

Odpowiedzi na postawione powyżej problemy staramy się uzyskać badając szeregi logarytmicznych notowań indeksu WIG20 oraz indeksów największych giełd światowych, jak również cechujące się mniejszą zmiennością szeregi kursów walutowych. Wykorzystujemy różne częstotliwości poczynając od danych 5 minutowych, a kończąc na danych dziennych. Niektóre badania przeprowadzone zostały na podstawie symulacji procesów dyfuzji lub procesów dyfuzji ze skokami z zadanymi parametrami.

Układ pracy jest następujący. Rozdział pierwszy poświęcony jest przedstawieniu podstawowych pojęć stosowanych w dalszej części pracy. Przedstawiamy tutaj proces Markowa oraz jego szczególny przypadek – proces dyfuzji. Największy nacisk przykładamy do martyngałów, martyngałów lokalnych oraz semimartyngałów. Te ostatnie są sumą martyngału lokalnego i procesu ze skończoną wariacją. Dodatkowo omawiamy pojęcia związane z analizą stochastyczną, z której wywodzi się pojęcie stochastycznego równania różniczkowego, wykorzystywanego w pracy do modelowania logarytmicznych procesów cen. Omawiamy także dyskretyzacje Eulera oraz Milsteina pozwalające symulować dyskretne trajektorie stochastycznego równania różniczkowego w sytuacji braku możliwości wskazania jawnej postaci jego rozwiązania. Ponadto przedstawione są wyniki badania oceniającego jakość przybliżenia trajektorii stochastycznego równania różniczkowego przez wcześniej wspomniane schematy dyskretyzacji.

Przedmiotem rozważań w rozdziale drugim jest zmienność instrumentów finansowych i ekonometryczna analiza jej własności. Zestawiamy tu podstawowe pojęcia i twierdzenia dotyczące zmienności zrealizowanej. Omawiamy również problem estymowania zmienności zrealizowanej w sytuacji braku notowań poza godzinami funkcjonowania giełdy. Ponadto przedstawiamy teorię wariacji wielopotęgowych oraz wiążące się z tym pojęciem centralne twierdzenia graniczne, które leżą u podstaw konstrukcji testów na występowanie skoków w procesach cen instrumentów finansowych.

Rozdział trzeci dotyczy estymacji parametrów modeli dyfuzji. Omawiane są modele wykorzystane w dalszej części pracy do opisu badanych szeregów czasowych – model Vašička, model Coxa, Ingersolla i Rossa oraz model CKLS będący uogólnieniem dwóch poprzednich wprowadzony przez Chana i in. (1992). W rozdziale tym przedsta-

wione są metody szybkiej estymacji parametrów dyfuzji począwszy od najbardziej elementarnych bazujących na funkcji wiarygodności aż po nowoczesną dwukrokową metodę estymacji wprowadzoną przez Phillipsa i Yu (2007) na którą położony zostanie główny nacisk. Proponujemy także algorytm estymacji modeli zmienności stochastycznej bazujący na metodzie Phillipsa i Yu, wykorzystujący estymator wariancji chwilowej zdefiniowany przez Renò (2008). W ostatnim podrozdziale omawiamy metodę wyznaczania prognoz za pomocą modeli dyfuzji.

Rozdział czwarty poświęcony jest modelom dyfuzji ze skokami, które oprócz zwykłych zaburzeń losowych opisanych funkcją dyfuzji, uwzględniają możliwość wystąpienia skoków. Motywacją do opisu procesów cen instrumentów finansowych za pomocą takich modeli stała się silna leptokurtoza występująca w większości szeregów zwrotów logarytmicznych. W rozdziale rozważamy bazujące na różnicach wariacji zrealizowanej oraz wariacji dwupotęęgowej testy na występowanie skoków wprowadzone przez Barndorff-Nielsen i Shepharda (2004b, 2006) a także alternatywne podejście zaprezentowane przez Jianga i Oomena (2008) oparte na różnicy tzw. wariacji swapowej i wariacji zrealizowanej. Ponadto przedstawiamy wprowadzone przez Andersena i in. (2005) definicje przesuniętych wariacji potęgowych, które zastosowane do testów Barndorff-Nielsen i Shepharda jako estymatory wariacji scałkowanej oraz kwartyczości minimalizują wpływ zaburzeń mikrostruktury na wyniki testu. Niezależnie omawiamy wprowadzone przez Lee i Mykland (2007) nieparametryczne testy na występowanie skoków. W ostatnim podrozdziale przedstawiamy wyniki badań mocy i rozmiaru testów, które stały się pomocne przy interpretacji wyników testów przeprowadzonych w rozdziale szóstym.

W rozdziale piątym omówiona jest klasyczna już metoda estymacji procesów dyfuzji za pomocą algorytmu GMM (Generalized Method of Moments – uogólniona metoda momentów) wprowadzona przez Hansena (1982). Uogólniona metoda momentów bazuje na porównaniu momentów wyznaczonych na podstawie próby z nieobserwowalnymi teoretycznymi wartościami momentów przy założeniu pewnego znanego rozkładu. Istotnym zagadnieniem z punktu widzenia jakości estymacji jest dobór odpowiedniej macierzy wag. Stosujemy wprowadzoną przez Neweya i Westa (1987) Gallanta (1987) oraz Andrewsa (1991) macierz odwrotną do macierzy gęstości spektralnych. Uogólniona metoda momentów posłuży nam do porównania otrzymanych oszacowań parametrów dla danych szeregów czasowych z oszacowaniami otrzymanymi za pomocą dwukrokowej metody Phillipsa i Yu.



W rozdziale szóstym przeprowadzamy estymacje parametrów modeli dyfuzji dla szeregów notowań giełdowych oraz kursów walutowych. Do estymacji użyliśmy danych dziennych oraz intraday, których wykorzystanie stało się możliwe dzięki dostępności danych wysokiej częstotliwości. Główny nacisk kładziemy na estymację metodą Phillipsa i Yu. Jakość otrzymanych za pomocą tego podejścia oszacowań parametrów oceniamy wykorzystując w tym celu symulacje Monte Carlo. Ponadto badamy wpływ częstotliwości danych na wyniki estymacji. Parametry modeli dyfuzji szacujemy również za pomocą uogólnionej metody momentów wyznaczając jednocześnie jej błąd standardowy i przeprowadzamy test na istotność oszacowań parametrów. Przeprowadzamy także estymację modelu zmienności stochastycznej typu Hestona (1993) za pomocą algorytmu zaproponowanego w rozdziale 3. Niezależnie wyznaczamy prognozy ex-post średniej warunkowej z modeli dyfuzji oszacowanych za pomocą dwóch wcześniej wspomnianych metod oraz z modeli ARIMA i ARIMA-GARCH. Jakość prognoz została oceniona za pomocą standardowych miar błędów.

Ostatni rozdział dotyczy testowania skoków w procesach cen logarytmicznych. Występowanie skoków w notowaniach dziennych weryfikujemy na różnych poziomach istotności za pomocą testów Barndorff-Nielsen i Shepharda oraz Jianga i Oomena, obliczając takie charakterystyki, jak wariancja zrealizowana, zrealizowane wariacje dwupotęęgowe, czy wariancja swapowa na podstawie danych intraday. Dodatkowo wyznaczamy korelogramy dla szeregów zwrotów różnej częstotliwości, a za pomocą rozkładu asymptotycznego i bootstrapowego weryfikujemy istotność autokorelacji różnych stopni, by na jej podstawie określić występowanie efektów mikrostruktury w zwrotach danej częstotliwości. W sytuacji wykrycia takiego zjawiska wykorzystujemy przesunięte wariacje wielopotęęgowe jako estymatory wariancji scałkowanej i kwartyczności, aby zmniejszyć jego wpływ na liczbę wykrytych skoków. Występowanie skoków w notowaniach śróddziennych weryfikujemy za pomocą testów Lee i Myklanda (2007). Odrzucenie hipotezy zerowej o niewystępowaniu skoków daje motywację do podjęcia próby modelowania szeregów czasowych za pomocą procesów dyfuzji ze skokami, które bardziej precyzyjnie opisują badane szeregi.

Wyniki badań zostały otrzymane za pomocą oprogramowania powstającego w znaczących światowych ośrodkach ekonometrycznych takich jak University of Exeter, Maastricht University czy Purdue University udostępnianego do celów naukowych. Za zgodą autorów programy te były często modyfikowane do potrzeb niniejszej pracy. Znaczna część programów została napisana przez autora rozprawy w językach Ox 5.0 i

Matlab 7.0. Dotyczą one m.in. estymacji parametrów modeli dyfuzji metodą Philipsa i Yu, oceny efektywności tej metody, wyznaczania prognoz z modeli dyfuzji, testów na występowanie skoków metodą Jianga i Oomena itp. Wykorzystane do badań szeregi czasowe zostały opracowane na podstawie danych historycznych pochodzących z serwisu internetowego stooq.pl.

Niniejsza rozprawa doktorska powstała w wyniku realizacji projektu badawczego promotorskiego nr N N111 109235 MNiSW pt. „*Badanie finansowych szeregów czasowych za pomocą modeli z czasem ciągłym*”.

Szczególne podziękowania kieruję do mojego promotora prof. dr. hab. Ryszarda Domana za zainteresowanie mnie problematyką wykorzystania modeli z czasem ciągłym do opisu i prognozowania procesów finansowych, wprowadzenie mnie w tę ciekawą tematykę, chętnie okazywaną wszechstronną pomoc, cenne wskazówki, a przede wszystkim za cierpliwość i wyrozumiałość. Dziękuję również uczestnikom Seminarium z Ekonometrii Finansowej na Wydziale Matematyki i Informatyki UAM za pomocne porady oraz wskazanie przydatnych źródeł, a także prof. Francesco Zirilliemu z Dipartimento di Matematica “G. Castelnuovo” Università di Roma “La Sapienza” za przydatne uwagi dotyczące technicznej strony estymacji modeli z czasem ciągłym.

# 1. Pojęcia wstępne

Niniejszy rozdział poświęcony zostanie procesom dyfuzji, za pomocą których będziemy opisywać ewolucję logarytmów cen instrumentów finansowych, oraz zagadnieniom związanym z nimi bezpośrednio. Pojęcie procesu dyfuzji pochodzi od Roberta Browna, który w roku 1827 wykorzystał go do opisu ruchu cząsteczki w procesie przenikania cząsteczek jednej substancji w obręb drugiej na skutek chaotycznego ruchu molekuł. Z czasem okazało się, że procesy te znajdują coraz szersze zastosowanie. Wykorzystano je między innymi w psychologii do opisu cech pamięci rozpoznawczej (Ratcliff, 1978) i czasów reakcji (Ratcliff i in. 1999) do budowy dynamicznych modeli decyzyjnych (Busemeyer i Townsend, 1993) modeli kategoryzacji (Nosofsky, Palmeri 1997, Ashby 2000) czy w mechanice materiałów (Matychak i in., 1998). Wielkim obszarem zastosowań modeli dyfuzji są finanse. W roku 1900 Bachelier wykorzystał proces dyfuzji, nazywany dziś arytmetycznym ruchem Browna, do opisu ceny akcji. Proces ten stał się motywacją do wprowadzenia pojęcia stochastycznego równania różniczkowego. Szybki rozwój narzędzi analizy stochastycznej na przełomie lat czterdziestych i pięćdziesiątych umożliwił coraz szersze stosowanie modeli dyfuzji w praktyce. W latach siedemdziesiątych Black, Scholes oraz Merton wykorzystali je do wyceny opcji kupna i sprzedaży. Nieco później modele dyfuzji zaczęto wykorzystywać do opisu krótkoterminowej stopy procentowej. Silny związek z pojęciem stochastycznego równania różniczkowego mają martyngały, którym poświęcimy drugi podrozdział. Zdefiniujemy również semimartyngał, który ma szczególne znaczenie w teorii procesów dyfuzji ze skokami. W wielu przypadkach nie istnieją jawne rozwiązania stochastycznych równań różniczkowych. Dlatego w ostatnim podrozdziale przedstawione zostaną dyskretyzacje Eulera i Milsteina, pozwalające w takiej sytuacji znaleźć przybliżoną postać rozwiązania stochastycznego równania różniczkowego.

## 1.1. Procesy dyfuzji

Zdefiniowanie procesu dyfuzji wymaga wprowadzenia kilku definicji z zakresu rachunku prawdopodobieństwa. Zaczniemy od pojęcia procesu stochastycznego.

**Definicja 1.1.1.** Niech  $\mathbb{T}$  będzie podzbiorem liczb rzeczywistych interpretowanych jako momenty czasu. *Procesem stochastycznym* nazywamy rodzinę zmiennych  $(X_t)_{t \in \mathbb{T}}$ , określonych na tej samej przestrzeni probabilistycznej  $(\Omega, \mathcal{F}, \mathbb{P})$ . Jeśli  $\mathbb{T}$  jest zbiorem przeliczalnym, to mówimy o *procesie z czasem dyskretnym*. Jeśli  $\mathbb{T}$  jest przedziałem (zwykle  $[0, T]$  lub  $[0, \infty)$ ), to mówimy o *procesie z czasem ciągłym*.

W przypadku czasu dyskretnego zmienne tworzące proces będziemy oznaczali przez  $X_n$ ,  $n \in \{0, 1, 2, \dots\}$ , natomiast w przypadku czasu ciągłego używana będzie również notacja  $X(t)$ .

Spośród procesów stochastycznych szczególne znaczenie mają procesy Markowa, w których ewolucja w przyszłości zależy jedynie od stanu w danej chwili. Proces taki definiuje się w następujący sposób.

**Definicja 1.1.2.** Niech  $\mathcal{B}$  oznacza  $\sigma$ -algebrę zbiorów borelowskich na prostej. Proces stochastyczny  $(X_t)_{t \in \mathbb{T}}$ , nazywamy *procesem Markowa*, jeżeli dla każdej liczby naturalnej  $n$ , dowolnego zbioru  $B \in \mathcal{B}$ , dowolnego ciągu chwil  $t_1 < t_2 < \dots < t_n$  oraz dowolnego ciągu liczb rzeczywistych  $x_1, x_2, \dots, x_{n-1}$ , zachodzi równość

$$\mathbb{P}(X_{t_n} \in B \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) = \mathbb{P}(X_{t_n} \in B \mid X_{t_{n-1}} = x_{n-1}).$$

**Definicja 1.1.3.** *Prawdopodobieństwem przejścia* procesu  $(X_t)_{t \in \mathbb{T}}$  ze stanu  $x$  w chwili  $s$  do zbioru borelowskiego  $B$  w chwili  $t$  nazywamy funkcję

$$\mathbb{P}(B \mid x, s, t-s) = \begin{cases} \mathbb{P}(X_t \in B \mid X_s = x) & \text{dla } t-s > 0 \\ \chi_B(x) & \text{dla } t=s \end{cases},$$

gdzie  $\chi_B(x)$  jest funkcją charakterystyczną zbioru  $B$ .

Dla ustalonych  $x$ ,  $t$  i  $s \leq t$  funkcja  $\mathbb{P}(\cdot \mid x, s, t-s)$  jest miarą probabilistyczną określoną na  $\sigma$ -algebrze zbiorów borelowskich. Funkcję  $p(\cdot \mid x, s, t-s)$  będącą jej gęstością nazywamy *funkcją przejścia* lub *gęstością przejścia*. Szczególną klasą procesów Markowa są procesy dyfuzji, wykorzystane pierwotnie do opisu rozprzestrzeniania się cząsteczek jednej substancji w głąb drugiej na skutek bezładnego ruchu cieplnego cząsteczek materii.

**Definicja 1.1.4.** Proces Markowa  $(X_t)_{t \in \mathbb{T}}$ , o ciągłych trajektoriach nazywamy *procesem dyfuzji*, jeśli dla dowolnych  $s, t \in \mathbb{T}$ , takich że  $s < t$ , spełnione są warunki

$$(i) \quad \lim_{t \rightarrow s^+} \frac{1}{t-s} \int_{|y-x| > \varepsilon} p(y | x; s; t-s) dy = 0,$$

(ii) istnieje funkcja dryfu  $\mu(s, x)$  taka, że

$$\mu(s, x) = \lim_{t \rightarrow s^+} \frac{1}{t-s} \int_{|y-x| \leq \varepsilon} (y-x) p(y | x; s; t-s) dy,$$

(iii) istnieje funkcja dyfuzji  $\sigma(s, x)$  taka, że

$$\sigma(s, x) = \lim_{t \rightarrow s^+} \frac{1}{t-s} \int_{|y-x| \leq \varepsilon} (y-x)^2 p(y | x; s; t-s) dy.$$

## 1.2. Martyngaly

Niech  $(\Omega, \mathcal{F}, \mathbb{P})$  będzie przestrzenią probabilistyczną oraz niech  $(\mathcal{F}_t)_{t \in \mathbb{T}}$  będzie niemalejącą rodziną  $\sigma$ -ciał  $\mathcal{F}_t \subset \mathcal{F}$ , czyli taką, dla której  $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ , o ile  $s < t$  i  $s, t \in \mathbb{T}$ . Każde  $\sigma$ -ciało  $\mathcal{F}_t$  można interpretować jako rodzinę zdarzeń, o których zajściu lub nie zajściu wiadomo, o ile doświadczenie będzie obserwowane do chwili  $t$ .

**Definicja 1.2.1.** Dla ustalonej przestrzeni probabilistycznej  $(\Omega, \mathcal{F}, \mathbb{P})$  *uzupełnieniem  $\sigma$ -ciała  $\mathcal{F}$  względem  $\mathbb{P}$*  nazywamy  $\sigma$ -ciało  $\mathcal{F}^{\mathbb{P}}$  takich zbiorów  $A \subset \Omega$ , dla których istnieją zbiory  $A_1, A_2 \in \mathcal{F}$  takie, że  $A_1 \subset A \subset A_2$  i  $P(A_2 \setminus A_1) = 0$ .

W rachunku prawdopodobieństwa zaniedbuje się zwykle zdarzenia o prawdopodobieństwie równym zero. Dlatego zakłada się, że przestrzenie probabilistyczne są zupełne, tzn., że  $\sigma$ -ciało  $\mathcal{F}$  jest równe swojemu uzupełnieniu  $\mathcal{F}^{\mathbb{P}}$ .

**Definicja 1.2.2.** Każdą niemalejącą rodzinę  $\sigma$ -ciał  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$  nazywamy *filtracją*.

Na podstawie powyższej definicji można zauważyć, że filtracja może być interpretowana jako historia procesu. W przypadku procesów z czasem ciągłym będziemy przyjmować, że filtracja spełnia tzw. warunki zwykłe. Aby wyjaśnić, kiedy filtracja je spełnia, musimy najpierw zdefiniować filtrację zupełną.

**Definicja 1.2.3.** Filtrację  $\mathbb{F}$  nazywamy *zupelną*, gdy każde  $\sigma$ -ciało rodziny  $\mathcal{F}$  jest zupełne. Filtrację  $\mathcal{F}$  nazywamy *prawostronnie ciągłą*, gdy

$$\mathcal{F}_t = \bigcap_{s < t} \mathcal{F}_s$$

dla każdego  $t \in \mathbb{T}$ .

**Definicja 1.2.4.** Mówimy, że filtracja  $\mathbb{F}$  spełnia *warunki zwykłe*, gdy jest prawostronnie ciągła oraz zupełna. W dalszym ciągu będziemy zakładać, że rozważane filtracje spełniają warunki zwykłe.

Ważnym pojęciem związanym z filtracją jest pojęcie procesu adaptowanego do filtracji. Mówi o nim poniższa definicja.

**Definicja 1.2.5.** Proces  $X_t$  nazywamy *adaptowanym do filtracji  $\mathbb{F}$* , jeśli dla dowolnego  $t \in \mathbb{T}$ , zmienna losowa  $X_t$  jest  $\mathcal{F}_t$ -mierzalna.

**Definicja 1.2.6.** Proces  $X_t$  adaptowany do filtracji  $\mathbb{F}$  nazywamy *martyngałem* (odpowiednio *nadmartyngałem*, *podmartyngałem*) względem miary  $\mathbb{P}$  i filtracji  $\mathbb{F}$ , jeśli

- (i)  $\mathbb{E}(|X_t|) < \infty$ ,
- (ii) dla dowolnego  $s < t$  mamy

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s \text{ (odpowiednio } \mathbb{E}(X_t | \mathcal{F}_s) \leq X_s, \mathbb{E}(X_t | \mathcal{F}_s) \geq X_s \text{)}.$$

Z definicji tej wynika, że proces  $X_t$  jest martyngałem (odpowiednio nadmartyngałem, podmartyngałem) względem miary  $\mathbb{P}$  i filtracji  $\mathcal{F}$ , jeśli warunkowa wartość oczekiwana procesu w chwili przyszłej  $t$  pod warunkiem znajomości wszystkich wcześniejszych stanów procesu do chwili  $s$  (znajomości historii do chwili  $s$ ) jest równa (odpowiednio mniejsza równa, większa równa) wartości samego procesu w chwili  $s$ . Ponadto, jak łatwo zauważyć, każdy martyngał jest zarówno nadmartyngałem jak i podmartyngałem.

### 1.2.1. Martyngały lokalne

Często w modelowaniu finansowym wykorzystuje się martyngały lokalne. Są one niezbędne do zdefiniowania semimartyngałów. Na początek określimy pojęcie momentów stopu nazywanych także momentami Markowa. Moment stopu związany jest z mode-

lami finansowymi dającymi się zinterpretować jako gra sprawiedliwa i pozwala na przeanalizowanie bardzo ważnego zagadnienia – chwili przerwania gry.

**Definicja 1.2.7.** *Momentem stopu lub momentem Markowa* względem filtracji  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$  nazywamy zmienną losową  $\tau : \Omega \rightarrow \mathbb{T} \cup \{+\infty\}$  spełniającą warunek:

$$\{\tau \leq t\} \in \mathcal{F}_t,$$

dla każdego  $t \in \mathbb{T}$ . Ponadto, jeśli  $X$  jest procesem stochastycznym, a  $\tau$  jest momentem stopu, to proces  $X^\tau$  określony wzorem  $X_t^\tau = X_{\min\{t, \tau\}}$  nazywamy *procesem zatrzymanym w  $\tau$* .

Bardzo ważnym pojęciem w analizie stochastycznej jest tzw. lokalizacja. Przypuśćmy, że mamy klasę procesów  $K$ . Oznaczmy przez  $K_{loc}$  klasę lokalną, tzn. rodzinę procesów  $X$ , dla których istnieje ciąg momentów stopu  $\tau_n$  niemalejący i rozbieżny do  $\infty$  taki, że  $X^{\tau_n} \in K$ , tzn. procesy zatrzymane w  $\tau_n$  należą do  $K$ . Taki ciąg momentów nazywamy ciągiem lokalizującym. Klasa procesów  $K$ , w zależności od potrzeb, może być przyjmowana w rozmaity sposób. Może ona zatem być klasą procesów ograniczonych, procesów o wahanu skończonym lub wahanu całkowalnym itp. Jedną z ważniejszych klas lokalnych jest klasa martyngałów lokalnych zdefiniowanych poniżej.

**Definicja 1.2.8.** Niech  $(\Omega, \mathcal{F}, \mathbb{P})$  będzie przestrzenią probabilistyczną i niech  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$  będzie filtracją. Ponadto niech  $(X_t)_{t \in \mathbb{T}}$  będzie procesem adaptowanym do  $\mathbb{F}$ . Wówczas mówimy, że  $X$  jest  *$\mathbb{F}$ -martyngałem lokalnym*, jeżeli istnieje ciąg momentów stopu  $\tau_k$  taki, że

- (i)  $\tau_k$  jest rosnący,
- (ii)  $\tau_k$  jest rozbieżny do  $\infty$ ,
- (iii)  $X_t^{\tau_k} \mathcal{X}_{\{\tau_k > 0\}}$  jest  $\mathbb{F}$ -martyngałem.

Łatwo zauważyć, że każdy martyngał jest martyngałem lokalnym. Aby to stwierdzić, wystarczy przyjąć jako ciąg lokalizujący  $\tau_k = k$  dla  $k \in \mathbb{N}$ . Stwierdzenie odwrotne nie jest prawdziwe, ponieważ istnieją martyngały lokalne, które nie są martyngałami. Przykłady takiego martyngału lokalnego są przedstawione m.in w monografiach Jakubowskiego i in. (2003) (przykład 2.25, str. 106) lub Baina (2008) (rozdział 5.2 str. 9).

Wobec tego naturalnym staje się pytanie jakie warunki musi spełnić martyngał lokalny aby być martyngałem. Odpowiedzią na nie jest następujące twierdzenie (patrz Jakubowski i in. 2003).

**Twierdzenie 1.2.1.** Niech  $X$  będzie martyngałem lokalnym. Wówczas zachodzą następujące implikacje:

- (i) Jeśli funkcja  $X^{(t)} = \sup_{s \leq t} |X_s|$  jest całkowalna dla każdego  $t$ , to proces  $X$  jest martyngałem.
- (ii) Jeżeli  $\mathcal{F}_0 = \{\emptyset, \Omega\}$  i proces  $X$  jest ograniczony z dołu, to  $X$  jest nadmartyngałem.
- (iii) Jeżeli  $\mathcal{F}_0 = \{\emptyset, \Omega\}$  i proces  $X$  jest całkowalny, to  $X$  jest nadmartyngałem.

### 1.2.2. Semimartyngały

Semimartyngał jest procesem, który ma szczególne znaczenie w analizie stochastycznej. Semimartyngały tworzą największą klasę procesów, dla których zdefiniowana jest całka stochastyczna. Ponadto semimartyngałem jest również proces będący rozwiązaniem stochastycznego równania różniczkowego Itô, o ile funkcje dryfu i dyfuzji są adaptowalne. Każdy podmartyngał i nadmartyngał, który jest prawostronnie ciągły z lewostronnymi granicami, jest jednocześnie semimartyngałem. Zarysy dowodów tych faktów zostały przedstawione m.in. przez Prottera (2004). Semimartyngał może być rozumiany jako suma martyngału lokalnego oraz procesu z lokalnie ograniczoną wariancją, co odzwierciedla poniższa definicja.

**Definicja 1.2.8.** Niech  $(\Omega, \mathcal{F}, \mathbb{P})$  będzie przestrzenią probabilistyczną oraz niech  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  będzie filtracją w  $\mathcal{F}$ . Proces  $(X_t)_{t \geq 0}$  nazywamy  $\mathbb{F}$ -semimartyngałem, jeśli można go rozłożyć na sumę

$$X_t = M_t + A_t,$$

gdzie  $M$  jest martyngałem lokalnym, a  $A$  jest procesem adaptowanym do  $\mathbb{F}$ , mającym trajektorie prawostronnie ciągłe z lewostronnymi granicami (càdlàg, od francuskiego *continue à droite, limitée à gauche*) oraz lokalnie ograniczoną wariancję.



Przynależność do klasy semimartyngałów będziemy oznaczać jako  $X \in SM$ . Pewną szczególną klasą semimartyngałów są ciągle semimartyngały,  $SM^c$ .

**Definicja 1.2.9.** Niech  $(\Omega, \mathcal{F}, \mathbb{P})$  będzie przestrzenią probabilistyczną oraz niech  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  będzie filtracją w  $\mathcal{F}$ . Proces  $(X_t)_{t \geq 0}$  adaptowany do  $\mathbb{F}$  nazywamy *ciągłym  $\mathbb{F}$ -semimartyngałem*, jeśli można go rozłożyć na sumę

$$X_t = M_t^c + A_t^c,$$

gdzie  $M^c$  jest ciągłym martyngałem lokalnym, a  $A^c$  jest ciągłym procesem adaptowanym do  $\mathbb{F}$  z lokalnie ograniczoną wariacją.

Ważniejsze własności semimartyngałów opisuje poniższe twierdzenie.

**Twierdzenie 1.2.2.** Dla klasy semimartyngałów zachodzą następujące własności

- (i) Jeśli rodzina procesów  $X^i \in SM$  dla  $i = 1, \dots, n$ , to

$$X = \sum_{i=1}^n a_i X^i \in SM.$$

- (ii) Jeśli  $X^1 \in SM$  oraz  $X^2 \in SM$ , to  $X = X^1 \cdot X^2 \in SM$ .
- (iii) Jeśli  $X \in SM$  jest procesem o wartościach w  $\mathbb{R}^m$  natomiast funkcja  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  ma ciągłe pochodne pierwszego i drugiego stopnia, to  $f(X)$  jest semimartyngałem.
- (iv) (twierdzenie Strickera) Niech  $X$  będzie  $\mathbb{F}$ -semimartyngałem. Jeżeli  $\mathbb{G}$  jest podfiltracją  $\mathbb{F}$  i  $X$  jest adaptowany do  $\mathbb{G}$ , to  $X$  jest  $\mathbb{G}$ -semimartyngałem.

Dowody powyższych własności znajdują się w monografii Prottera (2005).

### 1.3. Elementy analizy stochastycznej

Początki analizy stochastycznej sięgają pierwszej połowy XX wieku, kiedy Norbert Wiener sformalizował pojęcie ruchu Browna. Zasadniczy rozwój dziedziny przypadł na lata pięćdziesiąte XX wieku. Wówczas japoński matematyk Kiyosi Itô wniósł znaczący wkład w teorię modelowania finansowego tworząc silne matematyczne podstawy dziedziny. W latach 1945-1952 napisał on kilka prac, w których wprowadził takie pojęcia, jak różniczka stochastyczna, całka stochastyczna i stochastyczne równanie różnic-

kowe. Pojęcia te w krótkim czasie znalazły zastosowania w psychologii, fizyce, mechanice, energetyce, inżynierii, a nawet biologii. Szybko okazało się również, że mogą być bardzo przydatne we współczesnych finansach. Szczególne znaczenie mają tutaj stochastyczne równania różniczkowe, które stosuje się między innymi do opisu rynków finansowych oraz do modelowania stóp procentowych. Podejście to może także posłużyć do wyceny instrumentów pochodnych, jak chociażby w modelu Blacka-Scholesa (1973) Mertona (1973) czy Blacka (1976) gdzie zakłada się, że cena instrumentu podstawowego jest modelowana za pomocą stochastycznego równania różniczkowego. Ponadto w modelu Mertona (1973) dodatkowo przyjmuje się, że stopa procentowa opisana jest stochastycznym równaniem różniczkowym. Kolejnym przykładem zastosowania stochastycznego równania różniczkowego jest model Hestona (1993) w którym proces ceny instrumentu podstawowego opisany jest za pomocą modelu zmienności stochastycznej. Bardzo pomocnym narzędziem pozwalającym na znalezienie analitycznego rozwiązania niektórych stochastycznych równań różniczkowych jest lemat Itô. W konsekwencji rozwiązania uzyskane z jego pomocą umożliwiają konstrukcję wyżej wymienionych modeli wyceny opcji. W niniejszym rozdziale omówione zostaną narzędzia analizy stochastycznej wykorzystywane w dalszej części pracy do modelowania kursów walutowych i logarytmicznych notowań indeksów giełdowych.

### 1.3.1. Błądzenie losowe i ruch Browna

Błądzenie losowe jest jednym z najbardziej elementarnych procesów stochastycznych. Stosujemy je do konstrukcji znacznie bardziej złożonych pojęć, takich jak na przykład ruch Browna. Z kolei ruch Browna jest niezbędny do zdefiniowania całki stochastycznej oraz stochastycznych równań różniczkowych.

**Definicja 1.3.1.** *Błądzeniem losowym z odstępem  $\delta$*  nazywamy proces stochastyczny indeksowany liczbami,  $\delta, 2\delta, 3\delta, \dots$ , gdzie  $\delta > 0$ , opisany następującym równaniem rekurencyjnym

$$B_{n\delta} = B_{(n-1)\delta} + z_{n\delta}.$$

z warunkiem początkowym  $B_0 = 0$ , gdzie zmienne losowe  $z_{n\delta}$  tworzą ciąg niezależnych zmiennych losowych o jednakowym rozkładzie normalnym z wartością oczekiwaną  $\mu = 0$  oraz wariancją  $\sigma^2 = \delta$ .

Ruch Browna intuicyjnie można rozumieć jako błądzenie losowe, w którym wariancja  $\delta$  zakłócenia losowego  $z_{n\delta}$  staje się wielkością nieskończenie małą  $dt$ . Formalna definicja jest następująca.

**Definicja 1.3.2.** *Standardowym ruchem Browna lub procesem Wienera nazywamy proces stochastyczny  $(B_t)_{t \geq 0}$  spełniający następujące warunki:*

- (i)  $B_0 = 0$ ,
- (ii) Przyrosty procesu  $B$  są niezależne, czyli dla dowolnego  $n$  i dowolnego ciągu  $0 < t_1 < \dots < t_n$ , zmienne losowe

$$B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$$

są niezależne,

- (iii) Dla dowolnych  $0 \leq s < t$  przyrost  $B_t - B_s$  ma rozkład gaussowski:

$$B_t - B_s \sim N(0, t - s),$$

- (iv) Proces  $B$  jest ciągły, tzn. prawie wszystkie realizacje procesu  $B$  są funkcjami ciągłymi.

### 1.3.2. Całka stochastyczna Itô

Całka stochastyczna, której postać przypomina całkę Riemanna-Stieltjesa, różni się od niej dość zasadniczo. Jak wskażemy w dalszym ciągu, wszystkie trajektorie ruchu Browna mają nieskończoną wariację. Zatem klasyczna całka Riemanna-Stieltjesa nie może być stosowana do definiowania całki dowolnego procesu stochastycznego względem ruchu Browna. Na szczęście okazuje się, że trajektorie ruchu Browna mają skończoną kwadratową wariację, o czym mówi kolejne twierdzenie. Fakt ten umożliwia zdefiniowanie całki stochastycznej Itô.

**Twierdzenie 1.3.1 (o skończonej wariacji kwadratowej ruchu Browna).** Niech  $[a, b]$  będzie podprzedziałem przedziału  $[0, \infty)$  i niech

$$\pi_n := \{a = t_0^n < t_1^n < t_2^n < \dots < t_{m_n}^n = b\}$$

będzie ciągiem normalnym podziałów (tj. jego średnica

$$\delta(\pi_n) := \max_{0 \leq k \leq m_n - 1} |t_{k+1}^{(n)} - t_k^{(n)}| \rightarrow 0 \text{ przy } n \rightarrow \infty.$$

Wtedy

$$\sum_{k=0}^{m_n-1} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2 \rightarrow b - a$$

w  $\mathbb{L}^2(\Omega)$ .

**Twierdzenie 1.3.2 (o nieskończonej wariacji ruchu Browna).** Wszystkie trajektorie ruchu Browna mają nieskończoną wariację z prawdopodobieństwem równym 1.

Aby zdefiniować całkę stochastyczną, konieczne będzie wprowadzenie definicji procesu schodkowego.

**Definicja 1.3.3.** Adaptowany do filtracji  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  proces stochastyczny  $G \in \mathbb{L}^2(0, T)$  nazywamy procesem schodkowym, gdy istnieje podział

$$0 = t_0 < t_1 < t_2 < \dots < t_n = T$$

przedziału  $[0, T]$  taki, że

$$G_t = G_{t_k} \text{ dla } t_{k-1} < t \leq t_k \text{ gdzie } k = 1, 2, \dots, n.$$

**Definicja 1.3.4.** Niech  $G \in \mathbb{L}^2(0, T)$  będzie procesem schodkowym. Wtedy liczbę

$$\int_0^T G_t dB_t = \sum_{k=0}^{n-1} G_{t_k} (B_{t_{k+1}} - B_{t_k})^2$$

nazywamy *całką stochastyczną Itô procesu schodkowego  $G$  na przedziale  $[0, T]$* .

W celu zdefiniowania całki stochastycznej dla dowolnego procesu stochastycznego  $G \in \mathbb{L}^2(0, T)$ , przytoczymy następujące twierdzenie (Bain, 2008).

**Twierdzenie 1.3.3.** Jeżeli adaptowany do filtracji  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  proces  $G \in \mathbb{L}^2(0, T)$ , to istnieje ciąg ograniczonych procesów schodkowych  $G^{(n)} \in \mathbb{L}^2(0, T)$  taki, że

$$\mathbb{E} \int_0^T |G_t^{(n)} - G_t|^2 dt \rightarrow 0$$

przy  $n \rightarrow \infty$ .

**Definicja 1.3.5.** Niech  $G \in \mathbb{L}^2(0, T)$  będzie dowolnym procesem stochastycznym adaptowanym do filtracji  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  i niech  $G^{(n)}$  będzie ciągiem ograniczonych procesów schodkowych w przestrzeni  $\mathbb{L}^2(0, T)$  takim, że

$$\mathbb{E} \int_0^T |G_t^{(n)} - G_t|^2 dt \rightarrow 0$$

przy  $n \rightarrow \infty$ . Wtedy liczbę

$$\int_0^T G_t dB_t = \lim_{n \rightarrow \infty} \int_0^T G_{t_k}^{(n)} dB_t$$

nazywamy *całką stochastyczną Itô procesu  $G$  na przedziale  $[0, T]$* .

Własności całki stochastycznej opisuje następujący lemat (Bain, 2008, Shreve, 2004).

**Lemat 1.3.1.** Załóżmy, że  $G, H \in \mathbb{L}^2(0, T)$  są dowolnymi procesami stochastycznymi na przedziale  $[0, T]$  adaptowanymi do filtracji  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ . Wówczas całka stochastyczna Itô ma następujące własności:

- (i)  $\int_0^T (aG_t + bH_t) dB_t = a \int_0^T G_t dB_t + b \int_0^T H_t dB_t$  dla dowolnych  $a, b \in \mathbb{R}$ ,
- (ii)  $\mathbb{E} \left( \int_0^T G_t dB_t \right) = 0$ ,
- (iii)  $\mathbb{E} \left( \int_0^T G_t dB_t \right)^2 = \mathbb{E} \left( \int_0^T G_t^2 dt \right)$ ,
- (iv)  $\mathbb{E} \int_0^T G_t dB_t \int_0^T H_t dB_t = \mathbb{E} \int_0^T G_t H_t dt$ .

### 1.3.3. Stochastyczne równanie różniczkowe Itô

Naturalną konsekwencją pojęcia całki stochastycznej Itô są stochastyczne równania różniczkowe. Przyjmujemy następujące oznaczenia:

- (i) Niech  $B$  będzie adaptowanym do filtracji  $(\mathcal{F}_t)_{t \geq 0}$  ruchem Browna i niech  $Z$  będzie zmienną losową niezależną od  $B$ ,
- (ii) Niech

$$\mu : \mathbb{R} \times [0, T] \rightarrow \mathbb{R} \text{ oraz } \sigma : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$$

będą funkcjami borelowsko mierzalnymi.

**Definicja 1.3.6.** Załóżmy, że spełnione są warunki (i) i (ii) oraz niech  $X_t$  będzie adaptowanym do filtracji  $(\mathcal{F}_t)_{t \geq 0}$  procesem stochastycznym o ciągłych trajektoriach. *Stochastycznym równaniem różniczkowym Itô* nazywamy równanie postaci

$$\begin{cases} dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t \\ X_0 = Z \end{cases} \quad (1.1)$$

Ponadto funkcję  $\mu(\cdot, \cdot)$  nazywamy funkcją dryfu, a  $\sigma(\cdot, \cdot)$  funkcją dyfuzji.

Kolejna definicja będzie dotyczyła pojęcia rozwiązania stochastycznego równania różniczkowego Itô.

**Definicja 1.3.7.** Niech dana będzie przestrzeń probabilistyczna  $(\Omega, \mathcal{F}, \mathbb{P})$  z filtracją  $(\mathcal{F}_t)_{t \geq 0}$ . Mówimy, że proces stochastyczny  $X$  jest *rozwiązaniem stochastycznego równania różniczkowego (1.1)* dla  $t \in [0, T]$ , gdy zachodzą następujące warunki:

- (i)  $X$  jest adaptowany do filtracji  $(\mathcal{F}_t)_{t \geq 0}$ ,
- (ii)  $\mathbb{P}\left(\int_0^T \mu(X_s, s)ds < \infty\right) = 1$ ,
- (iii)  $\mathbb{P}\left(\int_0^T \sigma^2(X_s, s)ds < \infty\right) = 1$ ,
- (iv)  $X$  wyraża się wzorem

$$X_t = Z + \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dB_s$$

prawie na pewno dla wszystkich  $t \in [0, T]$ .

Poniżej przytoczymy twierdzenie mówiące o istnieniu i jednoznaczności rozwiązania stochastycznego równania różniczkowego (1.1).

**Twierdzenie 1.3.6 (o istnieniu i jednoznaczności rozwiązania).** Niech  $\mu : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$  oraz  $\sigma : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$  będą funkcjami spełniającymi następujące globalne warunki Lipschitza:

Istnieje stała  $K > 0$  taka, że dla dowolnych  $x, y \in \mathbb{R}$  i  $t \in [0, T]$  zachodzą nierówności

$$\begin{aligned} |\mu(x, t) - \mu(y, t)| &\leq K |x - y|, \\ |\sigma(x, t) - \sigma(y, t)| &\leq K |x - y| \end{aligned}$$

oraz

$$|\mu(x, t)| + |\sigma(x, t)| \leq K(1 + |x|).$$

Załóżmy ponadto, że  $Z$  jest zmienną losową taką, że

$$\mathbb{E}(|Z|^2) < \infty$$

oraz  $Z$  jest niezależna od ruchu Browna  $B_t$ . Wtedy istnieje dokładnie jedno rozwiązanie  $X \in \mathbb{L}^2(0, T)$  stochastycznego równania różniczkowego (1.1).

Powyższe twierdzenie jest udowodnione między innymi w monografiach Baina (2008) oraz Prottera (2005).

Naturalnym wydaje się pytanie o związek rozwiązania stochastycznego równania różniczkowego i procesu dyfuzji. Procesem dyfuzji jest ono wtedy, gdy funkcje dryfu i dyfuzji są ciągle względem zmiennej  $t$ . W celu znalezienia postaci rozwiązania stochastycznego niekiedy wykorzystuje się fakt istnienia różniczki stochastycznej dla transformacji procesu  $X$ . Różniczkę tą wyznacza się z wprowadzonego przez Itô (1951) wzoru nazywanego wzorem Itô.

**Twierdzenie 1.3.8 (Lemat Itô).** Niech  $X(t)$  będzie procesem opisanym równaniem (1.1). Dla dowolnej funkcji  $f(X_t, t) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^+)$  zachodzi następujący wzór

$$\begin{aligned} df(X_t, t) &= \left[ \frac{\partial f}{\partial t}(X_t, t) + \mu(X_t, t) \frac{\partial f}{\partial X}(X_t, t) + \frac{1}{2} \frac{\partial^2 f}{\partial X^2}(X_t, t) \sigma(X_t, t) \right] dt \\ &\quad + \sigma(X_t, t) \frac{\partial f}{\partial X}(X_t, t) dB_t \end{aligned}$$

prawie na pewno dla wszystkich  $t \in [0, T]$ .

Dowodu powyższego lematu nie podajemy. Jest on przedstawiony w podręcznikach Baina (2008), Prottera (2005) oraz Werona i Werona (1999).

## 1.4. Schematy dyskretyzacji stochastycznych równań różniczkowych Itô

Pomimo istnienia i jednoznaczności rozwiązania stochastycznego równania różniczkowego, w wielu przypadkach nie ma możliwości wskazania jego jawnej postaci. Wobec tego konieczne jest wyznaczenie przybliżonej formy rozwiązania. Ponadto notowania wykorzystywane do estymacji dostępne są w pewnych ustalonych odstępach czasowych. W tej sytuacji konieczne jest zastąpienie równania dyfuzji równaniem różnicowym. W niniejszej pracy wykorzystane zostaną w tym celu dwa schematy dyskretyzacji – Eulera i Milsteina. Pierwszy z nich zaproponowany przez Maruyamę (1955) polega na aproksymacji procesu

$$\begin{cases} dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t \\ X_0 = Z \end{cases} \quad (1.2)$$

równaniem

$$\begin{cases} X_n^{(\delta)} - X_{n-1}^{(\delta)} = \mu(X_{n-1}^{(\delta)}, t)\delta + \sigma(X_{n-1}^{(\delta)}, t)\sqrt{\delta}\varepsilon_n^{(\delta)} \\ X_0 = Z \end{cases}, \quad (1.3)$$

gdzie  $\varepsilon_n^{(\delta)}$  dla  $n=0,1,\dots$  jest standardowym białym szumem, przy czym dla procesu  $X_n^{(\delta)}$  spełniona jest równość  $X_t^{(\delta)} = X_{[t/\delta]}^{(\delta)}$ , gdzie  $[x]$  oznacza część całkowitą liczby  $x$ . Dyskretyzacja Eulera pozwala także znaleźć warunkowy rozkład zmiennej  $X_t$  względem  $X_0$ , w sytuacji gdy nie jest możliwe znalezienie analitycznego rozwiązania równania (1.2). Do wykorzystania dyskretyzacji Eulera uprawnia nas następujące twierdzenie, które także mówi o tempie zbieżności  $X_n^{(\delta)}$  do  $X_t$ .

**Twierdzenie 1.4.1** (por. Kloden i Platen 1992, Kostrzewski 2006b). Załóżmy, że  $\mathbb{E}(Z^2) < \infty$  oraz, że istnieje stała  $K > 0$  taka, że dla dowolnych  $t_1, t_2 \in [0, T]$  i  $x, y \in \mathbb{R}$  funkcje dryfu i dyfuzji procesu  $X_t$  spełniają następujące warunki:

- (i)  $|\mu(x, t_1) - \mu(y, t_1)| + |\sigma(x, t_1) - \sigma(y, t_1)| \leq K|x - y|$ ,
- (ii)  $|\mu(x, t_1)| + |\sigma(x, t_1)| \leq K(1 + |x|)$ ,
- (iii)  $|\mu(x, t_1) - \mu(x, t_2)| + |\sigma(x, t_1) - \sigma(x, t_2)| \leq K(1 + |x|)|t_1 - t_2|^{1/2}$ .



Wtedy istnieje taka stała  $C$ , że  $\mathbb{E}|X_T - X_T^{(\delta)}| < C\delta^{1/2}$  oraz  $\mathbb{E}(X_t - X_t^{(\delta)})^2 < O(\delta)$  dla  $t < T$ .

Na podstawie twierdzenia 1.4.1, proces  $X_n^{(\delta)}$  zbiega w silnym sensie do  $X_t$  z rzędem  $\nu = 0,5$ . Milstein (1974,1978) zaproponował nieco bardziej zaawansowany schemat dyskretyzacji. Schemat ten polega na aproksymacji równania (1.2) za pomocą wyrażenia

$$\begin{cases} X_n^{(\delta)} - X_{n-1}^{(\delta)} = \mu(X_{n-1}^{(\delta)}, t)\delta + \sigma(X_{n-1}^{(\delta)}, t)\sqrt{\delta}\varepsilon_n^{(\delta)} + \frac{1}{2}\sigma(X_{n-1}^{(\delta)}, t)\frac{\partial\sigma(X, t)}{\partial X}\Big|_{X=X_{n-1}^{(\delta)}}\delta(\varepsilon_n^{(\delta)})^2 \\ X_0 = Z \end{cases} \quad (1.4)$$

Jak łatwo zauważyć, w sytuacji gdy funkcja dyfuzji zależy wyłącznie od czasu, schemat Milsteina jest równoważny schematowi Eulera. Poniższe twierdzenie mówi nam o tempie zbieżności  $X_n^{(\delta)}$  do  $X_t$ .

**Twierdzenie 1.4.2** (por. Kloden i Platen 1992, Kostrzewski 2006b). Załóżmy, że  $\mathbb{E}(Z^2) < \infty$  oraz, że istnieje stała  $K > 0$  taka, że dla dowolnych  $t_1, t_2 \in [0, T]$  i  $x, y \in \mathbb{R}$  funkcje dryfu i dyfuzji procesu  $X_t$  spełniają następujące warunki:

- (i)  $|\mu^*(x, t_1) - \mu^*(y, t_1)| \leq K|x - y|$ ,
- (ii)  $|\sigma(x, t_1) - \sigma(y, t_1)| \leq K|x - y|$ ,
- (iii)  $\left| \sigma(x, t_1)\frac{\partial\sigma(X, t_1)}{\partial X}\Big|_{X=x} - \sigma(y, t_1)\frac{\partial\sigma(X, t_1)}{\partial X}\Big|_{X=y} \right| \leq K|x - y|$ ,
- (iv)  $|\mu^*(x, t_1)| + \left| \sigma(x, t_1)\frac{\partial\mu(X, t_1)}{\partial X}\Big|_{X=x} \right| \leq K(1 + |x|)$ ,
- (v)  $|\sigma(x, t_1)| + \left| \sigma(y, t_1)\frac{\partial\sigma(X, t_1)}{\partial X}\Big|_{X=y} \right| \leq K(1 + |x|)$ ,
- (vi)  $\left| \sigma^2(x, t_1)\frac{\partial^2\sigma(X, t_1)}{\partial X^2}\Big|_{X=x} \right| \leq K(1 + |x|)$ ,
- (vii)  $|\mu^*(x, t_1) - \mu^*(x, t_2)| \leq K(1 + |x|)|t_1 - t_2|^{1/2}$ ,
- (viii)  $|\sigma(x, t_1) - \sigma(x, t_2)| \leq K(1 + |x|)|t_1 - t_2|^{1/2}$ ,

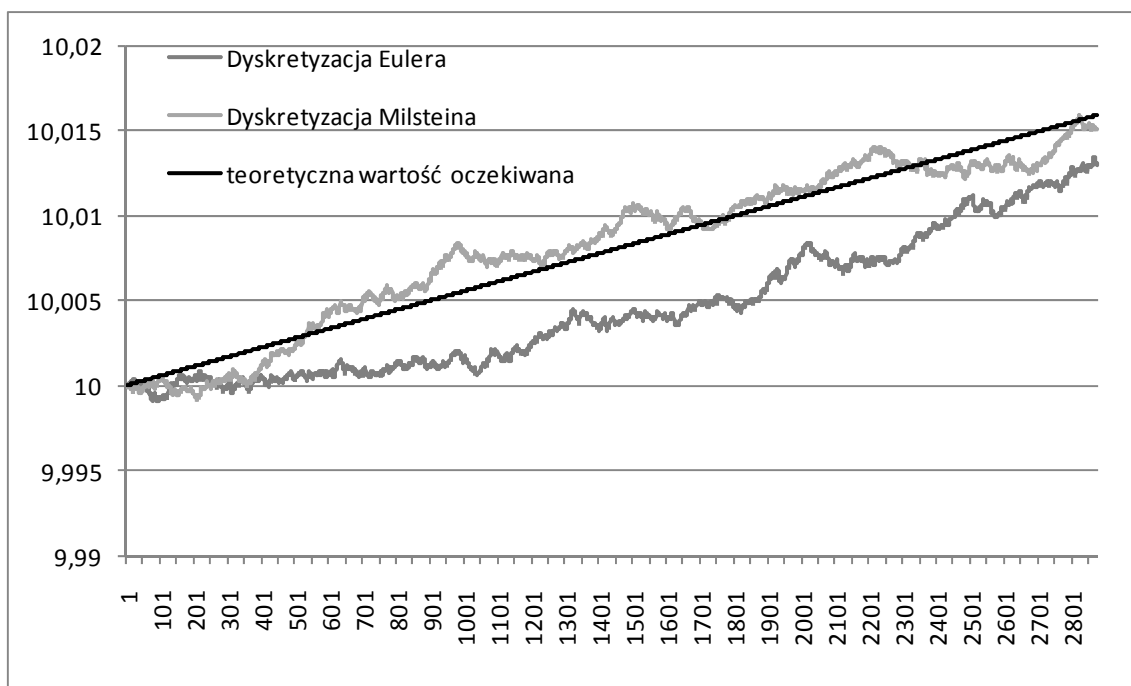
$$(ix) \left| \sigma(x, t_1) \frac{\partial \sigma(X, t_1)}{\partial X} \Big|_{X=x} - \sigma(x, t_2) \right| \leq K(1 + |x|) |t_1 - t_2|^{1/2},$$

$$\text{gdzie } \mu^*(x, t) = \mu(x, t) - \frac{1}{2} \sigma(x, t) \frac{\partial \sigma(X, t)}{\partial X} \Big|_{X=x}.$$

Wówczas istnieje taka stała  $C$ , że  $\mathbb{E} |X_T - X_T^{(\delta)}| < C\delta$  oraz  $\mathbb{E}(X_t - X_t^{(\delta)})^2 < O(\delta^2)$  dla  $t < T$ .

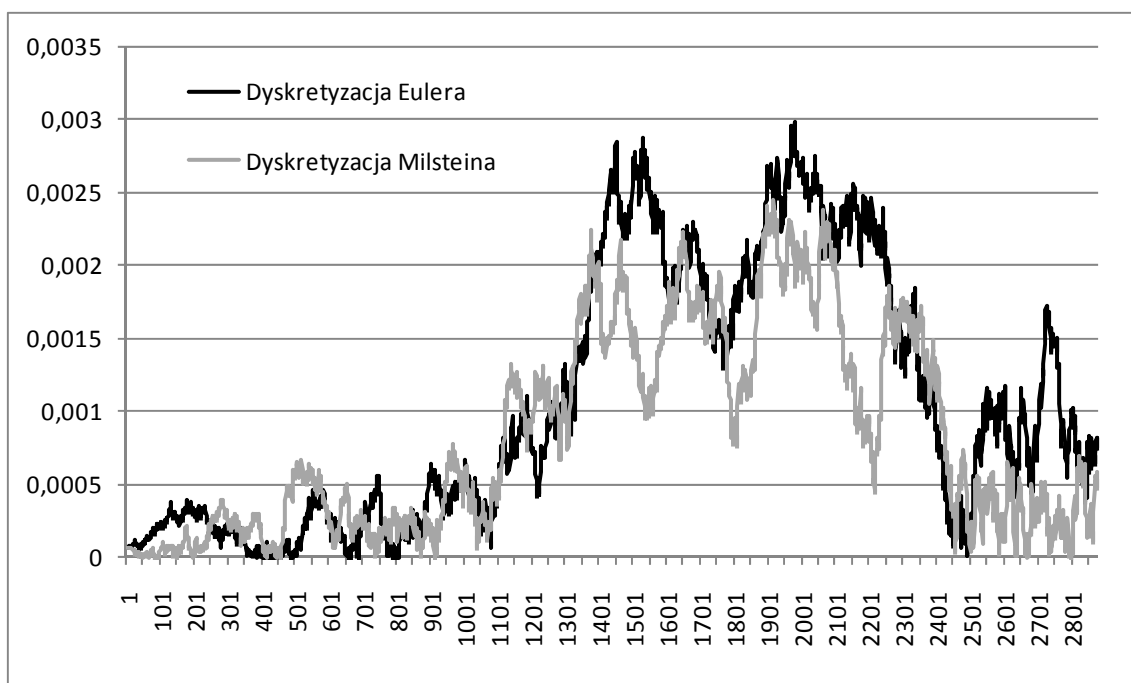
Z twierdzenia 1.4.2 wynika, że wyznaczony z równania (1.4) proces  $X_n^{(\delta)}$  jest zbieżny w silnym sensie do  $X_t$  z rzędem  $\nu = 1$ . Należy zaznaczyć, że wyższe rzędy zbieżności można uzyskać przez uwzględnienie kolejnych wyrazów wzoru Taylora. Schematy takie rozważane są między innymi w pracach Saito i Mitsui (1996) oraz Platena (1999).

W celu oceny jakości przybliżenia trajektorii stochastycznego równania różniczkowego przez wyżej opisane schematy dyskretyzacji, przeprowadziliśmy symulację 10000 trajektorii stochastycznego równania Coxa, Ingersolla i Rossa (1985) (Patrz rozdział 3.3.2) z parametrami  $\kappa = 0,1$ ,  $\mu = 12$ ,  $\sigma = 0,5$  i warunkiem początkowym  $X_0 = 10$  za pomocą schematu Eulera i Milsteina dla 10 min. odstępów czasowych. Na podstawie symulacji wyznaczyliśmy średnie i wariancje we wszystkich momentach czasowych i porównaliśmy je z wartościami teoretycznymi. Na poniższym wykresach przedstawiona teoretyczna wartość oczekiwana rozważanego modelu oraz średnie empiryczne wyznaczone na podstawie 10000 trajektorii otrzymanych za pomocą badanych schematów dyskretyzacji.



Rys. 1.1. Teoretyczna wartość oczekiwana rozważanego modelu oraz średnie empiryczne wyznaczone na podstawie 10000 trajektorii otrzymanych za pomocą badanych schematów dyskretyzacji

Na kolejnym wykresie przedstawione są różnice wariancji empirycznych i wariancji teoretycznych rozważanego modelu. Wariancja teoretyczna została wyznaczona ze wzoru (3.3).



Rys. 1.2. Różnice wariancji empirycznych wyznaczonych na podstawie 10000 trajektorii otrzymanych za pomocą badanych schematów dyskretyzacji i wariacji teoretycznej rozważanego modelu

Jak łatwo zauważyć, różnice średnich jak i wariancji teoretycznych oraz empirycznych są bardzo niewielkie w przypadku zastosowania obydwu schematów dyskretyzacji. Zgodnie z oczekiwaniami dyskretyzacja Milsteina przybliża charakterystyki ciągłego procesu opisanego stochastycznym równaniem różniczkowym nieco lepiej niż dyskretyzacja Eulera. Jakość otrzymanego przybliżenia jest na tyle zadowalająca, że nie ma sensu stosować wspomnianych wcześniej schematów o wyższym rzędzie zbieżności. Biorąc pod uwagę, że czasochłonność symulacji trajektorii za pomocą dyskretyzacji Milsteina jest tylko nieznacznie wyższa niż w przypadku zastosowania dyskretyzacji Eulera, w pracy stosować będziemy głównie drugi schemat.

## 2. Analiza ekonometryczna zmienności instrumentów finansowych

Zmienność ceny instrumentu finansowego może być rozumiana jako miara niepewności co do przyszłych zmian ceny. Pojęcie to jest wykorzystywane we wszystkich teoriach dotyczących finansów i podejmowania decyzji. Najbardziej naturalnymi miarami tej niepewności są wariancja i odchylenie standardowe.

Analizę stochastyczną zmienności można prowadzić w dwóch ujęciach, statycznym i dynamicznym. Pierwsze spośród nich zwane także historycznym zakłada, że zwroty historyczne nie zawierają użytecznych informacji prognostycznych. Zmienność jest wyliczana wówczas na podstawie wcześniejszych zwrotów tak jakby były one realizacjami zmiennych niezależnych. W praktyce sprowadza to się do klasycznego estymatora wariancji. W drugim podejściu przyjmuje się założenie, że istnieje możliwość przewidywania przyszłych wartości zwrotów w oparciu o ich przeszłe wartości oraz opóźnione wartości innych zmiennych objaśniających. Szczególne znaczenie ma tutaj występowanie nieliniowych zależności między tymi wielkościami. Koncepcja i specyfikacja modelu, w którym wykorzystano tego typu relację pojawiła się po raz pierwszy w pracy Engle'a (1982) dotyczącej zmienności stóp inflacji w Wielkiej Brytanii. Zaproponowany w pracy model autoregresyjnej heteroskedaktyczności warunkowej ARCH( $p$ ) opiera się na założeniu istnienia autokorelacji pomiędzy kwadratami zwrotów. W kolejnych latach model ten podlegał licznym rozszerzeniom i modyfikacjom znajdując zastosowania w modelowaniu zmienności finansowych i makroekonomicznych szeregów czasowych.

Alternatywnym podejściem do zaproponowanego przez Engle'a modelu ARCH( $p$ ) są modele zmienności stochastycznej. Motywacji ich powstania należy dopatrywać się w powszechnie stosowanym po dzień dzisiejszy modelu Blacka-Scholesa (1973) którego cechuje nierealistyczne założenie o stałości zmienności. Założenie to stało się głównym powodem wprowadzenia przez Hulla i White'a (1987) modelu wyceny opcji, w którym zmienność opisana jest stochastycznym równaniem różniczkowym.

W ostatnich latach ogromne znaczenie zyskały dane wysokiej częstotliwości i wykorzystanie ich do wyznaczenia wariancji zrealizowanej będącej estymatorem nie-

obserwowalnej zmienności instrumentów finansowych. W niniejszym rozdziale przedstawimy uogólnienia tego typu obiektu, wariację potęgową i wielopotęgową, będące odpowiednio, sumą potęg i iloczynów potęg bezwzględnych zwrotów. Pojęcia te mogą być bardzo przydatne przy ekonometrycznej analizie zmienności instrumentów finansowych. Wariacja dwupotęgowa jest estymatorem wariancji scałkowanej w modelach zmienności stochastycznej. Pojęcie wariacji potęgowej i dwupotęgowej może być wykorzystane do testowania występowania skoków w szeregach cen logarytmicznych.

## 2.1. Pojęcia wstępne

Niech  $X$  będzie procesem cen logarytmicznych instrumentu finansowego zdefiniowanym na zupełnej przestrzeni probabilistycznej  $(\Omega, \mathcal{F}, \mathbb{P})$ . Jeśli  $h > 0$  oznacza ustalony okres czasu, np. dzień, to  $i$ -ty zwrot o częstotliwości  $h$  dany jest wzorem

$$y_i = y_{ih,h} = X_{ih} - X_{(i-1)h}.$$

Przyjmijmy teraz, że w ciągu  $i$ -tego okresu obserwowanych jest  $M$  notowań. Oznaczmy  $\delta = hM^{-1}$ . Wówczas  $j$ -ty zwrot wysokiej częstotliwości w  $i$ -tym okresie zdefiniowany jest wzorem

$$\begin{aligned} y_{j,i} = y_{(i-1)h+j\delta,\delta} &= X\left((i-1)h + \frac{hj}{M}\right) - X\left((i-1)h + \frac{h(j-1)}{M}\right) \\ &= X((i-1)h + j\delta) - X((i-1)h + (j-1)\delta). \end{aligned}$$

W dalszych rozdziałach pracy nie zawsze będziemy zakładać, że proces cen logarytmicznych ma ciągle trajektorie. Wobec tego ważne jest jednoznaczne zdefiniowanie procesów jednostronnie ciągłych oraz występujących w procesach skali nieciągłości, które będziemy nazywać *skokami*.

**Definicja 2.1.1.** Załóżmy, że dla każdego  $t \in (0, T)$  istnieją granice jednostronne  $X_{t-} = \lim_{\tau \rightarrow t, \tau < t} X_\tau$  i  $X_{t+} = \lim_{\tau \rightarrow t, \tau > t} X_\tau$ . Mówimy że proces  $X_t$  jest *prawostronnie ciągły z lewostronnymi granicami (càdlàg)*, jeśli  $X_t = X_{t+}$ . Podobnie, proces  $X_t$  jest *lewostronnie ciągły z prawostronnymi granicami (càglàd)* jeśli  $X_t = X_{t-}$ .

Przyjmijmy, że  $X$  jest procesem prawostronnie ciągłym z lewostronnymi granicami. Wówczas wartość skoku w chwili  $t$  jest równa

$$c(t) = \Delta X_t = X_t - X_{t-}.$$

W celu wprowadzenia takich pojęć, jak wariacja zrealizowana czy wariacje wielopotęgowe, przypomnijmy pojęcie wariacji kwadratowej (por. Barndorff-Nielsen i Shephard 2004c).

**Definicja 2.1.2.** Mówimy, że wyrażenie  $[X]_t$  jest *procesem wariacji kwadratowej procesu*  $X$ , jeśli

$$[X]_t = \int_0^t (dX_s)^2 = \text{plim}_{n \rightarrow \infty} \left[ \sum_{j=0}^{m_n-1} (X_{t_{j+1}^{(n)}} - X_{t_j^{(n)}})^2 \right],$$

gdzie  $\pi_n = \{0 = t_1^{(n)} < t_2^{(n)} < \dots < t_{m_n}^{(n)} = t\}$  jest ciągiem normalnym podziałów natomiast symbol plim oznacza granicę według prawdopodobieństwa.

**Definicja 2.1.3.** (Barndorff-Nielsen i Shephard 2004c) Proces  $X$  jest nazywany *ciągłym semimartynagałem zmienności stochastycznej* ( $SVSM^c$ ), jeśli daje się przedstawić w postaci

$$X = \alpha^* + m^*, \tag{2.1}$$

gdzie  $\alpha^*$  jest ciągłym procesem dryfu,  $m^*$  jest procesem zmienności stochastycznej ( $SV$ ) określonym wzorem

$$m^*(t) = \int_0^t \sigma(s) dB_s,$$

gdzie  $B_s$  jest standardowym ruchem Browna,  $\sigma(t)$  jest ściśle dodatnim i prawostronnie ciągłym z lewostronnymi granicami (càdlàg) procesem zmienności chwilowej, lokalnie odgraniczonym od zera, dla którego proces wariacji skalkowanej, definiowany jako

$$\sigma^{2*}(t) = \int_0^t \sigma^2(s) ds,$$

przyjmuje wartości skończone dla dowolnego  $t < \infty$ .

Można zauważyć, że zachodzi następująca zależność  $SVSM^c \subset SM^c \subset SM$ . Jednocześnie należy zwrócić uwagę, że dla wszystkich modeli zmienności stochastycznej  $\sigma^{2*}(t) = [X]_t$ .

Przyjmijmy, że proces cen logarytmicznych instrumentu finansowego  $X \in SVSM^c$  jest semimartyngałem. Równanie (2.1) można zapisać w postaci

$$dX = d\alpha^* + dm^*.$$

Wówczas, jeżeli  $m^*$  jest martyngałem, to prognozowalność procesu  $\alpha^*$  oznacza, że

$$\mathbb{E}(dX_t | \mathcal{F}_t) = d\alpha^*(t),$$

gdzie  $(\mathcal{F}_t)$  jest filtracją naturalną procesu  $\alpha^*(t)$ . Wobec tego funkcja dryfu może być zapisana w postaci następującej całki z oczekiwanych zwrotów chwilowych:

$$\alpha^*(t) = \int_0^t \mathbb{E}(dX_s | \mathcal{F}_s).$$

Ponadto, jeżeli  $m^*$  jest martyngałem ciągłym, który jest lokalnie całkowny z kwadratem, to

$$[X]_t = \int_0^t \text{var}(dX_s | \mathcal{F}_s) = \int_0^t \text{var}(dm^*(s) | \mathcal{F}_s),$$

co oznacza, że wariacja kwadratowa jest równa scałkowanej wariancji zwrotów chwilowych.

Jak wiadomo,  $[X]_t$  można wyrazić w postaci

$$[X]_t = [m^{*c}]_t + \sum_{0 < s < t} \{\Delta X_s\}^2 = [m^{*c}]_t + [X^d]_t,$$

gdzie  $\Delta X_t = X_t - X_{t-}$  jest skokiem (Protter 2005). Z powyższej równości wynika, że wariację kwadratową  $[X]_t$  można rozłożyć na sumę wariacji kwadratowych  $[m^{*c}]_t$  oraz  $[X^d]_t$ . Daje to możliwość testowania istnienia skoków poprzez sprawdzenie, czy  $[X]_t = [m^{*c}]_t$ . Spostrzeżenie to jest podstawą wprowadzonych przez Barndorff-Nielsen i Shepharda (2006) testów na występowanie skoków.

## 2.2. Wariancja zrealizowana

Dostępność notowań o dużej częstotliwości oznacza, że możemy wykorzystać sumę kwadratów zwrotów intraday jako pewną miarę zmienności instrumentu finansowego zwaną *wariancją zrealizowaną*.



W modelu zmienności stochastycznej logarymiczne ceny instrumentów finansowych spełniają następujące stochastyczne równanie różniczkowe

$$dX_t = [\mu + \beta\sigma^2(t)]dt + \sigma(t)dB_t, \quad (2.2)$$

gdzie  $\sigma^2(t)$  jest wariancją chwilową, dla której każda trajektoria jest całkowalna na dowolnym przedziale skończonym oraz proces  $\sigma^2(t)$  jest stacjonarny i niezależny od ruchu Browna  $B_t$ . Ponadto,  $\mu$  oznacza dryf, a  $\beta$  premię za ryzyko. W takiej sytuacji, warunkowy rozkład zwrotu  $y_i = y_{ih,h}$  dany jest wzorem

$$y_i | \sigma_i^2 \sim N(\mu h + \beta\sigma_i^2, \sigma_i^2),$$

gdzie  $\sigma_i^2$  zdefiniowane jako

$$\sigma_i^2 = \sigma^{2*}(ih) - \sigma^{2*}[(i-1)h],$$

nazywane jest *wariancją aktualną w i-tym przedziale*, przy czym

$$\sigma^{2*}(t) = \int_0^t \sigma^2(s)ds.$$

Oznaczmy przez  $\xi$ ,  $\omega^2$  i  $r$  odpowiednio, średnią, wariancję i funkcję autokorelacji procesu  $\sigma^2(t)$ . Wówczas

$$\begin{aligned} \mathbb{E}(\sigma_i^2) &= \xi h, \\ \text{var}(\sigma_i^2) &= 2\omega^2 r^{**}(h), \\ \text{cov}(\sigma_i^2, \sigma_{i+s}^2) &= \omega^2 \diamond r^{**}(hs), \end{aligned} \quad (2.3)$$

gdzie

$$\diamond r^{**}(s) = r^{**}(s+h) - 2r^{**}(s) + r^{**}(s-h)$$

oraz

$$\begin{aligned} r^*(t) &= \int_0^t r(s)ds, \\ r^{**}(t) &= \int_0^t r^*(s)ds. \end{aligned}$$

Założmy ponownie, że w okresie długości  $h$  dostępnych jest  $M$  notowań intraday. Wówczas suma kwadratów zwrotów intraday w  $i$ -tym okresie długości  $h$  wyrażona wzorem

$$[X_\delta]_i = \sum_{j=1}^M y_{j,i}^2 = \sum_{j=1}^M \{X_{(i-1)h+\delta j} - X_{(i-1)h+\delta(j-1)}\}^2,$$

jest oszacowaniem wariancji aktualnej  $\sigma_i^2$  i nazywamy ją *wariancją zrealizowaną w przedziale*  $[(i-1)h, i\hbar)$ . Najczęściej  $h$  odpowiada okresowi jednego dnia. Wówczas wariancją zrealizowaną nazywamy *dzienną wariancją zrealizowaną w dniu*  $i$ .

W pracach Andersena i in. (2000) oraz Andersena i in. (2001) zbadano empiryczne własności szeregu  $[X_\delta]_i$  zarówno dla rynku walutowego, jak i rynku papierów wartościowych. Z badań tych wynika, że  $[X_\delta]_i$  jest zadowalająco dokładnym estymatorem wariancji  $\sigma_i^2$ .

### 2.2.1. Zależności pomiędzy wariancją zrealizowaną a zmiennością aktualną

Barndorff-Nielsen i Shephard (2004a) zbadali formalne własności różnic między wariancją zrealizowaną a zmiennością aktualną oraz własności samej wariancji zrealizowanej. Wariancję zrealizowaną możemy wyrazić w następujący sposób

$$[X_\delta]_{i,h} = \sigma_i^2 + u_i,$$

gdzie  $u_i$  oznacza błąd estymacji wariancji zrealizowanej. Ponadto wiadomo, że jeżeli  $M \rightarrow \infty$ , to  $[X_\delta]_i \rightarrow \sigma_i^2$  prawie na pewno. Rozważmy zatem przypadek skończonego  $M$ . Dla uproszczenia przyjmijmy, że  $\mu$  i  $\beta$  w równaniu (2.2) są równe 0. Wówczas średnia warunkowa zmiennej  $u_i$  względem  $\sigma_i^2$  jest równa 0, co oznacza, że wariancja zrealizowana jest nieobciążonym estymatorem wariancji aktualnej. Mamy zatem

$$\begin{aligned} \mathbb{E}([X_\delta]_{i,h}) &= \hbar \xi, \\ \text{var}([X_\delta]_{i,h}) &= \text{var}(\sigma_i^2) + \text{var}(u_i), \\ \text{cov}([X_\delta]_{i,h}, [X_\delta]_{(i+s),h}) &= \text{cov}(\sigma_i^2, \sigma_{i+s}^2). \end{aligned}$$

Ponadto

$$\sigma_{j,i}^2 = \sigma^{2*}[(i-1)\hbar + \delta j] - \sigma^{2*}[(i-1)\hbar + \delta(j-1)].$$

Wobec tego  $u_i$  ma taki sam rozkład, jak

$$\sum_{j=1}^M \sigma_{j,i}^2 (\varepsilon_{j,i}^2 - 1),$$

gdzie  $\varepsilon_{j,i}$  jest ciągiem zmiennych niezależnych o rozkładzie  $N(0,1)$  i niezależnych od  $\sigma_{j,i}^2$ . Załóżmy, że wszystkie elementy szeregu  $\sigma_{j,i}^2$  mają ten sam rozkład graniczny.

Jak łatwo zauważyć  $\mathbb{E}(\varepsilon_{j,i}^2 - 1) = 0$ , a więc

$$\text{var}(\varepsilon_{j,i}^2 - 1) = \mathbb{E}(\varepsilon_{j,i}^2 - 1)^2 = \mathbb{E}(\varepsilon_{j,i}^4 - 2\varepsilon_{j,i}^2 + 1) = 3 - 2 + 1 = 2.$$

Wobec tego przy założeniu, że wszystkie elementy szeregu  $\sigma_{j,n}^2$  mają ten sam rozkład graniczny co  $\sigma_{1,n}^2$ , mamy

$$\text{var}(u_i) = 2M \mathbb{E}[(\sigma_{1,i}^2)^2] = 2M[\text{var}(\sigma_{1,i}^2) + \mathbb{E}^2(\sigma_{1,i}^2)]. \quad (2.4)$$

Na mocy równania (2.3) wnioskujemy, że

$$\mathbb{E}(\sigma_{1,i}^2) = \frac{\hbar \xi}{M} = \delta \xi \quad \text{ i } \quad \text{var}(\sigma_{1,i}^2) = 2\omega^2 r^{**} \left( \frac{\hbar}{M} \right) = 2\omega^2 r^{**}(\delta). \quad (2.5)$$

Barndorff-Nielsen i Shephard (2004a) wyznaczyli także rozkład graniczny błędu oszacowania wariancji zrealizowanej. Jest on opisany następującym twierdzeniem granicznym.

**Twierdzenie 2.2.1.** Załóżmy, że model zmienności stochastycznej opisany jest równaniem (2.2). Ponadto przyjmijmy, że  $\sigma_t^2$  ma lokalnie ograniczoną wariację (tzn. trajektorie są ograniczone na dowolnym zwartym podprzedziale przedziału  $[0, \infty)$ ). Wtedy dla dowolnego dodatniego  $\hbar$  i  $M \rightarrow \infty$

$$\frac{u_i}{\sqrt{2 \sum_{j=1}^M (\sigma_{j,i}^2)^2}} \rightarrow N(0,1),$$

$$\text{gdzie } \sigma_{j,i}^2 = \sigma^{2*} \left[ (i-1)\hbar + \frac{\hbar j}{M} \right] - \sigma^{2*} \left[ (i-1)\hbar + \frac{\hbar(j-1)}{M} \right].$$

Ponadto

$$\hbar^{-1} M \sum_{j=1}^M (\sigma_{j,i}^2)^2 \rightarrow \sigma_i^{4*}$$

prawie na pewno, gdzie  $\sigma_i^{4*}$  nazywane jest *kwartycznością* i opisywane wzorem

$$\sigma_i^{4*} = \int_{(i-1)\hbar}^{i\hbar} \sigma^4(s) ds.$$

### 2.2.2. Estymatory wariancji zrealizowanej

Stosowana przez Andersena i Bollersleva (1998a) dzienna wariancja zrealizowana została określona jako suma kwadratów zwrotów intraday o ustalonych częstotliwościach. W wyżej wymienionej pracy analiza dotyczyła rynku walutowego i opierała się na 288 pięciominutowych zwrotach. Jeśli rozważymy wariancję zrealizowaną dla instrumentów notowanych na giełdzie, pojawia się trudność związana z tym, że giełda funkcjonuje tylko w określonych godzinach, a więc zwroty intraday poza tym okresem nie są znane. W takim przypadku wariancja zrealizowana może być rozważana jako suma kwadratów zwrotów śróddziennych i kwadratu zwrotu nocnego, tzn.

$$[X_{\delta}^{(1)}]_i = y_{0,i}^2 + \sum_{j=1}^M y_{j,i}^2,$$

gdzie  $M$  jest liczbą notowań intraday o rozważanej częstotliwości,  $y_{j,i}$  jest  $j$ -tym zwrotem intraday w  $i$ -tym dniu oraz  $y_{0,i}$  jest zwrotem nocnym uprzedzającym  $i$ -ty dzień.

Powyższy estymator nie daje niestety zbyt dobrego oszacowania wariancji zrealizowanej. Wynika to z faktu, że zwrot nocny jest pewnym szczególnym przypadkiem. Długi okres jakiemu on odpowiada powoduje znaczne obciążenie estymatora szumem. Na polskim rynku finansowym jest to wyraźnie zauważalne, ponieważ wiele informacji gospodarczych jest podawanych dopiero po godzinie 16.10, gdy giełda już nie pracuje. Ponadto uregulowania KPWiG obligują spółki do podawania istotnych informacji dopiero po zamknięciu giełdy. Również giełdy w Stanach Zjednoczonych pracują w okresie gdy Giełda Papierów Wartościowych w Warszawie jest zamknięta, a więc inwestorzy otrzymują informacje z za oceanu poza godzinami funkcjonowania giełdy. Aby uniknąć obciążenia, Andersen i in. (2001) określają miarę wariancji zrealizowanej jako sumę zwrotów śróddziennych bez uwzględnienia zwrotu nocnego. Wówczas mamy

$$[X_{\delta}^{(2)}]_i = \sum_{j=1}^M y_{j,i}^2. \quad (2.6)$$

Miara ta w powyższej pracy jest nazywana zmiennością dzienną. Praktyka pokazuje, że podejście to daje niekiedy mocno zaniżone estymacje wariancji zrealizowanej, dlatego Martens (2002) zaproponował pomnożyć  $[X_{\delta}^{(2)}]_k$  przez  $(1+c)$ , gdzie  $c$  jest pewną stałą dodatnią. Koopman i in. (2005) zaproponowali by za stałą  $c$  podstawić  $\tilde{\sigma}_{co}^2/\tilde{\sigma}_{oc}^2$ , gdzie  $\tilde{\sigma}_{co}^2$  oznacza empiryczną wariancję zwrotu między kursem zamknięcia a kursem otwarcia natomiast  $\tilde{\sigma}_{oc}^2$  jest empiryczną wariancją zwrotu dziennego tj.

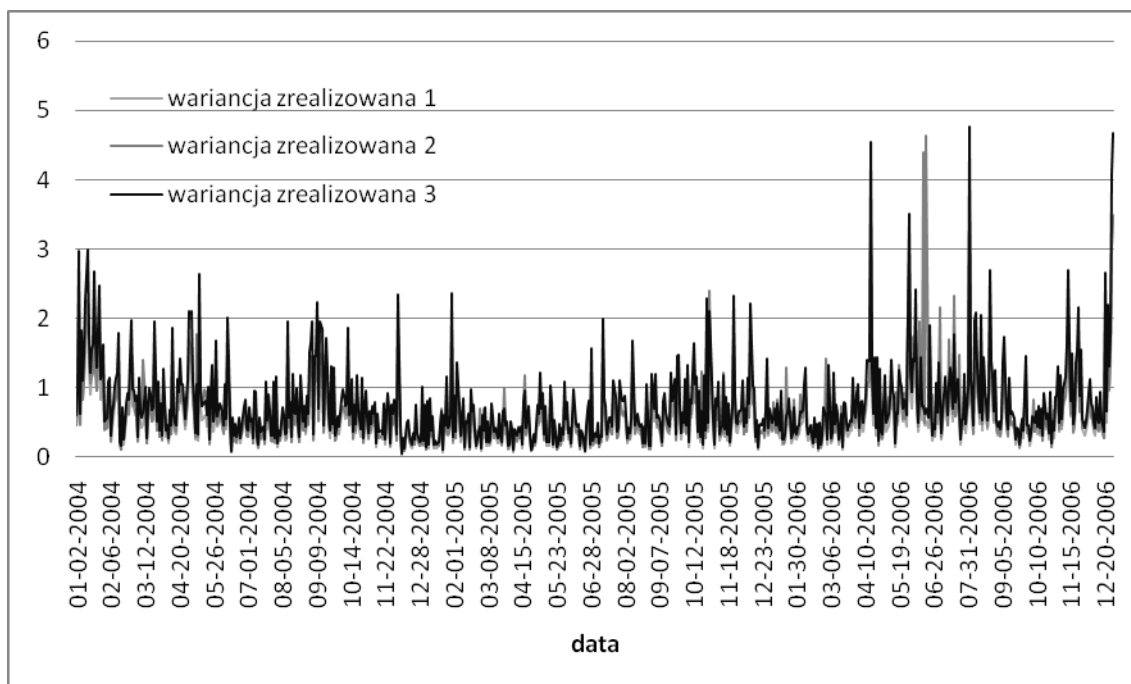
$$\tilde{\sigma}_{co}^2 = Var(r_{t,0}) \text{ oraz } \tilde{\sigma}_{oc}^2 = Var\left(\sum_{t=1}^N r_{t,n}\right).$$

Wówczas wariancja zrealizowana opisana jest za pomocą wzoru

$$[X_{\delta}^{(3)}]_i = \left(1 + \frac{\tilde{\sigma}_{co}^2}{\tilde{\sigma}_{oc}^2}\right) [X_{\delta}^{(2)}]_i.$$

Nieco inne podejście można znaleźć w pracy Hansena i Lunde'a (2002) którzy wyznaczają  $\tilde{\sigma}_{oc}^2$  jako średnią ze wszystkich zmienności zrealizowanych obliczanych ze wzoru (2.6). Inne podejście do rozwiązania problem braku notowań w ciągu nocy zostało zaproponowane przez Areala i Taylora (2002) którzy użyli różnych wag dla zwrotów intraday i zwrotu nocnego. Wagi te są uzależnione od ilorazu  $\sigma_{co}^2/\sigma_{oc}^2$ . Problematyka skalowania wariancji została szerzej omówiona w pracy Hansena i Lunde'a (2005).

Na rysunku 2.1 przedstawiona została wariancja zrealizowana indeksu WIG20 od dnia 2 stycznia 2004 do 29 grudnia 2006 uzyskana za pomocą trzech przedstawionych powyżej estymatorów.



Rys. 2.1. Wariancja zrealizowana indeksu WIG20 od dnia 2 stycznia 2004 do 29 grudnia 2006 w zależności od wykorzystanego estymatora

Stałe  $1 + \sigma_{co}^2 / \sigma_{oc}^2$  wyznaczyliśmy dla badanych w pracy szeregów indeksów giełdowych oraz kursów walutowych. Są one przedstawione w poniższej tabeli. Wartości przyjęte przez nie są dosyć zróżnicowane. Niekiedy bardzo nieznacznie przekraczają 1, a w wielu przypadkach osiągają wartości bliskie 1,5.

Tabela 2.1 Wartości współczynnika  $1 + \sigma_{co}^2 / \sigma_{oc}^2$  w szeregach indeksów CAC40, DAX, NASDAQ, WIG20 oraz kursów walutowych EURPLN i USDPLN.

Szereg notowań	CAC40	DAX	NASDAQ	WIG20	EURPLN	USDPLN
stała	1,330687	1,025948	1,264046	1,326895	1,2222	1,249934

### 2.3. Wariacje wielopotęgowe

Jak już wyżej wspomnieliśmy, dane wysokiej częstotliwości znajdują zastosowanie w szacowaniu zmienności cen logarymicznych, poprzez wyznaczenie jej estymatora – wariancji zrealizowanej. W tym podrozdziale zajmiemy się analizą uogólnienia tego pojęcia – zrealizowanej wariacji potęgowej i wielopotęgowej, które są odpowiednio sumą potęg i iloczynów potęg bezwzględnych zwrotów. Wskażemy także ich zastosowanie do wyznaczania skoków w procesach cen logarymicznych.

### 2.3.1. Wariacja potęgowa

Pojęcie wariacji potęgowej zostało wprowadzone przez Barndorff-Nielsen i Shepharda (2003). Załóżmy, że w okresie  $[(i-1)h, ih)$  dostępnych jest  $M$  notowań w odstępach czasowych  $\delta$ . Rozważmy przypadek, w którym  $\delta \rightarrow 0$ . Wówczas *wariację potęgową* dla  $r \geq 0$  definiujemy w następujący sposób

$$[X]^{[r]}(t) = \text{plim}_{\delta \rightarrow 0} \sum_{j=1}^{\lfloor t/\delta \rfloor} |y_j(\delta)|^r,$$

o ile granica istnieje, gdzie  $y_j(\delta) = X(j\delta) - X((j-1)\delta)$ . Jeśli  $\delta > 0$ , to wariację potęgową można przybliżyć *zrealizowaną wariacją potęgową*, która została już wprowadzona przez Barndorff-Nielsen i Shepharda (2003) i wyraża się wzorem

$$[X_\delta]^{[r]}(t) = \sum_{j=1}^{\lfloor t/\delta \rfloor} |y_j(\delta)|^r.$$

Unormowane wersje wariacji potęgowej oraz zrealizowanej wariacji potęgowej dane są następującymi wzorami

$$\begin{aligned} \{X\}^{\{r\}}(t) &= \text{plim}_{\delta \rightarrow 0} \delta^{1-r/2} \sum_{j=1}^{\lfloor t/\delta \rfloor} |y_j(\delta)|^r, \\ \{X_\delta\}^{\{r\}}(t) &= \delta^{1-r/2} \sum_{j=1}^{\lfloor t/\delta \rfloor} |y_j(\delta)|^r. \end{aligned}$$

Jak łatwo zauważyć, podstawiając  $r = 2$ , uzyskujemy wariancję zrealizowaną, natomiast dla  $r = 1$ , wprowadzoną przez Andersena i Bollersleva (1998b) zrealizowaną wariację bezwzględną. Ponadto dla wariacji potęgowej zachodzi następujące twierdzenie, którego dowód można znaleźć w pracy Barndorff-Nielsen i Shepharda (2003).

**Twierdzenie 2.3.1.** Jeżeli  $X \in SVSM^c$  oraz  $\alpha^*$  i  $\sigma$  są niezależne od ruchu Browna  $B_t$ , to

$$\{X\}^{\{r\}}(t) = \mu_r \int_0^t \sigma^r(s) ds,$$

gdzie

$$\mu_r = \mathbb{E} |u|^r = 2^{r/2} \frac{\Gamma\left(\frac{1}{2}(r+1)\right)}{\Gamma\left(\frac{1}{2}\right)}$$

dla  $r > 0$  oraz  $u \sim N(0,1)$ .

### 2.3.2. Wariacja dwupotęgową

Pojęcie wariacji dwupotęgowej jest uogólnieniem pojęcia opisanej powyżej wariacji potęgowej. Proces wariacji kwadratowej istnieje zawsze, gdy  $X$  jest semimartyngelem. Warunek ten nie jest koniecznym dla wariacji dwupotęgowej. Wariacją dwupotęgową procesu  $X$  dla  $r \geq 0$  i  $s \geq 0$  nazywamy wyrażenie

$$\{X\}^{\{r,s\}}(t) = \text{plim}_{\delta \rightarrow 0} \delta^{1-(r+s)/2} \sum_{j=1}^{\lfloor t/\delta \rfloor - 1} |y_j(\delta)|^r |y_{j-1}(\delta)|^s,$$

o ile granica istnieje. Ponadto zachodzi następujące twierdzenie (Barndorff-Nielsen i Shephard 2004):

**Twierdzenie 2.3.2.** Jeżeli  $X \in SVSM^c$ ,  $\alpha^* = 0$  oraz  $\sigma$  jest niezależne od ruchu Browna  $B_t$ , to

$$\{X\}^{\{r,r\}}(t) = \mu_r^2 \int_0^t \sigma^{2r}(s) ds.$$

Ponieważ pojęcie wariacji dwupotęgowej leży u podstaw testowania skoków w procesach cen logarytmicznych, podajemy definicję nowej klasy procesów stochastycznych, jakimi są semimartyngały zmienności stochastycznej ze skokami.

**Definicja 2.3.1.** Proces  $X$  należy do klasy semimartyngałów zmienności stochastycznej ze skokami (SVJSM), jeśli

$$X_t = \alpha^*(t) + \int_0^t \sigma(s) dB_s + \int_0^t c(s) dq_s,$$

gdzie  $\alpha^*$  jest procesem z ciągłymi trajektoriami i ze skończoną wariacją,  $\sigma$  jest procesem z trajektoriami prawostronnie ciągłymi z lewostronnymi granicami,  $c(t)$  jest niezzerową zmienną losową reprezentującą wielkość skoku w chwili  $t$ , natomiast  $q_t$  jest pro-



cesem liczącym, wobec czego  $dq_t$  jest skokowym procesem przyjmującym wartość 1 w sytuacji, gdy w chwili  $t$  wystąpi skok i 0 w przeciwnym wypadku. Ponadto dla dowolnego skończonego  $t$  zachodzi

$$\int_0^t \sigma^2(s) ds < \infty \text{ oraz } q_t < \infty.$$

Jeśli  $q_t = 0$  dla dowolnego  $t > 0$ , to stosuje się notację  $X \in SVSM^c$ . Jeżeli natomiast  $\alpha^*(t) = 0$  dla dowolnego  $t > 0$ , to mówimy, że  $X$  jest *procesem zmienności stochastycznej ze skokami* ( $SVJ$ ). Jeżeli zarówno  $q_t = 0$  jak i  $\alpha^*(t) = 0$  dla dowolnego  $t > 0$ , to  $X$  jest procesem zmienności stochastycznej ( $X \in SV$ ).

Wariacja kwadratowa zdefiniowanego powyżej semimartynała zmienności stochastycznej ze skokami jest równa

$$[X]_t = \int_0^t \sigma^2(s) ds + \int_0^t c(s) dq_s,$$

gdzie  $dq_s$  jest procesem przyjmującym wartość 1, gdy występuje skok oraz 0 w przeciwnym wypadku. Z drugiej strony na podstawie twierdzenia 2.4.2 mamy, że jeżeli  $X$  jest procesem zmienności stochastycznej ze skokami i  $\sigma$  jest niezależne od ruchu Browna, to

$$\{X\}^{[1,1]}(t) = \mu_1^{-2} \int_0^t \sigma^2(s) ds.$$

Stąd  $\mu_1^{-2} \{X\}_t^{[1,1]}$  i  $[X]_t$  zwracają te same wartości z dokładnością do pewnego znanego czynnika w przypadku podstawowego modelu zmienności stochastycznej, ale inne w przypadku modelu dopuszczającego skoki. Na podstawie powyższych wzorów łatwo zauważyć, że

$$[X]_t - \mu_1^{-2} \{X\}_t^{[1,1]} = \int_0^t c^2(s) dq_s.$$

W praktyce, z uwagi na dostępność jedynie dyskretnego zbioru notowań, proces wariacji dwupotęgowej nie jest obserwowalny. Może on być estymowany przez proces zrealizowanej wariacji dwupotęgowej określony wzorem

$$\{X_\delta\}^{[r,s]}(t) = \delta^{1-(r+s)/2} \sum_{j=1}^{\lfloor t/\delta \rfloor - 1} |y_j(t)|^r |y_{j-1}(t)|^s.$$

Barndorff-Nielsen i Shephard (2004c) wykazali następujące twierdzenie.

**Twierdzenie 2.3.3.** Jeżeli  $X \in SVSM^c$ ,  $\alpha^* = 0$ ,  $\sigma$  jest niezależne od ruchu Browna  $B_t$  oraz spełniony jest następujący warunek regularności

$$\lim_{\delta \rightarrow 0} \delta^{1/2} \sum_{j=1}^M |\sigma^r(\eta_j) - \sigma^r(\xi_j)| = 0$$

dla pewnego  $r > 0$  i dowolnych  $\xi_j(\delta)$  i  $\eta_j(\delta)$  takich, że

$$0 \leq \xi_1 \leq \eta_1 \leq \delta \leq \xi_2 \leq \eta_2 \leq 2\delta \leq \dots \leq \xi_M \leq \eta_M \leq t,$$

to

$$\{X_\delta\}^{[r,r]}(t) = \mu_r^2 \int_0^t \sigma^{2r}(s) ds + o_p(\delta^{1/2}).$$

Niech

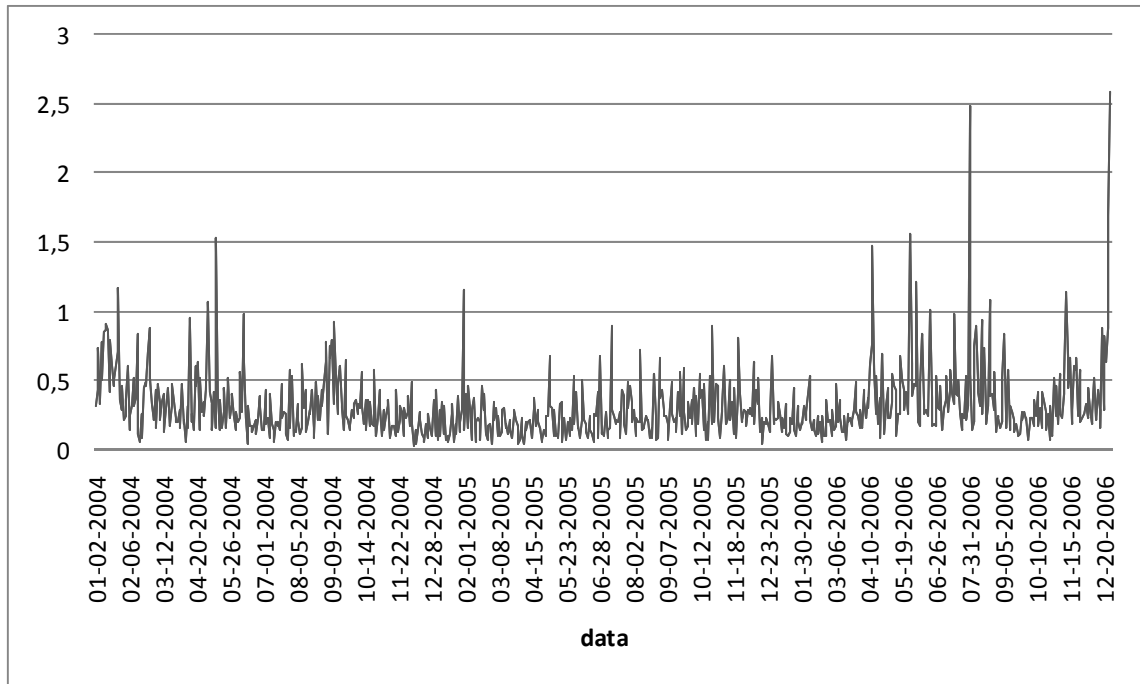
$$\{X_\delta\}_i^{[r,s]} = \{X_\delta\}^{[r,s]}(ih) - \{X_\delta\}^{[r,s]}((i-1)h)$$

oznacza przyrost zrealizowanej wariacji dwupotęęgowej w dniu  $i$ . Estymując dodatkowo wariację kwadratową za pomocą wariacji zrealizowanej mamy

$$[X_\delta]_i - \mu_1^{-2} \{X_\delta\}_i^{[1,1]} = \int_{(i-1)h}^{ih} c(t) dq_t.$$

Z powyższej równości wynika, że wariacja dwupotęęgowa może być wykorzystana do estymacji ciągłych i skokowych składników wariacji kwadratowej. Ponadto, jest użyteczna przy testowaniu hipotezy o ciągłości trajektorii.

Dzienne przyrosty zrealizowanej wariacji dwupotęęgowej  $\{X_\delta\}_i^{[1,1]}$  indeksu WIG20 od dnia 2 stycznia 2004 do 29 grudnia 2006 przedstawione są na rysunku 2.2.



Rys. 2.2. Dienne przyrosty zrealizowanej wariacji dwupotęowej z parametrami [1,1] indeksu WIG20 od dnia 2 stycznia 2004 do 29 grudnia 2006

### 2.3.3. Uogólnienie wariacji dwupotęowej

Barndorff-Nielsen i Shephard (2004b) wprowadzili wariację wielopotęową, która jest uogólnieniem zdefiniowanego w powyższych paragrafach pojęcia wariacji potęowej i dwupotęowej. Wariację wielopotęową oraz jej unormowaną wersję można zapisać za pomocą następujących wzorów:

$$[X]^{[r]}(t) = \text{plim}_{\delta \rightarrow 0} \sum_{j=1}^{[t/\delta]-m} |y_j(\delta)|^{r_1} |y_{j+1}(\delta)|^{r_2} \cdots |y_{j+m-1}(\delta)|^{r_m},$$

$$\{X\}^{[r]}(t) = \text{plim}_{\delta \rightarrow 0} \delta^{1-(r_1+r_2+\dots+r_m)/2} \sum_{j=1}^{[t/\delta]-m} |y_j(\delta)|^{r_1} |y_{j+1}(\delta)|^{r_2} \cdots |y_{j+m-1}(\delta)|^{r_m}.$$

Zrealizowany proces wariacji wielopotęowej można z kolei zapisać wzorem

$$[X_\delta]^{[r]}(t) = \sum_{j=1}^{[t/\delta]-m} |y_j(\delta)|^{r_1} |y_{j+1}(\delta)|^{r_2} \cdots |y_{j+m-1}(\delta)|^{r_m}$$

oraz w wersji unormowanej

$$\{X_\delta\}^{[r]}(t) = \delta^{1-(r_1+r_2+\dots+r_m)/2} [X_\delta]^{[r]}(t) = \delta^{1-(r_1+r_2+\dots+r_m)/2} \sum_{j=1}^{[t/\delta]-m} |y_j(\delta)|^{r_1} |y_{j+1}(\delta)|^{r_2} \cdots |y_{j+m-1}(\delta)|^{r_m},$$

gdzie  $[\mathbf{r}] = [r_1, r_2, \dots, r_m]$ .

Oznaczmy, odpowiednio, przez

$$\{X\}_i^{[\mathbf{r}]} = \{X\}^{[\mathbf{r}]}(i\hbar) - \{X\}^{[\mathbf{r}]}((i-1)\hbar),$$

oraz

$$\{X_\delta\}_i^{[\mathbf{r}]} = \{X_\delta\}^{[\mathbf{r}]}(i\hbar) - \{X_\delta\}^{[\mathbf{r}]}((i-1)\hbar)$$

przyrosty wariacji wielopotęgowej i zrealizowanej wariacji wielopotęgowej w dniu  $i$ . Analogicznie jak dla wariacji potęgowej zachodzi twierdzenie.

**Twierdzenie 2.3.4.** Jeżeli  $X \in SVSM^c$  oraz  $\alpha^*$  i  $\sigma$  są niezależne od ruchu Browna  $B_t$ , przy czym  $\max(r_1, r_2, \dots, r_m) < 2$ , to

$$\{X\}_i^{[\mathbf{r}]} = \left( \prod_{k=1}^m \mu_{r_k} \right) \int_{(i-1)\hbar}^{i\hbar} \sigma^{r_1+r_2+\dots+r_m}(t) dt.$$

Powyższe twierdzenie znajduje szerokie zastosowania. Szczególne znaczenie mają tutaj zrealizowana wariacja trójpotęgowa i czteropotęgowa, które opisane są odpowiednio wzorami

$$\{X\}^{[r,s,u]}(t) = \text{plim}_{\delta \rightarrow 0} \delta^{1-(r+s+u)/2} \sum_{j=1}^{[t/\delta]-2} |y_j(\delta)|^r |y_{j+1}(\delta)|^s |y_{j+2}(\delta)|^u \text{ dla } r, s, u > 0$$

i

$$\{X\}^{[r,s,u,v]}(t) = \text{plim}_{\delta \rightarrow 0} \delta^{1-(r+s+u+v)/2} \sum_{j=1}^{[t/\delta]-3} |y_j(\delta)|^r |y_{j+1}(\delta)|^s |y_{j+2}(\delta)|^u |y_{j+3}(\delta)|^v$$

dla  $r, s, u, v > 0$ .

Jeżeli podstawimy  $r = s = u = 2/3$  do wzoru na wariację trójpotęgową, to na mocy twierdzenia 2.3.4, otrzymujemy

$$\{X\}^{[2/3, 2/3, 2/3]}(t) = \text{plim}_{\delta \rightarrow 0} \sum_{j=1}^{[t/\delta]-2} |y_j(\delta)|^{2/3} |y_{j+1}(\delta)|^{2/3} |y_{j+2}(\delta)|^{2/3} = \mu_{2/3}^3 \int_0^t \sigma^2(t) dt.$$

Podobnie

$$\{X\}^{[1/2, 1/2, 1/2, 1/2]}(t) = \text{plim}_{\delta \rightarrow 0} \sum_{j=1}^{[t/\delta]-3} |y_j(\delta)|^{1/2} |y_{j+1}(\delta)|^{1/2} |y_{j+2}(\delta)|^{1/2} |y_{j+3}(\delta)|^{1/2} = \mu_{1/2}^4 \int_0^t \sigma^2(t) dt.$$

W ten sposób uzyskaliśmy dwa dodatkowe estymatory wariancji scałkowanej. Podobnie otrzymać można estymatory kwartyczności wykorzystywanej przez Barndorff-Nielsen i Shepharda (2006) przy konstrukcji statystyk testów na występowanie skoków. Jeżeli podstawimy  $r = s = u = 4/3$  do wzoru na wariację trójpotęgową, to otrzymamy

$$\{X\}^{[4/3,4/3,4/3]}(t) = \text{plim}_{\delta \rightarrow 0} \delta^{-1} \sum_{j=1}^{[t/\delta]-2} |y_j(\delta)|^{4/3} |y_{j+1}(\delta)|^{4/3} |y_{j+2}(\delta)|^{4/3} = \mu_{4/3}^3 \int_0^t \sigma^4(t) dt, \quad (2.7)$$

natomiast podstawiając  $r = s = u = v = 1$  do wzoru na wariację czteropotęgową, mamy

$$\{X\}^{[1,1,1,1]}(t) = \text{plim}_{\delta \rightarrow 0} \delta^{-1} \sum_{j=1}^{[t/\delta]-3} |y_j(\delta)|^1 |y_{j+1}(\delta)|^1 |y_{j+2}(\delta)|^1 |y_{j+3}(\delta)|^1 = \mu_1^4 \int_0^t \sigma^2(t) dt. \quad (2.8)$$

Te dwa podejścia zostały wykorzystane przez Huanga i Tauchena (2005) oraz Barndorff-Nielsen i Shepharda (2006) przy wyznaczaniu statystyk testów na występowanie skoków.

### 3. Wykorzystanie wariancji zrealizowanej do estymacji procesów stochastycznych z czasem ciągłym

Modele finansowych procesów stochastycznych z czasem ciągłym zakładają, że logarytmiczne ceny instrumentów finansowych są opisane przez stochastyczne równania różniczkowe. Podejście to, zapoczątkowane przez Bacheliera (1900) było rozwijane w kilku pracach w pierwszej połowie dwudziestego wieku. Niemniej jednak zastosowanie w finansach znalazło dopiero w latach siedemdziesiątych dwudziestego wieku. Wtedy bowiem okazało się, że przy założeniu zupełności rynku, modele z czasem ciągłym mogą być bardzo przydatne do konstrukcji strategii zabezpieczających i wyceny instrumentów pochodnych. Dużym problemem pozostaje wciąż szybka i precyzyjna estymacja parametrów procesów z czasem ciągłym.

Niebayesowskie metody estymacji parametrów są stosunkowo nowym zagadnieniem. W pracy Lánskiej (1979) przedstawiona jest metoda pozwalająca na estymację parametrów równania różniczkowego w sytuacji, gdy funkcja dyfuzji jest znana. Pierwsze elementarne metody estymacji parametrów funkcji dyfuzji zostały wprowadzone przez Florens-Zmirou (1989) i Yoshidę (1992). Metody te dają jednak bardzo niedokładne oszacowania parametrów. Aït Sahalia (1999,2002) wykorzystał do estymacji parametrów stochastycznego równania różniczkowego metodę największej wiarygodności, stosując wielomiany Hermite'a do wyznaczenia gęstości przejścia. Wyestymował on parametry sześciu różnych modeli dyfuzji dla szeregu miesięcznych notowań stopy procentowej FED. W późniejszej pracy Aït Sahalia (2008) zaproponował alternatywne podejście, wyznaczając gęstości przejścia z przybliżonych równań Fokkera, Plancka i Kołmogorowa.

W niniejszym rozdziale główny nacisk zostanie położony na dwuetapowe podejście wprowadzone przez Phillipsa i Yu (2007). Metoda ta pozwala estymować parametry stochastycznego równania różniczkowego postaci

$$dX_t = \mu(X_t, \theta_1)dt + \sigma(X_t, \theta_2)dB_t, \quad (3.1)$$

gdzie  $B_t$  jest standardowym ruchem Browna,  $\mu(X_t, \theta_1)$  jest funkcją dryfu,  $\sigma(X_t, \theta_2)$  jest funkcją dyfuzji, natomiast  $\theta = (\theta_1, \theta_2)'$  jest wektorem  $k_1 + k_2$  parametrów. W pierwszym kroku estymowane są parametry funkcji dyfuzji przy wykorzystaniu, rozwiniętej przez Barndorff-Nielsen i Shepharda (2004a) teorii centralnych twierdzeń granicznych dla wariancji zrealizowanej. W drugim etapie, przy danym oszacowaniu wektora parametrów  $\theta_2$ , konstruowana jest aproksymacyjna funkcja wiarygodności, za pomocą której estymuje się wektor parametrów funkcji dryfu. Dzienna wariancja zrealizowana instrumentu finansowego rozumiana jest tutaj jako suma kwadratów zwrotów śróddziennych o ustalonej częstotliwości.

### 3.1. Prezentacja wybranych modeli

Do modelowania logarytmicznych cen instrumentów finansowych wykorzystujemy popularne modele dyfuzji i zmienności stochastycznej. Większość z nich była pierwotnie stosowana głównie do modelowania stóp procentowych. Istotną cechą wszystkich wykorzystanych w pracy modeli jest zdefiniowana poniżej własność powracania procesu do średniej.

**Definicja 3.1.1.** Mówimy, że proces  $X_t$  *powraca do średniej*, jeśli istnieje skończona granica

$$\lim_{t \rightarrow \infty} \mathbb{E}(X_t | X_0 = x_0).$$

Najbardziej elementarnym modelem, który wykorzystamy do estymacji procesów finansowych jest model Vašíčka (1977).

#### 3.1.1. Model Vašíčka

Model Vašíčka (1977) zakłada, że proces modelowany jest przez następujące stochastyczne równanie różniczkowe

$$\begin{cases} dX_t = \kappa(\mu - X_t)dt + \sigma dB_t \\ X_{t_0} = x_0 \end{cases}, \quad (3.2)$$

gdzie parametry  $\kappa$ ,  $\mu$  i  $\sigma$  przyjmują wartości dodatnie. Do wartości parametru  $\mu$  zbiega wartość oczekiwana procesu  $X_t$  przy  $t \rightarrow \infty$ , natomiast wartość parametru  $\kappa$  może być interpretowana jako tempo powracania procesu do średniej. Parametr  $\sigma$  określa z kolei zmienność procesu  $X_t$ . Zasadniczą wadą modelu jest możliwość przyjęcia przez modelowany proces wartości ujemnych. Słabości tej nie posiada między innymi model Coxa, Ingersolla i Rossa, w którym funkcja dyfuzji uzależniona jest od bieżącej wartości procesu. Rozwiązaniem stochastycznego równania różniczkowego (3.2) jest następujący proces

$$X_t = \mu + (X_{t_0} - \mu)e^{-\kappa(t-t_0)} + \sigma \int_{t_0}^t e^{-\kappa(t-u)} dB_u.$$

Ponadto warunkowa wartość oczekiwana i wariancja w chwili  $t$  przy założeniu znanej filtracji  $\mathcal{F}_s$  dane są wzorami

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s e^{-\kappa(t-s)} + \mu(1 - e^{-\kappa(t-s)})$$

oraz

$$Var(X_t | \mathcal{F}_s) = \frac{\sigma^2}{\kappa} (1 - e^{-2\kappa(t-s)}). \quad (3.3)$$

### 3.1.2. Model Coxa, Ingersolla i Rossa

Cox, Ingersoll i Ross (1985) zaproponowali następujący model, nazywany w skrócie CIR, będący następującą modyfikacją modelu Vašíčka (1977)

$$\begin{cases} dX_t = \kappa(\mu - X_t)dt + \sigma\sqrt{X_t}dB_t, \\ X_{t_0} = x_0 \end{cases},$$

gdzie parametry  $\kappa$ ,  $\mu$  i  $\sigma$  przyjmują wartości dodatnie. Parametry te mają taką samą interpretację, jak w modelu Vašíčka. Umieszczenie pierwiastka kwadratowego w funkcji dyfuzji powoduje, że modelowany proces przyjmuje wartości nieujemne, jeśli  $2\kappa\mu > \sigma^2$ . Rozwiązaniem stochastycznego równania różniczkowego Coxa, Ingersolla i Rossa jest następujący proces

$$X_t = \mu + (X_{t_0} - \mu)e^{-\kappa(t-t_0)} + \sigma \int_{t_0}^t e^{-\kappa(t-u)} \sqrt{X_u} dB_u.$$



W przeciwieństwie do równania Vašíčka, rozwiązanie równania Coxa, Ingersolla i Rossa nie posiada postaci jawnej. Znana jest natomiast postać warunkowej wartości oczekiwanej i wariancji w chwili  $t$ , które dane są wzorami

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s e^{-\kappa(t-s)} + \mu(1 - e^{-\kappa(t-s)})$$

oraz

$$\text{Var}(X_t | \mathcal{F}_s) = X_s \frac{\sigma^2}{2\kappa} (e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}) + \frac{\mu\sigma^2}{2\kappa} (1 - e^{-\kappa(t-s)})^2.$$

### 3.1.3. Model CKLS

Chan i in. (1992) wprowadzili model, nazywany w skrócie od pierwszych liter nazwisk autorów *modelem CKLS*. Autorzy założyli, że proces modelowany jest przez następujące różniczkowe równanie stochastyczne

$$\begin{cases} dX_t = \kappa(\mu - X_t)dt + \sigma X_t^\beta dB_t \\ X_{t_0} = x_0 \end{cases}$$

gdzie  $\beta \geq 0$  nazywany jest parametrem elastyczności wariancji, natomiast parametry  $\kappa$ ,  $\mu$  i  $\sigma$  przyjmują wartości dodatnie i mają taką samą interpretację, jak w modelach Vašíčka oraz Coxa, Ingersolla i Rossa. Model ten jest uogólnieniem przedstawionych tu dotychczas modeli dyfuzji. Nakładając restrykcje na poszczególne parametry można uzyskać 7 różnych modeli dyfuzji. Podstawiając  $\gamma = 0$  i  $\gamma = 0,5$  otrzymujemy, odpowiednio, opisane powyżej modele Vašíčka i Coxa, Ingersolla i Rossa. Model wprowadzony przez Dothana (1978) otrzymamy ustalając wartości parametrów  $\kappa$ ,  $\mu$  na 0, oraz  $\beta$  na 1. Podstawiając  $\mu = 0$  oraz  $\beta = 1$  otrzymujemy model nazywany geometrycznym ruchem Browna, wykorzystany do opisu ceny akcji przez Blacka i Scholesa (1973). Podstawiając  $\gamma = 1$  otrzymamy model Brennana and Schwartz (1980), wykorzystany przez autorów przy konstrukcji modelu wyceny obligacji zamiennych oraz przez Courtadona (1982) przy wycenie opcji na obligacje. Ustalając  $\gamma = 1,5$ ,  $\kappa = 0$  oraz  $\mu = 0$  otrzymujemy wprowadzony przez Coxa, Ingersolla i Rossa (1980) model kontraktów kredytowych ze zmienną stopą procentową. Nakładając warunek  $\mu = 0$ , uzyskujemy wprowadzony przez Coxa (1975) model stałej elastyczności wariancji, nazywany w skrócie CEV.

### 3.1.4. Model Hestona

Stosowany powszechnie do wyceny opcji model Blacka-Scholesa (1973) zakłada, że cena instrumentu podstawowego opisana jest następującym stochastycznym równaniem różniczkowym

$$\begin{cases} dX_t = \mu X_t dt + \sigma X_t dB_t \\ X_{t_0} = x_0 \end{cases},$$

gdzie  $\mu$  i  $\sigma$  są stałe. W celu uwzględnienia zmian współczynnika zmienności  $\sigma$  w czasie, Heston (1993) wprowadził model wyceny opcji, będący rozwinięciem koncepcji Blacka-Scholesa, w którym cena instrumentu podstawowego opisana jest następującym modelem zmienności stochastycznej

$$\begin{cases} dX_t = \mu X_t dt + \sigma(t) X_t dB_t^{(1)}, \\ d\sigma^2(t) = \theta(\omega - \sigma^2(t))dt + \xi\sigma(t)dB_t^{(2)}, \\ \sigma^2(0) = \sigma_0^2, X_0 = x_0, \end{cases} \quad (3.4)$$

gdzie parametr  $\xi$  może być interpretowany jako współczynnik zmienności procesu  $\sigma^2(t)$ ,  $\omega$  jest wielkością, do której zbiega wartość oczekiwana procesu  $\sigma^2(t)$  przy  $t \rightarrow \infty$ , natomiast wartość parametru  $\theta$  może być interpretowana jako tempo powracania procesu  $\sigma^2(t)$  do średniej.

## 3.2. Estymacja parametrów stochastycznych równań różniczkowych Itô

Oznaczmy przez  $X$  proces logarytmicznych notowań danego instrumentu finansowego i przyjmijmy, że jest on opisany stochastycznym równaniem różniczkowym postaci (3.1). Załóżmy, że  $X_t$  jest szeregiem znanych logarytmicznych notowań danego instrumentu finansowego dla

$$t = \delta, 2\delta, \dots, M\delta, (M+1)\delta, \dots, 2M\delta, \dots, n\delta = T,$$

gdzie  $M$  jest liczbą notowań intraday w ciągu jednego dnia,  $n = KM$  jest liczbą wszystkich dostępnych notowań,  $\delta$  jest częstotliwością notowań, natomiast  $K$  jest liczbą dni. Ponadto, podobnie jak w rozdziale 2, przez  $h = M\delta$  oznaczymy długość jednego dnia.

### 3.2.1. Metody elementarne

Założmy, że funkcja dyfuzji nie zależy od żadnych nieznanymi parametrów, tzn.  $\sigma(X_t, \theta_2) = \sigma(X_t)$ . W takim przypadku możemy stosować dokładną postać logarytmicznej funkcji wiarygodności  $L(\theta_1)$ , wykorzystując w tym celu twierdzenia Girsanowa. Założmy, że miara probabilistyczna  $\mathbb{Q}$  jest równoważna mierze  $\mathbb{P}$ , natomiast  $B_t$  jest ruchem Browna ze względu na miarę  $\mathbb{P}$ . Oznaczmy

$$\gamma_t = \frac{\mu(X_t, \theta_1)}{\sigma(X_t)}.$$

Z twierdzenia Girsanowa wnosimy, że jeżeli spełniony jest warunek

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}_T = \exp\left(\int_0^T \gamma_s dB_s - \frac{1}{2} \int_0^T \gamma_s^2 ds\right), \quad (3.5)$$

to  $\tilde{B}_t = B_t - \int_0^t \gamma_s ds$  jest ruchem Browna względem miary probabilistycznej  $\mathbb{Q}$ . Bezpośrednio ze wzoru (3.5) mamy

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{P}} &= \exp\left(\int_0^T \frac{\mu(X_t, \theta_1)}{\sigma(X_t)} dB_t - \frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t)} dt\right) \\ &= \exp\left(\int_0^T -\frac{\mu(X_t, \theta_1)}{\sigma(X_t)} \left(d\tilde{B}_t - \frac{\mu(X_t, \theta_1)}{\sigma(X_t)} dt\right) - \frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t)} dt\right) \\ &= \exp\left(\frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t)} ds - \int_0^T \frac{\mu(X_t, \theta_1)}{\sigma(X_t)} d\tilde{B}_t\right). \end{aligned}$$

Ponieważ  $dX_t = \sigma(X_t) d\tilde{B}_t$ , otrzymany powyżej wzór możemy zapisać w następującej postaci

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t)} dt - \int_0^T \frac{\mu(X_t, \theta_1)}{\sigma^2(X_t)} dX_t\right).$$

Stąd

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \exp\left(\int_0^T \frac{\mu(X_t, \theta_1)}{\sigma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t)} dt\right).$$

Wobec tego

$$L(\theta_1) = \ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) = \int_0^T \frac{\mu(X_t, \theta_1)}{\sigma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t)} dt. \quad (3.6)$$

Założenie o dostępności ciągłej trajektorii danych nie jest realistyczne. W zastosowaniach finansowych dysponujemy zwykle szeregiem danych śróddziennych. Dlatego logarytmiczną funkcję wiarygodności (3.6) będziemy aproksymować następującym wyrażeniem

$$L_{IF}(\theta_1) = \sum_{j=2}^n \frac{\mu(X_{(j-1)\delta}, \theta_1)}{\sigma^2(X_{(j-1)\delta})} (X_{j\delta} - X_{(j-1)\delta}) - \frac{\delta}{2} \sum_{j=2}^n \frac{\mu^2(X_{(j-1)\delta}, \theta_1)}{\sigma^2(X_{(j-1)\delta})}.$$

W praktyce, założenie, że znana jest funkcja dyfuzji, również nie jest realistyczne. Okazuje się jednak, że parametry funkcji dryfu i dyfuzji można estymować oddzielnie (Bandi i Phillips 2007). Ponadto estymacja parametrów funkcji dyfuzji staje się stosunkowo łatwa, jeżeli założymy, że  $\sigma(X_t, \theta_2)$  jest funkcją jednorodną ze względu na  $\theta_2$ , tzn.

$$\sigma(X_t, \theta_2) = \theta_2 f(X_t).$$

Gdy znane jest oszacowanie  $\hat{\theta}_2$ , to funkcja wiarygodności wyznaczona za pomocą twierdzenia Girsanowa wyraża się wzorem

$$L_{IF}(\theta_1) = \int_0^T \frac{\mu(X_t, \theta_1)}{\sigma^2(X_t, \hat{\theta}_2)} dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(X_t, \theta_1)}{\sigma^2(X_t, \hat{\theta}_2)} dt. \quad (3.7)$$

Jeśli założyć, że funkcja dyfuzji nie jest zależna od parametrów, to rozwiązanie zagadnienia staje się trywialne, niemniej taki model nie najlepiej opisywałby proces. Model ten można łatwo zmodyfikować, przyjmując, że  $\sigma(X_t, \theta_2) = \theta_2 f(X_t)$ . Wówczas mamy

$$[X]_T = \int_0^T (dX_s)^2 = \int_0^T \theta_2^2 f^2(X_s) ds \approx \theta_2^2 \delta \sum_{j=2}^n f^2(X_{(j-1)\delta}).$$

Wobec tego

$$\hat{\theta}_2^2 = \frac{[X]_T}{\delta \sum_{j=2}^n f^2(X_{(j-1)\delta})}. \quad (3.8)$$

Następnie wystarczy wyestymować parametry funkcji dryfu za pomocą funkcji wiarygodności (3.2), która przyjmie teraz postać

$$L_{IF}(\theta_1) = \sum_{j=2}^n \frac{\mu(X_{(j-1)\delta}, \theta_1)}{\hat{\theta}_2^2 f^2(X_{(j-1)\delta})} (X_{j\delta} - X_{(j-1)\delta}) - \frac{\delta}{2} \sum_{j=2}^n \frac{\mu^2(X_{(j-1)\delta}, \theta_1)}{\hat{\theta}_2^2 f^2(X_{(j-1)\delta})}.$$

Wykorzystana w tym podejściu funkcji dyfuzji może zostać zastosowana do wielu modeli krótkoterminowej stopy procentowej, w szczególności do bardzo popularnego modelu CIR (Cox i in. 1985). Yoshida (1992) zmodyfikował estymator (3.8), przybliżając wariację kwadratową  $[X]_T$  uśrednioną wartością kwadratów zwrotów śróddziennych. Wówczas

$$\hat{\theta}_2^2 = \frac{1}{T} \sum_{j=2}^n \frac{(X_{j\delta} - X_{(j-1)\delta})^2}{f^2(X_{(j-1)\delta})}. \quad (3.9)$$

### 3.2.2. Metoda największej wiarygodności

Niech  $p(\cdot | \cdot, \theta)$  będzie gęstością przejścia procesu  $X_t$ . Wówczas dzięki własności Markowa procesu opisanego równaniem (3.1), logarytmiczna funkcja wiarygodności wyraża się wzorem

$$\sum_{i=2}^n \ln(p(X_{i\delta} | X_{(i-1)\delta}, \theta)). \quad (3.10)$$

Niestety bardzo często nie jest możliwe wyznaczenie gęstości przejścia w postaci jawnej i dlatego konieczne jest zastosowanie przybliżonej funkcji największej wiarygodności. Dla stochastycznych równań różniczkowych Vašíčka i Coxa, Ingersolla i Rossa, problem ten nie istnieje. W pracy Vašíčka (1977) znajdujemy wzór na gęstość przejścia, który możemy zapisać w następujący sposób

$$p(X_{i\delta} | X_{(i-1)\delta}, \theta) = \phi(m, v),$$

gdzie  $\phi(\cdot, \cdot)$  jest gęstością rozkładu normalnego natomiast

$$m = \mu + (X_{(i-1)\delta} - \mu)e^{-\kappa\delta} \quad \text{oraz} \quad v = \sigma^2 \frac{1 - e^{-2\kappa\delta}}{2\kappa}.$$

W celu uniknięcia nadmiaru zmiennoprzecinkowego, w zastosowaniach praktycznych najlepiej sprowadzić funkcję wiarygodności do następującej postaci

$$l = \sum_{i=2}^n \frac{X_{i\delta} - m}{2v} - \ln \sqrt{2\pi v},$$

gdzie  $m$  i  $v$  dane są takim samym wzorem, jak powyżej.

W modelu CIR gęstość przejścia dana jest wzorem

$$p(X_{i\delta} | X_{(i-1)\delta}, \theta) = ce^{-u-v} \left(\frac{v}{u}\right)^{q/2} I_q(2\sqrt{uv}),$$

gdzie

$$c = \frac{2\kappa}{\sigma^2(1 - e^{-\kappa\Delta})}, \quad u = cX_{(i-1)\delta}e^{-\kappa\Delta}, \quad v = cx, \quad q = \frac{2\kappa\mu}{\sigma^2} - 1,$$

a ponadto  $I_q(z)$  jest zmodyfikowaną funkcją Bessela pierwszego rodzaju rzędu  $q$ :

$$I_q(z) = \left(\frac{z}{2}\right)^q \sum_{k=0}^{\infty} \frac{(z^2/4)^k}{k!\Gamma(q+k+1)} = \frac{1}{\pi} \int_0^{\pi} e^{z \cos \varphi} \cos(q\varphi) d\varphi - \frac{\sin(q\varphi)}{\pi} \int_0^{\pi} e^{-z \cosh \varphi - q\varphi} dt.$$

### 3.2.3. Dwukrokowa metoda estymacji procesu dyfuzji

#### 3.2.3.1. Opis metody

Dwukrokowa metoda estymacji procesu dyfuzji została wprowadzona przez Phillipsa i Yu (2007). Autorzy zaproponowali następujące podejście. W pierwszym kroku zastosowali, wprowadzoną przez Jacoda (1994) a spopularyzowaną przez Barndorff-Nielsen i Shepharda (2004a), teorię twierdzeń granicznych dla wariancji zrealizowanej, do wyznaczenia równania regresji służącego do estymacji parametrów funkcji dyfuzji. W drugim kroku, sugerują estymację parametrów funkcji dryfu za pomocą metody wykorzystującej logarymiczną funkcję wiarygodności wyznaczoną za pomocą twierdzenia Girsanowa.

Z definicji wariancji zrealizowanej wynika, że jeżeli  $\delta \rightarrow 0$ , to

$$[X_{\delta}]_{i,h} \xrightarrow{p} [X]_{ih} - [X]_{(i-1)h}, \quad (3.11)$$

gdzie symbol  $\xrightarrow{p}$  oznacza zbieżność względem prawdopodobieństwa. Korzystając z twierdzenia 2.2.1, można wykazać, że

$$\frac{1}{r_k} \left\{ [X_{\delta}]_{i,h} - ([X]_{ih} - [X]_{(i-1)h}) \right\} \xrightarrow{d} N(0,1), \quad (3.12)$$

gdzie

$$r_i = \sqrt{\frac{2}{3} \sum_{j=2}^M (X_{(i-1)M+j\delta} - X_{(i-1)M+(j-1)\delta})^4}$$

dla  $k = 1, 2, \dots, K$ , przy czym symbol  $\xrightarrow{d}$  oznacza zbieżność według rozkładu.

W pracy Barndorff-Nielsen i Shepharda (2005) wprowadzono logarymiczną wersję powyższego twierdzenia, której tezę można zapisać w następujący sposób

$$\frac{1}{s_i} \left\{ \ln([X_\delta]_{i,h}) - \ln([X]_{ih} - [X]_{(i-1)h}) \right\} \xrightarrow{d} N(0,1), \quad (3.13)$$

gdzie

$$s_i = \min \left\{ \frac{r_i}{[X_\delta]_{i,h}}, \sqrt{\frac{2}{M}} \right\},$$

dla  $k$  z zakresu jak powyżej. Jednocześnie wykazano, że lepsze wyniki w próbach skończonych zostaną uzyskane, jeżeli do licznika we wzorze (3.13) dodamy  $\frac{1}{2} s_i^2$ .

Wektor  $\theta_2$  może być estymowany przez przybliżenie metodą najmniejszych kwadratów standaryzowanej wariancji zrealizowanej, zdefiniowanej jako

$$\frac{1}{r_i} [X_\delta]_{i,h}$$

przez standaryzowaną funkcję dyfuzji

$$\frac{1}{r_i} \left\{ [X]_{ih} - [X]_{(i-1)h} \right\} = \frac{1}{r_i} \int_{(i-1)h}^{ih} \sigma^2(X_s, \theta_2) ds \approx \frac{1}{r_i} \sum_{j=2}^M \sigma^2(X_{(i-1)M+(j-1)\delta}, \theta_2) \delta.$$

Wobec tego wektor parametrów  $\theta_2$  możemy wyznaczyć z następującego wzoru

$$\hat{\theta}_2 = \arg \min_{\theta_2} Q(\theta_2),$$

gdzie

$$Q(\theta_2) = \delta \sum_{i=1}^K \frac{1}{r_i^2} \left[ [X_\delta]_{i,h} - \sum_{i=2}^M \sigma^2(X_{(i-1)M+(j-1)\delta}, \theta_2) \delta \right]^2. \quad (3.14)$$

Korzystając ze wzoru (3.13), wektor parametrów  $\theta_2$  możemy oszacować w analogiczny sposób, przybliżając metodą najmniejszych kwadratów wyrażenie

$$\frac{1}{s_i} \ln\{[X_\delta]_{i,h}\},$$

funkcją

$$\frac{1}{s_i} \left\{ \ln\{[X]_{ih}\} - \ln\{[X]_{(i-1)h}\} + \frac{1}{2} s_i^2 \right\}.$$

Wówczas, po uwzględnieniu wyżej opisanej modyfikacji wzoru (3.13), mamy

$$\hat{\theta}_2 = \arg \min_{\theta_2} \tilde{Q}(\theta_2),$$

gdzie

$$\tilde{Q}(\theta_2) = \delta \sum_{k=1}^K \left[ \frac{\ln\{[X_\delta]_{i,h}\} - \ln\left\{ \sum_{j=2}^M \sigma^2(X_{(i-1)M+(j+1)\delta}, \theta_2) \delta \right\} + \frac{1}{2} s_i^2}{s_i} \right]^2.$$

W drugim kroku metody estymowany jest wektor parametrów dryfu przez maksymalizację logarytmicznej funkcji wiarygodności postaci

$$l_{AIF}(\theta_1) = \sum_{j=2}^n \frac{\mu(X_{(j-1)\delta}, \theta_1)}{\sigma^2(X_{(j-1)\delta}, \hat{\theta}_2)} (X_{j\delta} - X_{(j-1)\delta}) - \frac{\delta}{2} \sum_{j=2}^n \frac{\mu^2(X_{(j-1)\delta}, \theta_1)}{\sigma^2(X_{(j-1)\delta}, \hat{\theta}_2)}. \quad (3.15)$$

### 3.2.3.2. Symulacja Monte Carlo

W celu oceny jakości przedstawionej powyżej dwukrokowej metody stosujemy symulację Monte Carlo. Za pomocą dyskretyzacji Eulera zamieniamy równania dyfuzji na odpowiadające im równania różnicowe. Mają one następujące postacie:

$$X_n^{(\delta)} - X_{n-1}^{(\delta)} = \kappa(\mu - X_{n-1}^{(\delta)})\delta + \sigma\sqrt{\delta}\varepsilon_n^{(\delta)}, \text{ dla } n = 1, 2, \dots$$

z warunkiem początkowym  $X_0^{(\delta)}$  dla modelu Vašíčka,

$$X_n^{(\delta)} - X_{n-1}^{(\delta)} = \kappa(\mu - X_{n-1}^{(\delta)})\delta + \sigma\sqrt{X_{n-1}^{(\delta)}}\delta\varepsilon_n^{(\delta)}, \text{ dla } n = 1, 2, \dots$$

z warunkiem początkowym  $X_0^{(\delta)}$  dla modelu Coxa, Ingersolla i Rossa oraz

$$X_n^{(\delta)} - X_{n-1}^{(\delta)} = \kappa(\mu - X_{n-1}^{(\delta)})\delta + \sigma(X_{n-1}^{(\delta)})^\beta\sqrt{\delta}\varepsilon_n^{(\delta)}, \text{ dla } n = 1, 2, \dots$$



z warunkiem początkowym  $X_0^{(\delta)}$  dla modelu CKLS, gdzie  $\varepsilon_n^{(\delta)}$  dla  $n = 1, 2, \dots$  oznacza realizację standardowego białego szumu gaussowskiego. Zakładając, że  $X_0^{(\delta)}$  jest pierwszą znaną logarymiczną ceną instrumentu finansowego i podstawiając za parametry wyestymowane wcześniej wartości, generujemy 1000 trajektorii o takiej samej długości jak badany szereg, szacując nowe wartości parametrów dla każdej trajektorii.

Analogiczne badanie przeprowadzamy używając dyskretyzacji Milsteina. Łatwo zauważyć, że otrzymane za pomocą schematu Milsteina równanie różnicowe odpowiadające modelowi Vašíčka jest identyczne, jak otrzymane za pomocą schematu Eulera, ponieważ

$$\frac{\partial \sigma(X, t)}{\partial X} = 0.$$

Równania różnicowe odpowiadające stochastycznym równaniom różniczkowym Coxa, Ingersolla i Rossa oraz CKLS wyznaczone za pomocą dyskretyzacji Milsteina, przyjmą teraz postać

$$X_n^{(\delta)} - X_{n-1}^{(\delta)} = \kappa(\mu - X_{n-1}^{(\delta)})\delta + \sigma\sqrt{X_{n-1}^{(\delta)}}\delta\varepsilon_n^{(\delta)} + \frac{\sigma^2}{4}\delta(\varepsilon_n^{(\delta)})^2, \text{ dla } n = 1, 2, \dots$$

oraz

$$X_n^{(\delta)} - X_{n-1}^{(\delta)} = \kappa(\mu - X_{n-1}^{(\delta)})\delta + \sigma(X_{n-1}^{(\delta)})^\beta\sqrt{\delta}\varepsilon_n^{(\delta)} + \frac{\sigma^2\beta}{2}(X_{n-1}^{(\delta)})^{2\beta-1}\delta(\varepsilon_n^{(\delta)})^2, \text{ dla } n = 1, 2, \dots$$

### 3.3. Zastosowanie metody Phillipsa i Yu do estymacji modeli zmienności stochastycznej

W niniejszym podrozdziale prezentujemy procedurę, która pozwala na estymację za pomocą metody Phillipsa i Yu (2007) procesów zmienności stochastycznej postaci

$$\begin{cases} dX_t = \mu_1(X_t, \theta_1^{(1)})dt + \sigma_1(X_t, \sigma^2(t))dB_t^{(1)} \\ d\sigma^2(t) = \mu_2(\sigma^2(t), \theta_1^{(2)})dt + \sigma_2(\sigma^2(t), \theta_2^{(2)})dB_t^{(2)} \end{cases} \quad (3.16)$$

gdzie  $B_t^{(1)}$  i  $B_t^{(2)}$  są niezależnymi względem siebie ruchami Browna. Pewnym problemem pozostaje nieobserwalność procesu wariancji chwilowej  $\sigma^2(t)$ , który opisany jest

drugim równaniem w (3.16). Renò (2006) proponuje szacowanie  $\sigma^2(t)$  za pomocą następującego wyrażenia

$$\hat{\sigma}^2(t) = \frac{1}{\delta_t^- + \delta_t^+} \int_{t-\delta_t^-}^{t+\delta_t^+} \sigma^2(s) ds. \quad (3.17)$$

Znajdująca się z prawej strony równania wariancja scałkowana może z kolei być przybliżona odpowiadającą okresowi  $(1 - \delta_t^-, 1 + \delta_t^+)$  wariancją zrealizowaną. W szczególności można przyjąć, że  $\delta_t^+ = \delta_t^- = \delta_t$

Procedurę estymacji modelu (3.16) opiszemy za pomocą następującego algorytmu:

1. Należy oszacować wariancję chwilową korzystając z wzoru (3.17), wykorzystując możliwie wysoką częstotliwość danych intraday. Do szacowania wykonależy wykorzystać nie mniej niż 5 notowań.
2. Dla oszacowanego w ten sposób szeregu stosujemy opisaną w podrozdziale 3.2.3 dwukrokową metodę Phillipsa i Yu (2007). Otrzymujemy w ten sposób wektory oszacowań parametrów  $\theta_1^{(2)}$  oraz  $\theta_2^{(2)}$ .
3. Wektor parametrów  $\theta_1^{(1)}$  szacujemy, maksymalizując logarymiczną funkcję wiarygodności wprowadzoną przez Lánską (1979). Funkcja ta przyjmie w tym przypadku postać

$$l_{AIF}(\theta_1) = \sum_{j=2}^n \frac{\mu_1(X_{(j-1)\delta}, \theta_1^1)}{\sigma_1^2(X_t, \sigma_t^2)} (X_{j\delta} - X_{(j-1)\delta}) - \frac{\delta}{2} \sum_{j=2}^n \frac{\mu_1^2(X_{(j-1)\delta}, \theta_1^1)}{\sigma_1^2(X_t, \sigma_t^2)}.$$

Należy podkreślić, że postać modelu (3.16) jest uogólnieniem najczęściej pojawiających się w literaturze modeli zmienności stochastycznej, m.in. modelu Hestona (1993), wprowadzonego przez Nelsona (1990) modelu GARCH z czasem ciągłym oraz wykorzystywanego do opisu stóp LIBOR modelu SABR (Hagan i in. 2002). W celu oceny jakości przedstawionej powyżej metody stosujemy symulację Monte Carlo podobnie, jak robiliśmy to testując dwukrokową metodę Phillipsa i Yu. Model z czasem dyskretnym, odpowiadający rozważanemu przez nas modelowi Hestona zmienności stochastycznej, wyraża się wzorem

$$\begin{cases} X_n^{(\delta)} - X_{n-1}^{(\delta)} = \mu X_{n-1}^{(\delta)} \delta + \sigma_{n-1}^{(\delta)} X_{n-1}^{(\delta)} \sqrt{\delta} \varepsilon_n^{(\delta)}, \text{ dla } n = 1, 2, \dots \\ \sigma_n^{2(\delta)} - \sigma_{n-1}^{2(\delta)} = \theta(\omega - \sigma_{n-1}^{2(\delta)}) \delta + \xi \sigma_{n-1}^{(\delta)} \sqrt{\delta} \varepsilon_n^{(\delta)}, \text{ dla } n = 1, 2, \dots \end{cases}$$

w przypadku wykorzystania dyskretyzacji Eulera, oraz

$$\begin{cases} X_n^{(\delta)} - X_{n-1}^{(\delta)} = \mu X_{n-1}^{(\delta)} \delta + \sigma_{n-1}^{(\delta)} X_{n-1}^{(\delta)} \sqrt{\delta} \varepsilon_n^{(\delta)} + \frac{1}{2} (\sigma_{n-1}^{(\delta)})^2 X_{n-1}^{(\delta)} \delta (\varepsilon_n^{(\delta)})^2, \text{ dla } n = 1, 2, \dots \\ (\sigma_n^{(\delta)})^2 - (\sigma_{n-1}^{(\delta)})^2 = \theta (\omega - (\sigma_{n-1}^{(\delta)})^2) \delta + \xi \sigma_{n-1}^{(\delta)} \sqrt{\delta} \eta_n^{(\delta)} + \frac{\xi^2}{4} \delta (\eta_n^{(\delta)})^2, \text{ dla } n = 1, 2, \dots \end{cases}$$

jeśli wykorzystamy dyskretyzację Milsteina. Ponadto,  $\varepsilon_n^{(\delta)}$  oraz  $\eta_n^{(\delta)}$  dla  $n = 1, 2, \dots$  są realizacjami niezależnych względem siebie standardowych białych szumów gaussowskich. Zakładając, że  $X_0^{(\delta)}$  jest pierwszą znaną logarymiczną ceną instrumentu finansowego, a  $\sigma_0^{(\delta)}$  pierwszym dostępnym oszacowaniem zmienności chwilowej, generujemy 1000 trajektorii o takiej samej długości, jak badany szereg, podstawiając za parametry wyestymowane wcześniej wartości. Następnie szacujemy model dla każdej otrzymanej trajektorii.

### 3.4. Wyznaczenie jednodniowych prognoz ex-post z modeli dyfuzji

Oznaczmy przez  $\hat{X}_t(l)$  prognozę zmiennej  $X_{t+l}$ , odpowiadającą błędowi średniokwadratowemu jako funkcji straty. Oznacza to, że prognoza jest zmienną losową dobraną w taki sposób, by

$$\mathbb{E}[X_{t+l} - \hat{X}_t(l)]^2 \leq \min_g \mathbb{E}[X_t - g(X_t, \dots, X_1)]^2,$$

gdzie  $g(X_t, \dots, X_1)$  jest dowolną funkcją mierzalną względem  $\sigma$ -ciała generowanego przez informacje na temat procesu do momentu  $t$  włącznie. Zmienną  $\hat{X}_t(l)$  (lub jej realizację) nazywamy prognozą  $l$  kroków naprzód zmiennej  $X_t$ . Można wykazać, że

$$\hat{X}_t(h) = \mathbb{E}(X_{t+h} | \mathcal{F}_t).$$

Wobec tego, jeśli założymy, że badany proces opisany jest modelem dyfuzji, a długość jednego kroku równa jest  $h$ , to

$$\begin{aligned}
\hat{X}_t(h) &= \mathbb{E} \left[ X_0 + \int_0^{t+h\delta} \mu(X_s, s, \hat{\theta}_1) ds + \int_0^{t+h\delta} \sigma(X_s, s, \hat{\theta}_2) dB(s) \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[ X_t + \int_t^{t+h\delta} \mu(X_s, s, \hat{\theta}_1) ds + \int_t^{t+h\delta} \sigma(X_s, s, \hat{\theta}_2) dB(s) \right] \\
&= X_t + \int_t^{t+h\delta} \mu(X_s, s, \hat{\theta}_1) ds + \mathbb{E} \int_t^{t+h\delta} \sigma(X_s, s, \hat{\theta}_2) dB(s),
\end{aligned}$$

gdzie  $\hat{\theta}_1$  i  $\hat{\theta}_2$  są, odpowiednio, oszacowaniami wektorów parametrów funkcji dryfu i dyfuzji otrzymanych na bazie szeregu notowań do chwili  $s$  włącznie. Biorąc pod uwagę, że notowania dostępne są jedynie w stałych odstępach czasowych wartość prognozy  $\hat{X}_t(l)$  aproksymujemy przez następujące wyrażenie

$$\hat{X}_t^{(\delta)}(h) = X_t^{(\delta)} + \sum_{k=0}^{h-1} \mu(X_{t+k|n}^{(\delta)}, t, \hat{\theta}_1) \delta + \sum_{k=0}^{h-1} \sigma(X_{t+k|n}^{(\delta)}, t, \hat{\theta}_2) \sqrt{\delta} \varepsilon_{t+k}^{(\delta)}, \quad (3.18)$$

przy czym dla dowolnego  $t$  spełniona jest równość  $X_t^{(\delta)} = X_{[t/\delta]}$ . Cziryki i Kucherenko (2008) uzyskują oszacowania  $\hat{X}_t^{(\delta)}(l)$ , symulując niezależnie  $N$   $h$ -okresowych realizacji wektora innowacji  $(\varepsilon_{n+1}^{(\delta)}, \dots, \varepsilon_{n+k}^{(\delta)})$ , a następnie za pomocą każdego z nich wyznaczając  $h$ -okresową trajektorię (3.18) z warunkiem początkowym  $X_t^{(\delta)}$ . Za oszacowanie  $h$ -okresowej prognozy przyjmują średnią arytmetyczną z ostatnich wyrazów wszystkich trajektorii.

Jednokrokowe prognozy ex-post są otrzymywane przez dopasowanie modelu na podstawie danych do momentu  $T$  włącznie, a następnie obliczania na podstawie otrzymanego w ten sposób modelu, prognoz i reszt od chwili  $T+1$  do  $T+F$ , przy założeniu, że notowanie poprzedzające prognozę jest znane.

## 4. Charakterystyka procesów stochastycznych ze składową skokową

Badania empiryczne wskazują, że modele dyfuzji nie uwzględniają pewnych cech charakterystycznych szeregów czasowych. W szczególności nie są one w stanie opisać silnej leptokurtozy, jaka często pojawia się w szeregach zwrotów logarytmicznych. Dlatego w najnowszej literaturze proponuje się wykorzystanie szerszej klasy modeli zwanych procesami dyfuzji ze skokami. Uwzględniają, one oprócz zwykłych zaburzeń losowych opisanych funkcją dyfuzji, możliwość wystąpienia skoków. Jak wykazali Lahaye i in. (2007), skoki w procesach logarytmicznych poziomów indeksów giełdowych i kursów walutowych są silnie skorelowane z korzystnymi lub niekorzystnymi informacjami makroekonomicznymi. Często są one następstwem masowego zamykania pozycji przez inwestorów pod koniec dnia sesyjnego. Dlatego w takich pracach, jak Andersen i in. (2002) Andersen i in. (2003a) czy Chernov i in. (2003), logarytmiczna cena instrumentu finansowego jest opisana za pomocą ciągłego procesu dyfuzji oraz nieciągłego składnika opisującego skoki. Z punktu widzenia badań empirycznych niezwykle ważne są metody testowania skoków w procesach generujących finansowe szeregi czasowe. Aït-Sahalia (2002) zaproponował metodę weryfikacji występowania skoków bazującą na gęstościach przejścia. Carr i Wu (2003) wykazali, że występowanie skoków w danym szeregu czasowym można ocenić na podstawie zachowania się kursów opcji at-the-money out-of-the-money z krótkim czasem do wygaśnięcia. Johannes i in. (2004a, 2004b) zaproponowali podejście oparte na metodzie MCMC. Dosyć ciekawą metodę znajdujemy w pracy Mancini (2004). Autorka wprowadziła tzw. progi skoków, których wartość dążyła do 0, gdy liczba obserwacji w ciągu jednego dnia dąży do nieskończoności. Skoki są wykrywane, o ile ich wartość przekracza wartość progu.

W niniejszym rozdziale wykorzystamy testy wprowadzone przez Barndorff-Nielsen i Shepharda (2006) bazujące na różnicach pomiędzy wariacją dwupotęgową a wariacją zrealizowaną oraz test Jianga i Oomena (2008) oparty na różnicach pomiędzy tzw. wariacją swapową a wariacją zrealizowaną. Ponadto zastosujemy nieparametryczny test Lee i Mykland (2007) pozwalający na weryfikację występowania skoków dla każdego okresu intraday niezależnie.

Testowanie skoków w procesach cen logarytmicznych instrumentów finansowych, a w dalszej konsekwencji ich modelowanie za pomocą modeli w czasie ciągłym, pozwala na lepsze zrozumienie natury procesów i ich bardziej precyzyjny opis. Jak się okazuje, podejście to może też służyć dokładniejszej wycenie instrumentów pochodnych. Modele wyceny opcji oparte na procesach dyfuzji ze skokami zostały wprowadzone przez Kou (2002) oraz Kou i Wang (2004).

Modele dyfuzji ze skokami opisane są następującym stochastycznym równaniem różniczkowym

$$\begin{cases} dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB(t) + c(t)dq_t, \\ X_0 = Z \end{cases}, \quad (4.1)$$

gdzie proces  $\mu(X_t, t)$  ma lokalnie ograniczoną wariację,  $\sigma(X_t, t)$  jest procesem dodatnim i prawostronnie ciągłym z lewostronnymi granicami,  $B_t$  jest ruchem Browna,  $dq_t$  jest skokowym procesem przyjmujący wartość 1 w sytuacji, gdy w momencie  $t$  wystąpi skok i 0 w przeciwnym wypadku, natomiast  $c(t)$  jest wielkością skoku.

Składowa skokowa opisana jest za pomocą procesu Poissona  $q_t$ , będącego dyskretnym procesem z czasem ciągłym. Zatem  $q_0 = 0$  oraz dla dowolnych  $0 < s < t$  przyrost  $q_t - q_s$  jest niezależny od  $\mathcal{F}_s$  i ma rozkład Poissona z parametrem  $\lambda(t-s)$ :

$$P(q_t - q_s = k) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}, \text{ dla } k = 0, 1, 2, \dots$$

Parametr  $\lambda$  ( $\lambda \geq 0$ ) nazywany jest intensywnością procesu.

Jednookresowy  $i$ -ty logarytmiczny zwrot długości  $\hbar$  można przedstawić w następujący sposób

$$y_i = X_{i\hbar} - X_{(i-1)\hbar} = \int_{(i-1)\hbar}^{i\hbar} \mu(X_t, t)dt + \int_{(i-1)\hbar}^{i\hbar} \sigma(X_t, t)dB_t + \int_{(i-1)\hbar}^{i\hbar} c(t)dq_t.$$

Ostatnia całka z prawej strony równania kumuluje skoki, które wystąpiły pomiędzy chwilami  $(i-1)\hbar$  i  $i\hbar$ .

Dodatkowa składowa skokowa w równaniu (4.1) może być na przykład interpretowana jako reakcja na ogłoszenia makroekonomiczne. Potwierdzają to najnowsze wyniki badań. Andersen i in. (2003b) znaleźli istotną korelację pomiędzy ogłoszeniami makroekonomicznymi a skokami w kursach DEM/USD. Lahaye i in. (2007) wykazali silne powiązanie skoków w kursach walutowych, oraz różnego rodzaju kontraktach

futures z ogłoszeniami makroekonomicznymi. Andersen i in. (2007b) udowodnili, że skoki w procesach kursów walutowych, cen akcji i obligacji są następstwem różnic pomiędzy oczekiwanymi a rzeczywistymi danymi makroekonomicznymi.

#### 4.1. Testy Barndorff-Nielsena i Shepharda na występowanie skoków

W tym podrozdziale zostaną opisane wprowadzone przez Barndorff-Nielsena i Shepharda (2006) metody testowania występowania skoków. Koncepcja konstrukcji testów opiera się na rozumowaniu przedstawionym w podrozdziale 2.3.2. Wariancja zrealizowana w  $i$ -tym okresie długości  $h$ , będąca sumą podniesionych do kwadratów zwrotów o częstotliwości  $\delta$  w przedziale czasowym długości  $[(i-1)h, ih)$  jest estymatorem sumy zmienności scałkowanej oraz podniesionych do kwadratów skoków. Przyrost wariacji dwupotęgowej w tym przedziale będący sumą iloczynów następujących po sobie bezwzględnych zwrotów przybliży zmienność scałkowaną bez skoków. Wobec tego różnica między wariacją zrealizowaną i zrealizowaną wariacją dwupotęgową estymuje sumę kwadratów skoków. W celu weryfikacji występowania skoków w okresie  $i$  wykorzystane zostaną dwie statystyki – liniowa

$$\mu_1^{-2} \{X_\delta\}_i^{[1,1]} - [X_\delta]_i$$

oraz ilorazowa

$$\frac{\mu_1^{-2} \{X_\delta\}_i^{[1,1]}}{[X_\delta]_i}.$$

Wyprowadzone przez Barndorff-Nielsena i Shepharda (2006) rozkłady wyżej wymienionych statystyk, otrzymane przy wykorzystaniu teorii wariacji wielopotęgowej opisanej we wcześniejszym rozdziale, opisane są w poniższym twierdzeniu.

**Twierdzenie 4.1.** Przyjmijmy, że  $X \in SVSM^c$ . Załóżmy, że spełnione są następujące warunki:

- (i) Proces  $\sigma^2$  ma trajektorie odgraniczone od zera,
- (ii) Proces łączny  $(\alpha^*, \sigma)$  jest niezależny od ruchu Browna  $B$ .

Wówczas, jeśli  $\delta$  dąży monotonicznie do 0, to

$$G = \frac{\delta^{-1/2}(\mu_1^{-2}\{X_\delta\}_i^{[1,1]} - [X_\delta]_i)}{\sqrt{\int_0^t \sigma^4(s)ds}} \xrightarrow{d} N(0, \mathcal{G}) \quad (4.2)$$

oraz

$$H = \frac{\delta^{-1/2}\left(\frac{\mu_1^{-2}\{X_\delta\}_i^{[1,1]} - 1}{[X_\delta]_i}\right)}{\sqrt{\int_0^t \sigma^4(s)ds} / \left\{\int_0^t \sigma^2(s)ds\right\}^2} \xrightarrow{d} N(0, \mathcal{G}), \quad (4.3)$$

$$\text{gdzie } \mathcal{G} = \frac{\pi^2}{4} + \pi - 5.$$

Powyższy rezultat jest uogólnieniem wyniku Barndorff-Nielsen i Shepharda (2004c). W oryginale został on wykazany dla przypadku  $\alpha^* = 0$ . Wynik ten implikuje, że przy spełnieniu hipotezy alternatywnej o występowaniu skoków prawdziwe są wzory

$$\mu_1^{-2}\{X_\delta\}_i^{[1,1]} - [X_\delta]_i \xrightarrow{p} - \int_{(i-1)h}^{ih} c^2(t)dq(t) \leq 0 \quad (4.4)$$

oraz

$$\frac{\mu_1^{-2}\{X_\delta\}_i^{[1,1]} - 1}{[X_\delta]_i} \xrightarrow{p} - \frac{\int_{(i-1)h}^{ih} c^2(t)dq(t)}{\int_{(i-1)h}^{ih} \sigma^2(t)dt + \int_{(i-1)h}^{ih} c^2(t)dq(t)} \leq 0. \quad (4.5)$$

Z uwagi na nieobserwowalność kwartyczności, statystyki  $G$  i  $H$  nie mogłyby być stosowane w praktyce. W takiej sytuacji pomocne są własności wariacji wielopotęgowej. Na podstawie (2.7) i (2.8), mamy

$$\{X_\delta\}_i^{[4/3, 4/3, 4/3]} = \delta^{-1} \sum_{j=1}^{[h/\delta]-2} |y_{j,i}|^{4/3} |y_{j+1,i}|^{4/3} |y_{j+2,i}|^{4/3} \xrightarrow{p} \mu_{4/3}^3 \int_{(i-1)h}^{ih} \sigma^4(t)dt$$

oraz

$$\{X_\delta\}_i^{[1,1,1,1]} = \delta^{-1} \sum_{j=1}^{[h/\delta]-3} |y_{j,i}| |y_{j+1,i}| |y_{j+2,i}| |y_{j+3,i}| \xrightarrow{p} \mu_1^4 \int_{(i-1)h}^{ih} \sigma^4(t)dt.$$

Pierwszy z tych wyników został zastosowany przez Huanga i Tauchena (2005), drugi znajdujemy w pracach Barndorffa-Nielsen i Shepharda (2006) oraz Andersena i in.



(2003a). Obydwa estymatory są zgodne przy założeniu, że  $X \in SVJSM^c$ . Wobec tego możliwe jest wyznaczenie przybliżonych wartości statystyk testowych (4.2) i (4.3) testujących występowanie skoków w  $i$ -tym okresie długości  $h$ . Przyjmują one postać

$$\hat{G} = \frac{(\mu_1^{-2} \{X_\delta\}_t^{[1,1]} - [X_\delta]_t)}{\sqrt{\delta \mu_1^{-4} \{X_\delta\}_t^{[1,1,1,1]}}} \xrightarrow{d} N(0, \vartheta) \quad (4.6)$$

oraz

$$\hat{H} = \frac{1}{\sqrt{\delta \{X_\delta\}_t^{[1,1,1,1]} / \{\{X_\delta\}_t^{[1,1]}\}^2}} \left( \frac{\mu_1^{-2} \{X_\delta\}_t^{[1,1]} - [X_\delta]_t}{[X_\delta]_t} - 1 \right) \xrightarrow{d} N(0, \vartheta). \quad (4.7)$$

W pracy Huang i Tauchena (2005) znajdujemy pewną modyfikację stosowanych w powyższych statystykach estymatorów wariacji wielopotęgowych. W celu uniknięcia niedoszacowania wariacji wielopotęgowych, autorzy zaproponowali, by mnożyć ich estymatory przez odpowiedni czynnik skalujący. W ten sposób uzyskali następujące wzory:

$$\{X_\delta\}_i^{[1,1]} = \delta^{-1} \frac{t/\delta}{(t/\delta) - 1} \sum_{j=1}^{[t/\delta]-1} |y_{j,i}| |y_{j+1,i}| \xrightarrow{p} \mu_1^2 \int_{(i-1)h}^{ih} \sigma^2(t) dt,$$

$$\{X_\delta\}_i^{[4/3,4/3,4/3]} = \delta^{-1} \frac{t/\delta}{(t/\delta) - 2} \sum_{j=1}^{[t/\delta]-2} |y_{j,i}|^{4/3} |y_{j+1,i}|^{4/3} |y_{j+2,i}|^{4/3} \xrightarrow{p} \mu_{4/3}^3 \int_{(i-1)h}^{ih} \sigma^4(t) dt$$

oraz

$$\{X_\delta\}_i^{[1,1,1,1]} = \delta^{-1} \frac{t/\delta}{(t/\delta) - 3} \sum_{j=1}^{[t/\delta]-3} |y_{j,i}| |y_{j+1,i}| |y_{j+2,i}| |y_{j+3,i}| \xrightarrow{p} \mu_1^4 \int_{(i-1)h}^{ih} \sigma^4(t) dt.$$

Statystyki (4.6) i (4.7) odrzucają hipotezę zerową o niewystępowaniu skoków, jeżeli przyjmują wartość istotnie ujemną. Huang i Tauchen (2005) przedstawili znormalizowaną wersję statystyki (4.6). Przyjmuje on postać

$$\frac{(\mu_1^{-2} \{X_\delta\}_i^{[1,1]} - [X_\delta]_i)}{\sqrt{\delta \vartheta \mu_1^{-4} \{X_\delta\}_i^{[1,1,1,1]}}} \xrightarrow{d} N(0, 1).$$

Logarymiczna wersja powyższej statystyki została wyznaczona za pomocą metody delta przez Barndorff-Nielsen i Shepharda (2004c). Po znormalizowaniu wzór ten można zapisać w następującej postaci

$$\frac{\log(\mu_1^{-2} \{X_\delta\}_i^{[1,1]}) - \log([X_\delta]_i)}{\sqrt{\delta \mathcal{G} \{X_\delta\}_i^{[1,1,1,1]} / \{\{X_\delta\}_i^{[1,1]}\}^2}} \xrightarrow{d} N(0,1). \quad (4.8)$$

Barndorff-Nielsen i Shephard (2006) i Huang Tauchen (2005) sugerują, że statystyka (4.8) lepiej sprawdza się w zastosowaniach praktycznych. W szczególności, przy jej stosowaniu znika problem zbyt częstego odrzucenia hipotezy zerowej o nie występowaniu skoków. Statystyki te posłużą nam do weryfikacji występowania istotnych skoków w szeregach notowań dziennych.

Omówimy teraz znormalizowaną wersję statystyki ilorazowej. Wiąże się z tym pojęcie skoku relatywnego  $c_j^R$ , który może być interpretowany jako procentowy udział skoku w dziennej wariancji:

$$c_j^R = \frac{\mu_1^{-2} \{X_\delta\}_i^{[1,1]} - [X_\delta]_i}{\{X_\delta\}_i}.$$

Statystykę (4.7) można zatem zapisać w postaci

$$\frac{c_j^R}{\sqrt{\delta \{X_\delta\}_i^{[1,1,1,1]} / \{\{X_\delta\}_i^{[1,1]}\}^2}} \xrightarrow{d} N(0, \mathcal{G}).$$

Stosując nierówność Jensena, mamy

$$\frac{\{X_\delta\}_i^{[1,1,1,1]}}{\{\{X_\delta\}_i^{[1,1]}\}^2} \xrightarrow{p} \frac{\int_{(i-1)h}^{ih} \sigma^4(t) dt}{\left\{ \int_{(i-1)h}^{ih} \sigma^2(t) dt \right\}^2} \geq 1.$$

Stąd

$$\frac{c_j^R}{\sqrt{\delta \max\left[1, \{X_\delta\}_i^{[1,1,1,1]} / \{\{X_\delta\}_i^{[1,1]}\}^2\right]}} \xrightarrow{d} N(0, \mathcal{G}).$$

Znormalizowana wersja powyższej statystyki dana jest wzorem

$$\frac{c_j^R}{\sqrt{\delta \mathcal{G} \max\left[1, \{X_\delta\}_i^{[1,1,1,1]} / \{\{X_\delta\}_i^{[1,1]}\}^2\right]}} \xrightarrow{d} N(0,1). \quad (4.9)$$

Z uwagi na częstotliwość danych musimy przyjąć założenie, że w ciągu jednego dnia maksymalnie wystąpić może jeden skok. Ponadto, korzystając z wzoru (4.5), możemy wyznaczyć przybliżoną wartość bezwzględną skoku

$$|c_j| \approx \sqrt{[X_\delta]_i - \mu_1^{-2} \{X_\delta\}_i^{[1,1]}}. \quad (4.10)$$

## 4.2. Testowanie skoków za pomocą tzw. wariacji swapowej

Alternatywne podejście do problemu testowania występowania skoków zostało zaproponowane przez Jianga i Oomena (2008). Zasadnicza odmienność w stosunku do testów wprowadzonych przez Barndorff-Nielsen i Shepharda polega na tym, że zamiast różnicy między wariacją dwupotęgową a wariacją zrealizowaną bada się różnicę pomiędzy wprowadzoną przez autorów tzw. wariacją swapową a wariacją zrealizowaną.

Przyjmijmy, że proces cen logarymicznych opisany jest następującym stochastycznym równaniem różniczkowym

$$dX_t = (\mu_t - \lambda_t \eta_t - \frac{1}{2} \sigma^2(t))dt + \sigma(t)dB_t + c(t)dq_t, \quad (4.11)$$

gdzie  $\mu_t$  jest chwilowym dryfem,  $\sigma^2(t)$  jest chwilową wariacją w sytuacji nie wystąpienia skoku,  $q_t$  jest procesem liczącym ze skończoną chwilową intensywnością  $\lambda_t$  ( $0 \leq \lambda_t < \infty$ ), natomiast  $c(t)$  jest niezerową zmienną losową reprezentującą skok ceny ze średnią chwilową równą  $\exp(c(t)) - 1 = \eta_t$ . Oznaczmy przez  $S_t$  proces ceny ( $X_t = \ln S_t$ ). Wówczas, stosując lemat Itô dla równania (4.8), otrzymujemy, że

$$dS_t / S_t = (\mu_t - \lambda_t \eta_t)dt + \sigma(t)dB_t + (\exp c(t) - 1)dq_t.$$

Odejmując stronami równania (4.11) i (4.12) uzyskujemy

$$2(dS_t / S_t - dX_t) = \sigma^2(t)dt + (\exp c(t) - c(t) - 1)dq_t \quad (4.12)$$

lub w postaci scałkowanej

$$2 \int_0^T (dS_t / S_t - dX_t) = \int_0^T \sigma^2(t)dt + \int_0^T (\exp c(t) - c(t) - 1)dq_t.$$

Z punktu widzenia finansów, lewa strona równania może być zinterpretowana jako skumulowane przyrosty ze strategii delta-zabezpieczającej dwie krótkie pozycje w logarymicznym kontrakcie forward. Neuberger (1994) wykazał, że jeżeli w procesie ceny nie występują skoki, to portfel delta-zabezpieczający kontrakt logarymiczny może do-

skonale replikować wymianę wariancji. W przeciwnej sytuacji strategia replikująca będzie obciążona losowym i niepoddającym się zabezpieczeniu błędem równym

$$\int_0^T (\exp c(t) - c(t) - 1) dq_t. \quad (4.13)$$

Zależności te są podstawą konstrukcji testów na występowanie skoków. Jiang i Oomen (2008) wprowadzają pojęcie tzw. wariancji swapowej, która jest dyskretnym odpowiednikiem lewej strony równania (4.12), w granicznym przypadku jest równa wariancji scałkowanej i wyraża się wzorem

$$S_w V_M^i = 2 \sum_{j=1}^M (Y_{j,i} - y_{j,i}),$$

gdzie  $Y_{j,i}$  jest zdefiniowany jako

$$Y_{j,i} = (S_{(i-1)h+j\delta} - S_{(i-1)h+(j-1)\delta}) / S_{(i-1)h+j\delta}.$$

Wobec tego, wykorzystując fakt, że wariancja zrealizowana jest estymatorem wariacji kwadratowej procesu, czyli sumy wariacji scałkowanej i kwadratów skoków, mamy

$$\text{plim}_{M \rightarrow \infty} (S_w V_M^i - \{X_\delta\}_i) = \begin{cases} 0 & \text{gdy nie ma skoków w } [(i-1)h, ih] \\ 2 \int_{(i-1)h}^{ih} (\exp c(t) - \frac{1}{2} c^2(t) - c(t) - 1) dq_t & \text{poza tym.} \end{cases} \quad (4.14)$$

Różnice  $S_w V_M^i - [X_\delta]_i$  mogą być wykorzystane do zbadania występowania skoków w poszczególnych okresach, w sposób analogiczny, jak różnice  $\mu_1^{-2} \{X_\delta\}_i^{[1,1]} - [X_\delta]_i$  w testach Barndorff-Nielsen i Shepharda. Jednak w przeciwieństwie do nich, testy Jianga i Oomena są testami obustronnymi, gdyż wyrażenie  $S_w V_M^i - [X_\delta]_i$  może przyjmować wartości zarówno dodatnie, jak i ujemne. Wprowadzone przez Jianga i Oomena statystyki testujące występowanie skoków opisane są następującym twierdzeniem.

**Twierdzenie 4.2.1.** Załóżmy, że proces cen logarytmicznych jest opisany równaniem (4.11). Niech ponadto  $\mu_t$  będzie prognozowanym procesem z lokalnie ograniczoną wariacją,  $\sigma^2(t)$  będzie procesem ściśle dodatnim i prawostronnie ciągłym z lewostronnymi granicami oraz takim, że

$$\int_0^T \sigma^2(t) dt < \infty.$$

Wówczas, przy założeniu hipotezy  $H_0$  mówiącej, że w przedziale czasowym  $[(i-1)h, ih)$  nie występują skoki, dla poniższych statystyk przy  $\delta \rightarrow 0$  ( $\delta = h/M$ ) zachodzi:

(i) Statystyka liniowa

$$\frac{3M}{\sqrt{\mu_6 \int_{(i-1)h}^{ih} \sigma^6(t) dt}} (S_w V_M^i - [X_\delta]_i) \xrightarrow{d} N(0,1),$$

(ii) Statystyka logarytmiczna

$$\frac{3M \int_{(i-1)h}^{ih} \sigma^2(t) dt}{\sqrt{\mu_6 \int_{(i-1)h}^{ih} \sigma^6(t) dt}} (\ln S_w V_M^i - \ln [X_\delta]_i) \xrightarrow{d} N(0,1),$$

(iii) Statystyka ilorazowa

$$\frac{3M \int_{(i-1)h}^{ih} \sigma^2(t) dt}{\sqrt{\mu_6 \int_{(i-1)h}^{ih} \sigma^6(t) dt}} \left( 1 - \frac{[X_\delta]_i}{S_w V_M^i(T)} \right) \xrightarrow{d} N(0,1).$$

Analogicznie jak w przypadku statystyk wprowadzonych przez Barndorff-Nielsen i Shepharda (2006) praktyczne zastosowanie statystyk wymienionych w tezie twierdzenia 4.2.1 umożliwi przybliżenie wariancji scałkowanej zrealizowaną wariacją dwupotęgową oraz sekstyczności za pomocą ilorazu wariacji czteropotęgowej  $\{X_\delta\}_i^{[3/2, 3/2, 3/2, 3/2]}$  i  $\mu_{3/2}^4$  lub sześciopotęgowej  $\{X_\delta\}_i^{[1, 1, 1, 1, 1, 1]}$  i  $\mu_1^6$ . Stosując statystyki Jianga i Oomena w praktyce, podobnie jak w przypadku testów opisanych w poprzednim rozdziale, musimy przyjąć założenie, że w ciągu jednego dnia może wystąpić co najwyżej jeden skok. Na podstawie wzoru (4.13), łatwo zauważyć, że w przypadku odrzucenia hipotezy zerowej o niewystępowaniu skoków w danym przedziale czasowym, przybliżoną wartość skoku można wyznaczyć numerycznie z równania

$$S_w V_M^i - [X_\delta]_i = 2 \exp c(t) - c^2(t) - 2c(t) - 2. \quad (4.15)$$

### 4.3. Wpływ zaburzeń mikrostruktury rynku na wariacje wielopotęgowe

Jak wiadomo, rzeczywisty logarytmiczny proces ceny, a więc w konsekwencji także rzeczywiste zwroty tego procesu, są zaburzone takimi czynnikami, jak dostępność notowań w dyskretnych odstępach czasowych czy różnice w cenach oferowanych przez nabywcę i sprzedającego. Na skalę zaburzeń wpływ ma dobór odpowiedniej częstotliwości notowań. Z jednej strony, mogłoby się wydawać, że należałoby wykorzystać dane o możliwie wysokiej częstotliwości, z drugiej, okazuje się jednak, że takie dane są zanieczyszczone dużym szumem wynikającym z silnej autokorelacji zwrotów wysokiej częstotliwości. W pracach Aït-Sahalii, Mykland i Zhanga (2005) Bandiego i Russella (2004) oraz Zhanga, Aït-Sahalii i Mykland (2005) założono, że proces cen logarytmicznych wyraża się wzorem

$$X_t = X_t^* + \nu_t,$$

gdzie  $X_t^*$  jest rzeczywistym nieobserwowalnym procesem cen logarytmicznych a  $\nu$  jest ścisłym białym szumem. Wobec tego, zwrot w przedziale długości  $\delta$  można zapisać w następujący sposób

$$y_{t,\delta} = X_t - X_{t-\delta} = X_t^* - X_{t-\delta}^* + \nu_t - \nu_{t-\delta} = y_{t,\delta}^* + \eta_{t,\delta},$$

gdzie  $y_{t,\delta}^*$  jest rzeczywistym nieobserwowalnym zwrotem, natomiast  $\eta_{t,\delta}$  jest procesem MA(1). Silna autokorelacja szeregu  $\eta_{t,\delta}$  zwiększa dodatkowo powiązanie dowolnych następujących po sobie zwrotów  $r_{t+\delta,\delta}$  i  $r_{t+\delta(j-1),\delta}$ . Staje się to źródłem obciążenia estymatorów wariacji wielopotęgowych (Andersen i in. 2007a).

Aby uniknąć niepożądanych autokorelacji pomiędzy sąsiednimi zwrotami intraday, w pracy Andersena i in. (2007a) wprowadzono tzw. przesunięte wariacje wielopotęgowe. Przy ich konstrukcji zamiast iloczynów kolejnych zwrotów stosuje się iloczyny zwrotów przesuniętych względem siebie o dwie lub więcej pozycji. Dodatkowo, w celu uniknięcia niedoszacowania estymatorów wariacji wielopotęgowych, zostają one przeskalowane w ten sam sposób co w pracy Huanga i Tauchena (2005). Wówczas przesunięte o  $k$  pozycji, wariacje dwu-, trój-, i czteropotęgowe wyrażają się, odpowiednio, wzorami

$$\{X_\delta\}_{i,[k]}^{[1,1]} = \delta^{-1} \frac{t/\delta}{(t/\delta)-2} \sum_{j=1}^{\lfloor t/\delta \rfloor - 2} |y_{j,i}|^1 |y_{j+k+1,i}|^1,$$

$$\{X_\delta\}_{i,[k]}^{[4/3,4/3,4/3]} = \delta^{-1} \frac{t/\delta}{(t/\delta)-4} \sum_{j=1}^{\lfloor t/\delta \rfloor - 2k - 2} |y_{j,i}|^{4/3} |y_{j+k+1,i}|^{4/3} |y_{j+2k+2,i}|^{4/3}$$

oraz

$$\{X_\delta\}_{i,[k]}^{[1,1,1,1]} = \delta^{-1} \frac{t/\delta}{(t/\delta)-6} \sum_{j=1}^{\lfloor t/\delta \rfloor - 3k - 3} |y_{j,i}|^1 |y_{j+k+1,i}|^1 |y_{j+2k+2,i}|^1 |y_{j+3k+3,i}|^1.$$

W praktyce najczęściej stosuje się wariacje przesunięte o jedną pozycję. Wyrażają się one wzorami

$$\{X_\delta\}_{i,[1]}^{[1,1]} = \delta^{-1} \frac{t/\delta}{(t/\delta)-2} \sum_{j=1}^{\lfloor t/\delta \rfloor - 2} |y_{j,i}|^1 |y_{j+2,i}|^1,$$

$$\{X_\delta\}_{i,[1]}^{[4/3,4/3,4/3]} = \delta^{-1} \frac{t/\delta}{(t/\delta)-4} \sum_{j=1}^{\lfloor t/\delta \rfloor - 4} |y_{j,i}|^{4/3} |y_{j+2,i}|^{4/3} |y_{j+4,i}|^{4/3}$$

oraz

$$\{X_\delta\}_{i,[1]}^{[1,1,1,1]} = \delta^{-1} \frac{t/\delta}{(t/\delta)-6} \sum_{j=1}^{\lfloor t/\delta \rfloor - 6} |y_{j,i}|^1 |y_{j+2,i}|^1 |y_{j+4,i}|^1 |y_{j+6,i}|^1.$$

Zgodnie z tym, co wykazali Barndorff-Nielsen i Shephard (2004b), statystyka testowa (4.9) z zastosowanymi przesuniętymi wariacjami wielopotęgowymi jako estymatorami zmienności scałkowanej i kwartyczności zachowuje asymptotyczny rozkład normalny, dlatego w części empirycznej pracy wykorzystamy również ilorazową statystykę testową z przesuniętymi wariacjami wielopotęgowymi. Ponadto przeprowadzone przez Huang i Tauchena (2005) symulacje Monte Carlo dowodzą, że daje ona zaskakująco dobre rezultaty.

#### 4.4. Nieparametryczne testy na występowanie skoków

Inne podejście do testowania skoków zostało zaprezentowane przez Lee i Mykland (2007). Jego zaletą jest możliwość weryfikacji występowania skoków dla każdego okresu intraday niezależnie. Intuicja metody weryfikacji polega na ocenie wielkości danego zwrotu intraday i ocenie, czy jest ona na tyle mała, że może być następstwem samego

ruchu Browna, czy na tyle duża, że musiała zostać dodatkowo wygenerowana przez skok. Wprowadzone przez Lee i Mykland (2007) statystyki testowe, podobnie jak statystyki wprowadzone przez Barndorff-Nielsen i Shephard (2006) wykorzystują zrealizowaną wariację dwupotęgową jako miarę nie wrażliwą na skoki.

Statystyka testująca występowanie skoku w przedziale czasowym  $(i + (j-1)\delta, i + j\delta)$  jest zdefiniowana jako znormalizowany zwrot, tj. zwrot umniejszony o lokalną średnią i podzielony przez lokalne odchylenie standardowe:

$$L_{\mu}(i + j\delta) = \frac{y_{j,i} - \hat{m}(i + j\delta)}{\hat{\sigma}(i + j\delta)},$$

gdzie

$$\hat{m}(i + j\delta) = \frac{1}{K-1} \sum_{l=j-K+1}^{j-1} y_{l,i}$$

jest lokalną średnią,

$$\hat{\sigma}^2(i + j\delta) = \frac{1}{K-2} \sum_{l=j-K+2}^{j-1} |y_{l,i} - y_{l-1,i}|$$

jest pomnożoną przez  $\mu_1^2 / (K-2)$  zrealizowaną wariacją dwupotęgową, a  $K$  jest liczbą notowań intraday wykorzystanych do wyznaczenia lokalnej zmienności. W sytuacji gdy badany szereg czasowy opisany jest procesem stochastycznym z czasem ciągłym bez dryfu, przyjmujemy, że  $\hat{m}(i + j\delta) = 0$ . Statystyka  $L_{\mu}$  przy hipotezie  $H_0$  o niewystępowaniu skoku w przedziale czasowym  $(i + (j-1)\delta, i + j\delta)$  ma rozkład  $N(0, \pi/2)$ . Przy stosowaniu tego testu problemem pozostaje odpowiedni dobór wielkości  $K$ . Aby uniknąć efektu skoków w estymacji zmienności chwilowej,  $K$  musi być wystarczająco duże. Z drugiej strony, musi być ono mniejsze od liczby obserwacji w ciągu dnia. Przyjęte przez Lee i Mykland (2007)  $K = O_p(\delta^{\alpha})$ , gdzie  $-1 < \alpha < -0,5$ , spełnia te założenia. Autorzy ci jednocześnie sugerują przyjęcie  $K = \delta^{-0,5}$ , tak jak zostało to zrobione przez Lahaye i in. (2007).

Opisana statystyka testowa zachowuje się różnie w zależności od tego, czy hipoteza o istnieniu skoku jest przyjęta, czy odrzucona. W pierwszym przypadku statystyka przyjmuje typowe wartości dla rozkładu normalnego, w drugim jej wartości są bardzo duże. Dlatego autorzy zaproponowali modyfikację statystyki testowej, której zadaniem jest eliminacja takiej sytuacji. Zmodyfikowana statystyka testowa przyjmuje postać



$$\frac{\max |L_\mu(i + j\delta)| - C_n}{S_n} \rightarrow \psi,$$

gdzie  $\psi$  ma rozkład  $P(\psi \leq x) = \exp(-\exp(-x))$ ,

$$C_n = \frac{\sqrt{2 \log n}}{\mu_1} - \frac{\log \pi + \log(\log n)}{2\mu_1 \sqrt{2 \log n}}, \quad S_n = \frac{1}{\mu_1 \sqrt{2 \log n}}$$

oraz  $n$  jest liczbą obserwacji. Wobec powyższego na poziomie istotności  $\alpha$  hipotezę zerową możemy odrzucić jeśli

$$\hat{L}_\mu = \frac{|L_\mu(i + j\delta)| - C_n}{S_n} > \beta, \quad (4.16)$$

gdzie  $\beta = -\ln(-\ln(1 - \alpha))$ .

W sytuacji, gdy badany proces nie posiada dryfu, a więc, gdy przyjmiemy, że  $\hat{m}(i + j\delta) = 0$ , wówczas statystykę  $\hat{L}_\mu$  będziemy oznaczać przez  $\hat{L}$ . Wyznaczone z powyższego równania wartości krytyczne dla różnych poziomów istotności dane są w poniższej tabeli

**Tabela 4.1** Wartości krytyczne rozkładu statystki  $\hat{L}_\mu$  dla różnych poziomów istotności.

Poziom istotności	Wartość krytyczna	Poziom istotności	Wartość krytyczna
0,05	2,97	0,001	6,91
0,01	4,6	0,05	7,6
0,5	5,29	0,0001	9,21

## 4.5. Badanie mocy i rozmiaru testów na występowanie skoków

Chociaż przedstawione w poprzednich podrozdziałach testy mają silną podstawę teoretyczną, wyjątkowo ważnym zagadnieniem jest ocena ich przydatności w praktyce. W tym celu przeprowadzimy symulacje, które posłużą nam do sprawdzenia stopnia w jakim wariancja zrealizowana przybliży zmienność scałkowaną, a odpowiednie wariancje wielopotęgowe przybliżają kwartyczość i sekstyczość w zależności od częstotliwości danych. Ponadto symulacje te wykorzystamy do oszacowania mocy i rozmiaru

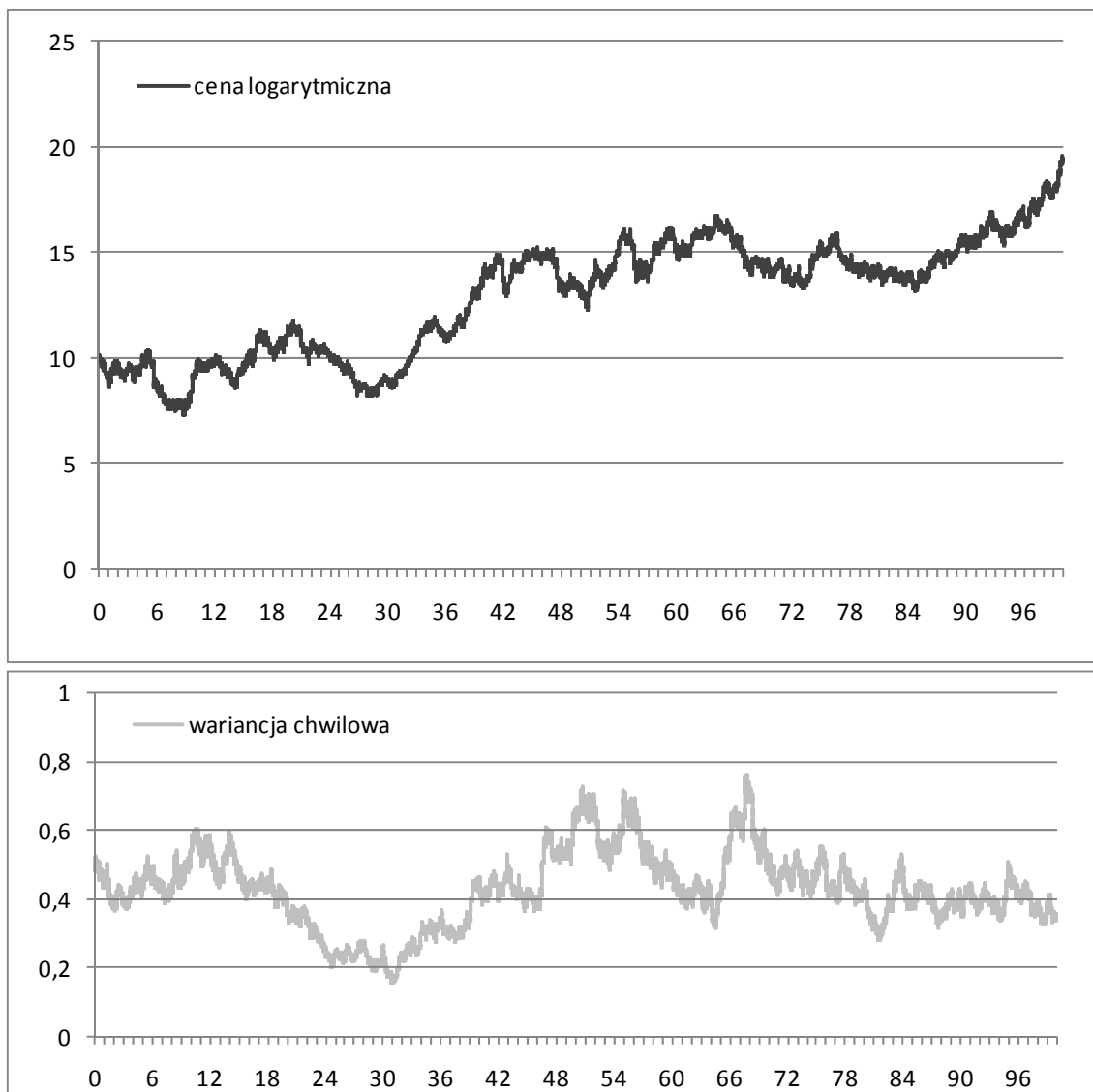
testów na skoki i dokonamy oceny wpływu częstotliwości danych na te wielkości. Rozważymy dwa procesy z czasem ciągłym. Pierwszy z nich, wprowadzony przez Nelsona (1990), opisany jest za pomocą następującego modelu z czasem ciągłym

$$\begin{cases} dX_t = \sigma(t)dB_t^{(1)}, \\ d\sigma^2(t) = \theta[\omega - \sigma^2(t)]dt + \xi\sigma^2(t)dB_t^{(2)}, \\ X_0 = x_0, \sigma^2(0) = \sigma_0^2, \end{cases} \quad (4.17)$$

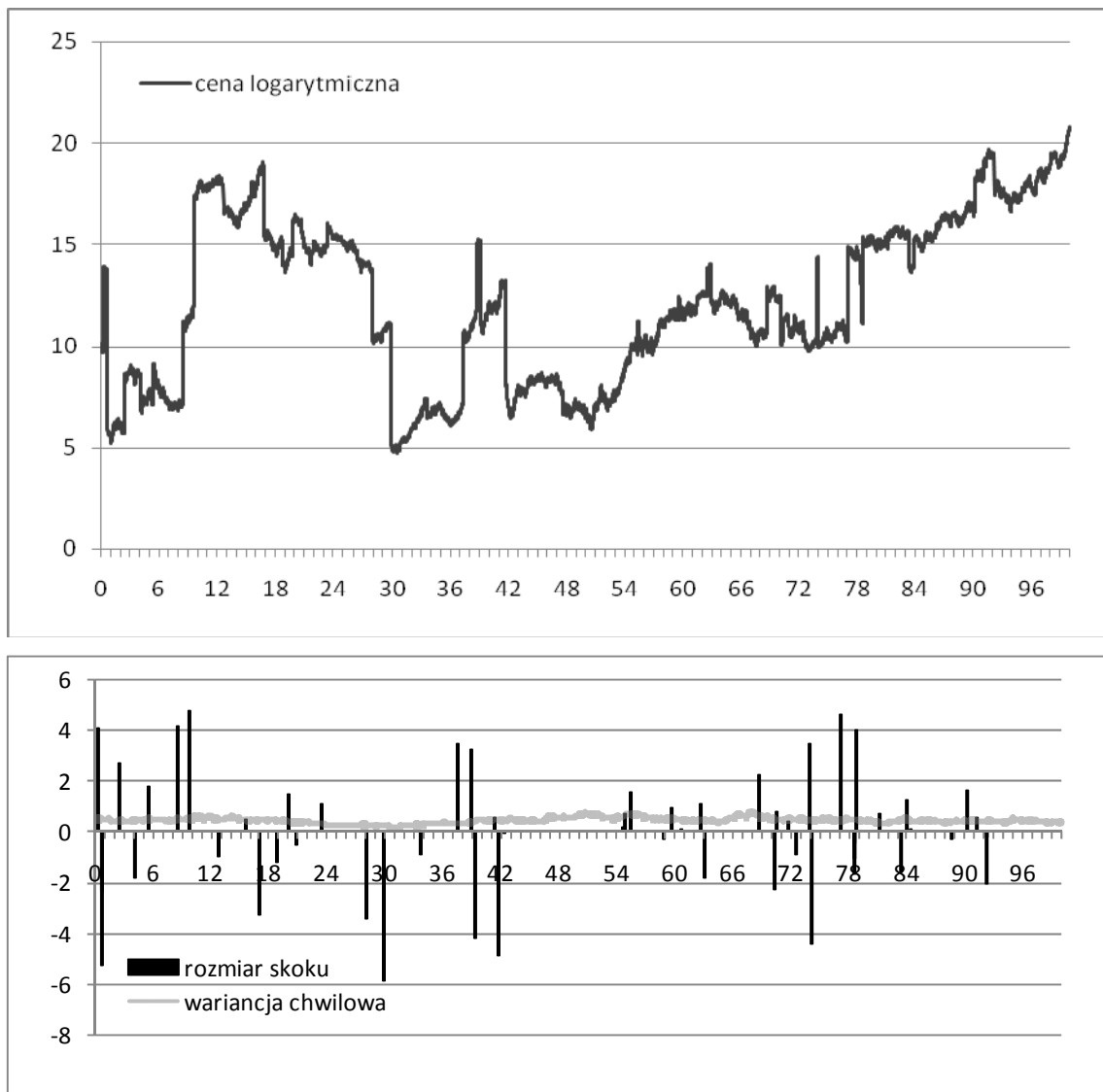
gdzie  $B_t^{(1)}$  i  $B_t^{(2)}$  są niezależnymi względem siebie ruchami Browna. Proces  $X_t$  będący rozwiązaniem układu (4.17) nazywany jest *procesem GARCH z czasem ciągłym*. Drugi z procesów, to *proces GARCH z czasem ciągłym ze skokami*. Opisuje go następujący układ równań

$$\begin{cases} dX_t = \sigma(t)dB_t^{(1)} + c(t)dq_t, \\ d\sigma^2(t) = \theta[\omega - \sigma^2(t)]dt + \xi\sigma^2(t)dB_t^{(2)}, \\ c(t) \sim N(0, \sigma_j^2) \\ X_0 = x_0, \sigma^2(0) = \sigma_0^2, \end{cases} \quad (4.18)$$

gdzie  $q_t$  jest opisany procesem Poissona z parametrem  $\lambda$ . Odpowiadające okresowi 100 dni trajektorie procesów (4.17) oraz (4.18) zostały uzyskane za pomocą dyskretyzacji Milsteina, przy wykorzystaniu jednodominutowych odstępów czasowych. Na rysunku 4.1 przedstawione są symulowane trajektorie procesu (4.1) z parametrami  $\theta = 0,035$ ,  $\omega = 0,635$ ,  $\xi = 0,144$  i wartościami początkowymi  $x_0 = 10$ ,  $\sigma_0^2 = 0,5$ . Trajektorie procesu (4.18) z tymi samymi parametrami, przy dodatkowym założeniu, że  $\sigma_j^2 = 3$  oraz, że oczekiwana liczba skoków w ciągu dnia wynosi 0,5, przedstawione są na rysunku 4.2.

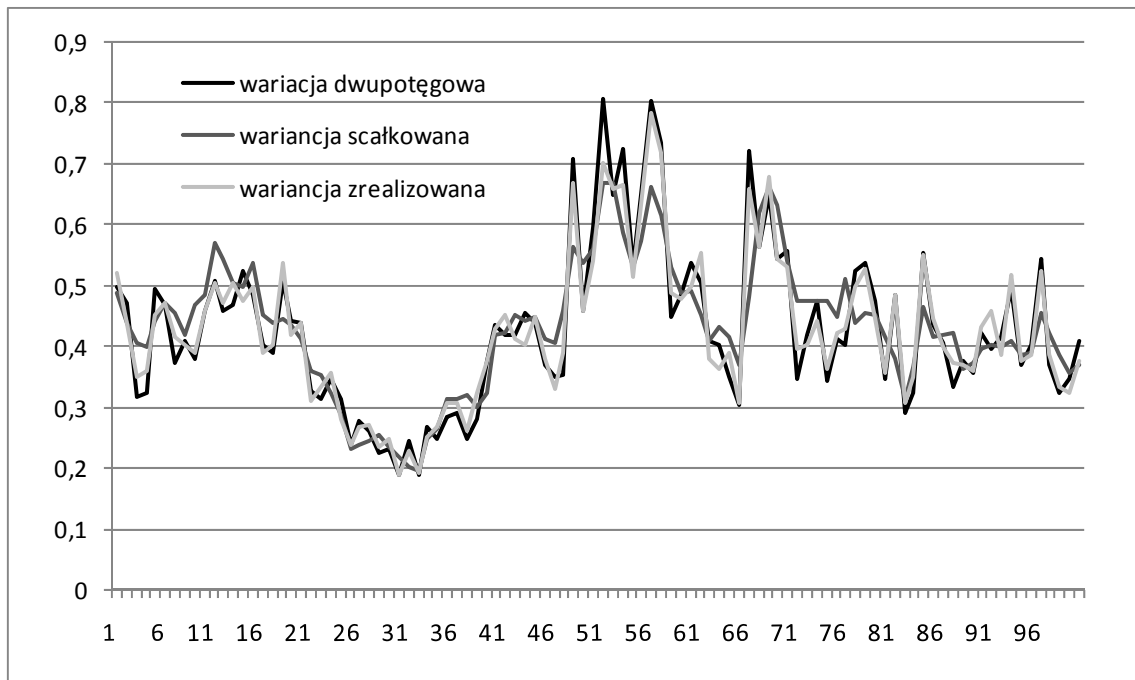


Rysunek 4.1. Symulowana trajektoria procesu opisanego  $X_t$  równaniem (4.17) z parametrami  $\theta = 0,035$ ,  $\omega = 0,065$  i  $\xi = 0,144$  oraz wartościami początkowymi  $x_0 = 10$  i  $\sigma_0^2 = 0,5$ . Poniżej trajektoria wariancji chwilowej  $\sigma^2(t)$ .



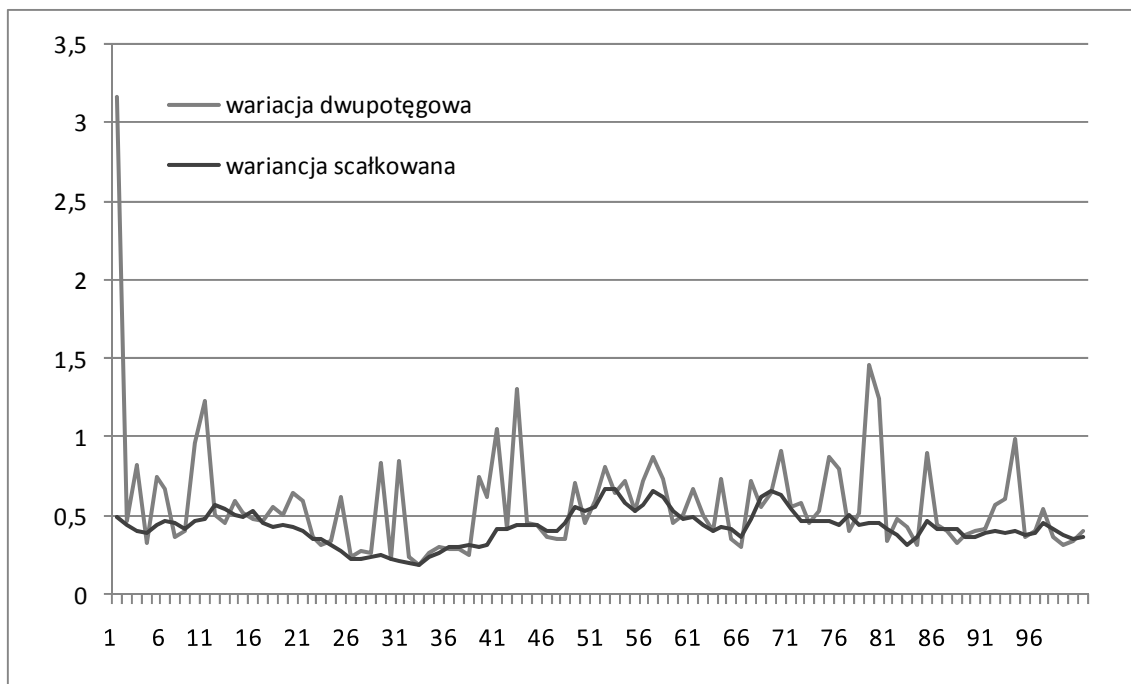
Rysunek 4.2. Symulowana trajektoria procesu opisanego  $X_t$  równaniem (4.17) z parametrami  $\theta = 0,035$ ,  $\omega = 0,065$  i  $\xi = 0,144$  i  $\sigma_J^2 = 3$  oraz wartościami początkowymi  $x_0 = 10$  i  $\sigma_0^2 = 0,5$ . Poniżej trajektoria wariancji chwilowej  $\sigma^2(t)$  oraz rozmiar skoku.

W sytuacji nie występowania skoków w procesie, zarówno wariancja dwupotęgorowa pomnożona przez  $\mu_1^{-2}$ , jak i wariancja zrealizowana są estymatorami wariancji scałkowanej (rozdział 2.3.2). Na rysunku 4.3 przedstawiona jest wariancja zrealizowana i wariancja dwupotęgorowa wraz z wariancją scałkowaną procesu opisanego równaniem (4.17), otrzymane na podstawie danych 5-minutowych.

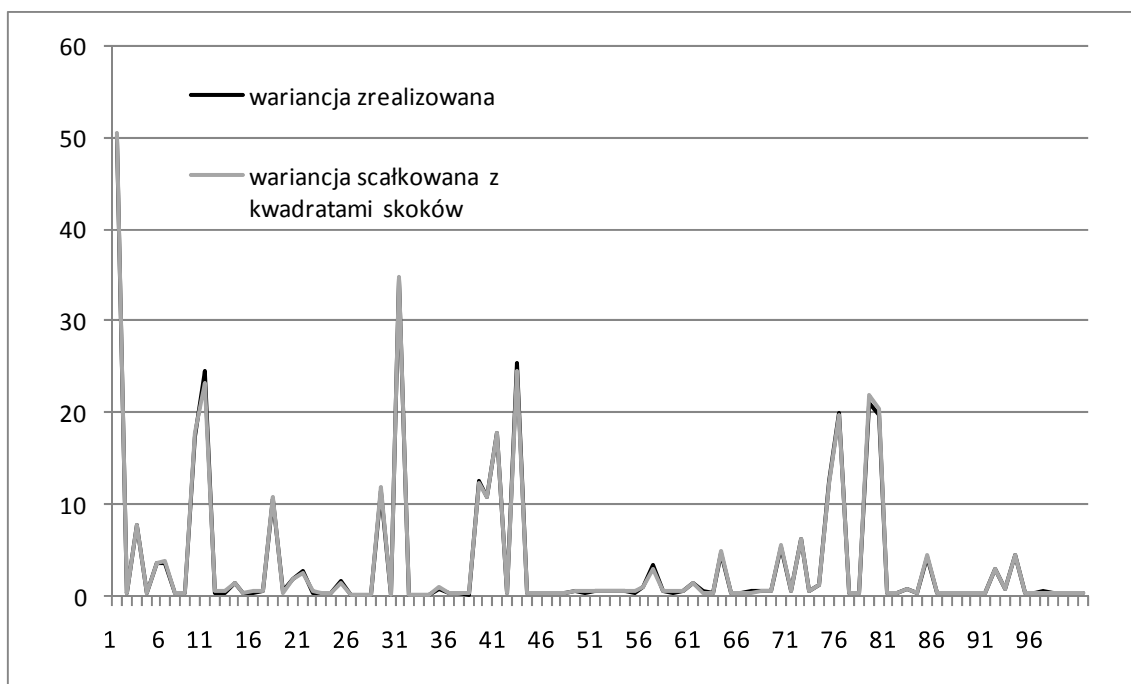


**Rysunek 4.3.** Wariancja scałkowana wraz z wyznaczoną na podstawie danych 5-minutowych wariancją zrealizowaną oraz pomnożoną przez  $\mu_1^{-2}$  wariancją dwupotęgową procesu opisanego równaniem (4.17).

Jak widać na rysunku (4.3), w sytuacji braku skoków w procesie cen logarytmicznych wariancja dwupotęgową oraz wariancja zrealizowana przybliżają wariancję scałkową równie dobrze. Rozbieżności w wartościach dwóch estymatorów są bardzo niewielkie, co przemawia za małym prawdopodobieństwem popełnienia błędu I rodzaju przy stosowaniu testów Barndorff-Nielsen i Shepharda, a więc za małym rozmiarem tych testów. Jeżeli w procesie stochastycznym występują skoki, wówczas pomnożona przez  $\mu_1^{-2}$  wariancja dwupotęgową przybliży zmienność scałkową, natomiast wariancja zrealizowana jest estymatorem wariancji scałkowej wraz z kwadratami skoków (rozdział 2.3.2). Na rysunkach 4.4 i 4.5 przedstawione są pomnożona przez  $\mu_1^{-2}$  wariancja dwupotęgową wraz z wariancją scałkową procesu (4.18) oraz wariancja zrealizowana wraz z wariancją scałkową z kwadratami skoków procesu (4.18). Jak wynika z wykresów, w sytuacji gdy w procesie cen logarytmicznych występują skoki, wariancja dwupotęgową nie jest już tak dobrym estymatorem wariancji scałkowej, jak w przypadku ich braku. Skoki wpływają na zawyżenie oszacowania wariancji scałkowej, co może mieć negatywny wpływ na moc testu. Jednocześnie obserwujemy, że wariancja zrealizowana bardzo dobrze przybliży wariancję scałkową z dodanymi kwadratami skoków.



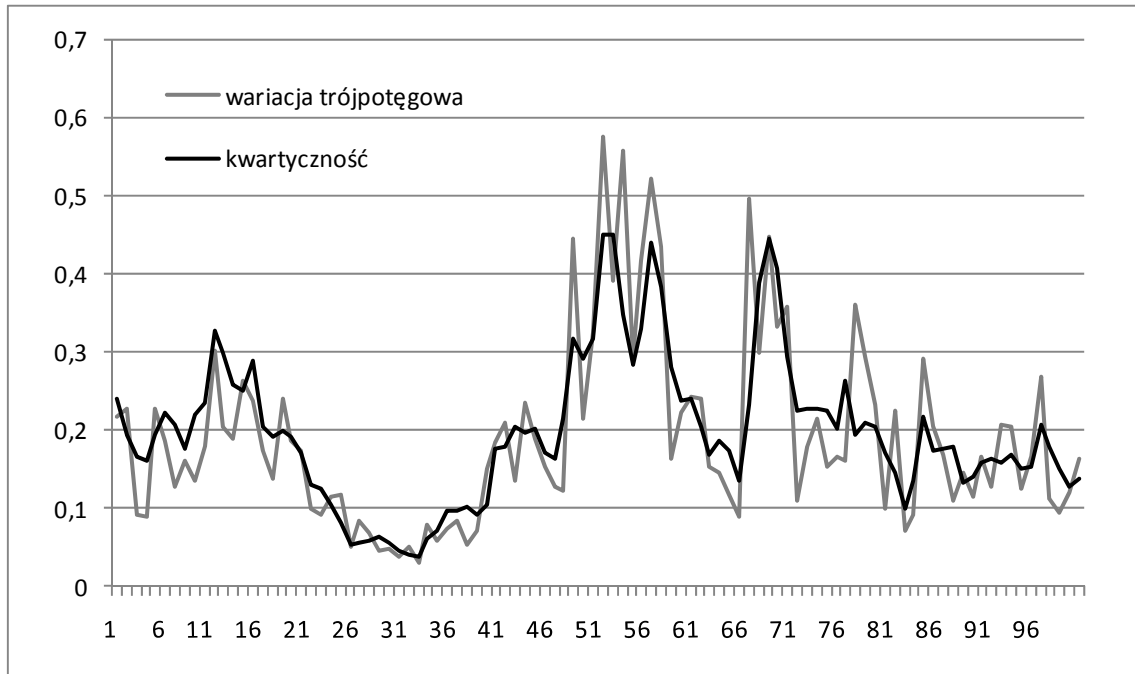
Rysunek 4.4. Wariacja dwupotęgową (pomnożona przez  $\mu_1^{-2}$ ) wraz z wariancją scałkowaną procesu opisanego równaniem (4.18).



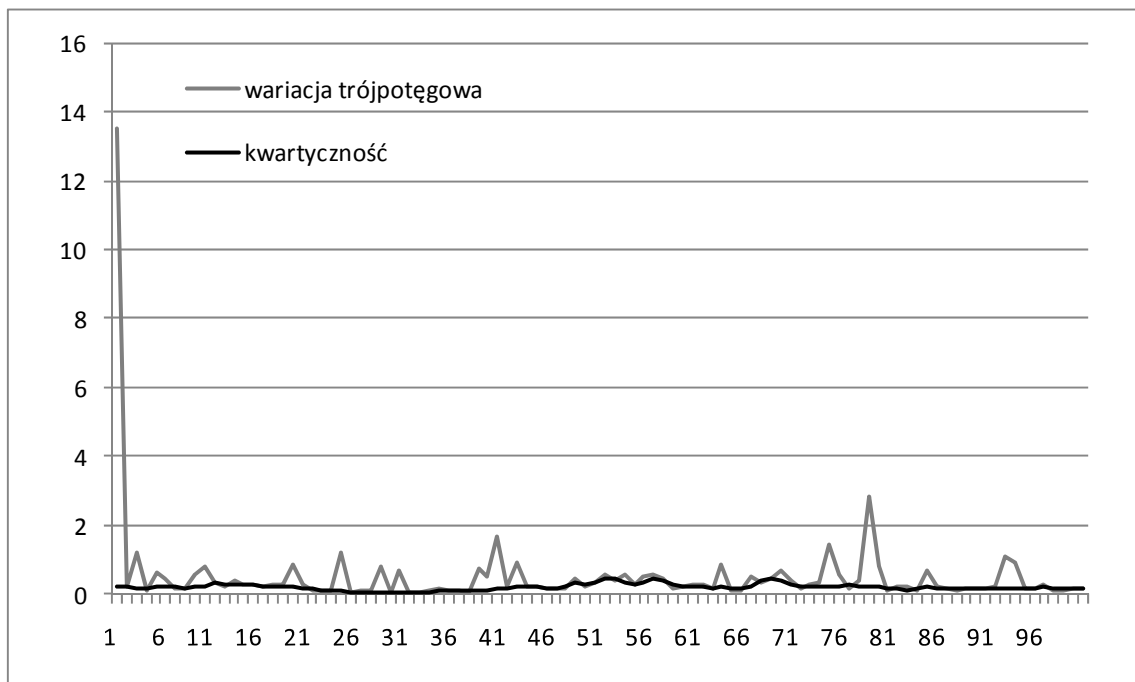
Rysunek 4.5. Wariancja scałkowana z dodanymi kwadratami skoków oraz wyznaczona na podstawie danych 5-minutowych wariancja zrealizowana procesu opisanego równaniem (4.18).

Pomnożona przez  $\mu_1^{-4}$  wariacja trójpotęgowa  $\{X_\delta\}_i^{[4/3,4/3,4/3]}$  jest estymatorem kwarcywności (rozdział 4.1). Własność tę wykorzystuje się do konstrukcji wprowadzonych przez Barndorff-Nielsen i Shepharda (2006) statystyk na występowanie skoków. Ry-

sunki 4.6 i 4.7 przedstawiają pomnożoną przez  $\mu_1^{-4}$  wariację trójpotęgową  $\{X_\delta\}_i^{[4/3, 4/3, 4/3]}$  zestawioną z kwartycznością wyznaczoną dla procesu opisanego układem równań (4.17) i (4.18).



Rysunek 4.6. Wyznaczona na podstawie danych 5-minutowych wariacja trójpotęgowa pomnożona przez  $\mu_1^{-4}$  wraz z kwartycznością procesu opisanego równaniem (4.17).



Rysunek 4.7. Wyznaczona na podstawie danych 5-minutowych wariacja trójpotęgowa pomnożona przez  $\mu_1^{-4}$  wraz z kwartycznością procesu opisanego równaniem (4.18).

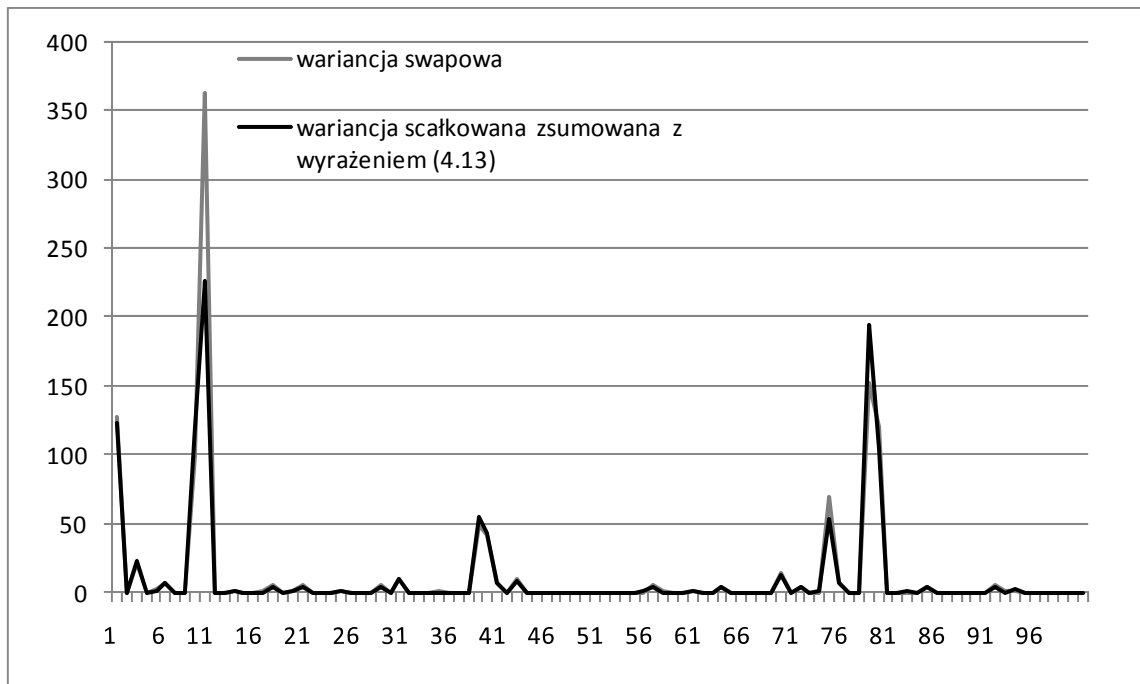
Jak obserwujemy na rysunku 4.7, w sytuacji występowania skoków w procesie cen logarytmicznych, wariancja trójpotęgowa znacząco zawyża oszacowanie kwartyczności. Wobec tego statystyki  $\hat{G}$ ,  $\hat{H}$  oraz  $\hat{J}$  są niedoszacowanymi estymatorami statystyk  $G$ ,  $H$  oraz  $J$ . Jest to na pewno najslabsza strona zaproponowanego przez Barndorff-Nielsen i Shepharda podejścia do testowania skoków.

Wykorzystywana w testach Jianga i Oomena wariancja swapowa wraz z wariancją scałkowaną i wariancją zrealizowaną przykładowej trajektorii procesu (4.17) oraz wariancja swapowa wraz z wariancją scałkowaną zsumowaną z wyrażeniem (4.13) dla trajektorii procesu (4.18) są przedstawione na rysunkach 4.8 i 4.9.



**Rysunek 4.8.** Wyznaczona na podstawie danych 5-minutowych wariancja swapowa i zrealizowana wraz z wariancją scałkowaną dla przykładowej trajektorii procesu opisanego równaniem (4.17).



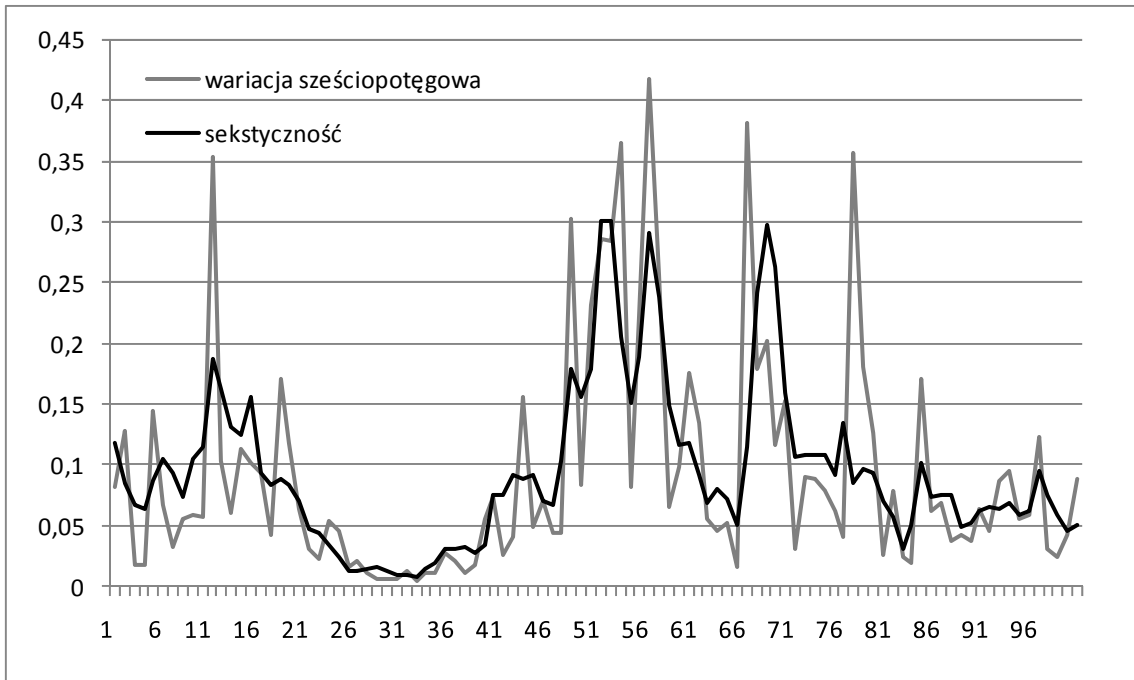


Rysunek 4.9. Wyznaczona na podstawie danych 5-minutowych wariancja swapowa wraz ze zsumowaną z wyrażeniem (4.13) wariancją skalowaną dla przykładowej trajektorii procesu opisanego równaniem (4.18).

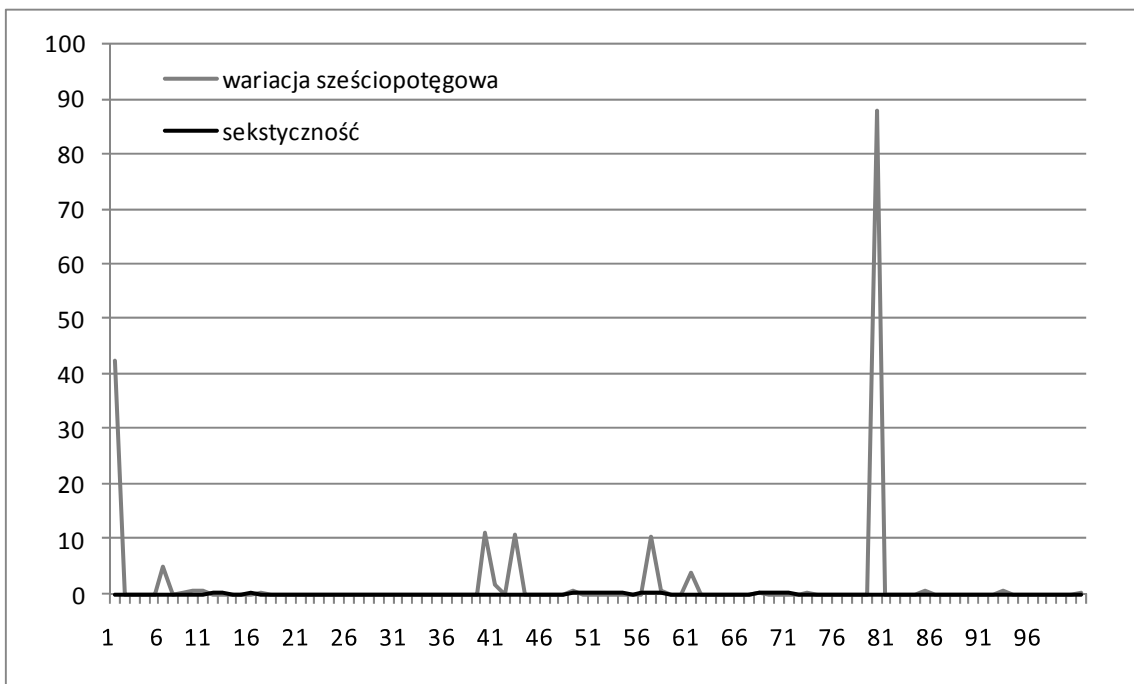
Z rysunków 4.8 i 4.9 widzimy, że w sytuacji, gdy w procesie cen logarytmicznych nie występują skoki, wariancja swapowa jest równie dobrym estymatorem wariancji skalowanej, co wariacja zrealityzowana. Jeśli natomiast proces cen logarytmicznych zawiera skoki, wówczas wariancja swapowa dość dobrze przybliża wariancję skalowaną zsumowaną z wyrażeniem (4.13).

Z rozdziału 4.2 wiemy, że pomnożona przez  $\mu_1^{-6}$  wariacja sześciopotęgorowa  $\{X_\delta\}_i^{[1,1,1,1,1,1]}$  jest estymatorem sekstyczności. Własność ta jest wykorzystana przez Jiana i Oomena (2008) do konstrukcji statystyk na występowanie skoków. Na rysunkach 4.10 i 4.11 przedstawiona jest pomnożona przez  $\mu_1^{-6}$  wariacja sześciopotęgorowa  $\{X_\delta\}_i^{[1,1,1,1,1,1]}$  zestawiona z sekstycznością wyznaczoną dla procesu opisanego układem równań (4.16) i (4.17).

Jak widać z rysunku 4.10 i 4.11, w sytuacji gdy w procesie cen logarytmicznych pojawiają się skoki, wariacja sześciopotęgorowa może znacznie zawyżyć oszacowanie sekstyczności. Jeżeli w procesie nie ma skoków, to również dokładność oszacowania pozostawia dużo do życzenia. Brak możliwości uzyskania dokładnego oszacowania sekstyczności jest największą wadą testu i głównym czynnikiem mogącym negatywnie wpłynąć na jego rozmiar i moc.



Rysunek 4.10. Wyznaczona na podstawie danych 5-minutowych pomnożona przez  $\mu_1^{-6}$  wariacja sześciopotęgorowa wraz z seksstycznością procesu opisanego równaniem (4.17).



Rysunek 4.11. Wyznaczona na podstawie danych 5-minutowych pomnożona przez  $\mu_1^{-4}$  wariacja sześciopotęgorowa pomnożona przez  $\mu_1^{-6}$  wraz z seksstycznością procesu opisanego równaniem (4.18).

Aby ocenić rozmiar testu, symulujemy trajektorie procesu cen logarymicznych dla 10000 dni za pomocą modelu (4.17) z parametrami  $\theta = 0,035$ ,  $\omega = 0,635$  oraz  $\xi = 0,144$  i wartościami początkowymi  $x_0 = 10$  i  $\sigma_0^2 = 0,5$ . Do symulacji wykorzystujemy dyskretyzację Milsteina dla jednoninutowych odstępów czasowych. Następnie testujemy występowanie skoków niezależnie we wszystkich dniach, wykorzystując różne statystyki testowe, częstotliwości zwrotów oraz poziomy istotności. Rozmiar testu szacujemy jako stosunek odrzuceń hipotezy  $H_0$  do liczby wszystkich badanych dni. W celu oceny mocy testu dokonujemy symulacji procesów dla okresu 10000 dni za pomocą modelu opisanego równaniem (4.18) z tymi samymi parametrami, umieszczając w losowym momencie każdego dnia jeden skok, którego wartość opisana jest rozkładem  $N(0, \sigma_k^2)$  gdzie  $\sigma_k^2 = 3$ . Dla zmodyfikowanej trajektorii przeprowadzamy testy na występowanie skoków, podobnie jak robimy to wyznaczając rozmiar testu. Moc testu jest wówczas szacowana jako stosunek odrzuceń hipotezy  $H_0$  do liczby wszystkich badanych dni.

W tabelach 4.2-4.4 zamieszczone są oszacowania rozmiaru i mocy testów Barndorff-Nielsen i Shepharda.

**Tabela 4.2. Rozmiar i moc liniowego testu na skoki Barndorff-Nielsen i Shepharda w zależności od poziomu ufności i zastosowanych estymatorów zmienności skalkowanej i kwartyczości.**

Częstotliwość notowań	Estymatory zm. scalk. i kwartyczości	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	BV, TQ	26,41%	18,58%	15,95%	12,05%	93,07%	91,74%	91,38%	90,56%
	BV*, TQ*	6,97%	2,23%	1,44%	0,56%	89,45%	87,63%	87,06%	85,69%
	BV**, TQ**	3,17%	0,84%	0,53%	0,21%	88,65%	86,51%	85,78%	84,46%
10 min.	BV, TQ	34,67%	27,80%	25,45%	21,07%	92,54%	91,39%	90,99%	90,00%
	BV*, TQ*	7,72%	2,92%	2,09%	0,93%	86,71%	84,38%	83,54%	82,16%
	BV**, TQ**	3,52%	1,04%	0,67%	0,29%	85,59%	82,92%	82,24%	80,71%
30 min.	BV, TQ	44,37%	40,25%	38,61%	35,66%	90,46%	89,51%	89,24%	88,67%
	BV*, TQ*	10,05%	4,85%	3,76%	2,19%	80,70%	77,43%	76,45%	74,39%
	BV**, TQ**	4,50%	1,43%	1,03%	0,45%	79,22%	75,80%	74,69%	72,60%
60 min.	BV, TQ	51,07%	47,71%	46,75%	44,56%	92,86%	92,22%	91,97%	91,68%
	BV*, TQ*	12,76%	7,33%	5,88%	3,92%	83,18%	80,58%	79,75%	77,84%
	BV**, TQ**	5,78%	2,31%	1,67%	0,99%	82,24%	79,73%	78,73%	76,88%
120 min.	BV, TQ	57,05%	54,75%	53,95%	52,25%	88,02%	87,31%	87,12%	86,54%
	BV*, TQ*	15,76%	10,63%	9,03%	7,09%	67,73%	62,92%	61,44%	58,57%
	BV**, TQ**	7,08%	3,70%	2,87%	1,71%	65,68%	61,12%	59,56%	56,77%

**Tabela 4.3. Rozmiar i moc ilorazowego testu na skoki Barndorff-Nielsen i Shepharda w zależności od poziomu ufności i zastosowanych estymatorów zmienności skalkowanej oraz kwartyczności.**

Częstotliwość notowań	Estymatory zm. scałk. i kwartyczności	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	BV, TQ	7,22%	2,07%	1,32%	0,50%	89,50%	87,52%	86,84%	85,36%
	BV*, TQ*	6,27%	1,79%	1,13%	0,44%	89,21%	87,30%	86,61%	85,18%
	BV**, TQ**	2,70%	0,66%	0,43%	0,14%	88,34%	86,04%	85,26%	83,72%
10 min.	BV, TQ	8,03%	2,81%	1,87%	0,68%	86,76%	84,16%	83,22%	81,66%
	BV*, TQ*	6,81%	2,33%	1,49%	0,51%	86,36%	83,75%	82,96%	81,24%
	BV**, TQ**	2,79%	0,75%	0,46%	0,22%	85,18%	82,16%	81,28%	79,53%
30 min.	BV, TQ	11,05%	4,50%	3,28%	1,62%	81,01%	77,13%	75,61%	73,11%
	BV*, TQ*	8,65%	3,52%	2,58%	1,16%	79,88%	76,06%	74,76%	71,94%
	BV**, TQ**	3,25%	0,83%	0,55%	0,22%	78,29%	74,08%	72,67%	70,09%
60 min.	BV, TQ	14,74%	7,26%	5,52%	3,25%	83,79%	80,33%	79,19%	76,41%
	BV*, TQ*	10,80%	5,11%	3,91%	2,37%	82,13%	79,06%	77,64%	74,85%
	BV**, TQ**	3,54%	1,01%	0,70%	0,29%	81,20%	77,64%	76,17%	73,87%
120 min.	BV, TQ	19,42%	11,40%	9,38%	6,36%	70,12%	63,43%	61,08%	56,98%
	BV*, TQ*	13,34%	7,69%	6,38%	4,34%	65,52%	59,46%	57,28%	53,24%
	BV**, TQ**	3,92%	1,10%	0,77%	0,27%	63,03%	56,59%	54,27%	50,03%

**Tabela 4.4. Rozmiar i moc logarytmicznego testu na skoki Barndorff-Nielsen i Shepharda w zależności od poziomu ufności i zastosowanych estymatorów zmienności skalkowanej oraz kwartyczności.**

Częstotliwość notowań.	Estymatory zm. scałk. i kwartyczności	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	BV, TQ	5,90%	1,37%	0,81%	0,21%	89,13%	86,97%	86,26%	84,46%
	BV*, TQ*	5,20%	1,18%	0,65%	0,20%	88,91%	86,78%	86,01%	84,29%
	BV**, TQ**	1,92%	0,66%	0,43%	0,14%	87,79%	86,04%	85,26%	83,72%
10 min.	BV, TQ	6,25%	1,60%	0,87%	0,15%	86,16%	83,10%	82,21%	80,15%
	BV*, TQ*	5,22%	1,24%	0,71%	0,12%	85,66%	82,81%	81,81%	79,80%
	BV**, TQ**	1,88%	0,75%	0,46%	0,22%	84,05%	82,16%	81,28%	79,53%
30 min.	BV, TQ	7,04%	1,94%	1,07%	0,28%	79,20%	73,96%	72,10%	68,05%
	BV*, TQ*	5,48%	1,43%	0,82%	0,21%	78,06%	73,05%	71,00%	67,14%
	BV**, TQ**	1,58%	0,83%	0,55%	0,22%	76,05%	74,08%	72,67%	70,09%
60 min.	BV, TQ	8,33%	2,30%	1,28%	0,34%	81,25%	75,38%	73,10%	67,89%
	BV*, TQ*	5,87%	1,61%	0,86%	0,22%	79,85%	74,15%	71,65%	66,63%
	BV**, TQ**	1,48%	1,01%	0,70%	0,29%	78,33%	77,64%	76,17%	73,87%
120 min.	BV, TQ	9,23%	2,51%	1,43%	0,22%	61,85%	49,05%	43,53%	30,11%
	BV*, TQ*	6,17%	1,73%	0,86%	0,14%	57,76%	45,25%	39,64%	27,52%
	BV**, TQ**	0,93%	1,10%	0,77%	0,27%	54,42%	56,59%	54,27%	50,03%

Zauważmy, że dla wszystkich trzech testów im mniejsza jest częstotliwość danych, tym rozmiar testu jest większy, natomiast moc mniejsza. Sugeruje to konieczność stosowania możliwie największej częstotliwości notowań przy testowaniu skoków w szeregach

czasowych, pamiętając jednak, że dla rzeczywistych szeregów czasowych przy wysokiej częstotliwości danych pojawiają się silne efekty mikrostruktury mogące wyraźnie zniekształcać wyniki badań. Rozmiar testu jest zdecydowanie największy dla statystyki liniowej, co potwierdza sugestię Barndorff-Nielsena i Shepharda (2006) oraz Huanga i Tauchena (2005) że statystyka ta może istotnie zawyżać liczbę wykrytych skoków. Przy zastosowaniu nieprzeskalowanej wariacji dwupotęęgowej, wartość ta jest tak duża, że dyskwalifikuje tę statystykę w zastosowaniach praktycznych. Test ilorazowy i logarytmiczny ma natomiast nieco mniejszą moc. Testy z przeskalowanymi wariacjami dwu i trójpotęgowymi mają nieznacznie mniejszy rozmiar i moc. Efekt ten jest następstwem mniejszych wartości statystyk testowych. Chociaż estymatory te teoretycznie powinny lepiej przybliżać zmienność scałkowaną i kwartyczość, jak widać ich zastosowanie w powyższych statystykach nie ma istotnego wpływu na jakość testów. Godny uwagi jest natomiast ogromny spadek rozmiaru testu w sytuacji zastosowania przesuniętych wariacji dwu- i trójpotęgowych, przy jednocześnie niewielkim osłabieniu mocy testu. W największym stopniu jest to widoczne dla testu logarytmicznego. Wynik ten jest potwierdzeniem rezultatów otrzymanych przez Huanga i Tauchena (2005), wskazujących na zaskakująco dobrą jakość tego testu.

Analogiczne badania przeprowadzamy dla testu Jianga i Oomena (2008). W tabelach 4.5-4.7 zamieszczone są oszacowania rozmiaru i mocy testów.

**Tabela 4.5. Rozmiar i moc liniowego testu Jianga i Oomena na skoki, w zależności od poziomu ufności i zastosowanych estymatorów zmienności scałkowanej i sekstyczności.**

Częstotliwość notowań	Estymatory sekstyczności	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	SS	7,42%	2,21%	1,35%	0,52%	94,18%	93,54%	93,39%	92,92%
10 min.	SS	9,07%	3,09%	2,06%	0,94%	92,78%	91,87%	91,51%	90,97%
30 min.	SS	14,30%	6,96%	5,39%	3,31%	90,05%	88,31%	87,83%	86,96%
60 min.	SS	11,74%	4,90%	3,44%	1,66%	88,11%	86,37%	85,71%	84,57%
120 min.	SS	34,23%	25,80%	23,61%	19,71%	87,37%	85,10%	84,37%	82,65%

**Tabela 4.6. Rozmiar i moc ilorazowego testu Jianga i Oomena na skoki, w zależności od poziomu ufności i zastosowanych estymatorów zmienności skalkowanej i sekstyczności.**

Częstotliwość notowań.	Estymatory zm. scałk. i seks tyczności	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	BV, SS	7,15%	1,81%	1,08%	0,35%	94,06%	93,39%	93,18%	92,73%
	BV*, SS	7,16%	1,81%	1,08%	0,35%	94,06%	93,39%	93,18%	92,73%
10 min.	BV, SS	8,31%	2,47%	1,60%	0,52%	92,56%	91,48%	91,10%	90,47%
	BV*, SS	8,35%	2,48%	1,60%	0,52%	92,57%	91,49%	91,10%	90,47%
30 min.	BV, SS	11,79%	4,92%	3,55%	1,75%	89,46%	87,48%	86,88%	85,83%
	BV*, SS	11,80%	4,92%	3,56%	1,75%	89,47%	87,49%	86,88%	85,84%
60 min.	BV, SS	17,47%	9,69%	7,90%	4,85%	87,16%	84,76%	83,93%	82,34%
	BV*, SS	17,51%	9,70%	7,91%	4,85%	87,16%	84,76%	83,93%	82,34%
120 min.	BV, SS	28,39%	19,54%	17,14%	13,61%	85,75%	82,60%	81,53%	79,24%
	BV*, SS	28,43%	19,55%	17,15%	13,65%	85,76%	82,60%	81,53%	79,24%

**Tabela 4.7. Rozmiar i moc logarytmicznego testu Jianga i Oomena na skoki, w zależności od poziomu ufności i zastosowanych estymatorów zmienności skalkowanej i sekstyczności.**

Częstotliwość notowań.	Estymatory zm. scałk. i kwartyczności	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	BV, SS	7,14%	1,82%	1,06%	0,35%	94,06%	93,37%	93,18%	92,73%
	BV*, SS	7,14%	1,82%	1,07%	0,35%	94,06%	93,38%	93,18%	92,73%
10 min.	BV, SS	8,26%	2,43%	1,56%	0,51%	92,58%	91,49%	91,10%	90,45%
	BV*, SS	8,27%	2,44%	1,58%	0,51%	92,58%	91,49%	91,10%	90,45%
30 min.	BV, SS	11,74%	4,87%	3,44%	1,66%	89,46%	87,47%	86,94%	85,81%
	BV*, SS	11,74%	4,90%	3,44%	1,66%	89,47%	87,48%	86,94%	85,81%
60 min.	BV, SS	17,29%	9,59%	7,66%	4,81%	86,99%	84,82%	84,09%	82,30%
	BV*, SS	17,30%	9,61%	7,67%	4,81%	87,02%	84,82%	84,09%	82,30%
120 min.	BV, SS	28,16%	19,36%	16,97%	13,23%	85,68%	82,69%	81,49%	79,18%
	BV*, SS	28,17%	19,38%	16,98%	13,23%	85,69%	82,69%	81,49%	79,18%

Podobnie jak dla testów Barndorff-Nielsen i Shepharda, dla testów Jianga i Oomena można zauważyć wpływ częstotliwości notowań, wykorzystanych do oszacowania wariacji zrealizowanej, wariacji dwupotęgowej i sekstyczności, na rozmiar i moc testów. Dla wszystkich trzech testów obserwujemy wzrost rozmiaru i spadek mocy wraz ze zmniejszającą się częstotliwością. Jak łatwo zaobserwować, statystyka ilorazowa i logarytmiczna ma dla zdecydowanej większości przypadków większą moc i mniejszy rozmiar od statystyki liniowej. Jednocześnie można zauważyć, że wielkości te są bardzo zbliżone dla statystyki ilorazowej i logarytmicznej. Ponadto widoczny jest znikomy wpływ zastosowanego estymatora wariacji skalkowanej na wyniki testu.

Porównując wyniki dwóch alternatywnych podejść do testowania skoków, można zauważyć, że testy Barndorff-Nielsena mają w większości badanych przypadków mniejszy rozmiar, natomiast na korzyść testów Jianga i Oomena przemawia nieco większa moc. Oznacza to, że nie można jednoznacznie powiedzieć, które testy są lepsze i należałoby je w pierwszej kolejności zweryfikować w praktyce. Można oczekiwać, że testy Barndorff-Nielsena mogą w pewnym stopniu niedoszacowywać liczbę wykrywanych skoków w szeregach czasowych, natomiast testy Jianga i Oomena mogą ją przeszacowywać.

Ponieważ test Lee i Myklanda (2007) bada niezależnie skoki we wszystkich okresach śróddziennych, wobec tego w celu dokonania jego oceny zastosowaliśmy nieco inne podejście. Na początku, wykorzystując dyskretyzację Milsteina dla jednonumitowych odstępów czasowych, przeprowadzamy symulację procesu dobierając długość trajektorii w taki sposób, aby otrzymać 100000 notowań śróddziennych. Symulujemy niezależnie dwa procesy z czasem ciągłym. Pierwszy bez dryfu opisany jest układem równań

$$\begin{cases} dX_t = \sigma(t)dB_t^{(1)} + c(t)dq_t \\ d\sigma^2(t) = \theta[\omega - \sigma^2(t)]dt + \xi\sigma^2(t)dB_t^{(2)} \\ c(t) \sim N(0, \sigma_J^2) \\ X_0 = x_0, \sigma^2(0) = \sigma_0^2, \end{cases} \quad (4.19)$$

gdzie  $q_t$  jest procesem Poissona z parametrem  $\lambda$ . Za wszystkie parametry podstawiamy takie same wartości jak przy procesie danym równaniem (4.17). Za dodatkowe parametry  $\sigma_J^2$  i  $\lambda$  podstawiamy odpowiednio 3 i 0,001. Drugi proces różni się od poprzedniego funkcją dryfu w równaniu cen logarytmicznych. Opisany jest następującym układem równań

$$\begin{cases} dX_t = \kappa(\mu - X(t))dt + \sigma(t)dB_t^{(1)} + c(t)dq_t \\ d\sigma^2(t) = \theta[\omega - \sigma^2(t)]dt + \xi\sigma^2(t)dB_t^{(2)} \\ c(t) \sim N(0, \sigma_J^2) \\ X_0 = x_0, \sigma^2(0) = \sigma_0^2 \end{cases} \quad (4.20)$$

z takimi samymi parametrami, jak proces dany równaniem (4.19). Za parametry funkcji dryfu podstawiliśmy odpowiednio wartości 0,1 i 10,5. W tabelach 4.12 i 4.13 znajdują się oszacowania rozmiaru i mocy testów Lee i Myklanda (2007) wyznaczone przy wykorzystaniu symulacji procesów opisanych równaniami (4.19) i (4.20).

**Tabela 4.12. Rozmiar testu Lee i Mykland na skoki, w zależności od poziomu ufności i zastosowanych estymatorów zmienności scalkowanej i seks tyczności, wyznaczony przy wykorzystaniu symulacji procesu opisanego równaniem (4.19).**

Częstotliwość notowań.	Statystyka testowa	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	93,04%	93,04%	92,64%	92,45%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	93,24%	93,04%	92,84%	92,25%
10 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	85,12%	84,91%	84,51%	83,89%
	$\hat{L}$	0,10%	0,00%	0,00%	0,00%	85,93%	85,12%	85,02%	84,30%
30 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	76,50%	74,76%	74,18%	72,75%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	76,33%	74,80%	74,11%	72,54%
60 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	62,58%	60,18%	59,28%	57,70%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	62,49%	59,92%	59,09%	57,48%
120 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	43,14%	40,20%	39,03%	36,27%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	42,99%	40,11%	38,93%	36,41%

**Tabela 4.13. Rozmiar testu Lee i Mykland na skoki, w zależności od poziomu ufności i zastosowanych estymatorów zmienności scalkowanej i seks tyczności, wyznaczony przy wykorzystaniu symulacji procesu opisanego równaniem (4.20).**

Częstotliwość notowań.	Estymatory zm. scalk. i kwartyczności	Rozmiar testu				Moc testu			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
5 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	75,28%	74,47%	74,31%	73,82%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	75,12%	74,31%	74,31%	73,82%
10 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	80,18%	78,90%	78,62%	77,43%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	80,09%	79,27%	78,72%	77,43%
30 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	74,59%	73,27%	72,58%	71,06%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	74,59%	73,20%	72,45%	70,96%
60 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	60,37%	58,28%	57,60%	55,72%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	60,28%	58,36%	57,65%	55,65%
120 min.	$\hat{L}_\mu$	0,00%	0,00%	0,00%	0,00%	42,54%	39,63%	38,64%	36,29%
	$\hat{L}$	0,00%	0,00%	0,00%	0,00%	42,49%	39,69%	38,57%	36,21%



Na podstawie wyników symulacji przedstawionych w tabelach 4.12 i 4.13, łatwo zauważyć, że test Lee i Myklanda praktycznie nigdy nie odrzuca prawdziwej hipotezy zerowej, co niewątpliwie jest jego ogromną zaletą. Podobny rezultat otrzymali Lee i Mykland (2007) którzy zastosowali prostszy model do symulacji. Sytuacja ta sugeruje możliwość zastosowania stosunkowo niskich poziomów ufności, chociaż z tabeli 4.8 i 4.9 wynika, że nawet dla bardzo wysokich poziomów ufności moc testu nie zmniejsza się znacząco. Jest ona stosunkowo duża przy użyciu danych wysokiej częstotliwości, niemniej dla danych 60 i 120 minutowych drastycznie spada, co w rezultacie przełoży się na znaczne zaniżenie liczby wykrytych skoków. Ponadto, dość zaskakującym spostrzeżeniem może być fakt, że statystyki  $\hat{L}_\mu$  i  $\hat{L}$  zastosowane zarówno do symulacji procesu z dryfem, jak i bez dryfu, mają niemalże równą moc. Sugeruje to możliwość zastosowania prostszego w implementacji i mniej czasochłonnego testu ze statystyką  $\hat{L}$ , bez względu na charakterystykę procesu.

## 5. Uogólniona metoda momentów

Wprowadzona przez Hansena (1982) Uogólniona Metoda Momentów nazywana w skrócie GMM (Generalized Method of Moments) powstała przez uogólnienie opracowanej przez Pearsona pod koniec XIX wieku metody momentów. Metoda momentów służyła do estymacji parametrów takich jak średnia, wariancja, mediana itp. Polegała na porównaniu momentów wyznaczonych z próby z nieobserwowalnymi teoretycznymi wartościami momentów przy założeniu pewnego znanego rozkładu zmiennych. W ten sposób otrzymujemy układ równań, którego rozwiązaniem są oszacowania szukanych parametrów. Metoda wprowadzona przez Hansena dotyczy estymacji parametrów znacznie szerszego zbioru modeli statystycznych, w tym także modeli dyfuzji. W tym celu wykorzystali ją m.in. Chan i in. (1992). Inne przykłady jej zastosowań tej metody można znaleźć w podręcznikach Gourieroux i Jasiak (2001) Hamiltona (1994) czy Verbecka (2004). Jej zasadnicza idea jest zbliżona do metody momentów i polega na doborze parametrów rozważanego modelu w taki sposób, aby momenty modelu były możliwie bliskie momentom z próby. Dodatkowo wprowadza się macierz wag, której zadaniem jest określenie względnego znaczenia każdego momentu.

Załóżmy, że  $\{X_t : t = 1, 2, \dots\}$  jest wielowymiarowym procesem z czasem dyskretnym i że dysponujemy skończoną liczbą jego realizacji. Model jest określony jako warunkowy wektor warunków na momenty

$$\mathbb{E}[g(X_{t+1}, \theta_0) | \mathcal{F}_t] = 0,$$

gdzie funkcja  $g$  jest  $L$ -wymiarową funkcją wektorową a  $\theta_0$  jest nieznanym wektorem parametrów należącym do przestrzeni parametrów  $\mathbb{P} \subset \mathbb{R}^p$ . Ponadto przyjmujemy założenie, że jeśli  $\theta$  należy do przestrzeni parametrów, to

$$\mathbb{E}[g(X_{t+1}, \theta) | \mathcal{F}_t] = 0 \text{ tylko wtedy, gdy } \theta = \theta_0.$$

Sama wartość wektora parametrów  $\theta_0$  nie wystarczy do opisu funkcji wiarygodności. Do konstrukcji modelu opisującego generowanie danych potrzebne są także inne parametry.

Wybermy teraz  $K$  zmiennych instrumentalnych  $z_{1,t}, \dots, z_{K,t}$  i zastąpmy  $L$  warunkowych restrykcji zbiorem  $KL$  brzegowych warunków na momenty

$$\mathbb{E}[z_{k,t}g_l(X_{t+1}, \theta_0)] = 0 \text{ dla } k = 1, \dots, K \text{ oraz } l = 1, \dots, L.$$

Oznaczając  $z_t = (z_{1,t}, \dots, z_{K,t})'$ , zbiór brzegowych warunków na momenty możemy zapisać w postaci

$$\mathbb{E}[g(X_{t+1}, \theta_0) \otimes z_t] = 0. \quad (5.1)$$

gdzie symbol  $\otimes$  oznacza iloczyn Kroneckera. Uogólniona metoda momentów polega na aproksymowaniu powyższych warunków na momenty i rozwiązaniu otrzymanego w ten sposób układu równań ze względu na  $\theta_0$ . Ze względu na wartości  $KL$  oraz  $p$  musimy rozważyć trzy przypadki rozwiązania powyższego problemu.

- (i) *Niedostateczna identyfikacja.* Jeśli liczba instrumentów jest mniejsza od liczby parametrów ( $KL < p$ ), to nie ma dostatecznie wielu równań, by móc jednoznacznie wyznaczyć wartości parametrów.
- (ii) *Dokładna identyfikacja.* Liczba instrumentów równa jest liczbie parametrów.
- (iii) *Nadmierna identyfikacja.* Jeśli liczba instrumentów jest większa od liczby parametrów ( $KL > p$ ), to równań jest więcej niż poszukiwanych parametrów, wobec czego nie możemy znaleźć dokładnego rozwiązania. Przybliżone rozwiązanie potrafimy znaleźć aproksymacyjnie.

Załóżmy, że  $(z_t, X_{t+1})$  jest procesem stacjonarnym. Wartość oczekiwana (5.1) może być przybliżona wyrażeniem

$$\frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta_0) \otimes z_t.$$

**Definicja 5.1.1.** Jeżeli  $KL > p$ , to estymator wektora parametrów  $\theta_0$  oparty na warunkach momentów (5.1) wektorze zmiennych instrumentalnych  $z_t$  oraz dodatnio określonej macierzy wag  $W$  dany jest wzorem

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} J_T,$$

gdzie

$$J_T = \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right]' W \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right]. \quad (5.2)$$

Powyższy estymator jest zgodny, a jego efektywność zależy od macierzy wag. Fakt ten orzeka następujące twierdzenie

**Twierdzenie 5.1.1.** Estymator  $\hat{\theta}$  jest zgodny, a ponadto zachodzi następujący warunek

$$\sqrt{T}(\hat{\theta} - \theta_0) = N(0, \Sigma(W)),$$

gdzie

$$\begin{aligned} \Sigma(W) = & \left\{ \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta_0) \otimes z_t \right]' \mathbb{W} \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta_0) \otimes z_t \right] \right\}^{-1} \\ & \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta_0) \otimes z_t \right]' \mathbb{W} V_{as} \left[ \frac{1}{\sqrt{T}} \sum_{t=1}^T g(X_{t+1}, \theta_0) \otimes z_t \right] \mathbb{W} \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta_0) \otimes z_t \right] \\ & \left\{ \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta_0) \otimes z_t \right]' \mathbb{W} \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta_0) \otimes z_t \right] \right\}^{-1} \end{aligned}$$

i  $V_{as}$  oznacza graniczną wariancję dla dostatecznie dużego  $T$ .

## 5.1. Macierz wag

Jak wiadomo, w sytuacji gdy dysponujemy taką samą liczbą warunków na momenty jak parametrów, wówczas parametry są doskonale dopasowane, natomiast funkcja (5.2) przyjmuje dla nich wartość 0. Jednakże, gdy warunków na momenty jest więcej niż parametrów, konieczne jest wprowadzenie macierzy wag, która określa względny wpływ poszczególnych momentów na dany parametr. Hansen (1982) zaproponował wykorzystanie odwróconej asymptotycznej macierzy kowariancji jako macierzy wag

$$W^* = \left\{ V_{as} \left[ \frac{1}{\sqrt{T}} \sum_{t=1}^T g(X_{t+1}, \theta_0) \otimes z_t \right] \right\}^{-1} = S^{-1} \quad (5.3)$$

Otrzymane w ten sposób oszacowania parametrów mają najmniejszą wariancję równą

$$\Sigma(W^*) = \left\{ \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta) \otimes z_t \right]' \mathbb{W} \mathbb{E} \left[ \frac{\partial g}{\partial \theta'}(X_{t+1}, \theta) \otimes z_t \right] \right\}^{-1}.$$

Intuicyjnie, większą wagę otrzymują warunki momentów z mniejszą niepewnością. Ze względu na nieznaną wartość rzeczywistego wektora parametrów optymalna macierz wag nie jest znana. Jej przybliżoną postać można wyznaczyć za pomocą następującej procedury iteracyjnej

$$\begin{aligned}
 W_0 &= I_{KL} \\
 \hat{\theta}_1 &= \arg \min_{\theta} \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right]' W_0 \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right], \\
 W_1 &= W^*(\hat{\theta}_1) \\
 \hat{\theta}_2 &= \arg \min_{\theta} \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right]' W_1 \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right], \\
 &\dots \\
 \hat{\theta}_n &= \arg \min_{\theta} \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right]' W_{n-1} \left[ \frac{1}{T} \sum_{t=1}^T g(X_{t+1}, \theta) \otimes z_t \right],
 \end{aligned}$$

gdzie  $I_m$  oznacza macierz jednostkową stopnia  $m$ . Prowadząc iterację do  $n$ -tego kroku, otrzymujemy tzw.  $n$ -stopniową metodę GMM. Metodę należy stosować tak długo, aż funkcja celu przyjmie relatywnie małe wartości. Pewnym problemem jest duża wrażliwość metody na skalowanie danych. W celu uniknięcia tej niedogodności Cliff (2003) proponuje przyjęcie za macierz  $W_0$  macierzy  $[I_L \otimes Z'Z]^{-1}$ , gdzie  $Z$  jest macierzą, której  $t$ -tym wierszem jest  $z_t$ .

Macierz kowariancji  $S$  może być także rozumiana jako macierz gęstości spektralnych dla częstotliwości 0. Istnieje wiele podejść do jej estymacji. Newey i West (1987) wykorzystali w tym celu jądro Bartletta. W ogólnej formie estymator macierzy  $S$  można zapisać w postaci (por. Burnside, Eichenbaum 1994, Den Haan, Levin 1996)

$$\hat{S} = \sum_{j=-(T-1)}^{T-1} k\left(\frac{j}{B_T}\right) S_j,$$

gdzie

$$S_j = \begin{cases} \frac{1}{T} \sum_{t=j+1}^T [g(X_{t+1}, \theta) \otimes z_t][g(X_{t-j+1}, \theta)' \otimes z_t'] & \text{dla } j \geq 0 \\ \frac{1}{T} \sum_{t=j+1}^T [g(X_{t+1}, \theta) \otimes z_t][g(X_{t-j+1}, \theta)' \otimes z_t'] & \text{dla } j < 0 \end{cases}$$

oraz  $k(x)$  jest funkcją wag określoną za pomocą jądra Bartletta danego wzorem

$$k(x) = \begin{cases} 1 - |x| & \text{dla } |x| \leq 1 \\ 0 & \text{poza tym.} \end{cases}$$

Ponadto  $B_T$  jest parametrem wygładzania. Gallant (1987) zmodyfikował postać macierzy kowariancji zastępując funkcję wag  $k(x)$  jądrem Parzena wyznaczonym wzorem

$$k_p(x) = \begin{cases} 1 - 6x^2 - 6|x|^3 & \text{dla } 0 \leq |x| \leq 0,5, \\ 2(1 - |x|)^3 & \text{dla } 0,5 \leq |x| \leq 1, \\ 0 & \text{poza tym.} \end{cases}$$

Andrews (1991) natomiast zaproponował funkcję wykorzystanie jądra kwadratowo-spektralnego zdefiniowanego przez

$$k_{QS}(x) = \frac{25}{12\pi^2 x^2} \left( \frac{\sin(6\pi x / 5)}{6\pi x / 5} - \cos(6\pi x / 5) \right).$$

Jeszcze inne podejścia do wyznaczenia funkcji  $k(x)$  można znaleźć w pracach Hansena (1982) oraz Hansena i Hodricka (1980) gdzie wykorzystano w tym celu jądra obcięte. Pewnym problemem pozostaje odpowiedni dobór parametru wygładzania  $B_T$ . Andrews (1991) proponuje automatyczną procedurę doboru optymalnych wartości  $B_T$  dla różnego rodzaju funkcji wagowych. Wykazał on, że optymalna wartość  $B_T$  w sensie błędu średniokwadratowego równa jest

$$B_T^* = \begin{cases} 1,1447(\alpha(1)T)^{1/3} & \text{dla funkcji wag } k(x), \\ 2,6614(\alpha(2)T)^{1/5} & \text{dla funkcji wag } k_p(x), \\ 1,3221(\alpha(2)T)^{1/5} & \text{dla funkcji wag } k_{QS}(x), \end{cases}$$

gdzie

$$\alpha(q) = \frac{2(\text{vec}f^{(q)})'W(\text{vec}f^{(q)})}{\text{tr}W(I + K)f(0) \otimes f(0)}$$

jest funkcją zależną od nieznannej macierzy gęstości spektralnych danej wzorem

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \Omega_j \exp(-ij\lambda),$$

której  $q$ -tą uogólnioną pochodną można przedstawić w postaci

$$f^{(q)} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} |j|^q \Omega_j,$$

gdzie  $\Omega_j = \mathbb{E}[u_t(\hat{\theta})u_{t-j}(\hat{\theta})']$ , natomiast  $u_t(\hat{\theta})$  jest zależnym od wektora oszacowań parametrów wektorem reszt z modelu, natomiast  $K$  spełnia zależność  $\text{vec}(A') = K\text{vec}(A)$ , gdzie  $A$  jest macierzą postaci  $(I_M \ 0_{M \times J-M})$ , gdzie  $J$  jest wymiarem procesu  $X_t$ , a  $M \leq J$  jest wymiarem wektora momentów. Wartość  $\alpha(q)$  jest uzależniona od macierzy  $W$ . W najprostszym przypadku za macierz tą przyjmuje się macierz jednostkową.

Andrews and Monahan (1992) zaproponowali, by przed zastosowaniem powyższej procedury usunąć zależności liniowe z reszt  $u_t(\hat{\theta})$  za pomocą filtra  $\text{AR}(b)$ . W tym celu metodą najmniejszych kwadratów estymuje się model

$$u_t(\hat{\theta}) = \sum_{k=1}^b A_k u_{t-k}(\hat{\theta}) + e_t.$$

Andrews (1991) wyznaczył przybliżone wartości  $\alpha(1)$  oraz  $\alpha(2)$  w zależności od oszacowań współczynnika korelacji  $\hat{\rho}_l$  i wariancji  $\hat{\sigma}_l^2$  otrzymanych z modelu  $\text{AR}(1)$  dla każdego spośród  $M$  momentów. Dane są one wzorami

$$\hat{\alpha}(1) = \frac{\sum_{l=1}^M w_l \frac{4\hat{\rho}_l^2 \hat{\sigma}_l^4}{(1-\hat{\rho}_l)^6 (1+\hat{\rho}_l)^2}}{\sum_{l=1}^M w_l \frac{\hat{\sigma}_l^4}{(1-\hat{\rho}_l)^4}},$$

$$\hat{\alpha}(2) = \frac{\sum_{l=1}^M w_l \frac{4\hat{\rho}_l^2 \hat{\sigma}_l^4}{(1-\hat{\rho}_l)^8}}{\sum_{l=1}^M w_l \frac{\hat{\sigma}_l^4}{(1-\hat{\rho}_l)^4}}.$$

## 5.2. Warunki momentów

Warunki momentów ustalają takie wartości parametrów, przy których średnie funkcji  $g(X_{t+1}, \theta_0)$  przyjmują wartość 0. Bardzo proste ograniczenie można otrzymać w oparciu o średnią szeregu czasowego  $X_t$ . Jeżeli

$$\mathbb{E}(X_t) = \mu,$$

to otrzymujemy warunek

$$\mathbb{E}(X_t - \mu) = 0.$$

Inne ograniczenie możemy otrzymać wykorzystując wariancję szeregu  $X_t$ . Mamy

$$\mathbb{E}(X_t - \mu)^2 = \sigma^2.$$

Wobec tego otrzymujemy wektor warunków momentów

$$\mathbb{E} \begin{bmatrix} X_t - \mu \\ (X_t - \mu)^2 - \sigma^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Załóżmy, że dysponujemy teraz dwoma szeregami czasowymi  $X_t$  i  $Y_t$ . Oznaczmy przez  $\mu_X$ ,  $\mu_Y$ ,  $\sigma_X^2$ ,  $\sigma_Y^2$  i  $\sigma_{XY}^2$  odpowiednio średnie, wariancje oraz kowariancję szeregów  $X_t$  i  $Y_t$ . Wówczas możemy otrzymać następujący wektor warunków na momenty

$$\mathbb{E} \begin{bmatrix} X_t - \mu_X \\ Y_t - \mu_Y \\ (X_t - \mu_X)^2 - \sigma_X^2 \\ (Y_t - \mu_Y)^2 - \sigma_Y^2 \\ (X_t - \mu_X)(Y_t - \mu_Y) - \sigma_{XY}^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

### 5.3. Estymacja parametrów modeli dyfuzji

Metodę momentów zamierzamy wykorzystać do estymacji parametrów stochastycznych równań różniczkowych. Bierzemy pod uwagę te same modele, które rozważaliśmy w rozdziale trzecim. Zauważmy, że średnie i wariancje różniczki stochastycznej dla odpowiednich modeli dane są wzorami

$$\mathbb{E}(dX_t) = \kappa(\mu - X_t)dt \quad \text{i} \quad \text{Var}(dX_t) = \sigma^2 dt$$

dla modelu Vašíčka,

$$\mathbb{E}(dX_t) = \kappa(\mu - X_t)dt \quad \text{i} \quad \text{Var}(dX_t) = \sigma^2 X_t dt$$

dla modelu Coxa, Ingersolla i Rossa, oraz

$$\mathbb{E}(dX_t) = \kappa(\mu - X_t)dt \quad \text{i} \quad \text{Var}(dX_t) = \sigma^2 X_t^{2\beta} dt$$



dla modelu stałej elastyczności wariancji. Z uwagi na dostępność danych w dyskretnych odstępach czasowych, musimy zastąpić różniczkę  $dX_t$  przez zwrot  $y_t = X_t - X_{t-\delta}$ . Ponadto wybieramy wektor zmiennych instrumentalnych  $z_t = [1 \ X_t]$ . Wówczas dla każdego z rozważanych modeli dyfuzji otrzymujemy cztery momenty, które możemy zapisać w postaci

$$\begin{bmatrix} y_t - \kappa(\mu - X_t)\delta \\ (y_t - \kappa(\mu - X_t)\delta)^2 - \sigma^2\delta \end{bmatrix} \otimes [1 \ X_t]'$$

dla modelu Vašíčka,

$$\begin{bmatrix} y_t - \kappa(\mu - X_t)\delta \\ (y_t - \kappa(\mu - X_t)\delta)^2 - \sigma^2 X_t \delta \end{bmatrix} \otimes [1 \ X_t]'$$

dla modelu Coxa, Ingersolla i Rossa oraz

$$\begin{bmatrix} y_t - \kappa(\mu - X_t)\delta \\ (y_t - \kappa(\mu - X_t)\delta)^2 - \sigma^2 X_t^{2\beta} \delta \end{bmatrix} \otimes [1 \ X_t]'$$

dla modelu CKLS.

## 6. Modelowanie indeksów giełdowych i kursów walutowych

Niniejszy rozdział poświęcony jest ilustracji zastosowania procesów dyfuzji do modelowania indeksów giełdowych oraz kursów walutowych. W tym celu wykorzystaliśmy przede wszystkim metodę Phillipsa i Yu. Ponadto porównaliśmy otrzymane oszacowania parametrów dla różnej częstotliwości danych śróddziennych, a niezależnie przeprowadziliśmy estymację na podstawie danych dziennych. W celu oceny jakości otrzymanych oszacowań parametrów dokonaliśmy symulacji 1000 trajektorii z częstotliwością taką samą, jak w przypadku badanego szeregu i z parametrami zgodnymi z otrzymanymi za pomocą estymacji. Następnie niezależnie estymowaliśmy parametry dla każdej trajektorii, a średnie ich wartości porównaliśmy z otrzymanymi dla badanego szeregu.

Do modelowania procesów wykorzystaliśmy niezależnie trzy opisane w rozdziale trzecim modele dyfuzji. Programy do estymacji i prognozowania procesów finansowych zostały napisane w języku Matlab 7.0. Wykorzystują one funkcję optymalizacyjną `fmincon`, która bazuje na algorytmie optymalizacyjnym Newtona. W każdym kroku algorytmu, przybliżenie Hessianu oraz jego odwrotności jest uaktualniane, odpowiednio, za pomocą algorytmu BFGS (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970) i DFP (Davidon, 1991). W celu lepszego dopasowania modeli w niektórych przypadkach, badane szeregi zostały pomnożone przez arbitralnie dobrany skalar równy 100.

Zastosowaliśmy również niezależnie, bardziej klasyczne podejście do estymacji parametrów modeli dyfuzji – opisaną w rozdziale 5 uogólnioną metodę momentów. W tym celu użyliśmy pakietu procedur napisanych w języku Matlab przez Cliffa (2003). Wykorzystuje on dwa podejścia do estymacji macierzy wag – klasyczne zaproponowane przez Hansena (1982) oraz podejście bazujące na macierzach gęstości spektralnej, których zastosowanie ma na celu otrzymanie dokładniejszych oszacowań parametrów. Ponieważ specyfikacja modeli dyfuzji w procedurach Cliffa jest nieco inna od przyjętej przez nas, otrzymane oszacowania parametrów oraz ich błędy standardowe transformowaliśmy za pomocą metody delta. Umożliwiło to konfrontację wyników z otrzymanymi już wcześniej za pomocą metody Phillipsa i Yu (2007). Zaletą uogólnionej metody momentów, podobnie jak metody Phillipsa i Yu, są niskie koszty estymacji

modeli z czasem ciągłym. Metody te pozwalają wyestymować model w czasie nie dłuższym niż 10 sekund, podczas gdy powszechnie stosowane metody Bayesowskie potrzebują na to wiele godzin, co jest istotną przeszkodą w wykorzystywaniu ich w celach praktycznych. Ogromne różnice w prędkości wynikają z faktu, że, jak wykazał Kostrzewski (2006a), estymacja parametrów bayesowską metodą MCMC, cechującą się wysokim procentem akceptacji wymaga od 8000 do 100000 realizacji, natomiast estymacja parametrów dwukrokową metodą Phillipsa i Yu sprowadza się do dwóch zagadnień minimalizacji, a uogólniona metoda momentów do jednego bardziej rozbudowanego. Z uwagi na to, że proponowane przez nas metody wykorzystują także dane wysokiej częstotliwości, zbiór informacji, na których opiera się estymacja jest o wiele większy niż we wcześniejszych metodach, w których parametry estymowane są na bazie notowań dziennych. Sugeruje to większą dokładność estymacji, co zostanie zweryfikowane przez prowadzone badania.

Dodatkowo wyznaczamy prognozy na podstawie dopasowanych modeli dyfuzji, których oszacowania parametrów otrzymano za pomocą obydwu metod estymacji. Porównujemy również ich jakość, a także konfrontujemy je z prognozami otrzymanymi za pomocą powszechnie stosowanych dyskretnych modeli parametrycznych.

## **6.1. Modelowanie danych dziennych za pomocą modeli z czasem ciągłym**

W niniejszym podrozdziale przedstawimy wyniki modelowania procesów cen logarytmicznych generujących dane dzienne. Parametry funkcji dyfuzji zostały oszacowane za pomocą metody Phillipsa i Yu – przez minimalizację funkcji (3.11) oraz oszacowanie parametrów średniej metodą największej wiarygodności zaproponowaną przez Láńską (1979). Oszacowania parametrów otrzymujemy również niezależnie, wykorzystując uogólnioną metodę momentów, stosując dwa niezależne podejścia do estymacji macierzy wag.

### **6.1.1. Prezentacja danych**

Estymacje przeprowadzamy na podstawie dziennych logarytmicznych poziomów indeksu WIG20 oraz wybranych indeksów z największych światowych giełd a także logarytmicznych kursów walutowych EURPLN oraz USDPLN. Wszystkie dane pochodzą

z okresu od 2 stycznia 2001 do 29 grudnia 2006. Statystyki opisowe badanych szeregów przedstawione są w tabeli 6.1.

**Tabela 6.1. Statystyki opisowe dla dziennych logarytmicznych poziomów indeksów giełdowych oraz dziennych logarytmicznych kursów walutowych (pomnożonych przez 100).**

Szereg Notowań	liczba obs.	średnia	odch. std.	skośność	kurtoza	min.	maks.
CAC40	1534	8,3138	0,19922	-0,1759	2,0605	7,7845	8,6993
DAX	1528	8,40231	0,24511	-0,3745	2,3597	7,6976	8,824
NASDAQ	1554	7,55651	0,1818	-0,8033	3,0256	7,0158	7,9583
WIG20	1546	7,4555	0,34958	0,4318	2,0328	6,8979	8,1745
EURPLN	1547	1,3977	0,08424	0,2545	2,4275	1,2099	1,5875
USDPLN	1547	1,2952	0,11555	-0,4426	1,6828	1,0496	1,4844

### 6.1.2. Omówienie wyników empirycznych

Przy wykorzystaniu algorytmów optymalizacyjnych istotną rolę odgrywa odpowiedni dobór wartości startowych oszacowań parametrów. Aby uzyskać prawidłowe oszacowania, ważnym jest by wartości te były możliwie bliskie rozwiązaniu optymalnemu. Jako wartości startowe parametrów dryfu zdecydowaliśmy się przyjąć  $\kappa = 0$  oraz  $\mu$  równe średniej z próby. Za wartość startową dla  $\sigma$  w modelach Vašíčka oraz CIR przyjmujemy oszacowanie tego parametru za pomocą estymatora wprowadzonego przez Yoshidę, natomiast w modelu CKLS, dla  $\sigma$  przyjmujemy oszacowanie otrzymane uprzednio dla modelu CIR, a dla  $\beta$ , 0,5. Okresem, dla którego liczymy wariancję zrealizowaną jest 1 miesiąc. W tabelach 6.2-6.4 zaprezentowane są oszacowania parametrów modeli Vašíčka, CIR i CKLS otrzymane dla dziennych logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursów walutowych EURPLN i USDPLN za pomocą metody Phillipsa i Yu.

**Tabela 6.2. Oszacowania parametrów modelu Vašíčka dla dziennych logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz kursów walutowych EURPLN i USDPLN.**

Szereg Notowań log.	$\kappa$	$\mu$	$\sigma$
CAC40	0,034582	8,3644	0,11064
DAX	0,042929	8,4174	0,12291
NASDAQ	0,0000258	7,5829	0,12692
WIG20	0,0012536	7,4681	0,16780
EURPLN	0,05541	1,3962	0,61923
USDPLN.	0,01663	1,0182	0,07361

**Tabela 6.3. Oszacowania parametrów modelu Coxa. Ingersolla i Rossa dla dziennych logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz kursów walutowych EURPLN i USDPLN.**

Szereg notowań log.	$\kappa$	$\mu$	$\sigma$
CAC 40	0,03683	8,3589	0,03826
DAX	0,044431	8,4169	0,042108
NASDAQ	0,081467	7,5632	0,045667
WIG20	0,00776	7,3712	0,052088
EURPLN	0,033182	1,3977	0,04953
USDPLN	0,017271	1,0285	0,06397

**Tabela 6.4. Oszacowania parametrów modelu stałej elastyczności wariancji dla dziennych logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz kursów walutowych EURPLN i USDPLN.**

Szereg notowań log.	$\kappa$	$\mu$	$\sigma$	$\beta$
CAC 40	0,033606	8,3669	0,17815	-0,22647
DAX	0,041253	8,4177	0,43424	-0,59029
NASDAQ	0,081748	7,5632	0,72314	-0,85398
WIG20	0,010103	6,557	0,4757	-0,51843
EURPLN	0,047126	1,3937	0,13663	-2,3071
USDPLN	0,016646	1,0185	0,072919	0,033903

Niestety, oszacowania parametru  $\beta$  w modelu stałej elastyczności wariancji przyjmują wartości ujemne dla wszystkich badanych procesów poza procesem logarytmicznych kursów USDPLN. Jest to sprzeczne ze specyfikacją modelu. Dlatego estymację wykonaliśmy dodatkowo dla danych pomnożonych przez stałą równą 100. Oszacowania parametrów modelu CKLS dla rozważanych szeregów pomnożonych przez 100 przedstawione są w tabeli 6.5

**Tabela 6.5. Oszacowania parametrów modelu stałej elastyczności wariancji dla dziennych logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz kursów walutowych EURPLN i USDPLN.**

Szereg notowań log.	$\kappa$	$\mu$	$\sigma$	$\beta$
CAC40	0,03622	836,04	0,94563	0,36542
DAX	0,043841	841,71	1,5752	0,304
NASDAQ	0,00004	758,9	4,2643	0,16329
WIG20	0,001243	747,39	2,5869	0,28153
EURPLN	0,055173	139,62	7,7537	-0,045971
USDPLN	0,0036699	129,4	3,5784	0,1474

Oszacowania parametrów otrzymane dla szeregu pomnożonego przez stałą są zgodne ze specyfikacją modelu. Wyjątkiem pozostaje jedynie oszacowanie parametru  $\beta$  otrzymane dla szeregu EURPLN, które przyjmuje niewielką co do wartości bezwzględ-

nej wartość ujemną. Nietypowe wartości parametru  $\beta$  otrzymane dla szeregu cen logarytmicznych mogą wynikać z problemów dobrania optymalnych wartości początkowych, co prowadzi do wskazania przez algorytm optymalizacyjny odległych od prawidłowego rozwiązania minimów lokalnych. Oszacowania parametru  $\mu$  we wszystkich przypadkach oscylują w granicach średniej z próby, jedynie dla procesu USDPLN w modelu CIR i CKLS przyjmują wartość niższą. Otrzymane oszacowania parametrów  $\kappa$  sugerują, że wszystkie procesy wykazują dodatnie tempo zbieżności do średniej. W tabelach 6.6-6.10 przedstawione są wyniki symulacji Monte Carlo, które przeprowadziliśmy w celu oceny jakości otrzymanych oszacowań.

**Tabela 6.6. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia symulacji Monte Carlo dla modelu Vašička wykonane dla dziennych cen logarytmicznych. Trajektorie otrzymane za pomocą dyskretyzacji Eulera.**

Szereg Zwrotów	$\kappa$	$\mu$	$\sigma$
CAC40	0,030872 (0,05107)	8,2733 (0,85916)	0,09944 (0,00294)
DAX	0,031424 (0,05064)	8,3314 (0,87869)	0,11013 (0,00333)
NASDAQ	0,041843 (0,05799)	7,5533 (0,40102)	0,11378 (0,00343)
WIG20	0,056302 (0,07403)	7,4683 (0,10174)	0,15037 (0,00449)
EURPLN	0,06596 (0,06948)	1,4117 (0,41286)	0,05552 (0,00163)
USDPLN.	0,0023573 (0,00311)	1,016 (0,02828)	0,065954 (0,00199)

**Tabela 6.7. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia symulacji Monte Carlo dla modelu CIR wykonane dla dziennych cen logarytmicznych. Trajektorie otrzymane za pomocą dyskretyzacji Eulera.**

Szereg Zwrotów	$\kappa$	$\mu$	$\sigma$
CAC 40	0,077122 (0,06193)	8,4632 (2,0276)	0,0343 (0,0010553)
DAX	0,027166 (0,047448)	833,73 (92,091)	0,37846 (0,011192)
NASDAQ	0,052826 (0,069508)	775,57 (254,68)	0,41004 (0,012235)
WIG20	0,046332 (0,066019)	855,26 (803,37)	0,467 (0,014003)
EURPLN	0,064507 (0,07274)	2,7698 (5,2235)	0,044401 (0,00131)
USDPLN.	0,010319 (0,027487)	102,45 (14,948)	0,57312 (0,01736)

**Tabela 6.8. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia symulacji Monte Carlo dla modelu CIR wykonane dla dziennych cen logarytmicznych. Trajektorie otrzymane za pomocą dyskretyzacji Milsteina.**

Szereg Zwrotów	$\kappa$	$\mu$	$\sigma$
CAC 40	0,078272 (0,066706)	8,3314 (2,1105)	0,034284 (0,0010245)
DAX	0,02743 (0,047531)	8,3453 (0,97326)	0,37749 (0,001139)
NASDAQ	0,052355 (0,065201)	7,7931 (2,8697)	0,040941 (0,0012372)
WIG20	0,052036 (0,072821)	7,8626 (4,9835)	0,046689 (0,0014052)
EURPLN	0,061181 (0,073508)	2,8813 (5,2243)	0,044339 (0,0019033)
USDPLN	0,0090655 (0,02377)	1,0234 (0,14048)	0,057352 (0,0017218)

**Tabela 6.9. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia symulacji Monte Carlo dla modelu CKLS wykonane dla pomnożonych przez 100 dziennych cen logarytmicznych. Trajektorie otrzymane za pomocą dyskretyzacji Eulera.**

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$	$\beta$
CAC 40	0,053615 (0,067071)	893,82 (630,74)	0,93884 (0,026205)	0,35025 (0,0089899)
DAX	0,030938 (0,051703)	835,19 (146,81)	1,5643 (0,0036686)	0,28863 (0,0042302)
NASDAQ	0,045635 (0,0676)	834,65 (648,85)	4,2523 (0,0069057)	0,14703 (0,0042931)
WIG20	0,051708 (0,067554)	927,62 (991,36)	2,577 (0,53681)	0,26589 (0,004315)
EURPLN	0,06261 (0,066085)	287,76 (499,03)	5,1335 (0,18961)	0,015907 (0,001663)
USDPLN	0,055983 (0,070474)	232,4 (525,33)	3,5303 (0,29025)	0,13031 (0,042122)

**Tabela 6.10. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia symulacji Monte Carlo dla modelu CKLS wykonane dla pomnożonych przez 100 dziennych cen logarytmicznych. Trajektorie otrzymane za pomocą dyskretyzacji Milsteina.**

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$	$\beta$
CAC 40	0,061082 (0,067625)	882,82 (651,41)	0,93893 (0,010118)	0,35012 (0,0048638)
DAX	0,0269 (0,047802)	824,56 (134,93)	1,5643 (0,0036402)	0,28874 (0,0041468)
NASDAQ	0,04576 (0,064299)	849,42 (738,89)	4,2524 (0,0071401)	0,14711 (0,004532)
WIG20	0,051807 (0,067854)	865,52 (896,67)	2,5744 (0,073482)	0,26598 (0,011381)
EURPLN	0,062156 (0,067663)	303,62 (508,58)	5,1267 (0,19686)	0,015983 (0,0017298)
USDPLN	0,055478 (0,071179)	248,91 (572,82)	3,5586 (0,24301)	0,1271 (0,027372)

Przedstawione w tabelach 6.6-6.10 wyniki symulacji Monte Carlo mogą sugerować dosyć dobre dopasowanie modeli do badanych szeregów. Średnie oszacowań parametrów są stosunkowo bliskie rzeczywistym parametrom symulowanych trajektorii, co przemawia za dobrym dopasowaniem modeli do badanych szeregów cen logarytmicznych. Wyjątkiem pozostaje szereg EURPLN, dla którego oszacowanie parametru  $\beta$  było ujemne. Również średnia oszacowań parametru  $\mu$  w modelach CIR i CKLS dla logarytmicznego kursu USDPLN jest zawyżona. Odchylenia standardowe przyjmują niewielkie wartości dla parametrów funkcji dyfuzji, natomiast dla parametrów dryfu często są one dosyć duże. Jednocześnie można zauważyć, że zastosowany schemat dyskretyzacji nie ma istotnego wpływu na wyniki symulacji, co potwierdzałoby wnioski przedstawione w rozdziale 1.3.4.

Parametry rozważanych modeli dyfuzji dla badanych procesów estymujemy za pomocą opisaną w rozdziale 5 uogólnioną metody momentów. Wykorzystujemy dwa rodzaje macierzy wag – wyznaczoną na podstawie wzoru 5.1 oraz odwróconą macierz gęstości spektralnej, z jądrem Bartletta oraz zaproponowanym przez Andrewsa (1991) automatycznym doбором  $B_T$  podobnie, jak zostało to zrobione w pracy Burnside'a i Eichenbauma (1994). We wszystkich przypadkach stosujemy metodę 2-stopniową, ponieważ zastosowanie kolejnych wyższych rzędów nie miało zauważalnego wpływu na wyniki. Wraz z oszacowaniami parametrów przedstawiliśmy ich błędy standardowe oraz za pomocą  $t$ -statystyki oceniliśmy istotność parametrów. W tabelach 6.11-6.16 przedstawione są wyniki estymacji otrzymane dla badanych procesów cen logarytmicznych.



Tabela 6.11. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla dziennych logarytmicznych poziomów notowań indeksu CAC40 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	0,77537	0,5144	0,13193	-0,10636	0,54637	0,84568
	$\mu$	8,3007	0,016013	0	8,1772	1,2844	0
	$\sigma$	0,21922	0,0074562	0	0,16926	0,011029	0
CIR	$\kappa$	0,78272	0,51441	0,12832	-0,10702	0,54654	0,84478
	$\mu$	8,3	0,015698	0	8,1667	1,3701	0
	$\sigma$	0,075559	0,0026212	0	0,058477	0,0038484	0
CKLS	$\kappa$	0,41333	0,51608	0,42331	0,41333	0,56579	0,46517
	$\mu$	8,3644	0,06804	0	8,3644	0,031674	0
	$\sigma$	5,6219	1,9953e+7	1	5,5273	4,0987e+7	1
	$\beta$	-4,8316	1,6875e+6	1	-4,8234	3,497e+6	1

Tabela 6.12. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla dziennych logarytmicznych poziomów notowań indeksu DAX wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	0,98503	0,42709	0,021221	0,29613	0,48814	0,54417
	$\mu$	8,3802	0,011515	0	8,4126	0,10643	0
	$\sigma$	0,2596	0,0074667	0	0,21296	0,011752	0
CIR	$\kappa$	0,99642	0,42713	0,019785	0,29529	0,48817	0,54535
	$\mu$	8,3814	0,011245	0	8,4223	0,10342	0
	$\sigma$	0,088682	0,0026104	0	0,072836	0,0040696	0
CKLS	$\kappa$	0,51308	0,42989	0,23285	0,51308	0,49207	0,29724
	$\mu$	8,4174	0,042271	0	8,4174	0,035067	0
	$\sigma$	5,4216	1,5493e+6	1	5,3016	2,9638e+6	1
	$\beta$	-4,2256	1,3524e+5	0,99997	-4,2147	2,6185e+5	0,99999

Tabela 6.13. Oszacowania parametrów modeli Vašička, CIR i CKLS dla dziennych logarytmicznych poziomów notowań indeksu NASDAQ wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	0,57683	0,56768	0,30973	0,91407	0,65302	0,16179
	$\mu$	7,5811	0,033124	0	7,5811	0,011474	0
	$\sigma$	0,25871	0,0086133	0	0,26122	0,012449	0
CIR	$\kappa$	0,55553	0,56773	0,32798	0,91274	0,65303	0,1624
	$\mu$	7,5802	0,035672	0	7,5827	0,011484	0
	$\sigma$	0,09368	0,0031515	0	0,094545	0,0045812	0
CKLS	$\kappa$	0,97411	0,57283	0,089232	0,97411	0,65404	0,13659
	$\mu$	7,5632	0,011435	0	7,5632	0,010508	0
	$\sigma$	4,5291	1,5539e+5	0,99998	4,5291	2,5248e+5	0,99999
	$\beta$	-3,7814	17051	0,9998	-3,7814	27711	0,99988

Tabela 6.14. Oszacowania parametrów modeli Vašička, CIR i CKLS dla dziennych logarytmicznych poziomów notowań indeksu WIG20 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	-0,12749	0,26647	0,63241	-0,046178	0,29652	0,87626
	$\mu$	6,5984	3,7383	0,077746	5,8391	119,53	0,96104
	$\sigma$	0,23064	0,0051101	0	0,2289	0,0061883	0
CIR	$\kappa$	-0,13399	0,26647	0,61514	-0,018692	0,29697	0,94982
	$\mu$	6,6412	3,1012	0,032392	4,2921	2638,9	0,9987
	$\sigma$	0,084197	0,0018792	0	0,083224	0,0022749	0
CKLS	$\kappa$	-0,11478	0,26662	0,6669	-0,11478	0,2995	0,70159
	$\mu$	6,5059	5,5178	0,23855	6,5059	7,471	0,38399
	$\sigma$	1,7382	1,7331	0,31606	1,7382	2,3165	0,45315
	$\beta$	-1,007	0,49855	0,0025455	-1,007	0,66423	0,023415

Tabela 6.15. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla dziennych logarytmicznych kursów EURPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	0,21788	0,44822	0,62697	0,20648	0,38127	0,5882
	$\mu$	1,4042	0,032674	0	1,3017	0,067235	0
	$\sigma$	0,097497	0,003455	0	0,090876	0,002815	0
CIR	$\kappa$	0,11382	0,44825	0,79958	0,12246	0,38095	0,74791
	$\mu$	1,3954	0,11933	0	1,2052	0,48349	0,012778
	$\sigma$	0,081746	0,002957	0	0,075631	0,002392	0
CKLS	$\kappa$	0,66227	0,46658	0,15598	0,66227	0,48771	0,17469
	$\mu$	1,3962	0,003522	0	1,3962	0,003085	0
	$\sigma$	0,19205	0,037792	0	0,19205	0,079422	0,015715
	$\beta$	-2,0846	0,63495	0	-2,0846	1,1772	0,028276

Tabela 6.16. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla dziennych logarytmicznych kursów USDPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	0,024785	0,37506	0,94732	-0,027861	0,38533	0,94237
	$\mu$	-0,87534	1082,4	0,99935	2,9628	518,91	0,99545
	$\sigma$	0,10916	0,0027687	0	0,11048	0,0034796	0
CIR	$\kappa$	-0,082606	0,37537	0,82585	-0,1662	0,38557	0,66649
	$\mu$	1,9573	9,329	0,83385	1,5535	0,36673	0
	$\sigma$	0,095012	0,0024565	0	0,096997	0,0031672	0
CKLS	$\kappa$	0,19848	0,38145	0,6029	0,19848	0,4075	0,62627
	$\mu$	1,0178	0,33384	0,0023378	1,0178	0,41927	0,015319
	$\sigma$	0,1295	0,0086007	0	0,1295	0,011701	0
	$\beta$	-0,68743	0,26642	0	-0,68743	0,39275	0,0025408

Przedstawione w tabelach 6.11-6.16 wyniki sugerują bardzo słabe dopasowanie modelu CKLS do badanych szeregów cen logarytmicznych. We wszystkich przypadkach otrzymujemy ujemne wartości oszacowań parametru  $\beta$  podobnie, jak działo to się

w przypadku zastosowania metody Phillipsa i Yu. Dla indeksów giełdowych, błędy standardowe oszacowań tego parametru oraz parametru  $\sigma$  są bardzo duże, przez co oszacowania te są statystycznie nieistotne. Niestety zabiegi, które przyniosły poprawę jakości oszacowań w przypadku metody Phillipsa i Yu, nie przyniosły jej w przypadku uogólnionej metody momentów. Modele CIR i Vašíčka dopasowują się znacznie lepiej do badanych szeregów niż model CKLS, jednak błędy standardowe oszacowań parametrów w wielu przypadkach pozostają nadal dość duże. Dla wszystkich badanych szeregów, poza szeregiem notowań logarytmicznych indeksu DAX oszacowanie parametru  $\kappa$  pozostaje nieistotne. Oszacowania parametrów  $\mu$  i  $\sigma$  są statycznie istotne. Oszacowanie parametru  $\mu$  przyjmuje zwykle wartość bliską średniej z próby, chociaż dla procesów logarytmicznych poziomów indeksu WIG20 jest od niej odległe, a błędy standardowe przyjmują duże wartości. Ponadto obserwujemy, że zastosowana macierz wag nie ma istotnego wpływu na wyniki estymacji. W przeważającej większości przypadków oszacowania parametrów oraz ich błędy standardowe przyjmują zbliżone wartości bez względu na zastosowaną macierz wag.

## **6.2. Modelowanie procesów na podstawie danych ultra wysokiej częstotliwości**

W niniejszym podrozdziale podejmujemy próbę estymacji parametrów modeli Vašíčka, CIR oraz CKLS, wykorzystując w tym celu dane wysokiej częstotliwości. Jednocześnie zamierzamy ocenić wpływ częstotliwości notowań na wyniki estymacji, a także ocenić, w jakim stopniu poprawiła się jakość otrzymanych oszacowań parametrów w stosunku do oszacowań otrzymanych na podstawie danych dziennych.

### **6.2.1. Prezentacja danych**

Do badania wykorzystaliśmy notowania intraday indeksów CAC40, DAX, NASDAQ i WIG20 z okresu od 3 czerwca 2006 do 29 grudnia 2006 oraz kursów walutowych EURPLN i USDPLN z tego samego okresu. Notowania śróddzienne dostępne są w pięciominutowych odstępach czasowych, natomiast kursy walutowe w dziesięciominutowych. W przypadku indeksu WIG20 do estymacji parametrów wykorzystane zostały dane pięcio-, dziesięcio-, dwudziesto- i czterdziestominutowe, natomiast w przypadku

pozostałych indeksów, dane pięcio- i dziesięciominutowe oraz półgodzinne. Parametry modeli dyfuzji dla kursów walutowych szacowaliśmy na bazie danych dziesięciominutowych, półgodzinnych oraz godzinnych. W tabeli 6.17 przedstawione są statystyki opisowe wszystkich rozważanych szeregów notowań i kursów logarytmicznych.

**Tabela 6.17. Statystyki opisowe szeregów logarytmicznych poziomów indeksów giełdowych oraz logarytmicznych kursów różnej częstotliwości.**

Szereg Notowań	liczba obs.	średnia	odch. std.	skośność	kurtoza	min.	maks.
CAC 40 5min.	13056	8,5587	0,39808	-0,3621	2,1680	8,4583	8,6221
CAC 40 10min.	6528	8,5588	0,3981	-0,3620	2,1682	8,4588	8,622
CAC 40 30min.	2176	8,5588	0,39818	-0,3614	2,1693	8,4601	8,6221
DAX 5min.	13056	8,7032	0,55348	-0,0141	1,8083	8,589	8,7987
DAX 10min.	6528	8,7032	0,55338	-0,0140	1,8084	8,5899	8,7982
DAX 30min.	2176	8,7032	0,5536	-0,0144	1,8091	8,5906	8,7981
NASDAQ 5min.	9228	7,7229	0,61788	-0,1699	1,6133	7,6076	7,8117
NASDAQ 10min.	4614	7,7229	0,61793	-0,1702	1,6134	7,6076	7,8116
NASDAQ 30min.	1538	7,7229	0,61809	-0,1717	1,6140	7,6076	7,8116
WIG20 5 min.	10080	8,0367	0,46201	0,4003	2,1727	7,9491	8,1402
WIG20 10 min.	5040	8,0367	0,46204	0,4001	2,1724	7,9497	8,1402
WIG20 20 min.	2520	8,0367	0,4621	0,4019	2,1727	7,9504	8,1402
WIG20 40 min.	1210	8,0367	0,46195	0,4033	2,1748	7,9533	8,1402
EURPLN 10 min.	6804	1,3608	0,18671	0,1946	2,0333	1,3281	1,4061
EURPLN 30 min.	2268	1,3608	0,18657	0,19403	2,0306	1,329	1,4051
EURPLN 60 min.	1134	1,3608	0,18652	0,19107	2,0281	1,32938	1,4043
USDPLN 10 min.	6804	1,1128	0,31255	-0,55225	2,2467	1,0481	1,1721
USDPLN 30 min.	2268	1,1128	0,31251	-0,55147	2,2454	1,0481	1,1721
USDPLN 60 min.	1134	1,1128	0,31248	-0,55341	2,2437	1,049	1,1711

Łatwo zaobserwować, że szeregi czasowe bez względu na wykorzystaną częstotliwość mają bardzo zbliżone statystyki opisowe, wobec czego oszacowania parametrów modeli opisujących procesy generujące te szeregi powinny być również zbliżone.

## 6.2.2. Omówienie wyników empirycznych

W tabelach 6.12-6.17 zaprezentowane są oszacowania parametrów modeli Vašička, CIR i CKLS dla procesu logarytmicznych poziomów indeksów giełdowych oraz logarytmicznych kursów walutowych w zależności od wykorzystanej częstotliwości. Użyliśmy niezależnie dwóch rodzajów estymatorów wariancji zrealizowanej  $[X_\delta^2]$  oraz  $[X_\delta^3]$ . Wartości startowe przyjmowaliśmy w taki sam sposób, jak dla danych dziennych. Wyniki estymacji otrzymane za pomocą dwukrokowej metody Phillipsa i Yu zamieszczone są w tabelach 6.18-6.21.

**Tabela 6.18. Oszacowania parametrów modelu Vašíčka dla procesów logarytmicznych poziomów indeksów giełdowych, w zależności od częstotliwości danych wykorzystanych do estymacji.**

Est. wariancji zrealizowanej	[ $X_s^2$ ]			[ $X_s^3$ ]		
	$\kappa$	$\mu$	$\sigma$	$\kappa$	$\mu$	$\sigma$
Szereg notowań log.						
CAC40 5min.	0,01421	8,6185	7,6554	0,01445	8,618	9,6937
CAC40 10min.	0,01452	8,6175	7,0727	0,01450	8,6178	9,0012
CAC40 30min.	0,01456	8,6184	6,2694	0,01415	8,6193	7,9243
DAX 5min.	0,00270	9,1338	7,1615	0,00317	9,077	7,8021
DAX 10min.	0,00282	9,116	6,3962	0,00323	9,0693	6,9923
DAX 30min.	0,00276	9,1321	5,2144	0,00279	9,1273	5,8054
NASDAQ 5min.	0,00535	7,8756	7,8256	0,00559	7,8732	9,9976
NASDAQ 10min.	0,00569	7,8675	7,9012	0,00585	7,8667	9,7907
NASDAQ 30min.	0,00568	7,8667	5,0041	0,00592	7,8652	8,4877
WIG20 5 min.	0,04811	8,058	13,415	0,05210	8,0563	17,554
WIG20 10 min.	0,04735	8,0587	13,414	0,04820	8,0579	16,326
WIG20 20 min.	0,04682	8,0584	11,766	0,04698	8,0584	15,648
WIG20 40 min.	0,04661	8,0591	11,11	0,04589	8,0589	12,564

**Tabela 6.19. Oszacowania parametrów modelu Vašíčka dla procesów logarytmicznych kursów walutowych, w zależności od częstotliwości danych wykorzystanych do estymacji.**

Est. wariancji zrealizowanej	[ $X_s^2$ ]			[ $X_s^3$ ]		
	$\kappa$	$\mu$	$\sigma$	$\kappa$	$\mu$	$\sigma$
Szereg notowań log.						
EURPLN 10min.	0,04716	1,3503	5,2326	0,04688	1,3503	5,7738
EURPLN 30min.	0,04152	1,3494	4,3609	0,04381	1,3496	4,9996
EURPLN 60min.	0,03967	1,349	3,6433	0,04245	1,3492	4,0199
USDPLN 10min.	0,01847	1,0734	6,6913	0,01944	1,0741	7,7978
USDPLN 30min.	0,01771	1,0718	5,6871	0,01861	1,0724	6,9305
USDPLN 60min.	0,01687	1,0697	3,3782	0,01781	1,0706	5,6939

**Tabela 6.20. Oszacowania parametrów modelu Coxa, Ingersolla i Rossa dla procesów logarytmicznych poziomów indeksów, w zależności od częstotliwości danych wykorzystanych do estymacji.**

Est. wariancji zrealizowanej	[ $X_s^2$ ]			[ $X_s^3$ ]		
	$\kappa$	$\mu$	$\sigma$	$\kappa$	$\mu$	$\sigma$
Szereg notowań log.						
CAC40 5min.	0,01427	8,6181	0,02613	0,01453	8,6177	0,03309
CAC40 10min.	0,01449	8,6177	0,02414	0,01458	8,6174	0,03098
CAC40 30min.	0,01455	8,6184	0,0214	0,0142	8,619	0,02704
DAX 5min.	0,00272	9,1319	0,02421	0,00312	9,0828	0,02636
DAX 10min.	0,00283	9,1139	0,02163	0,00322	9,071	0,0236
DAX 30min.	0,00277	9,1313	0,01762	0,00276	9,1317	0,01959
NASDAQ 5min.	0,00532	7,8764	0,02812	0,00558	7,8738	0,35878
NASDAQ 10min.	0,00566	7,8684	0,02838	0,00585	7,8668	0,35115
NASDAQ 30min.	0,00566	7,8675	0,01792	0,00590	7,8655	0,30453
WIG20 5 min.	0,05246	8,0566	0,05630	0,05211	8,0563	0,06174
WIG20 10 min.	0,04817	8,0580	0,04994	0,04842	8,0577	0,05749
WIG20 20 min.	0,04690	8,0582	0,04523	0,04698	8,0584	0,05511
WIG20 40 min.	0,04587	8,0589	0,05139	0,04595	8,0588	0,04417

**Tabela 6.21. Oszacowania parametrów modelu Coxa, Ingersolla i Rossa dla procesów logarytmicznych kursów walutowych, w zależności od częstotliwości danych wykorzystanych do estymacji.**

Est. wariancji zrealizowanej	$[X_{\sigma}^2]$			$[X_{\sigma}^3]$		
	$\kappa$	$\mu$	$\sigma$	$\kappa$	$\mu$	$\sigma$
Szereg notowań log.						
EURPLN 10min.	0,04452	1,3501	0,04406	0,04688	1,3503	0,04964
EURPLN 30min.	0,04227	1,3504	0,03760	0,04384	1,3496	0,04297
EURPLN 60min.	0,03945	1,3489	0,31348	0,04249	1,3492	0,03462
USDPLN 10min.	0,02291	1,0893	0,06358	0,01945	1,0741	0,07472
USDPLN 30min.	0,01748	1,0893	0,05402	0,01870	1,0726	0,06683
USDPLN 60min.	0,01657	1,0691	0,03147	0,01784	1,0707	0,05476

Próba modelowania badanych procesów cen logarytmicznych przyniosła podobne efekty, jak w przypadku danych dziennych. Dlatego również w tym przypadku zdecydowaliśmy się wykorzystać szeregi notowań logarytmicznych pomnożone przez 100. W tabelach 6.22-6.23 przedstawione są oszacowania parametrów otrzymane na podstawie przeskalowanych szeregów cen logarytmicznych

**Tabela 6.22. Oszacowania parametrów modelu stałej elastyczności wariancji dla procesów logarytmicznych poziomów indeksów pomnożonych przez 100, w zależności od częstotliwości danych wykorzystanych do estymacji.**

Est. wariancji zrealizowanej	$[X_{\sigma}^2]$				$[X_{\sigma}^3]$			
	$\kappa$	$\mu$	$\sigma$	$\beta$	$\kappa$	$\mu$	$\sigma$	$\beta$
Szereg notowań log.								
CAC40 5min.	0,01423	861,86	0,79041	0,33597	0,01443	861,82	0,81795	0,34478
CAC40 10min.	0,01444	861,8	0,71105	0,34001	0,01451	861,79	0,93337	0,31552
CAC40 30min.	0,01454	861,86	0,59247	0,34919	0,01415	861,93	0,90283	0,30006
DAX 5min.	0,00270	913,27	0,8888	0,30791	0,00311	908,34	1,2013	0,27434
DAX 10min.	-0,00032	508,07	0,84166	0,29901	0,00321	907,18	1,0196	0,28206
DAX 30min.	0,00276	913,16	0,61154	0,31574	0,00276	913,22	0,99593	0,25762
NASDAQ 5min.	0,00533	787,62	0,84298	0,33469	0,00559	787,35	1,0561	0,31999
NASDAQ 10min.	0,00567	786,81	0,83711	0,33703	0,00586	786,65	1,0011	0,32451
NASDAQ 30min.	0,00569	786,66	6,7136	-0,04430	0,00591	786,52	0,93774	0,31306
WIG20 5 min.	0,05156	805,62	1,8623	0,30433	0,05202	805,63	2,3953	0,27599
WIG20 10 min.	0,04823	805,79	1,2654	0,35263	0,04813	805,79	1,6962	0,3172
WIG20 20 min.	0,04692	805,84	0,80735	0,40048	0,04690	805,84	2,3243	0,26347
WIG20 40 min.	0,04671	805,9	0,73773	0,40487	0,04589	805,89	0,4396	0,47955

**Tabela 6.23. Oszacowania parametrów modelu stałej elastyczności wariancji dla procesów logarytmicznych kursów walutowych pomnożonych przez 100, w zależności od częstotliwości danych wykorzystanych do estymacji.**

Est. wariancji zrealizowanej	$[X_{\sigma}^2]$				$[X_{\sigma}^3]$			
	$\kappa$	$\mu$	$\sigma$	$\beta$	$\kappa$	$\mu$	$\sigma$	$\beta$
Szereg notowań log.								
EURPLN 10min.	0,04674	135,02	0,12127	0,76677	0,04678	135,03	0,17347	0,69359
EURPLN 30min.	0,04109	134,93	0,07633	0,8253	0,04366	134,95	0,11895	0,7411
EURPLN 60min.	0,03932	134,89	0,06846	0,80956	0,0423	134,92	0,07901	0,77996
USDPLN 10min.	0,01956	107,43	2,0261	0,26423	0,01958	107,43	2,3365	0,23411
USDPLN 30min.	0,01840	107,19	0,13109	0,82119	0,01836	107,18	0,09847	0,8833
USDPLN 60min.	0,01759	107	0,07823	0,88963	0,01767	107,02	0,15954	0,73803

Oszacowania parametrów otrzymane dla indeksów giełdowych przyjmują zbliżone wartości niezależnie od częstotliwości notowań, co przemawia za dobrym dopasowaniem się modelu do szeregu danych. Jedynie oszacowania parametrów zaznaczone na szaro odbiegają znacząco od oszacowań otrzymanych dla innych częstotliwości. Przypuszczalnie jest to następstwem błędów algorytmu minimalizacyjnego, który zwrócił wartości oszacowań dla jednego z minimów lokalnych. Wraz ze spadkiem częstotliwości obserwujemy tylko nieznaczny spadek oszacowania parametru  $\sigma$  w przypadku modeli Vašička i CIR oraz wartości  $\sigma X_t^{\beta}$  w przypadku modelu CKLS. Może być to konsekwencją niższego wpływu efektów mikrostruktury na dane o mniejszej częstotliwości. Jednocześnie sugeruje to możliwość wykorzystania danych intraday o mniejszej częstotliwości do modelowania bez utraty precyzji. Mniejsza częstotliwość znacząco skraca czas estymacji oraz obniża koszty opracowywania danych. Oszacowania parametru  $\sigma$  w modelach Vašička i CIR oraz wartość  $\sigma X_t^{\beta}$  w modelu CKLS są znacznie wyższe dla indeksu WIG20 niż dla trzech pozostałych indeksów, co świadczy o większym wpływie czynników losowych na logarytmiczny poziom indeksu. Dla wszystkich trzech modeli oszacowania parametru  $\kappa$ , który interpretujemy jako tempo powrotu trajektorii do średniej, przyjmują bardzo małe wartości, lecz co ważne, są one praktycznie we wszystkich przypadkach dodatnie. Dla procesów logarytmicznych kursów walutowych, wraz ze zmniejszającą się częstotliwością, możemy obserwować nieco szybszy spadek oszacowania parametru  $\sigma$  w modelach Vašička i CIR oraz wartości  $\sigma X_t^{\beta}$  w modelu CKLS. Porównując oszacowania parametrów otrzymane przy wykorzystaniu dwóch różnych estymatorów wariancji zrealizowanej zauważamy, że oszacowanie dla  $\sigma$  w modelach Vašička i CIR jest większe, jeśli jako estymator wariancji zrealizowanej



przyjmujemy  $[X_\delta^3]$ . Różnice te są najmniejsze dla szeregu logarytmicznych notowań indeksu DAX, gdzie stała  $1 + \tilde{\sigma}_{co}^2 / \tilde{\sigma}_{oc}^2$  tylko nieznacznie przekracza 1. Wobec tego możemy wyciągnąć wniosek, że zastosowanie mającego tendencje do zaniżania wartości estymatora wariancji zrealizowanej  $[X_\delta^2]$  istotnie wpływa na wyniki estymacji.

W tabelach 6.24-6.35 przedstawione są wyniki symulacji Monte Carlo, które przeprowadziliśmy w celu oceny jakości otrzymanych oszacowań.

**Tabela 6.24. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu Vašíčka wykonanych dla różnej częstotliwości poziomów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_\delta^2]$ .**

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
CAC40 5min.	0,040009 (0,036214)	8,9529 (1,5488)	7,5063 (0,050323)
CAC40 10min.	0,043242 (0,037061)	9,0144 (1,6889)	6,7875 (0,072452)
CAC40 30min.	0,040628 (0,034978)	8,858 (1,2932)	5,4349 (0,15617)
DAX 5min.	0,03287 (0,036152)	11,496 (5,6447)	7,0194 (0,048367)
DAX 10min.	0,04185 (0,034591)	8,8001 (1,2933)	6,1391 (0,065253)
DAX 30min.	0,031773 (0,033339)	7,288 (3,4361)	4,5251 (0,13719)
NASDAQ 5min.	0,040655 (0,034843)	8,0384 (1,565)	7,6251 (0,059335)
NASDAQ 10min.	0,045483 (0,040589)	8,2009 (1,9815)	7,4768 (0,10495)
NASDAQ 30min.	0,038527 (0,038807)	7,1112 (2,7815)	4,1011 (0,18121)
WIG20 5 min.	0,039305 (0,036974)	8,1009 (2,5399)	13,075 (0,11327)
WIG20 10 min.	0,039741 (0,034945)	7,7639 (2,0993)	12,722 (0,16239)
WIG20 20 min.	0,040464 (0,035085)	7,8108 (1,9854)	10,478 (0,2547)
WIG20 40 min.	0,039572 (0,036092)	7,5948 (2,5554)	8,425 (0,54055)

Tabela 6.25. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu Vašíčka wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^2]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
EURPLN 10 min.	0,033329 (0,040395)	2,3795 (1,9045)	5,0326 (0,052553)
EURPLN 30 min.	0,034245 (0,037572)	2,0677 (1,5138)	3,8178 (0,10427)
EURPLN 60 min.	0,04293 (0,036905)	1,4745 (0,31698)	2,6345 (0,24849)
USDPLN 10 min.	0,037295 (0,037121)	0,97496 (0,75379)	6,4363 (0,063819)
USDPLN 30 min.	0,0377 (0,037104)	0,94761 (0,63479)	4,9867 (0,13655)
USDPLN 60 min.	0,039954 (0,036842)	1,0124 (0,51207)	2,44 (0,20867)

Tabela 6.26. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu Vašíčka wykonanych dla różnej częstotliwości poziomów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^3]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
CAC40 5min.	0,041098 (0,035894)	9,0605 (1,9229)	9,503 (0,067489)
CAC40 10min.	0,041034 (0,035004)	9,0948 (1,988)	8,6375 (0,091415)
CAC40 30min.	0,040419 (0,03613)	9,0641 (1,75550)	6,8736 (0,19586)
DAX 5min.	0,026457 (0,036268)	14,764 (9,1651)	7,6494 (0,05266)
DAX 10min.	0,039794 (0,037619)	9,9085 (3,2729)	6,7092 (0,068176)
DAX 30min.	0,03622 (0,035829)	7,8406 (3,1034)	5,0418 (0,14739)
NASDAQ 5min.	0,043201 (0,038856)	8,4279 (2,6255)	9,734 (0,077892)
NASDAQ 10min.	0,044962 (0,039426)	8,5185 (2,4412)	9,2606 (0,12113)
NASDAQ 30min.	0,044206 (0,038928)	8,2622 (2,0514)	6,9804 (0,31115)
WIG20 5 min.	0,041486 (0,035544)	7,7475 (1,5165)	17,112 (0,14776)
WIG20 10 min.	0,042406 (0,036981)	7,541 (1,9898)	15,47 (0,20761)
WIG20 20 min.	0,040634 (0,035639)	7,5025 (2,0086)	13,922 (0,33793)
WIG20 40 min.	0,03814 (0,040051)	7,7227 (2,9027)	9,4971 (0,62587)

Tabela 6.27. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu Vašíčka wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_\delta^3]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
EURPLN 10 min.	0,037555 (0,037218)	2,1217 (1,6872)	5,5514 (0,055303)
EURPLN 30 min.	0,035873 (0,038576)	2,1821 (1,6328)	4,3841 (0,11583)
EURPLN 60 min.	0,036138 (0,036567)	1,9933 (1,2573)	2,8948 (0,22974)
USDPLN 10 min.	0,038344 (0,034235)	1,1262 (0,54623)	7,5043 (0,074683)
USDPLN 30 min.	0,037953 (0,037402)	1,0133 (0,78768)	6,0672 (0,16931)
USDPLN 60 min.	0,040533 (0,039533)	0,93881 (0,52813)	4,0967 (0,34256)

Tabela 6.28. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CIR wykonanych dla różnej częstotliwości poziomów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_\delta^2]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
CAC40 5min.	0,039197 (0,034158)	9,0423 (1,8104)	0,025618 (0,0001717)
CAC40 10min.	0,04181 (0,038391)	8,8904 (1,4591)	0,023153 (0,00025806)
CAC40 30min.	0,042625 (0,036118)	8,8533 (1,3278)	0,018526 (0,00080634)
DAX 5min.	0,03709 (0,035923)	10,29 (4,1061)	0,023732 (0,00016534)
DAX 10min.	0,041947 (0,034761)	8,7671 (1,1185)	0,02075 (0,00020483)
DAX 30min.	0,031257 (0,033663)	7,1591 (3,518)	0,015289 (0,00064675)
NASDAQ 5min.	0,041857 (0,036316)	8,0327 (1,5492)	0,027388 (0,00022832)
NASDAQ 10min.	0,047644 (0,042565)	8,1315 (1,8504)	0,026885 (0,00035803)
NASDAQ 30min.	0,0022959 (0,0096411)	7,2022 (0,52259)	0,0014603 (0,00013781)
WIG20 5 min.	0,042292 (0,037621)	8,0667 (2,1691)	0,054858 (0,00042537)
WIG20 10 min.	0,044727 (0,037637)	7,8219 (1,7138)	0,04737 (0,00060959)
WIG20 20 min.	0,037652 (0,036148)	8,1553 (2,8351)	0,04026 (0,00099126)
WIG20 40 min.	0,041018 (0,037025)	7,4508 (1,996)	0,034247 (0,013568)

Tabela 6.29 Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CIR wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^2]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
EURPLN 10 min.	0,035511 (0,037743)	2,2022 (1,7162)	0,042373 (0,00042745)
EURPLN 30 min.	0,035496 (0,039089)	2,1136 (1,7173)	0,032965 (0,00088395)
EURPLN 60 min.	0,043702 (0,036673)	1,4997 (0,40503)	0,016024 (0,0111)
USDPLN 10 min.	0,04032 (0,034934)	1,9391 (3,2478)	0,061143 (0,0006617)
USDPLN 30 min.	0,035026 (0,037017)	0,90734 (0,81595)	0,047349 (0,0013314)
USDPLN 60 min.	0,039561 (0,03681)	1,0068 (0,40106)	0,015598 (0,011326)

Tabela 6.30. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CIR wykonanych dla różnej częstotliwości poziomów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^3]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
CAC40 5min.	0,042317 (0,036771)	9,0268 (1,86)	0,032435 (0,00022281)
CAC40 10min.	0,041504 (0,035109)	9,0187 (1,8152)	0,029734 (0,00032008)
CAC40 30min.	0,041952 (0,037055)	8,9311 (1,5604)	0,023474 (0,00067388)
DAX 5min.	0,02445 (0,032827)	14,327 (8,0781)	0,02584 (0,00017313)
DAX 10min.	0,038307 (0,036861)	9,9065 (3,1773)	0,022651 (0,00024758)
DAX 30min.	0,03374 (0,031691)	7,74 (3,0462)	0,016995 (0,00050089)
NASDAQ 5min.	0,043349 (0,038555)	8,3411 (2,2071)	0,034935 (0,00028902)
NASDAQ 10min.	0,044359 (0,040013)	8,5605 (2,5782)	0,033231 (0,00044848)
NASDAQ 30min.	0,043398 (0,039238)	8,3282 (2,1664)	0,024795 (0,0029216)
WIG20 5 min.	0,04286 (0,035653)	7,7439 (1,5866)	0,060173 (0,00023903)
WIG20 10 min.	0,043404 (0,03809)	7,6779 (1,4946)	0,027247 (0,00034886)
WIG20 20 min.	0,041538 (0,036258)	7,5529 (2,1024)	0,024513 (0,0006018)
WIG20 40 min.	0,03853 (0,03872)	8,0152 (2,9767)	0,01469 (0,0058945)

Tabela 6.31 Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CIR wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^3]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$
EURPLN 10 min.	0,037745 (0,037355)	2,0681 (1,6187)	0,047768 (0,00048492)
EURPLN 30 min.	0,03642 (0,040675)	2,2525 (1,7608)	0,037699 (0,00099726)
EURPLN 60 min.	0,032111 (0,034864)	2,0881 (1,4732)	0,016633 (0,012616)
USDPLN 10 min.	0,042776 (0,037317)	1,1251 (0,44662)	0,07186 (0,00075049)
USDPLN 30 min.	0,039413 (0,03811)	0,98184 (0,60712)	0,058532 (0,0016459)
USDPLN 60 min.	0,041023 (0,040393)	1,0008 (0,69052)	0,02587 (0,020003)

Tabela 6.32. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CKLS wykonanych dla różnej częstotliwości poziomów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^2]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$	$\beta$
CAC40 5min.	0,041818 (0,037084)	895,29 (158,42)	0,83707 (0,086189)	0,32531 (0,014834)
CAC40 10min.	0,041622 (0,037209)	895,46 (148,85)	0,72974 (0,12558)	0,33198 (0,024141)
CAC40 30min.	0,041622 (0,037209)	895,46 (148,85)	0,65405 (0,1948)	0,31955 (0,043042)
DAX 5min.	0,038856 (0,035721)	962,16 (311,69)	0,9353 (0,090317)	0,29803 (0,01337)
DAX 10min.	0,00070556 (0,0043726)	502,21 (150)	0,85102 (0,10834)	0,29234 (0,0182080)
DAX 30min.	0,038856 (0,037215)	951,24 (275,15)	0,65914 (0,17245)	0,28875 (0,037517)
NASDAQ 5min.	0,042443 (0,035049)	793,62 (119,5)	0,86646 (0,12507)	0,32797 (0,020505)
NASDAQ 10min.	0,043192 (0,036504)	811,25 (193,4)	0,85078 (0,17087)	0,32931 (0,029876)
NASDAQ 30min.	0,043192 (0,036504)	811,25 (193,4)	7,9139 (2,8138)	-0,11641 (0,34973)
WIG20 5 min.	0,041059 (0,041895)	861,12 (281,3)	1,9981 (0,48189)	0,29383 (0,033913)
WIG20 10 min.	0,040245 (0,036423)	782,34 (247,78)	1,3621 (0,43217)	0,34066 (0,045763)
WIG20 20 min.	0,042593 (0,038338)	748,43 (205,66)	0,84388 (0,35268)	0,38889 (0,061779)
WIG20 40 min.	0,040694 (0,039633)	741,24 (223,78)	0,82348 (0,5616)	0,29805 (0,14033)

Tabela 6.34. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CKLS wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^2]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$	$\beta$
EURPLN 10 min.	0,034529 (0,038717)	229,87 (190,21)	0,13781 (0,069605)	0,75879 (0,1035)
EURPLN 30 min.	0,036878 (0,03726)	195,25 (148,38)	0,096928 (0,061422)	0,79315 (0,13475)
EURPLN 60 min.	0,036878 (0,03726)	195,25 (148,38)	0,066254 (0,072625)	0,47128 (0,36504)
USDPLN 10 min.	0,039854 (0,034849)	118,62 (66,094)	2,3362 (1,4567)	0,26783 (0,13413)
USDPLN 30 min.	0,037268 (0,035275)	98,453 (72,402)	0,18515 (0,14246)	0,79054 (0,17853)
USDPLN 60 min.	0,038174 (0,037242)	98,328 (72,402)	0,081387 (0,10582)	0,51528 (0,42046)

Tabela 6.29. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CKLS wykonanych dla różnej częstotliwości poziomów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^3]$ .

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$	$\beta$
CAC40 5min.	0,039906 (0,038147)	906,44 (178,34)	0,87456 (0,11265)	0,33305 (0,018243)
CAC40 10min.	0,040953 (0,035733)	910,37 (187,41)	0,95984 (0,17004)	0,30732 (0,024833)
CAC40 30min.	0,040953 (0,033573)	912,21 (193,88)	0,97727 (0,30713)	0,27417 (0,04554)
DAX 5min.	0,037752 (0,041106)	1152,4 (518,46)	1,2633 (0,13784)	0,26471 (0,014126)
DAX 10min.	0,043992 (0,03639)	919,77 (223,16)	1,038 (0,16114)	0,27476 (0,021371)
DAX 30min.	0,039264 (0,033653)	872,54 (189,83)	1,0708 (0,30704)	0,23127 (0,040564)
NASDAQ 5min.	0,039259 (0,033104)	811,31 (174,51)	1,0828 (0,17101)	0,31403 (0,022922)
NASDAQ 10min.	0,045107 (0,039641)	822,38 (205,6)	1,0322 (0,25342)	0,31592 (0,035675)
NASDAQ 30min.	0,038126 (0,038264)	816,47 (178,34)	1,095 (0,47584)	0,27153 (0,063099)
WIG20 5 min.	0,04254 (0,038029)	835,5 (297,85)	2,5055 (0,64538)	0,27009 (0,037406)
WIG20 10 min.	0,04239 (0,038741)	813,2 (224,96)	1,814 (0,57823)	0,30639 (0,046935)
WIG20 20 min.	0,041771 (0,037556)	796,49 (236,34)	2,6518 (1,1574)	0,24023 (0,065626)
WIG20 40 min.	0,042794 (0,038452)	791,49 (221,32)	0,45192 (0,3216)	0,36037 (0,17379)

**Tabela 6.35. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu CKLS wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów. Jako estymator wariancji zrealizowanej wykorzystano  $[X_{\delta}^3]$ .**

Szereg zwrotów	$\kappa$	$\mu$	$\sigma$	$\beta$
EURPLN 10 min.	0,033676 (0,038408)	231,03 (193,55)	0,1995 (0,098961)	0,68325 (0,10416)
EURPLN 30 min.	0,039291 (0,037919)	195,81 (141,48)	0,15105 (0,094003)	0,70788 (0,13474)
EURPLN 60 min.	0,038203 (0,036118)	188,34 (117,83)	0,075515 (0,082959)	0,44979 (0,35297)
USDPLN 10 min.	0,039424 (0,038104)	98,855 (76,317)	0,13776 (0,10304)	0,84908 (0,17162)
USDPLN 30 min.	0,038145 (0,037728)	98,648 (65,078)	0,14649 (0,11287)	0,841 (0,17874)
USDPLN 60 min.	0,037472 (0,036728)	98,597 (61,364)	0,17065 (0,22018)	0,42539 (0,35111)

W zdecydowanej większości przypadków średnie oszacowań są bardzo bliskie oszacowaniom parametrów przedstawionym w tabelach 6.12-6.17. Rozbieżności możemy zaobserwować jedynie dla parametrów dryfu. Wynika to z faktu, że funkcja dryfu dla badanych procesów przyjmuje małe wartości. Jest to również przyczyna stosunkowo dużych wartości odchyłeń standardowych funkcji dryfu. Jednocześnie obserwujemy, że częstotliwość danych nie ma wpływu na wartości średnich i odchyłeń standardowych. Jedynie wartość odchyłeń standardowych oszacowań parametrów  $\sigma$  w modelach Vašíčka i CIR zwiększa się wraz ze spadkiem częstotliwości, ale nawet dla danych o najniższej częstotliwości pozostaje ona dosyć mała. Wyniki symulacji świadczą o dobrym dopasowaniu modeli do procesów cen logarytmicznych generujących przez dane wysokiej częstotliwości. Zarazem potwierdza się postawiony już wcześniej wniosek mówiący, że stosowanie danych bardzo wysokiej częstotliwości nie ma istotnego wpływu na jakość estymacji.

Podobnie jak dla procesów generujących dane dzienne, parametry modeli dyfuzji estymowaliśmy również za pomocą uogólnionej metody momentów. W tym przypadku również zastosowaliśmy metodę dwustopniową, jako, że oszacowania otrzymane za pomocą metod wyższych rzędów były niemal identyczne. W tabelach 6.36-6.54 przedstawione są oszacowania parametrów otrzymane dla procesów cen logarytmicznych generujących dane różnej częstotliwości. W tabelach brakuje wyników estymacji modelu CKLS dla półgodzinnych poziomów logarytmicznych indeksu CAC40 oraz

logarytmicznych kursów EURPLN dowolnej częstotliwości. W tych czterech przypadkach algorytm optymalizacyjny nie wykazał zbieżności.

**Tabela 6.36. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 5-min. logarytmicznych poziomów indeksu CAC40 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .**

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	6,8971	4,1691	0,098083	6,0989	4,6268	0,18747
	$\mu$	8,5893	0,0009277	0	8,6057	0,0014558	0
	$\sigma$	0,11935	0,0028264	0	0,10916	0,0027075	0
CIR	$\kappa$	6,9445	4,1691	0,095795	6,1278	4,6268	0,18538
	$\mu$	8,5892	0,0009123	0	8,6056	0,0014398	0
	$\sigma$	0,040766	0,0009669	0	0,037255	0,0009251	0
CKLS	$\kappa$	3,5727	4,1993	0,3949	3,5727	4,651	0,44241
	$\mu$	8,6185	0,0071211	0	8,6185	0,0061096	0
	$\sigma$	3,4183	1674,8	0,99837	3,4081	2004	0,99864
	$\beta$	-2,4504	228,23	0,98969	-2,449	273,59	0,9914

**Tabela 6.37. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 10-min. logarytmicznych poziomów indeksu CAC40 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .**

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	3,4485	2,0845	0,098098	4,2478	4,8059	0,37679
	$\mu$	8,5893	0,0009277	0	8,6124	0,0037803	0
	$\sigma$	0,084394	0,0019986	0	0,11143	0,0029323	0
CIR	$\kappa$	3,4723	2,0845	0,09581	4,2558	4,8059	0,37589
	$\mu$	8,5892	0,0009123	0	8,6124	0,0037651	0
	$\sigma$	0,028826	0,0006837	0	0,038037	0,0010022	0
CKLS	$\kappa$	1,7863	2,0996	0,39491	3,6251	4,8079	0,45089
	$\mu$	8,6185	0,0071211	0	8,618	0,0060884	0
	$\sigma$	2,8212	56,974	0,96051	5,294	360,95	0,9883
	$\beta$	-1,7792	9,4069	0,80856	-2,115	31,728	0,93432



Tabela 6.38. Oszacowania parametrów modeli Vašíčka, CIR dla 30-min. logarytmicznych poziomów CAC40 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	5,7503	4,1442	0,17089	2,7934	4,9865	0,57535
	$\mu$	8,5957	0,0015409	0	8,624	0,012392	0
	$\sigma$	0,11844	0,0036152	0	0,11223	0,0036039	0
CIR	$\kappa$	5,7828	4,1442	0,23138	2,7855	4,9865	0,57644
	$\mu$	8,5956	0,0015174	0	8,6242	0,012533	0
	$\sigma$	0,040459	0,0012367	0	0,038317	0,0012326	0

Tabela 6.39. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla 5-min. logarytmicznych poziomów indeksu DAX wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	3,7961	3,1142	0,22288	2,9042	3,4272	0,39677
	$\mu$	8,7789	0,0059563	0	8,8348	0,021224	0
	$\sigma$	0,12411	0,0030059	0	0,11066	0,002854	0
CIR	$\kappa$	3,8471	3,1143	0,21674	2,9318	3,4271	0,3923
	$\mu$	8,7782	0,0057261	0	8,8341	0,020607	0
	$\sigma$	0,042019	0,0010202	0	0,037424	0,0009670	0
CKLS	$\kappa$	0,67776	3,1405	0,82913	0,67776	3,4481	0,84417
	$\mu$	9,1338	4,0462	0,024003	9,1338	4,4643	0,040773
	$\sigma$	2,665	10691	0,9998	2,6559	13227	0,99984
	$\beta$	-2,8475	1854,6	0,99856	-2,8459	2298,3	0,99884

Tabela 6.40. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla 10-min. logarytmicznych poziomów indeksu DAX wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	3,9401	3,1171	0,20626	2,2311	3,5053	0,52448
	$\mu$	8,7758	0,0052602	0	8,8617	0,054762	0
	$\sigma$	0,12474	0,0036987	0	0,11062	0,0034285	0
CIR	$\kappa$	3,9906	3,1171	0,20051	2,2481	3,5053	0,52132
	$\mu$	8,7751	0,0050689	0	8,8611	0,053471	0
	$\sigma$	0,04223	0,0012555	0	0,037415	0,001162	0
CKLS	$\kappa$	0,70773	3,1568	0,82262	0,70773	3,5191	0,84062
	$\mu$	9,116	3,4508	0,0082695	9,116	3,9178	0,020005
	$\sigma$	2,3503	1325,1	0,99858	4,4328	3092,1	0,99886
	$\beta$	-2,2869	260,62	0,99147	-2,5797	321,91	0,99237

Tabela 6.41. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 30-min. logarytmicznych poziomów indeksu DAX wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	3,8647	3,0156	0,20013	2,1132	3,4869	0,54454
	$\mu$	8,7772	0,0052839	0	8,8623	0,059397	0
	$\sigma$	0,1221	0,0042762	0	0,11297	0,0043821	0
CIR	$\kappa$	3,9217	3,0156	0,19358	2,1298	3,4869	0,54138
	$\mu$	8,7764	0,0050552	0	8,8615	0,057947	0
	$\sigma$	0,041343	0,0014514	0	0,03822	0,0014865	0
CKLS	$\kappa$	0,69313	3,0917	0,82263	0,69313	3,5098	0,84347
	$\mu$	9,1321	3,721	0,014197	9,1321	4,3646	0,036528
	$\sigma$	4,3261	73,704	0,9532	3,8783	79,004	0,96085
	$\beta$	-1,732	7,875	0,77687	-1,6811	9,4039	0,81661

Tabela 6.42. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 5-min. logarytmicznych poziomów indeksu NASDAQ wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	-1,0765	2,9117	0,71162	-2,0188	3,283	0,53863
	$\mu$	7,529	0,30355	0	7,6875	0,015285	0
	$\sigma$	0,12845	0,0035074	0	0,11464	0,002935	0
CIR	$\kappa$	-1,1278	2,9117	0,69851	-2,0731	3,2829	0,52774
	$\mu$	7,5387	0,25189	0	7,6893	0,014027	0
	$\sigma$	0,046153	0,0012641	0	0,04114	0,0010555	0
CKLS	$\kappa$	1,343	2,9348	0,64725	1,343	3,3563	0,68907
	$\mu$	7,8755	0,12953	0	7,8755	0,13003	0
	$\sigma$	2,4916	4648,7	0,99957	5,0321	12027	0,99967
	$\beta$	-2,8195	913,05	0,9971	-3,1626	1166,9	0,9975

Tabela 6.43. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 10-min. logarytmicznych poziomów indeksu NASDAQ wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	-0,019792	2,9739	0,99469	-0,58681	3,3203	0,85972
	$\mu$	-2,9402	2,5672e+6	1	7,5038	1,887	0
	$\sigma$	0,13052	0,0034876	0	0,11972	0,0033527	0
CIR	$\kappa$	-0,051268	2,9739	0,98625	-0,62046	3,3202	0,85177
	$\mu$	3,6165	56752	0,99995	7,517	1,5142	0
	$\sigma$	0,0469	0,001257	0	0,042977	0,0012068	0
CKLS	$\kappa$	1,4286	2,9828	0,632	1,4286	3,3505	0,66984
	$\mu$	7,8675	0,10783	0	7,8675	0,10429	0
	$\sigma$	3,5857	597,35	0,99521	10,306	2202,5	0,99627
	$\beta$	-2,404	81,521	0,97159	-2,9197	104,39	0,97387

Tabela 6.44. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 30-min. logarytmicznych poziomów indeksu NASDAQ wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	2,4622	2,9859	0,40972	0,97332	3,4746	0,77942
	$\mu$	7,8036	0,015221	0	7,9429	0,55531	0
	$\sigma$	0,13087	0,0042187	0	0,12069	0,0039627	0
CIR	$\kappa$	2,4839	2,9859	0,4056	0,97087	3,4747	0,77996
	$\mu$	7,8031	0,01482	0	7,9446	0,56656	0
	$\sigma$	0,047027	0,0015207	0	0,043329	0,0014293	0
CKLS	$\kappa$	1,4271	2,9937	0,63363	1,4271	3,4776	0,68159
	$\mu$	7,8667	0,10776	0	7,8667	0,11059	0
	$\sigma$	5,9759	72,909	0,93469	5,1839	70,242	0,94118
	$\beta$	-1,9672	5,9699	0,67946	-1,897	6,6181	0,71727

Tabela 6.45. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla 5-min. logarytmicznych poziomów indeksu WIG20 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	11,346	7,0745	0,1088	12,026	6,2088	0,052759
	$\mu$	8,0591	0,0010257	0	8,0534	0,0006264	0
	$\sigma$	0,23182	0,0051336	0	0,22289	0,0057047	0
CIR	$\kappa$	11,299	7,0745	0,11028	11,997	6,2087	0,053336
	$\mu$	8,0591	0,0010354	0	8,0533	0,0006290	0
	$\sigma$	0,081727	0,0018119	0	0,0785	0,0020119	0
CKLS	$\kappa$	12,913	7,0843	0,068366	12,913	6,2153	0,037757
	$\mu$	8,0563	0,0007569	0	8,0563	0,0005571	0
	$\sigma$	14,87	2053,3	0,99422	42,781	6926,8	0,99507
	$\beta$	-2,676	66,27	0,96178	-3,1824	77,623	0,96216

Tabela 6.46. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla 10-min. logarytmicznych poziomów indeksu WIG20 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	11,391	6,803	0,094107	12,081	6,1904	0,051035
	$\mu$	8,0593	0,0009436	0	8,0529	0,0006106	0
	$\sigma$	0,22265	0,0054915	0	0,21406	0,0060915	0
CIR	$\kappa$	11,372	6,803	0,094658	12,075	6,1904	0,051159
	$\mu$	8,0593	0,0009471	0	8,0528	0,0006109	0
	$\sigma$	0,07849	0,0019384	0	0,075398	0,0021491	0
CKLS	$\kappa$	12,075	6,8057	0,076074	12,075	6,191	0,037769
	$\mu$	8,058	0,0008209	0	8,058	0,0006351	0
	$\sigma$	8,221	235,01	0,9721	8,0847	271,93	0,97628
	$\beta$	-2,0081	13,719	0,85494	-1,9999	16,125	0,8768

Tabela 6.47. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 20-min. logarytmicznych poziomów indeksu WIG20 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	10,236	6,6989	0,12664	11,413	6,043	0,05906
	$\mu$	8,0619	0,0011908	0	8,0514	0,0006559	0
	$\sigma$	0,21974	0,0069741	0	0,21238	0,0072741	0
CIR	$\kappa$	10,189	6,6988	0,12839	11,397	6,0429	0,059396
	$\mu$	8,062	0,0012036	0	8,0512	0,0006574	0
	$\sigma$	0,07747	0,0024617	0	0,074813	0,0025668	0
CKLS	$\kappa$	11,746	6,7204	0,08062	11,746	6,0475	0,052212
	$\mu$	8,0584	0,0008530	0	8,0584	0,0006476	0
	$\sigma$	5,0412	73,022	0,94497	3,591	57,719	0,9504
	$\beta$	-1,5569	6,9513	0,76733	-1,3937	7,7055	0,8059

Tabela 6.48. Oszacowania parametrów modeli Vašička, CIR i CKLS dla 40-min. logarytmicznych poziomów indeksu WIG20 wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T(B)$ .

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašička	$\kappa$	12,169	6,7733	0,072629	12,593	6,2089	0,042757
	$\mu$	8,0586	0,0008093	0	8,0539	0,0005895	0
	$\sigma$	0,21909	0,0081884	0	0,21926	0,0079706	0
CIR	$\kappa$	12,183	6,7733	0,072313	12,628	6,2094	0,042198
	$\mu$	8,0586	0,0008071	0	8,0536	0,0005857	0
	$\sigma$	0,077258	0,0028895	0	0,07733	0,0028143	0
CKLS	$\kappa$	11,931	6,7782	0,078608	11,931	6,2296	0,055682
	$\mu$	8,0591	0,0008493	0	8,0591	0,0006970	0
	$\sigma$	4,5686	3,2606e+7	1	4,5686	3,7817e+7	1
	$\beta$	-4,6124	3,4256e+6	1	-4,6124	3,9731e+6	1

Tabela 6.49. Oszacowania parametrów modeli Vašíčka, CIR i CKLS dla 10-min logarytmicznych kursów EURPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	16,4	5,1605	0,0014897	15,49	5,1327	0,002554
	$\mu$	1,3535	4,3008e-5	0	1,35	3,4011e-5	0
	$\sigma$	0,070939	0,0034422	0	0,064022	0,0014932	0
CIR	$\kappa$	16,179	5,1605	0,0017251	15,348	5,1323	0,0027952
	$\mu$	1,3535	4,4198e-5	0	1,35	3,4624e-5	0
	$\sigma$	0,060934	0,0029441	0	0,055112	0,0012816	0

Tabela 6.50. Oszacowania parametrów modeli Vašíčka, CIR dla 30-min logarytmicznych kursów EURPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	15,814	4,9261	0,0013387	12,555	5,1661	0,015144
	$\mu$	1,3534	4,1481e-5	0	1,3487	5,4225e-5	0
	$\sigma$	0,067029	0,0033742	0	0,05848	0,0018261	0
CIR	$\kappa$	15,641	4,926	0,0015108	12,509	5,1661	0,015515
	$\mu$	1,3534	4,2397e-5	0	1,3487	5,4572e-5	0
	$\sigma$	0,057609	0,0028834	0	0,050333	0,0015604	0

Tabela 6.51. Oszacowania parametrów modeli Vašíčka, CIR dla 60-min logarytmicznych kursów EURPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	14,352	4,8127	0,0029073	11,68	5,0803	0,021639
	$\mu$	1,3529	4,859e-5	0	1,348	6,1091e-5	0
	$\sigma$	0,06512	0,0033155	0	0,057945	0,0020743	0
CIR	$\kappa$	14,201	4,8125	0,0032172	11,636	5,0805	0,022137
	$\mu$	1,3528	4,9671e-5	0	1,348	6,1491e-5	0
	$\sigma$	0,055953	0,0028333	0	0,049863	0,00177	0

**Tabela 6.52. Oszacowania parametrów modeli Vašíčka, CIR dla 10-min logarytmicznych kursów USDPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).**

Macierz wag		(A)			(B)		
Model	para- metr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vaší- čka	$\kappa$	5,1683	4,4824	0,24895	5,0048	4,5026	0,26639
	$\mu$	1,0771	0,0016979	0	1,0749	0,0015297	0
	$\sigma$	0,098827	0,0038279	0	0,094261	0,0030239	0
CIR	$\kappa$	5,0648	4,4824	0,25856	4,946	4,5025	0,27203
	$\mu$	1,0764	0,001806	0	1,0747	0,0015803	0
	$\sigma$	0,093818	0,0036218	0	0,090322	0,0028868	0
CKLS	$\kappa$	4,6926	4,4918	0,2962	4,6926	4,5066	0,29777
	$\mu$	1,0734	0,0023153	0	1,0734	0,0018321	0
	$\sigma$	0,076323	0,013355	0	0,076323	0,0097432	0
	$\beta$	2,3852	1,5048	0,21032	2,3852	1,332	0,157

**Tabela 6.53. Oszacowania parametrów modeli Vašíčka, CIR dla 30-min logarytmicznych kursów USDPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).**

Macierz wag		(A)			(B)		
Model	para- metr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vaší- čka	$\kappa$	4,3227	4,4093	0,327	1,7479	4,2524	0,68109
	$\mu$	1,0701	0,0029152	0	1,0341	0,034911	0
	$\sigma$	0,096928	0,0041194	0	0,08905	0,0032127	0
CIR	$\kappa$	4,3443	4,4093	0,3246	2,0054	4,2516	0,6372
	$\mu$	1,0702	0,002871	0	1,0419	0,0218	0
	$\sigma$	0,09214	0,0038912	0	0,085282	0,0030422	0
CKLS	$\kappa$	4,4962	4,4107	0,30813	4,4962	4,4274	0,30995
	$\mu$	1,0718	0,0025605	0	1,0718	0,0018135	0
	$\sigma$	0,06018	0,01318	0	0,06018	0,0099621	0
	$\beta$	4,3413	1,8216	0,017245	4,3413	1,6859	0,010087

**Tabela 6.54. Oszacowania parametrów modeli Vašíčka, CIR dla 60-min logarytmicznych kursów USDPLN wraz z błędami standardowymi oszacowań oraz  $p$ -wartościami. Macierz wag wyznaczona jest na podstawie wzoru 5.1 (A) oraz jąder Bartletta z automatycznym doбором  $B_T$  (B).**

Macierz wag		(A)			(B)		
Model	parametr	oszac. parametru	błąd std.	$p$ -wartość	oszac. parametru	błąd std.	$p$ -wartość
Vašíčka	$\kappa$	4,3227	4,4093	0,327	1,5627	4,1106	0,68109
	$\mu$	1,0701	0,0029152	0	1,0344	0,038083	0
	$\sigma$	0,096928	0,0041194	0	0,088662	0,0031096	0
CIR	$\kappa$	4,3443	4,4093	0,3246	1,8357	4,1111	0,6372
	$\mu$	1,0702	0,002871	0	1,0428	0,022261	0
	$\sigma$	0,09214	0,0038912	0	0,084757	0,0029317	0
CKLS	$\kappa$	4,4962	4,4107	0,30813	4,2703	4,3507	0,30995
	$\mu$	1,0718	0,0025605	0	1,0697	0,0018518	0
	$\sigma$	0,06018	0,01318	0	0,060805	0,011056	0
	$\beta$	4,3413	1,8216	0,035079	4,0437	1,8116	0,022793

Analizując oszacowania parametrów otrzymane dla procesów cen logarytmicznych otrzymane na podstawie danych śróddziennych, od razu zauważamy, że błędy standardowe są znacznie mniejsze niż w przypadku wykorzystania danych dziennych, przez co w wielu przypadkach wszystkie oszacowania parametrów modeli są istotne. Podobnie jak dla metody Phillipsa i Yu, oszacowania parametrów przyjmują zbliżone wartości, bez względu na zastosowaną częstotliwość. W przypadku pierwszej metody skala rozbieżności jest jednak mniejsza. Błędy standardowe oszacowania parametru  $\kappa$  często pozostają duże. Podobnie jak w przypadku danych dziennych, oszacowania parametrów funkcji dyfuzji w modelu stałej elastyczności wariancji dla logarytmicznych notowań indeksów giełdowych przyjmują wartości niezgodne z jego specyfikacją i pozostają nieistotne. Otrzymane oszacowania wskazują również na znikomy wpływ zastosowanej macierzy wag na wyniki estymacji.

### **6.3. Modelowanie procesów cen logarytmicznych za pomocą modeli zmienności stochastycznej.**

W celu oszacowania parametrów modeli zmienności stochastycznej wykorzystaliśmy opisany w podrozdziale 3.3. algorytm bazujący na dwukrokowej metodzie Phillipsa i Yu (2007). Do badania wykorzystaliśmy te same szeregi, co w podrozdziale 6.2. Szeregi zmienności chwilowej szacowaliśmy na bazie 5-minutowych zwrotów logarytmicznych dla indeksów giełdowych oraz 10-minutowych zwrotów logarytmicznych dla



kursów walutowych. Statystyki opisowe szeregów oszacowań zmienności chwilowej przedstawione są w tabeli 6.55.

**Tabela 6.55. Statystyki opisowe szeregów oszacowań zmienności chwilowej logarytmicznych poziomów indeksów giełdowych oraz logarytmicznych kursów walutowych.**

Szereg oszacowań wariacji chwilowej	liczba obs.	średnia	odch. std.	skośność	kurtoza	min.	maks.
CAC 40 10min.	6528	1,4561	2,9824	7,6502	106,58	0	73,922
CAC 40 30min.	2176	1,4561	2,0912	5,1959	49,366	0,025977	33,218
DAX 10min.	6528	1,5883	4,0288	11,868	229,89	0	117,17
DAX 30min.	2176	1,5888	2,7047	6,5743	70,082	0	45,617
NASDAQ 10min.	4614	1,9997	3,8835	11,22	258,91	0	121,85
NASDAQ 30min.	1538	1,9819	2,5685	5,7198	67,41	0	44,343
WIG20 20 min.	2520	6,4627	10,109	5,2575	50,486	0,0321	178,71
WIG20 40 min.	1210	6,4627	7,8921	4,7793	39,51	0,1456	103,05
EURPLN 30 min.	2268	27,006	86,517	25,685	772,95	0,39398	2670,9
EURPLN 60 min.	1134	27,006	119,87	36,941	1583,4	0,01633	5240,2
USDPLN 30 min.	2268	58,544	139,24	18,342	519,55	0,19496	4557,7
USDPLN 60 min.	1134	58,544	101,63	12,031	226,92	1,6743	2284,4

Szacując parametry modelu Hestona, jako wartości startowe parametrów dryfu dla procesu wariacji chwilowej zdecydowaliśmy się przyjąć  $\theta = 0$ , oraz  $\omega$  równe średniej szeregu oszacowań. Za wartość startową  $\xi$  przyjęliśmy oszacowanie tego parametru za pomocą estymatora wprowadzonego przez Yoshidę, natomiast w przypadku parametru  $\mu$ , średnią stopę wzrostu procesu cen logarytmicznych. Niestety dla szeregów CAC40, DAX i NASDAQ, dla których stosunek wariacji oszacowań zmienności chwilowej do jej średniej był znacznie większy niż w pozostałych przypadkach, algorytm minimalizacyjny nie wykazał zbieżności. W tabeli 6.56 zaprezentowane są oszacowania parametrów modelu Hestona otrzymane dla logarytmicznych poziomów indeksu WIG20 oraz logarytmicznych kursów walutowych EURPLN i USDPLN, otrzymane za pomocą algorytmu zaprezentowanego w rozdziale 3.3.

**Tabela 6.56. Estymatory parametrów modelu Hestona dla procesów logarytmicznych kursów walutowych, w zależności od częstotliwości danych wykorzystanych do estymacji**

Szereg notowań log.	Częstotliwość notowań log.	Częstotliwość oszacowań war. chwilowej	$\mu$	$\theta$	$\omega$	$\xi$
WIG20	5 min.	20min.	-0,0052964	8,8902	6,4572	111,06
	5 min.	40min.	-0,042903	4,5052	6,4499	71,596
EURPLN	10min.	30min.	8,0044e-5	11,417	25,463	267,88
	10min.	60min.	0,0006699	7,4744	26,965	130,89
USDPLN	10min.	30min.	0,0047375	12,558	58,626	447,32
	10min.	60min.	0,02966	7,4524	58,612	196,86

W tabeli 6.57 przedstawione są wyniki symulacji Monte Carlo, które przeprowadziliśmy w celu oceny jakości otrzymanych oszacowań.

**Tabela 6.57. Średnie (odchylenia standardowe) oszacowań parametrów otrzymanych w wyniku przeprowadzenia 1000 symulacji Monte Carlo dla modelu Hestona, wykonanych dla różnej częstotliwości kursów logarytmicznych, z parametrami odpowiadającymi wcześniej otrzymanym oszacowaniom dla poszczególnych szeregów.**

Szereg notowań log.	Częstotliwość notowań log.	Częstotliwość oszacowań war. chwilowej	$\mu$	$\theta$	$\omega$	$\xi$
WIG20	5 min.	20min.	-0,015834 (0,052523)	7,4307 (3,7905)	8,7 (5,3113)	98,965 (2,0181)
	5 min.	40min.	-0,018346 (0,02594)	5,7627 (4,6436)	7,6661 (6,5477)	63,448 (1,6618)
EURPLN	10min.	30min.	0,018895 (0,046434)	2,4895 (7,0156)	25,858 (66,866)	236,86 (5,3324)
	10min.	60min.	0,024743 (0,074922)	0,60098 (5,0833)	21,772 (17,416)	116,34 (2,7822)
USDPLN	10min.	30min.	0,016513 (0,03256)	6,7418 (5,4602)	45,565 (30,1)	396,64 (10,3102)
	10min.	60min.	0,012466 (0,021512)	4,6255 (4,3594)	58,109 (69,902)	175,87 (3,4594)

Jak łatwo zauważyć, odchylenia standardowe parametrów są znacznie większe niż w przypadku oszacowań parametrów modeli dyfuzji. Średnie oszacowań parametrów  $\theta$  i  $\mu$  również odbiegają od oszacowań otrzymanych dla wyjściowych szeregów. Jedynie oszacowania parametru zmienności  $\xi$  procesu  $\sigma^2(t)$  mają stosunkowo małe odchylenia standardowe. Przyczyny takiej sytuacji należy doszukiwać się w burzliwej dynamice szeregów szacowanej wariancji chwilowej. Dopasowanie modelu dyfuzji, a zwłaszcza jego funkcji dryfu do takiego szeregu, jest znacznie trudniejsze.

## 6.4. Prognozy średniej warunkowej

W podrozdziale tym przedstawiamy wyniki badań dotyczących oceny jakości jednodniowych prognoz ex-post otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu oraz za pomocą uogólnionej metody momentów. Dla wszystkich badanych w niniejszym rozdziale szeregów wyznaczyliśmy 100 prognoz ex-post, a do oceny ich jakości wykorzystaliśmy powszechnie stosowane miary błędów opisane w dodatku A. Uzyskanie małych wartości błędów świadczy o satysfakcjonującej jakości prognoz i może sugerować dobre dopasowanie modeli dyfuzji do procesów generujących dzienne ceny logarytmiczne. Dobra jakość prognoz przema-

wia również za stosowaniem modeli dyfuzji do prognozowania średniej warunkowej jako alternatywą dla parametrycznych modeli szeregów czasowych.

Błędy prognoz zostały porównane z błędami otrzymanymi dla prognoz pochodzących z dopasowanych do szeregów powszechnie stosowanych parametrycznych modeli średniej warunkowej, dobranych w taki sposób, aby oszacowania wszystkich parametrów były istotne. Rząd modelu został określony na podstawie kryterium informacyjnego Schwarz'a. Estymację parametrów tych modeli przeprowadziliśmy za pomocą programu Time Series Modelling 4.26.

#### 6.4.1. Testy na występowanie pierwiastka jednostkowego

Występowanie pierwiastków jednostkowych jest bardzo powszechnym zjawiskiem w szeregach cen oraz cen logarytmicznych. Wobec tego, jeżeli zamierzamy modelować szereg czasowy za pomocą modeli  $ARMA(p, q)$ , to koniecznym jest uprzednie zweryfikowanie występowania pierwiastków jednostkowych w procesach. W sytuacji pozytywnej weryfikacji ich występowania, konieczne jest modelowanie szeregu za pomocą modelu  $ARIMA$ . Model  $ARIMA(p, d, q)$  oznacza, że po  $d$ -krotnym zróżnicowaniu szeregu czasowego otrzymamy model  $ARMA(p, q)$ . Najbardziej popularnym testem na występowanie pierwiastka jednostkowego jest test Dickey'a-Fullera (1979). Jego podstawową wadą jest mała moc, wobec czego faworyzuje on hipotezę zerową o występowaniu pierwiastka jednostkowego w badanym szeregach czasowych. Przyjmijmy, że model dla szeregu czasowego dany jest wzorem

$$X_t = \rho X_{t-1} + \varepsilon_t, \quad (6.1)$$

gdzie  $\varepsilon_t$  jest białym szumem. Powyższe równanie może być rozszerzone do postaci zawierającej dryf lub trend deterministyczny. Wówczas może ono przyjąć następującą postać

$$X_t = \alpha_0 + \rho X_{t-1} + \varepsilon_t \quad (6.2)$$

lub

$$X_t = \alpha_0 + \alpha_1 t + \rho X_{t-1} + \varepsilon_t. \quad (6.3)$$

Hipoteza zerowa testu mówi, że badany szereg zawiera co najmniej jeden pierwiastek jednostkowy, natomiast hipoteza alternatywna o jego braku. Zauważmy, że równania (6.1), (6.2) i (6.3) możemy przedstawić w następującej formie:

$$\begin{aligned}\Delta X_t &= (\rho - 1)X_{t-1} + \varepsilon_t = \delta X_{t-1} + \varepsilon_t, \\ \Delta X_t &= \alpha_0 + (\rho - 1)X_{t-1} + \varepsilon_t = \alpha_0 + \delta X_{t-1} + \varepsilon_t, \\ \Delta X_t &= \alpha_0 + \alpha_1 t + (\rho - 1)X_{t-1} + \varepsilon_t = \alpha_0 + \alpha_1 t + \delta X_{t-1} + \varepsilon_t,\end{aligned}$$

gdzie  $\delta = \rho - 1$  oraz  $\Delta$  jest operatorem różnicy. Wobec tego test sprowadza się do zwerifikowania hipotezy zerowej mówiącej, że  $\delta = 0$ , przeciwko hipotezie alternatywnej stwierdzającej, że  $\delta < 0$ . Statystyka testowa ma postać

$$DF = \frac{\hat{\delta}}{s(\hat{\delta})},$$

gdzie  $s$  jest błędem standardowym oszacowania  $\hat{\delta}$ . Ma ona rozkład niestandardowy, którego kwantyle można znaleźć między innymi w monografii Fullera (1976) oraz Charemza i Deadmana (1997).

Rozwinięciem koncepcji powyższego testu jest rozszerzony test Dickeya-Fullera wprowadzony przez Saida i Dickeya (1984). W najbardziej ogólnej wersji, do opisu szeregu czasowego, stosuje się model autoregresyjny rzędu  $p$  z dryfem i trendem. Postać takiego modelu dana jest wzorem

$$\Delta X_t = \alpha_0 + \alpha_1 t + \delta X_{t-1} + \sum_{i=1}^p \beta_i \Delta X_{t-i} + \varepsilon_t.$$

Wartość  $p$  będziemy dobierać automatycznie na podstawie kryterium informacyjnego Schwarzera. Podobnie jak w standardowej wersji testu, za pomocą tej samej statystyki weryfikujemy hipotezę zerową mówiącą, że  $\delta = 0$ , przeciwko hipotezie stwierdzającej, że  $\delta < 0$ .

Alternatywnym testem na występowanie pierwiastka jednostkowego jest test Phillipsa-Perrona (1988) (por. Doman, Doman 2004). Niestety, moc tego testu jest niska podobnie, jak w przypadku wyżej omówionych testów Dickeya-Fullera, ale wykazuje on lepsze własności statystyczne w przypadku występowania autokorelacji składników losowych. Do opisu szeregu czasowego stosuje się model wyrażony za pomocą równania (6.3). Hipoteza zerowa testu zakłada, że  $\rho = 1$  oraz  $\alpha_1 = 0$ . Procedura testu opiera

się na estymacji regresji równania (6.3) na podstawie obserwacji  $X_1, X_2, \dots, X_T$ . Statystyka Phillipsa-Perrona będąca skorygowaną  $t$ -statystyką dla parametru  $\rho$  ma postać

$$Z_{t,T} = \frac{\hat{\sigma}_{u,T}}{\hat{\sigma}_T} \frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\sigma_T}} - \frac{[\hat{\sigma}_T - \hat{\sigma}_{u,T}] \hat{\sigma}_{\sigma_T} T}{2\hat{\sigma}_T (T-3)^{-1} \sqrt{\sum_{t=1}^T (\hat{u}_t)^2}}.$$

W powyższym wzorze  $\hat{\rho}_T$  jest estymatorem klasycznej metody najmniejszych kwadratów parametru  $\rho$ ,  $\hat{\sigma}_{\sigma_T}$  jest błędem standardowym dla  $\hat{\rho}_T$ ,  $\hat{u}_t$  dla  $t=1, 2, \dots, T$ , jest szeregiem reszt z estymowanej regresji. Ponadto  $\hat{\sigma}_T^2$  jest estymatorem Neweya-Westa (1987) wariancji długookresowej procesu  $u_t$  i przy założeniu, że istotnych jest tylko  $q$  pierwszych autokorelacji równy jest

$$\hat{\sigma}_T^2 = \hat{\sigma}_{u,T}^2 + \frac{2}{T} \sum_{j=1}^q \left[ 1 - \frac{j}{q+1} \right] \sum_{t=j+1}^T \hat{u}_t \hat{u}_{t-j},$$

gdzie  $\hat{\sigma}_{u,T}^2 = T^{-1} \sum_{t=1}^T \hat{u}_t^2$ .

Wartości krytyczne testu Phillipsa-Perrona są takie same, jak dla rozszerzonego testu Dickeya-Fullera. Ponieważ szeregi cen logarymicznych często zawierają pierwiastki jednostkowe, wobec tego w pierwszej kolejności zweryfikowaliśmy ich występowanie. Wykorzystaliśmy w tym celu rozszerzony test Dickeya-Fullera (Said, Dickey, 1984) oraz Phillipsa-Perrona (1988). W przeprowadzonych obliczeniach podobnie jak Perron (1988) przyjęliśmy  $q$  równe części całkowitej z wielkości  $4(T/100)^{-1/4}$ .

W tabeli 6.58 przedstawione są wartości zmodyfikowanej statystyki Dickeya-Fullera oraz Phillipsa-Perrona wyznaczone dla badanych szeregów czasowych.

**Tabela 6.58. Wartości statystyk rozszerzonego testu Dickeya-Fullera oraz Phillipsa-Perrona dla szeregów logarymicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarymicznych kursów EURPLN i USDPLN.**

Szereg notowań logarymicznych	Rozszerzony test Dickeya-Fullera		test Phillipsa-Perrona	
	Wartość statystyki	$p$ -wartość	Wartość statystyki	$p$ -wartość
CAC40	-1,51166	> 0,1	-1,53337	> 0,1
DAX	-1,13349	> 0,1	-1,18127	> 0,1
NASDAQ	-1,72328	> 0,1	-1,70064	> 0,1
WIG20	0,148923	< 0,05	0,263227	< 0,025
EURPLN	-1,37664	> 0,1	-1,40488	> 0,1
USDPLN	-0,682937	> 0,1	-0,551511	> 0,1

Jedynie dla indeksu WIG20 hipotezy zerowe rozszerzonego testu Dickeya-Fullera i testu Phillipsa Perrona zostały odrzucone. We wszystkich pozostałych przypadkach brakuje podstaw do odrzucenia hipotezy zerowej, wobec czego możemy wnioskować, że procesy te posiadają pierwiastki jednostkowe. Dlatego średnią warunkową modelowaliśmy za pomocą modeli  $ARIMA(p,1,q)$ . Niezależnie wykorzystaliśmy dwa rozkłady innowacji – Studenta oraz GED. Dodatkowo zweryfikowaliśmy występowanie efektu ARCH w badanych szeregach czasowych, stosując w tym celu test Engle’a (1982) oraz McLeod-Li (1983), który polega na wykorzystaniu testu autokorelacji Ljung-Boxa (1978) do kwadratów reszt z modelu liniowego. Odrzucenie hipotezy zerowej o warunkowej homoskedaktyczności badanych szeregów uzasadnia stosowanie modeli ARIMA-GARCH. W tabeli 6.59 przedstawione są wartości statystyki testowej Engle’a oraz ich  $p$ -wartości.

**Tabela 6.59. Wartości statystyki testu Engle’a wraz z  $p$ -wartościami dla szeregów reszt z modelu ARIMA dopasowanych do szeregów logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursów EURPLN i USDPLN.**

Szereg zwrotów	Model (rozkład innowacji)	test Engle’a		test McLeod-Li	
		Wartość statystyki	$p$ -wartość	Wartość statystyki	$p$ -wartość
CAC40	ARIMA(1,1,1) (Student)	61,2886	0	1334,4	0
	ARIMA(0,1,1) (GED)	61,5799	0	1332,95	0
DAX	ARIMA(1,1,1) (Student)	58,6881	0	1442	0
	ARIMA(1,1,1) (GED)	57,4231	0	1434,39	0
NASDAQ	ARIMA(0,1,2) (Student)	56,681	0	934,53	0
	ARIMA(0,1,2) (GED)	56,6771	0	934,638	0
EURPLN	ARIMA(1,1,1) (Student)	162,14	0	314,61	0
	ARIMA(1,1,1) (GED)	178,63	0	329,3	0
USDPLN	ARIMA(0,1,1) (Student)	40,4701	0	94,7376	0
	ARIMA(0,1,1) (GED)	40,5883	0	94,8459	0

Warunkowa homoskedaktyczność została silnie odrzucona we wszystkich przypadkach przez obydwa testy, co przemawia za modelowaniem i prognozowaniem szeregów za pomocą modeli ARIMA-GARCH.

#### 6.4.2. Ocena jakości prognoz

W tabelach 6.60-6.65 przedstawione są błędy prognoz ex-post otrzymane za pomocą 10000 symulacji z modelu Vašíčka, CIR oraz CKLS, z parametrami wyestymowanymi metodą Phillipsa, Yu (2007) i uogólnioną metodą momentów oraz za pomocą parame-

trycznych modeli ARIMA i ARIMA-GARCH. Oznaczenia błędów wyjaśnione są w dodatku A.

**Tabela 6.60. Wartości błędów prognoz, otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu, za pomocą uogólnionej metody momentów oraz za pomocą parametrycznych modeli średniej warunkowej. Logarytmiczne poziomy indeksu CAC40.**

Błąd	Metoda Phillipsa i Yu			Uogólniona metoda momentów		
	Vašiček	CIR	CKLS	Vašiček	CIR	CKLS
MSE	6,9069e-5	6,8895e-5	6,902e-5	6,8355e-5	6,817e-5	6,9836e-5
MedE	2,5575e-5	2,6669e-5	2,6053e-5	2,6771e-5	2,6696e-5	2,7293e-5
ME	0,00077931	0,00077539	0,00078897	0,00032694	0,00032105	0,0012119
AE	0,0062754	0,0062588	0,0062567	0,0062394	0,0062339	0,0062989
RMSE	0,0083108	0,0083003	0,0083078	0,0082677	0,0082565	0,0083568
MAPE	0,00072612	0,0007242	7,2396e-6	0,00072199	0,00072135	0,00072881
AMAPE	0,00036308	0,00036212	3,62e-6	0,000361	0,00036067	0,00036444
LL	9,2627e-7	9,2395e-7	9,2561e-7	9,1675e-7	9,1422e-7	9,3652e-7
Błąd	Modele parametryczne					
	ARIMA(1,1,1) (Student)	ARIMA(0,1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (Student)	ARIMA(1,1,1)- GARCH(1,1) (GED)		
MSE	7,0223e-5	-	7,0214e-5	7,0232e-5		
MedE	2,6884e-5	-	2,6873e-5	2,6895e-5		
ME	0,00092873	-	0,00092695	0,00093054		
AE	0,0063371	-	0,0063367	0,006338		
RMSE	0,0083799	-	0,0083794	0,0083805		
MAPE	0,00073332	-	0,00073327	0,00073342		
AMAPE	0,00036668	-	0,00036665	0,00036673		
LL	9,4187e-7	-	9,4175e-7	9,42e-7		

**Tabela 6.61. Wartości błędów prognoz, otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu, za pomocą uogólnionej metody momentów oraz za pomocą parametrycznych modeli średniej warunkowej. Logarytmiczne poziomy indeksu DAX.**

Błąd	Metoda Phillipsa i Yu			Uogólniona metoda momentów		
	Vašiček	CIR	CKLS	Vašiček	CIR	CKLS
MSE	7,7296e-5	7,7358e-5	7,7691e-5	7,8574e-5	7,8449e-5	8,0539e-5
MedE	2,6917e-5	2,7475e-5	2,742e-5	2,706e-5	2,9142e-5	3,0036e-5
ME	0,0014902	0,0014927	0,0014958	0,0017948	0,0017836	0,0023101
AE	0,0066737	0,0066797	0,0066842	0,0067379	0,0067449	0,0068758
RMSE	0,0087918	0,0087953	0,0088142	0,0088642	0,0088571	0,0089744
MAPE	0,00075446	0,00075514	7,5564e-6	0,00076169	0,00076248	0,00077721
AMAPE	0,00037726	0,0003776	3,7785e-6	0,00038089	0,00038128	0,00038867
LL	9,8999e-7	9,9078e-7	9,9505e-7	1,0063e-6	1,0047e-6	1,0312e-6
Błąd	Modele parametryczne					
	ARIMA(0,1,2) (Student)	ARIMA(0,1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (Student)	ARIMA(1,1,1)- GARCH(1,1) (GED)		
MSE	7,8051e-5	7,8615e-5	7,8058e-5	7,8058e-5		
MedE	2,7454e-5	2,7438e-5	2,7546e-5	2,7546e-5		
ME	0,0015387	0,0016234	0,0015409	0,0015409		
AE	0,0067188	0,0067622	0,0067197	0,0067197		
RMSE	0,0088346	0,0088665	0,0088351	0,0088351		
MAPE	0,00075963	0,00076455	0,00075973	0,00075973		
AMAPE	0,00037985	0,0003823	0,0003799	0,0003799		
LL	9,9984e-7	1,007e-6	9,9994e-7	9,9994e-7		

Tabela 6.62. Wartości błędów prognoz, otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu, za pomocą uogólnionej metody momentów oraz za pomocą parametrycznych modeli średniej warunkowej. Logarytmiczne poziomy indeksu NASDAQ.

Błąd	Metoda Phillipsa i Yu			Uogólniona metoda momentów		
	Vašiček	CIR	CKLS	Vašiček	CIR	CKLS
MSE	7,3033e-5	7,3078e-5	7,3089e-5	7,5079e-5	7,4249e-5	7,5569e-5
MedE	1,8082e-5	1,7335e-5	1,8443e-5	2,1237e-5	1,9876e-5	2,1606e-5
ME	0,055699	0,00064038	0,055435	0,0013926	0,0013442	0,0015984
AE	0,0062114	0,0062146	0,0062015	0,0064573	0,0064043	0,0065172
RMSE	0,0085459	0,0085486	0,0085492	0,0086648	0,0086168	0,008693
MAPE	0,00079535	0,00079575	7,9409e-6	0,00082676	0,00081998	0,00083443
AMAPE	0,00039763	0,00039783	3,9699e-6	0,00041338	0,00040998	0,00041722
LL	1,1983e-6	1,199e-6	1,1992e-6	1,2319e-6	1,2183e-6	1,24e-6
Modele parametryczne						
Błąd	ARIMA(0,1,2) (Student)		ARIMA(0,1,2) (GED)	ARIMA(0,1,2)- GARCH(1,1) (Student)	ARIMA(1,1,1)- GARCH(1,1) (GED)	
MSE	7,2567e-5		7,2567e-5	7,282e-5	7,282e-5	
MedE	1,7764e-5		1,7769e-5	1,8811e-5	1,8819e-5	
ME	0,056472		0,056466	0,055536	0,055532	
AE	0,0062011		0,006201	0,0062328	0,0062328	
RMSE	0,0085187		0,0085186	0,0085335	0,0085334	
MAPE	0,00079404		0,00079404	0,00079811	0,00079811	
AMAPE	0,00039697		0,00039697	0,00039901	0,00039901	
LL	1,1907e-6		1,1907e-6	1,1949e-6	1,1948e-6	

Tabela 6.63. Wartości błędów prognoz, otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu, za pomocą uogólnionej metody momentów oraz za pomocą parametrycznych modeli średniej warunkowej. Logarytmiczne poziomy indeksu WIG20.

Błąd	Metoda Phillipsa i Yu			Uogólniona metoda momentów		
	Vašiček	CIR	CKLS	Vašiček	CIR	CKLS
MSE	0,00018948	0,0001892	0,00018936	0,00018897	0,00018938	0,00018973
MedE	6,1233e-5	6,3602e-5	6,5722e-5	6,4217e-5	6,4506e-5	6,7774e-5
ME	0,00063017	0,00063866	0,058766	0,0001788	0,00032633	-0,0001169
AE	0,010608	0,010619	0,010609	0,010582	0,010595	0,010586
RMSE	0,013765	0,013755	0,013761	0,013747	0,013762	0,013774
MAPE	0,0013032	0,0013046	1,3034e-5	0,0013001	0,0013016	0,0013007
AMAPE	0,00065164	0,00065232	6,5172e-6	0,00065007	0,00065083	0,00065032
LL	2,8646e-6	2,8603e-6	2,8627e-6	2,8568e-6	2,8629e-6	2,8682e-6
Modele parametryczne						
Błąd	ARMA(0,2) (GED)			ARMA(1,1) (Student)		
MSE	0,01904131			0,019061629		
MedE	6,6946E-05			6,87936e-5		
ME	0,00035112			0,000398688		
AE	1,06889157			1,071935667		
RMSE	0,13799026			0,138063858		
MAPE	0,131329			0,1317012		
AMAPE	0,00065666			0,000658529		
LL	2,8789E-06			2,88198e-6		



Tabela 6.64. Wartości błędów prognoz, otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu, za pomocą uogólnionej metody momentów oraz za pomocą parametrycznych modeli średniej warunkowej. Logarytmiczne kursy EURPLN.

Błąd	Metoda Phillipsa i Yu			Uogólniona metoda momentów		
	Vašiček	CIR	CKLS	Vašiček	CIR	CKLS
MSE	1,5278e-5	1,5168e-5	-	1,5226e-5	1,5284e-5	1,5198e-5
MedE	4,5701e-6	4,2767e-6	-	4,4154e-6	4,2957e-6	4,7567e-6
ME	-7,1886e-6	-5,0222e-5	-	-8,6538e-6	3,8631e-5	-0,00013324
AE	0,0028836	0,0028222	-	0,0028301	0,0028265	0,0028352
RMSE	0,0039087	0,0038947	-	0,003902	0,0039095	0,0038984
MAPE	0,0021385	0,0020933	-	0,0020991	0,0020965	0,0021029
AMAPE	0,0010695	0,0010469	-	0,0010498	0,0010485	0,0010516
LL	8,4034e-6	8,3446e-6	-	8,3759e-6	8,408e-6	8,3602e-6
Modele parametryczne						
Błąd	ARIMA(0,1,1) (GED)	ARIMA(1,1,1)- (Student)	ARIMA(1,1,1)- GARCH(1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (GED)
MSE	1,4946e-5	1,495e-5	1,4948e-5	1,4948e-5	1,495e-5	1,495e-5
MedE	4,2942e-6	4,4347e-6	4,1881e-6	4,1881e-6	4,4659e-6	4,4659e-6
ME	-6,0594e-5	-6,1048e-5	-6,025e-5	-6,025e-5	-6,3148e-5	-6,3148e-5
AE	0,0028143	0,0028152	0,0028135	0,0028135	0,0028147	0,0028147
RMSE	0,003866	0,0038665	0,0038663	0,0038663	0,0038665	0,0038665
MAPE	0,0020871	0,0020877	0,0020866	0,0020866	0,0020874	0,0020874
AMAPE	0,0010438	0,0010441	0,0010436	0,0010436	0,001044	0,001044
LL	8,2212e-6	8,2229e-6	8,2223e-6	8,2223e-6	8,2228e-6	8,2228e-6

Tabela 6.65. Wartości błędów prognoz, otrzymanych za pomocą modeli dyfuzji z parametrami oszacowanymi metodą Phillipsa i Yu, za pomocą uogólnionej metody momentów oraz za pomocą parametrycznych modeli średniej warunkowej. Logarytmiczne kursy USDPLN.

Błąd	Metoda Phillipsa i Yu			Uogólniona metoda momentów		
	Vašiček	CIR	CKLS	Vašiček	CIR	CKLS
MSE	2,6047e-5	2,6108e-5	2,6014e-5	2,6067e-5	2,6135e-5	2,604e-5
MedE	9,8231e-6	1,0212e-5	9,7905e-6	9,4225e-6	8,9225e-6	1,0367e-5
ME	-0,00016563	-0,00016674	-0,00017965	4,2044e-5	0,00018846	-0,00011598
AE	0,003915	0,0039269	0,0039175	0,0038798	0,0038458	0,0039065
RMSE	0,51036	0,51096	0,51004	0,51055	0,51122	0,51029
MAPE	0,0036635	0,0036745	3,6659e-5	0,0036295	0,0035973	0,0036554
AMAPE	0,001833	0,0018385	1,8342e-5	0,0018164	0,0018006	0,0018291
LL	2,2826e-5	2,2876e-5	2,2793e-5	2,2842e-5	2,2907e-5	2,2815e-5
Modele parametryczne						
Błąd	ARIMA(0,1,1) (GED)	ARIMA(1,1,1) (Student)	ARIMA(1,1,1)- GARCH(1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (GED)	ARIMA(1,1,1)- GARCH(1,1) (GED)
MSE	2,6439e-5	2,6434e-5	2,6394e-5	2,6394e-5	2,6396e-5	2,6396e-5
MedE	1,052e-5	1,0504e-5	1,0524e-5	1,0524e-5	1,0487e-5	1,0487e-5
ME	-0,00024461	-0,00024575	-0,00024226	-0,00024226	-0,00024291	-0,00024291
AE	0,0039657	0,0039662	0,0039719	0,0039719	0,0039714	0,0039714
RMSE	0,51418	0,51414	0,51375	0,51375	0,51377	0,51377
MAPE	0,0037118	0,0037124	0,003718	0,003718	0,0037176	0,0037176
AMAPE	0,001857	0,0018573	0,0018601	0,0018601	0,0018599	0,0018599
LL	2,3178e-5	2,3174e-5	2,3146e-5	2,3146e-5	2,3148e-5	2,3148e-5

Przedstawione w tabelach 6.60-6.65 wartości błędów są zbliżone. Można jednak zaobserwować, że prognozy z modeli oszacowanych uogólnioną metodą momentów są istotnie gorsze dla procesu logarytmicznych poziomów indeksu NASDAQ, natomiast w mniejszym stopniu dla procesu logarytmicznych poziomów indeksu DAX oraz kursów logarytmicznych EURPLN. W przypadku indeksu CAC40, obserwujemy nieznacznie większe wartości błędów dla modeli oszacowanych metodą Phillipsa i Yu.

Błędy prognoz otrzymanych za pomocą modelu Vašíčka z parametrami oszacowanymi metodą Phillipsa i Yu są zwykle nieco mniejsze niż w przypadku prognoz otrzymanych za pomocą dwóch pozostałych modeli dyfuzji. Dla procesów logarytmicznych poziomów indeksów CAC40 i DAX, najlepszą jakość prognoz otrzymaliśmy za pomocą modelu Coxa, Ingersolla i Rossa. Dla pozostałych dwóch indeksów najlepsze prognozy otrzymujemy za pomocą modelu stałej elastyczności wariancji. Dla kursu USDPLN, modele Coxa-Ingersolla i Rossa oraz CKLS dają zbliżone prognozy. W przypadku modeli z parametrami oszacowanymi metodą momentów nie jesteśmy w stanie określić, zastosowanie którego z modeli prowadzi do otrzymania w rzeczywistości najlepszych prognoz. Można zauważyć, że błędy prognoz są nieco większe dla modelu CKLS, którego wartości oszacowań parametrów nie są zgodne ze specyfikacją modelu, chociaż dla procesu kursów logarytmicznych EURPLN jest odwrotnie.

Wartości błędów prognoz otrzymanych za pomocą modeli ARIMA-GARCH są nieco większe od uzyskanych za pomocą modeli dyfuzji. Jedynie dla szeregu EURPLN sytuacja wygląda przeciwnie. Dobra jakość prognoz otrzymanych za pomocą modeli dyfuzji może sugerować dobre dopasowanie modeli dyfuzji do badanych szeregów czasowych. Godny uwagi jest fakt, że w modelach dyfuzji zmienność uzależniona jest jedynie od czynnika losowego lub dodatkowo od bieżącej wartości procesu. W modelach ARIMA-GARCH zmienność jest opisana oddzielnym równaniem parametrycznym.

## **6.5. Podsumowanie**

Przeprowadzone w tym rozdziale badania sugerują, że metody szybkiej estymacji modeli dyfuzji zastosowane do szeregów logarytmicznych cen instrumentów finansowych pozwalają na ich szybkie i dokładne modelowanie i prognozowanie. W szczególności potwierdzają to estymacje przeprowadzone za pomocą metody Phillipsa i Yu. Jakość oszacowań parametrów uzyskana dla danych intraday jest znacznie lepsza niż dla da-

nych dziennych. Oszacowania parametrów uzyskane tą metodą są zbliżone bez względu na wykorzystaną częstotliwość danych śróddziennych, w zakresie stosowanym w tej pracy. Badania symulacyjne wskazują na ogół na mały błąd standardowy. Ponadto jakość prognoz średniej warunkowej otrzymanych za pomocą tej metody jest zbliżona lub nawet lepsza niż w powszechnie stosowanych modelach ARIMA-GARCH. Należy przy tym nadmienić, że sam proces estymacji modeli dyfuzji trwa zaledwie ułamek sekundy. Estymacja parametrów za pomocą bardziej klasycznej uogólnionej metody momentów trwa również bardzo krótko, niemniej jakość oszacowań w tym przypadku nie jest aż tak dobra. Oszacowania parametrów dla różnych częstotliwości danych przyjmują tutaj dość zróżnicowane wartości. Błędy standardowe oszacowań parametrów często pozostają dosyć duże. Ponadto, w wielu przypadkach oszacowania parametrów nie są zgodne ze specyfikacją danego modelu dyfuzji. Mimo słabych stron tej metody estymacji, jakość prognoz średniej warunkowej pozostaje na poziomie zbliżonym do jakości prognoz otrzymanych za pomocą modeli estymowanych metodą Phillipsa i Yu.

## **7. Testowanie skoków w procesach generujących notowania indeksów giełdowych i kursów walutowych**

W niniejszym rozdziale badamy występowanie skoków w procesach cen logarytmicznych z rynku kapitałowego i walutowego. Wykorzystujemy w tym celu testy Barndorff-Nielsen i Shepharda (2006) oraz Jianga i Oomena (2008). Występowanie skoków będziemy badać niezależnie we wszystkich okresach dziennych, natomiast do wyznaczenia estymatorów takich wielkości, jak wariancja scałkowana, kwartyczość czy sekstyczość, wykorzystywanych do konstrukcji statystyk testowych, posłużą nam notowania śróddzienne. Do weryfikacji występowania skoków stosujemy wszystkie trzy podejścia opisane w rozdziale 4. Wykorzystujemy przy tym różną częstotliwość danych. Ponadto posługujemy się statystyką testową z przesuniętymi wariancjami wielopotęgowymi w celu zminimalizowania efektów mikrostruktury, badając uprzednio autokorelację w szeregach danych. Występowanie skoków weryfikujemy również za pomocą nieparametrycznych testów Lee i Myklanda. Jako danych empirycznych używamy szeregów logarytmicznych notowań śróddziennych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 z okresu 02.01.2004 – 29.12.2006 (odpowiednio 771, 769, 755, 757 dni) oraz notowań śróddziennych logarytmicznych kursów walutowych EURPLN i USDPLN z tego samego okresu (odpowiednio 772 i 768 dni). Do wyznaczenia dziennej wariancji zrealizowanej i wariacji wielopotęgowych, wykorzystywanych do konstrukcji statystyk Barndorff-Nielsen i Shepharda oraz Jianga i Oomena na występowanie skoków, stosujemy, odpowiednio, w przypadku indeksów CAC40, DAX i NASDAQ zwroty pięcio- i dziesięciominutowe oraz półgodzinne, w przypadku indeksu WIG20, zwroty pięcio-, dziesięcio-, dwudziesto- i czterdziestominutowe, natomiast w przypadku kursów walutowych, zwroty dziesięciominutowe półgodzinne i godzinne. Za pomocą testów Barndorff-Nielsen i Shepharda oraz Jianga i Oomena testujemy występowanie skoków poszczególnych dniach. Testy Lee i Myklanda wykorzystywane są natomiast do wykrywania skoków niezależnie w każdym okresie intraday. Statystyki opisowe notowań i zwrotów logarytmicznych badanych szeregów przedstawione są w tabelach 7.1 i 7.2.

**Tabela 7.1. Podstawowe charakterystyki szeregów logarytmicznych notowań indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursów EURPLN i USDPLN.**

Szereg Notowań	liczba obs.	średnia	odch. std.	skośność	kurtoza	min.	maks.
CAC40 5 min.	78642	8,3678	0,1403	0,1905	1,5756	8,1475	8,6221
CAC40 10 min.	39321	8,3678	0,0197	0,1905	1,5756	8,1483	8,6221
CAC40 30 min.	13107	8,3678	0,1403	0,1905	1,5756	8,1483	8,6221
DAX 5 min.	78438	8,4710	0,1699	0,2508	1,6340	8,1943	8,7987
DAX 10 min.	39219	8,4710	0,1699	0,2508	1,6340	8,1943	8,7982
DAX 30 min.	13073	8,4711	0,1699	0,2507	1,6340	8,1952	8,7981
NASDAQ 5 min.	58890	7,6549	0,0708	0,0413	2,4999	7,4682	7,8117
NASDAQ 10 min.	29445	7,6549	0,0708	0,0412	2,4999	7,4682	7,8116
NASDAQ 30 min.	9815	7,6549	0,0709	0,0420	2,4986	7,4685	7,8116
WIG20 5 min.	57010	7,7233	0,2360	0,1767	1,4785	7,3633	8,1404
WIG20 10 min.	28506	7,7233	0,2360	0,1767	1,4785	7,3635	8,1375
WIG20 20 min.	14252	7,7233	0,2360	0,1766	1,4786	7,3650	8,1403
WIG20 40 min.	7126	7,7233	0,2360	0,1765	1,4783	7,3673	8,1403
EURPLN 10 min.	45324	1,4318	0,0803	0,6103	1,9500	1,3219	1,5967
EURPLN 30 min.	15108	1,4318	0,0803	0,6105	1,9504	1,3223	1,5944
EURPLN 60 min.	7054	1,4381	0,0795	0,5378	1,8601	1,3223	1,5943
USDPLN 10 min.	41472	1,1988	0,0850	0,6752	2,4777	1,0481	1,4021
USDPLN 30 min.	13824	1,1988	0,0849	0,6755	2,4783	1,0481	1,4017
USDPLN 60 min.	6912	1,1988	0,0849	0,6756	2,4786	1,0490	1,4002

**Tabela 7.2. Podstawowe charakterystyki zwrotów logarytmicznych indeksów CAC40, DAX, NASDAQ i WIG20 oraz kursów EURPLN i USDPLN.**

Szereg zwrotów	liczba obs.	średnia	odch. std.	skośność	kurtoza	min.	maks.
CAC40 5 min.	78642	0,0003	0,0651	-0,0837	14,717	-1,2192	1,1110
CAC40 10 min.	39321	0,05	0,0085	-0,1936	12,468	-1,6220	1,1127
CAC40 30 min.	13107	0,0016	0,1581	-0,5006	13,361	-2,1734	1,5818
DAX 5 min.	78438	0,0003	0,0778	-0,6293	39,174	-2,1689	1,7562
DAX 10 min.	39219	0,0007	0,1086	-0,7345	27,252	-2,2217	1,7285
DAX 30 min.	13073	0,0020	0,1858	-0,7233	14,243	-2,1403	1,6595
NASDAQ 5 min.	58890	-0,0002	0,0788	0,0411	6,6514	-0,8481	0,7394
NASDAQ 10 min.	29445	-0,0006	0,1149	0,0142	5,6694	-0,8602	0,7802
NASDAQ 30 min.	9815	-0,0017	0,2027	-0,0998	5,0834	-1,1064	1,1994
WIG20 5 min.	57009	0,0007	0,1301	-0,0770	8,9422	-1,6520	1,3182
WIG20 10 min.	28505	0,0009	0,1813	-0,1262	7,9641	-1,7970	1,6185
WIG20 20 min.	14251	0,0046	0,2930	-0,1824	23,570	-5,1875	4,6604
WIG20 40 min.	7125	0,0044	0,3254	-0,3173	7,9470	-2,6670	2,4845
EURPLN 10 min.	45323	-0,05	0,0796	0,4231	20,3282	-1,2877	1,4939
EURPLN 30 min.	15107	-0,0042	0,1682	0,0763	6,1152	-1,2446	1,0276
EURPLN 60 min.	7053	-0,0004	0,1301	0,4796	12,0013	-1,2027	1,5873
USDPLN 10 min.	41471	-0,0004	0,0908	-0,0405	11,3162	-1,1092	1,0199
USDPLN 30 min.	13823	-0,0042	0,1682	0,0763	6,1152	-1,2446	1,0276
USDPLN 60 min.	6911	-0,0012	0,1570	0,0522	7,9877	-1,4873	1,2129

Tabela 7.3 zawiera średnie i wariancje szeregów dziennych wariancji zrealizowanych (RV), przemnożonych przez  $\mu_1^{-2}$  wariacji dwupotęgowych (BV), przeskalowanych wariacji dwupotęgowych (BV\*), przemnożonych przez  $\mu_{4/3}^{-3}$  wariacji trójpotęgowych

(TQ), przeskalowanych wariacji oraz trójpotęgowych (TQ\*), obliczone dla szeregów notowań sróddziennych indeksów CAC40, DAX, NASDAQ i WIG20 różnych częstotliwości oraz kursów EURPLN i USDPLN. Jak łatwo zauważyć, wraz ze spadkiem częstotliwości średnie maleją, natomiast wariancje rosną.

**Tabela 7.3. Średnie ( $\mu$ ) i wariancje ( $\sigma^2$ ) wariacji zrealizowanej i zrealizowanych wariacji wielopotęgowych wyznaczonych na podstawie szeregów notowań intraday indeksów CAC40, DAX, NASDAQ i WIG20 oraz kursów intraday EURPLN i USDPLN różnych częstotliwości.**

Szereg notowań		RV	BV	BV*	TQ	TQ*
CAC40 5 min.	$\mu$	0,4147	0,3805	0,3843	0,6483	0,6612
	$\sigma^2$	0,1847	0,1692	0,1726	16,302	16,961
CAC40 10 min.	$\mu$	0,4338	0,3841	0,3918	0,5043	0,5249
	$\sigma^2$	0,1550	0,1476	0,1536	13,868	15,023
CAC40 30 min.	$\mu$	0,4281	0,3386	0,3598	0,2123	0,2407
	$\sigma^2$	0,2231	0,1202	0,1357	0,7365	0,9460
DAX 5 min.	$\mu$	0,6236	0,5036	0,5086	0,6632	0,6765
	$\sigma^2$	0,3253	0,1570	0,1601	3,8940	4,0513
DAX 10 min.	$\mu$	0,6082	0,4803	0,4899	0,5670	0,5902
	$\sigma^2$	0,3583	0,1529	0,1590	8,1341	8,8117
DAX 30 min.	$\mu$	0,5938	0,4352	0,4624	0,3353	0,3800
	$\sigma^2$	0,3962	0,1645	0,1857	0,9253	1,1885
NASDAQ 5 min.	$\mu$	0,4838	0,4375	0,4431	0,3476	0,3567
	$\sigma^2$	0,0871	0,0775	0,0796	0,4501	0,4741
NASDAQ 10 min.	$\mu$	0,5150	0,4619	0,4741	0,3420	0,3605
	$\sigma^2$	0,1041	0,0928	0,0978	0,3570	0,3967
NASDAQ 30 min.	$\mu$	0,5343	0,4449	0,4820	0,2622	0,3098
	$\sigma^2$	0,1450	0,1246	0,1462	0,4317	0,6029
WIG20 5 min.	$\mu$	1,2752	1,1457	1,1609	3,0644	3,1454
	$\sigma^2$	1,0817	0,8882	0,9108	104,84	110,29
WIG20 10 min.	$\mu$	1,2380	1,0888	1,1181	2,7487	2,8972
	$\sigma^2$	1,2607	1,0831	1,1394	172,79	191,47
WIG20 20 min.	$\mu$	1,1494	0,9622	1,0155	1,7996	2,55
	$\sigma^2$	1,1710	0,9411	1,0435	44,957	55,570
WIG20 40 min.	$\mu$	0,9951	0,7198	0,8041	0,7861	0,9899
	$\sigma^2$	1,0940	0,5731	0,7097	7,1329	11,181
EURPLN 10 min.	$\mu$	0,3321	0,2896	0,2946	0,2098	0,2160
	$\sigma^2$	0,0757	0,0479	0,0492	0,3467	0,3625
EURPLN 30 min.	$\mu$	0,3317	0,2581	0,2720	0,1274	0,1404
	$\sigma^2$	0,1056	0,0503	0,0553	0,1729	0,2087
EURPLN 60 min.	$\mu$	0,2772	0,2182	0,2428	0,0874	0,1046
	$\sigma^2$	0,0743	0,0567	0,0673	0,1489	0,1899
USDPLN 10 min.	$\mu$	0,4449	0,3828	0,3901	0,2599	0,2699
	$\sigma^2$	0,0765	0,0512	0,0532	0,2128	0,2295
USDPLN 30 min.	$\mu$	0,4437	0,3722	0,3941	0,2152	0,2421
	$\sigma^2$	0,1104	0,0745	0,0835	0,2070	0,2620
USDPLN 60 min.	$\mu$	0,4457	0,3438	0,3867	0,1492	0,1918
	$\sigma^2$	0,1441	0,0881	0,1115	0,1284	0,2122

## 7.1. Omówienie wyników empirycznych

W tabelach 7.4-7.6 przedstawiona jest liczba dni ze skokami wykrytych za pomocą statystyk Barndorff-Nielsen i Shepharda (2006) dla poziomów ufności 95%, 99%, 99,5% oraz 99,9%, w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20, w zależności od zastosowanego estymatora zmienności skalkowanej i kwartyczności.

We wszystkich przypadkach liczba wykrytych dni ze skokami jest znacznie większa w sytuacji zastosowania statystyki liniowej. Jest to empiryczne potwierdzenie wniosku postawionego przez Barndorff-Nielsen i Shepharda (2006) oraz Huanga Tau-chena (2005), mówiącego o zbyt częstym odrzuceniu hipotezy zerowej dla statystyki liniowej. Potwierdzają to również badania symulacyjne przeprowadzone w rozdziale 4, z których wynika, że rozmiar testu liniowego jest większy niż w przypadku innych statystyk testowych. Jednocześnie można zauważyć, że statystyka ilorazowa odrzuca hipotezę zerową znacznie rzadziej niż dwie pozostałe statystyki. Również stosowanie nieprzeskalowanych wariacji wielopotęgowych jako estymatorów zmienności skalkowanej i kwartyczności powoduje przeszacowanie liczby wykrytych skoków. Jest to następstwem sztucznego zaniżenia oszacowania tych wielkości, czego konsekwencją jest zawyżenie różnicy pomiędzy wariancją zrealizowaną a oszacowaniem zmienności skalkowanej lub logarytmami tych wielkości, na których oparte są statystyki testowe. Wniosek ten jest również potwierdzony przez badania symulacyjne w rozdziale 4. Rozmiar testu jest szczególnie duży dla statystyki liniowej. Nawet dla małej częstotliwości danych może on przekraczać 50%. Z kolei dla dwóch pozostałych statystyk i 120 minutowych częstotliwości danych wynosi on około 25%. Przeprowadzone badanie sugeruje, że najbardziej wiarygodne wyniki otrzymujemy stosując statystykę logarytmiczną z przeskalowanymi wariacjami wielopotęgowymi. Porównując zamieszczone w tabelach 7.5 i 7.6 charakterystyki zrealizowanych przeskalowanych wariacji dwu- oraz trójpotęgowych wyznaczonych dla różnych częstotliwości danych, można zauważyć, że ich wartości są największe, jeśli zastosujemy zwroty pięciominutowe. Może to być wynikiem zanieczyszczenia estymacji szumem spowodowanym autokorelacją sąsiednich zwrotów. Efekt ten uwidacznia się najsilniej w okresach dużej zmienności. Stanowiło to motywację do podjęcia próby testowania skoków z wykorzystaniem przedstawionych w rozdziale 4 przesuniętych wariacji wielopotęgowych, czego dokonamy w kolejnym podrozdziale. Łatwo również zauważyć, że stosowanie zbyt wysokich poziomów ufności

może prowadzić do zaniżenia liczby dni ze skokami, wobec czego 95% poziom ufności może wydawać się najbardziej właściwy.

**Tabela 7.4. Liczby dni ze skokami wykrytych za pomocą statystyki liniowej Barndorff-Nielsen i Shepharda w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursów walutowych EURPLN i USDPLN, w zależności od poziomu ufności, zastosowanych estymatorów zmienności skalkowanej i kwartyczości oraz częstotliwości danych.**

Szereg notowań	Estymatory zm. skalk. i kwartyczości	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	BV, TQ	223	156	128	92	29,00%	20,29%	16,64%	11,96%
	BV*, TQ*	213	138	119	88	27,70%	17,95%	15,47%	11,44%
CAC40 10 min.	BV, TQ	260	180	158	125	33,81%	23,41%	20,55%	16,25%
	BV*, TQ*	227	160	142	113	29,52%	20,81%	18,47%	14,69%
CAC40 30 min.	BV, TQ	306	215	196	168	39,79%	27,96%	25,49%	21,85%
	BV*, TQ*	244	179	168	143	31,73%	23,28%	21,85%	18,60%
DAX 5 min.	BV, TQ	374	272	248	201	48,51%	35,28%	32,17%	26,07%
	BV*, TQ*	356	257	237	192	46,17%	33,33%	30,74%	24,90%
DAX 10 min.	BV, TQ	327	231	207	174	42,41%	29,96%	26,85%	22,57%
	BV*, TQ*	288	208	193	161	37,35%	26,98%	25,03%	20,88%
DAX 30 min.	BV, TQ	346	279	249	209	44,88%	36,19%	32,30%	27,11%
	BV*, TQ*	299	224	203	173	38,78%	29,05%	26,33%	22,44%
NASDAQ 5 min.	BV, TQ	249	169	147	101	32,13%	21,81%	18,97%	13,03%
	BV*, TQ*	228	151	139	92	29,42%	19,48%	17,94%	11,87%
NASDAQ 10 min.	BV, TQ	237	151	129	95	30,58%	19,48%	16,65%	12,26%
	BV*, TQ*	203	126	109	76	26,19%	16,26%	14,06%	9,81%
NASDAQ 30 min.	BV, TQ	287	202	174	142	37,03%	26,06%	22,45%	18,32%
	BV*, TQ*	207	150	133	107	26,71%	19,35%	17,16%	13,81%
WIG20 5 min.	BV, TQ	265	171	142	105	35,01%	22,59%	18,76%	13,87%
	BV*, TQ*	232	149	130	92	30,65%	19,68%	17,17%	12,15%
WIG20 10 min.	BV, TQ	256	166	140	109	33,82%	21,93%	18,49%	14,40%
	BV*, TQ*	225	137	122	91	29,72%	18,10%	16,12%	12,02%
WIG20 20 min.	BV, TQ	269	203	188	158	35,54%	26,82%	24,83%	20,87%
	BV*, TQ*	217	169	156	134	28,67%	22,32%	20,61%	17,70%
WIG20 40 min.	BV, TQ	351	290	269	236	46,37%	38,31%	35,54%	31,18%
	BV*, TQ*	274	225	208	178	36,20%	29,72%	27,48%	23,51%
EURPLN 10 min.	BV, TQ	259	174	148	119	33,59%	22,57%	19,20%	15,43%
	BV*, TQ*	235	150	134	105	30,48%	19,46%	17,38%	13,62%
EURPLN 30 min.	BV, TQ	317	242	217	178	41,12%	31,39%	28,15%	23,09%
	BV*, TQ*	259	200	180	145	33,59%	25,94%	23,35%	18,81%
EURPLN 60 min.	BV, TQ	324	271	250	193	42,02%	35,15%	32,43%	25,03%
	BV*, TQ*	254	189	170	134	32,94%	24,51%	22,05%	17,38%
USDPLN 10 min.	BV, TQ	277	197	167	132	36,07%	25,65%	21,74%	17,19%
	BV*, TQ*	249	170	150	115	32,42%	22,14%	19,53%	14,97%
USDPLN 30 min.	BV, TQ	257	191	171	144	33,46%	24,87%	22,27%	18,75%
	BV*, TQ*	206	156	144	110	26,82%	20,31%	18,75%	14,32%
USDPLN 60 min.	BV, TQ	303	234	215	189	39,45%	30,47%	27,99%	24,61%
	BV*, TQ*	219	178	162	138	28,52%	23,18%	21,09%	17,97%



**Tabela 7.5. Liczby dni ze skokami wykrytych za pomocą statystyki logarytmicznej Barndorff-Nielsen i Shepharda w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursów walutowych EURPLN i USDPLN, w zależności od poziomu ufności, zastosowanych estymatorów zmienności scalkowanej i kwartyczności oraz częstotliwości danych.**

Szereg notowań	Estymatory zm. scalk. i kwartyczności	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	BV, TQ	214	126	108	73	29,00%	20,29%	16,64%	11,96%
	BV*, TQ*	192	119	97	64	27,70%	17,95%	15,47%	11,44%
CAC40 10 min.	BV, TQ	232	147	133	93	33,81%	23,41%	20,55%	16,25%
	BV*, TQ*	208	138	118	80	29,52%	20,81%	18,47%	14,69%
CAC40 30 min.	BV, TQ	268	181	161	124	39,79%	27,96%	25,49%	21,85%
	BV*, TQ*	208	150	134	107	31,73%	23,28%	21,85%	18,60%
DAX 5 min.	BV, TQ	357	245	218	173	48,51%	35,28%	32,17%	26,07%
	BV*, TQ*	329	236	207	159	46,17%	33,33%	30,74%	24,90%
DAX 10 min.	BV, TQ	293	201	181	137	42,41%	29,96%	26,85%	22,57%
	BV*, TQ*	260	187	167	123	37,35%	26,98%	25,03%	20,88%
DAX 30 min.	BV, TQ	324	228	196	156	44,88%	36,19%	32,30%	27,11%
	BV*, TQ*	268	183	163	131	38,78%	29,05%	26,33%	22,44%
NASDAQ 5 min.	BV, TQ	228	147	115	74	32,13%	21,81%	18,97%	13,03%
	BV*, TQ*	210	133	101	62	29,42%	19,48%	17,94%	11,87%
NASDAQ 10 min.	BV, TQ	212	117	102	65	30,58%	19,48%	16,65%	12,26%
	BV*, TQ*	177	106	80	58	26,19%	16,26%	14,06%	9,81%
NASDAQ 30 min.	BV, TQ	244	156	135	100	37,03%	26,06%	22,45%	18,32%
	BV*, TQ*	176	115	100	70	26,71%	19,35%	17,16%	13,81%
WIG20 5 min.	BV, TQ	242	140	120	66	31,97%	18,49%	15,85%	8,72%
	BV*, TQ*	218	128	104	60	28,80%	16,91%	13,74%	7,93%
WIG20 10 min.	BV, TQ	234	131	110	77	30,91%	17,31%	14,53%	10,17%
	BV*, TQ*	190	112	93	68	25,10%	14,80%	12,29%	8,98%
WIG20 20 min.	BV, TQ	244	165	153	113	32,23%	21,80%	20,21%	14,93%
	BV*, TQ*	200	143	129	91	26,42%	18,89%	17,04%	12,02%
WIG20 40 min.	BV, TQ	317	243	222	171	41,88%	32,10%	29,33%	22,59%
	BV*, TQ*	246	178	161	128	32,50%	23,51%	21,27%	16,91%
EURPLN 10 min.	BV, TQ	235	143	124	93	30,48%	18,55%	16,08%	12,06%
	BV*, TQ*	219	129	113	74	28,40%	16,73%	14,66%	9,60%
EURPLN 30 min.	BV, TQ	292	196	167	131	37,87%	25,42%	21,66%	16,99%
	BV*, TQ*	239	158	137	113	31,00%	20,49%	17,77%	14,66%
EURPLN 60 min.	BV, TQ	299	208	181	136	38,78%	26,98%	23,48%	17,64%
	BV*, TQ*	216	148	119	86	28,02%	19,20%	15,43%	11,15%
USDPLN 10 min.	BV, TQ	253	171	154	115	32,94%	22,27%	20,05%	14,97%
	BV*, TQ*	207	149	124	99	26,95%	19,40%	16,15%	12,89%
USDPLN 30 min.	BV, TQ	230	156	135	100	29,95%	20,31%	17,58%	13,02%
	BV*, TQ*	184	124	106	72	23,96%	16,15%	13,80%	9,38%
USDPLN 60 min.	BV, TQ	265	200	172	134	34,51%	26,04%	22,40%	17,45%
	BV*, TQ*	202	141	124	92	26,30%	18,36%	16,15%	11,98%

**Tabela 7.6. Liczby dni ze skokami wykrytych za pomocą statystyki ilorazowej Barndorff-Nielsen a i Shepharda w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursów walutowych EURPLN i USDPLN, w zależności od poziomu ufności, zastosowanych estymatorów zmienności scalkowanej i kwartyczości oraz częstotliwości danych.**

Szereg notowań	Estymatory zm. scalk. i kwartyczości	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	BV, TQ	186	99	81	52	24,19%	12,87%	10,53%	6,76%
	BV*, TQ*	176	90	75	46	22,89%	11,70%	9,75%	5,98%
CAC40 10 min.	BV, TQ	196	107	76	41	25,49%	13,91%	9,88%	5,33%
	BV*, TQ*	178	94	66	35	23,15%	12,22%	8,58%	4,55%
CAC40 30 min.	BV, TQ	179	81	58	30	23,28%	10,53%	7,54%	3,90%
	BV*, TQ*	150	65	46	24	19,51%	8,45%	5,98%	3,12%
DAX 5 min.	BV, TQ	321	210	179	117	41,63%	27,24%	23,22%	15,18%
	BV*, TQ*	300	201	169	112	38,91%	26,07%	21,92%	14,53%
DAX 10 min.	BV, TQ	246	159	124	79	31,91%	20,62%	16,08%	10,25%
	BV*, TQ*	223	140	109	74	28,92%	18,16%	14,14%	9,60%
DAX 30 min.	BV, TQ	222	118	86	35	28,79%	15,30%	11,15%	4,54%
	BV*, TQ*	182	97	70	30	23,61%	12,58%	9,08%	3,89%
NASDAQ 5 min.	BV, TQ	201	103	74	36	25,94%	13,29%	9,55%	4,65%
	BV*, TQ*	185	89	62	29	23,87%	11,48%	8,00%	3,74%
NASDAQ 10 min.	BV, TQ	160	67	50	25	20,65%	8,65%	6,45%	3,23%
	BV*, TQ*	134	57	43	21	17,29%	7,35%	5,55%	2,71%
NASDAQ 30 min.	BV, TQ	126	43	29	5	16,26%	5,55%	3,74%	0,65%
	BV*, TQ*	91	32	19	5	11,74%	4,13%	2,45%	0,65%
WIG20 5 min.	BV, TQ	208	99	71	38	27,48%	13,08%	9,38%	5,02%
	BV*, TQ*	190	89	60	34	25,10%	11,76%	7,93%	4,49%
WIG20 10 min.	BV, TQ	173	80	60	23	22,85%	10,57%	7,93%	3,04%
	BV*, TQ*	146	69	53	19	19,29%	9,11%	7,00%	2,51%
WIG20 20 min.	BV, TQ	169	83	54	22	22,32%	10,96%	7,13%	2,91%
	BV*, TQ*	138	65	44	15	18,23%	8,59%	5,81%	1,98%
WIG20 40 min.	BV, TQ	185	69	41	7	24,44%	9,11%	5,42%	0,92%
	BV*, TQ*	140	52	26	6	18,49%	6,87%	3,43%	0,79%
EURPLN 10 min.	BV, TQ	201	105	81	40	26,07%	13,62%	10,51%	5,19%
	BV*, TQ*	174	92	68	35	22,57%	11,93%	8,82%	4,54%
EURPLN 30 min.	BV, TQ	200	89	63	35	25,94%	11,54%	8,17%	4,54%
	BV*, TQ*	167	75	57	25	21,66%	9,73%	7,39%	3,24%
EURPLN 60 min.	BV, TQ	152	51	24	1	19,71%	6,61%	3,11%	0,13%
	BV*, TQ*	104	28	13	1	13,49%	3,63%	1,69%	0,13%
USDPLN 10 min.	BV, TQ	177	83	69	39	23,05%	10,81%	8,98%	5,08%
	BV*, TQ*	146	73	59	33	19,01%	9,51%	7,68%	4,30%
USDPLN 30 min.	BV, TQ	161	69	47	16	20,96%	8,98%	6,12%	2,08%
	BV*, TQ*	126	54	37	11	16,41%	7,03%	4,82%	1,43%
USDPLN 60 min.	BV, TQ	138	39	26	1	17,97%	5,08%	3,39%	0,13%
	BV*, TQ*	99	30	15	1	12,89%	3,91%	1,95%	0,13%

W tabelach 7.7-7.9 przedstawiona jest liczba skoków wykrytych za pomocą statystyk testowych Jianga i Oomena (2008) w procesach cen logarytmicznych indeksów CAC40, DAX, NASDAQ i WIG20.

**Tabela 7.7. Liczby dni ze skokami wykrytych za pomocą statystyki liniowej Jianga i Oomena w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursach EURPLN i USDPLN, w zależności od poziomu ufności i częstotliwości danych.**

Szereg notowań	Estymatory sekstyczności	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	SS	287	219	200	166	37,32%	28,48%	26,01%	21,59%
CAC40 10 min.	SS	311	235	218	171	40,44%	30,56%	28,35%	22,24%
CAC40 30 min.	SS	396	328	309	279	51,50%	42,65%	40,18%	36,28%
DAX 5 min.	SS	382	307	284	244	49,55%	39,82%	36,84%	31,65%
DAX 10 min.	SS	365	303	284	254	47,34%	39,30%	36,84%	32,94%
DAX 30 min.	SS	445	377	355	315	57,72%	48,90%	46,04%	40,86%
NASDAQ 5 min.	SS	274	194	167	121	35,35%	25,03%	21,55%	15,61%
NASDAQ 10 min.	SS	268	212	193	163	34,58%	27,35%	24,90%	21,03%
NASDAQ 30 min.	SS	354	300	280	246	45,68%	38,71%	36,13%	31,74%
WIG20 5 min.	SS	268	208	182	151	35,40%	27,48%	24,04%	19,95%
WIG20 10 min.	SS	325	249	225	195	42,93%	32,89%	29,72%	25,76%
WIG20 20 min.	SS	346	292	279	248	45,71%	38,57%	36,86%	32,76%
WIG20 40 min.	SS	419	365	340	311	55,35%	48,22%	44,91%	41,08%
EURPLN 10 min.	SS	301	224	199	164	39,04%	29,05%	25,81%	21,27%
EURPLN 30 min.	SS	398	337	318	278	51,62%	43,71%	41,25%	36,06%
EURPLN 60 min.	SS	353	297	281	247	45,78%	38,52%	36,45%	32,04%
USDPLN 10 min.	SS	331	253	226	186	43,10%	32,94%	29,43%	24,22%
USDPLN 30 min.	SS	323	258	235	200	42,06%	33,59%	30,60%	26,04%
USDPLN 60 min.	SS	346	292	266	212	45,05%	38,02%	34,64%	27,60%

**Tabela 7.8. Liczby dni ze skokami wykrytych za pomocą statystyki logarytmicznej Jianga i Oomea w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursach EURPLN i USDPLN, w zależności od poziomu ufności i częstotliwości danych.**

Szereg notowań	Estymatory zm. scalk. i sekstycznychności	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	BV, SS	258	189	173	121	33,55%	24,58%	22,50%	15,73%
	BV*, SS	261	193	175	125	33,94%	25,10%	22,76%	16,25%
CAC40 10 min.	BV, SS	266	190	159	124	34,59%	24,71%	20,68%	16,12%
	BV*, SS	275	196	167	127	35,76%	25,49%	21,72%	16,51%
CAC40 30 min.	BV, SS	355	269	249	213	46,16%	34,98%	32,38%	27,70%
	BV*, SS	372	290	264	226	48,37%	37,71%	34,33%	29,39%
DAX 5 min.	BV, SS	351	268	239	208	45,53%	34,76%	31,00%	26,98%
	BV*, SS	352	269	246	211	45,65%	34,89%	31,91%	27,37%
DAX 10 min.	BV, SS	335	258	237	192	43,45%	33,46%	30,74%	24,90%
	BV*, SS	340	263	242	196	44,10%	34,11%	31,39%	25,42%
DAX 30 min.	BV, SS	392	308	294	252	50,84%	39,95%	38,13%	32,68%
	BV*, SS	408	326	302	269	52,92%	42,28%	39,17%	34,89%
NASDAQ 5 min.	BV, SS	233	155	127	83	30,06%	20,00%	16,39%	10,71%
	BV*, SS	234	158	132	86	30,19%	20,39%	17,03%	11,10%
NASDAQ 10 min.	BV, SS	239	181	155	113	30,84%	23,35%	20,00%	14,58%
	BV*, SS	247	185	166	119	31,87%	23,87%	21,42%	15,35%
NASDAQ 30 min.	BV, SS	309	230	219	184	39,87%	29,68%	28,26%	23,74%
	BV*, SS	337	246	228	201	43,48%	31,74%	29,42%	25,94%
WIG20 5 min.	BV, SS	242	176	159	120	31,97%	23,25%	21,00%	15,85%
	BV*, SS	245	180	162	122	32,36%	23,78%	21,40%	16,12%
WIG20 10 min.	BV, SS	291	203	182	144	38,44%	26,82%	24,04%	19,02%
	BV*, SS	298	214	184	148	39,37%	28,27%	24,31%	19,55%
WIG20 20 min.	BV, SS	314	239	223	183	41,48%	31,57%	29,46%	24,17%
	BV*, SS	319	253	230	197	42,14%	33,42%	30,38%	26,02%
WIG20 40 min.	BV, SS	351	278	266	225	46,37%	36,72%	35,14%	29,72%
	BV*, SS	373	302	281	255	49,27%	39,89%	37,12%	33,69%
EURPLN 10 min.	BV, SS	253	189	161	128	32,81%	24,51%	20,88%	16,60%
	BV*, SS	257	191	165	130	33,33%	24,77%	21,40%	16,86%
EURPLN 30 min.	BV, SS	398	337	318	278	45,14%	35,93%	32,94%	27,63%
	BV*, SS	367	286	267	229	47,60%	37,09%	34,63%	29,70%
EURPLN 60 min.	BV, SS	293	232	211	166	38,00%	30,09%	27,37%	21,53%
	BV*, SS	319	252	235	201	41,37%	32,68%	30,48%	26,07%
USDPLN 10 min.	BV, SS	296	202	180	140	38,54%	26,30%	23,44%	18,23%
	BV*, SS	304	210	187	144	39,58%	27,34%	24,35%	18,75%
USDPLN 30 min.	BV, SS	283	203	184	139	36,85%	26,43%	23,96%	18,10%
	BV*, SS	300	218	196	155	39,06%	28,39%	25,52%	20,18%
USDPLN 60 min.	BV, SS	306	202	180	138	39,84%	26,30%	23,44%	17,97%
	BV*, SS	318	228	194	150	41,41%	29,69%	25,26%	19,53%

**Tabela 7.9. Liczby dni ze skokami wykrytych za pomocą statystyki ilorazowej Jianga i Oomena w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursach EURPLN i USDPLN, w zależności od poziomu ufności i częstotliwości danych.**

Szereg notowań	Estymatory zm. scałk. i sekstyczności	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	BV, SS	258	189	173	121	33,55%	24,58%	22,50%	15,73%
	BV*, SS	260	194	175	125	33,81%	25,23%	22,76%	16,25%
CAC40 10 min.	BV, SS	266	190	159	124	34,59%	24,71%	20,68%	16,12%
	BV*, SS	275	196	167	127	35,76%	25,49%	21,72%	16,51%
CAC40 30 min.	BV, SS	355	269	249	213	46,16%	34,98%	32,38%	27,70%
	BV*, SS	372	289	264	226	48,37%	37,58%	34,33%	29,39%
DAX 5 min.	BV, SS	351	268	239	208	45,53%	34,76%	31,00%	26,98%
	BV*, SS	352	269	246	211	45,65%	34,89%	31,91%	27,37%
DAX 10 min.	BV, SS	335	258	237	192	43,45%	33,46%	30,74%	24,90%
	BV*, SS	340	263	242	196	44,10%	34,11%	31,39%	25,42%
DAX 30 min.	BV, SS	392	308	294	252	50,84%	39,95%	38,13%	32,68%
	BV*, SS	408	326	302	269	52,92%	42,28%	39,17%	34,89%
NASDAQ 5 min.	BV, SS	233	155	127	83	30,06%	20,00%	16,39%	10,71%
	BV*, SS	234	158	132	86	30,19%	20,39%	17,03%	11,10%
NASDAQ 10 min.	BV, SS	239	181	155	113	30,84%	23,35%	20,00%	14,58%
	BV*, SS	247	185	166	119	31,87%	23,87%	21,42%	15,35%
NASDAQ 30 min.	BV, SS	309	230	219	184	39,87%	29,68%	28,26%	23,74%
	BV*, SS	337	246	228	201	43,48%	31,74%	29,42%	25,94%
WIG20 5 min.	BV, SS	242	176	159	120	31,97%	23,25%	21,00%	15,85%
	BV*, SS	245	180	162	122	32,36%	23,78%	21,40%	16,12%
WIG20 10 min.	BV, SS	291	203	182	144	38,44%	26,82%	24,04%	19,02%
	BV*, SS	298	214	184	148	39,37%	28,27%	24,31%	19,55%
WIG20 20 min.	BV, SS	314	239	223	183	41,48%	31,57%	29,46%	24,17%
	BV*, SS	319	253	230	197	42,14%	33,42%	30,38%	26,02%
WIG20 40 min.	BV, SS	351	277	266	225	46,37%	36,59%	35,14%	29,72%
	BV*, SS	373	302	281	255	49,27%	39,89%	37,12%	33,69%
EURPLN 10 min.	BV, SS	253	189	161	128	32,81%	24,51%	20,88%	16,60%
	BV*, SS	257	191	165	130	33,33%	24,77%	21,40%	16,86%
EURPLN 30 min.	BV, SS	398	337	318	278	51,62%	43,71%	41,25%	36,06%
	BV*, SS	367	286	267	229	47,60%	37,09%	34,63%	29,70%
EURPLN 60 min.	BV, SS	293	232	211	166	38,00%	30,09%	27,37%	21,53%
	BV*, SS	319	252	235	201	41,37%	32,68%	30,48%	26,07%
USDPLN 10 min.	BV, SS	296	202	181	140	38,54%	26,30%	23,57%	18,23%
	BV*, SS	304	210	187	144	39,58%	27,34%	24,35%	18,75%
USDPLN 30 min.	BV, SS	283	203	184	139	36,85%	26,43%	23,96%	18,10%
	BV*, SS	300	218	196	155	39,06%	28,39%	25,52%	20,18%
USDPLN 60 min.	BV, SS	306	202	180	138	39,84%	26,30%	23,44%	17,97%
	BV*, SS	318	228	194	150	41,41%	29,69%	25,26%	19,53%

Podobnie jak w przypadku statystyk Barndorff-Nielsen i Shepharda, statystyka liniowa wykrywa znacznie więcej skoków niż inne statystyki. Potwierdzają to również badania symulacyjne przeprowadzone w rozdziale 4. Statystyki logarytmiczna i ilorazowa wykrywają niemal identyczną liczbę skoków. Obserwujemy również, że i w tym przypadku, zastosowana częstotliwość danych ma istotny wpływ na liczbę wykrytych

dni ze skokami. Im mniejsza częstotliwość, tym większa liczba skoków wykrytych przez test. Na podstawie badań symulacyjnych, przeprowadzonych w rozdziale 4, możemy wywnioskować, że jest to następstwem rosnącego rozmiaru wraz ze spadkiem częstotliwości, przy nieznacznie zmniejszającej się mocy testu. Pomimo, że badania symulacyjne wskazały, że rodzaj estymatora zmienności scałkowanej i kwartyczności nie ma wpływu na rozmiar i moc testu, obserwujemy pewne niewielkie różnice w liczbie dni ze skokami wykrytych przez statystyki z przeskalowanymi i nieprzeskalowanymi estymatorami zmienności scałkowanej i kwartyczności. Liczba skoków wykrytych za pomocą odpowiadających sobie statystyk Barndorff-Nielsen i Shepharda (2006) oraz Jianga i Oomena (2008) przy zadanej częstotliwości i poziomie ufności jest w większości przypadków dość zbliżona, co zwiększa wiarygodność przeprowadzonych testów.

W tabeli 7.10 przedstawiona jest liczba skoków wykrytych za pomocą statystyk Lee i Myklanda w badanych procesach poziomów logarytmicznych indeksów giełdowych.

**Tabela 7.10. Liczby skoków wykrytych za pomocą statystyki Lee i Myklanda (2007) w procesach logarytmicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 oraz logarytmicznych kursach EURPLN i USDPLN, w zależności od poziomu ufności, zastosowanych estymatorów zmienności scałkowanej i kwartyczności oraz częstotliwości danych.**

Szereg notowań	Bez funkcji dryfu				Z funkcją dryfu			
	0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min.	224	162	150	113	217	167	144	110
CAC40 10 min.	106	89	78	68	104	89	79	63
CAC40 30 min.	58	35	32	28	58	36	31	27
DAX 5 min.	355	289	261	210	343	286	261	213
DAX 10 min.	178	144	133	104	174	146	129	103
DAX 30 min.	77	62	59	46	77	61	53	45
NASDAQ 5 min.	116	83	72	49	112	82	66	48
NASDAQ 10 min.	53	33	27	21	55	34	31	19
NASDAQ 30 min.	25	18	16	12	27	20	16	12
WIG20 5 min.	142	105	87	68	144	98	87	68
WIG20 10 min.	65	47	39	24	66	46	38	23
WIG20 20 min.	45	28	23	13	42	27	21	14
WIG20 40 min.	25	21	19	15	24	22	20	16
EURPLN 10 min.	143	103	94	72	145	101	91	70
EURPLN 30 min.	75	56	53	41	75	56	50	41
EURPLN 60 min.	27	20	17	9	26	20	16	11
USDPLN 10 min.	141	116	108	83	139	113	102	141
USDPLN 30 min.	54	44	36	28	54	43	38	30
USDPLN 60 min.	29	19	16	9	27	19	16	29

Jak łatwo zauważyć, statystyki  $\hat{L}$  oraz  $\hat{L}_\mu$  wykrywają niemal identyczną liczbę skoków pomimo, że oszacowania parametrów otrzymane w podrozdziale 6.1 mogą sugerować, iż badane procesy cen logarytmicznych posiadają istotną funkcję dryfu. Wniosek ten potwierdzają badania przeprowadzone w podrozdziale 4.5. Liczba wykrytych skoków maleje wraz ze zmniejszającą się częstotliwością. Badania symulacyjne, przedstawione w rozdziale 4.5, wykazały, że moc testu szybko maleje wraz ze zmniejszającą się częstotliwością danych. Efekt ten jest najsilniejszy dla logarytmicznych kursów walutowych, gdzie wartość skoku w stosunku do bieżącego kursu logarytmicznego jest najmniejsza. Wobec tego, gdy stosujemy tę statystykę testową, szczególnie ważne jest użycie danych o odpowiednio wysokiej częstotliwości. Należy jednak przy tym pamiętać o efektach mikrostruktury, które są charakterystyczne dla tego typu danych.

## 7.2. Testowanie skoków a mikrostruktura rynku

Tak jak opisywaliśmy w rozdziale 4.3, szeregi zwrotów o wysokiej częstotliwości są często zanieczyszczone szumem mikrostruktury, który możemy zaobserwować poprzez wykrycie w szeregach zwrotów wysokiej częstotliwości istotnej autokorelacji niskich rzędów. Zjawisko to ma znaczący wpływ na estymację wariacji wielopotęgowych, a więc także na wartości statystyk testowych. W niniejszym podrozdziale przeprowadzimy test istotności autokorelacji różnego stopnia w szeregach zwrotów śróddziennych. W sytuacji występowania wysokich wartości współczynników autokorelacji, zgodnie z zaleceniem Barndorff-Nielsen i Shepharda (2006), zastosujemy przesunięte wariacje wielopotęgowe. Testowanie istotności autokorelacji rzędu  $h$  umożliwia nam następujące twierdzenie (np. Shumway i Stoffer 2006).

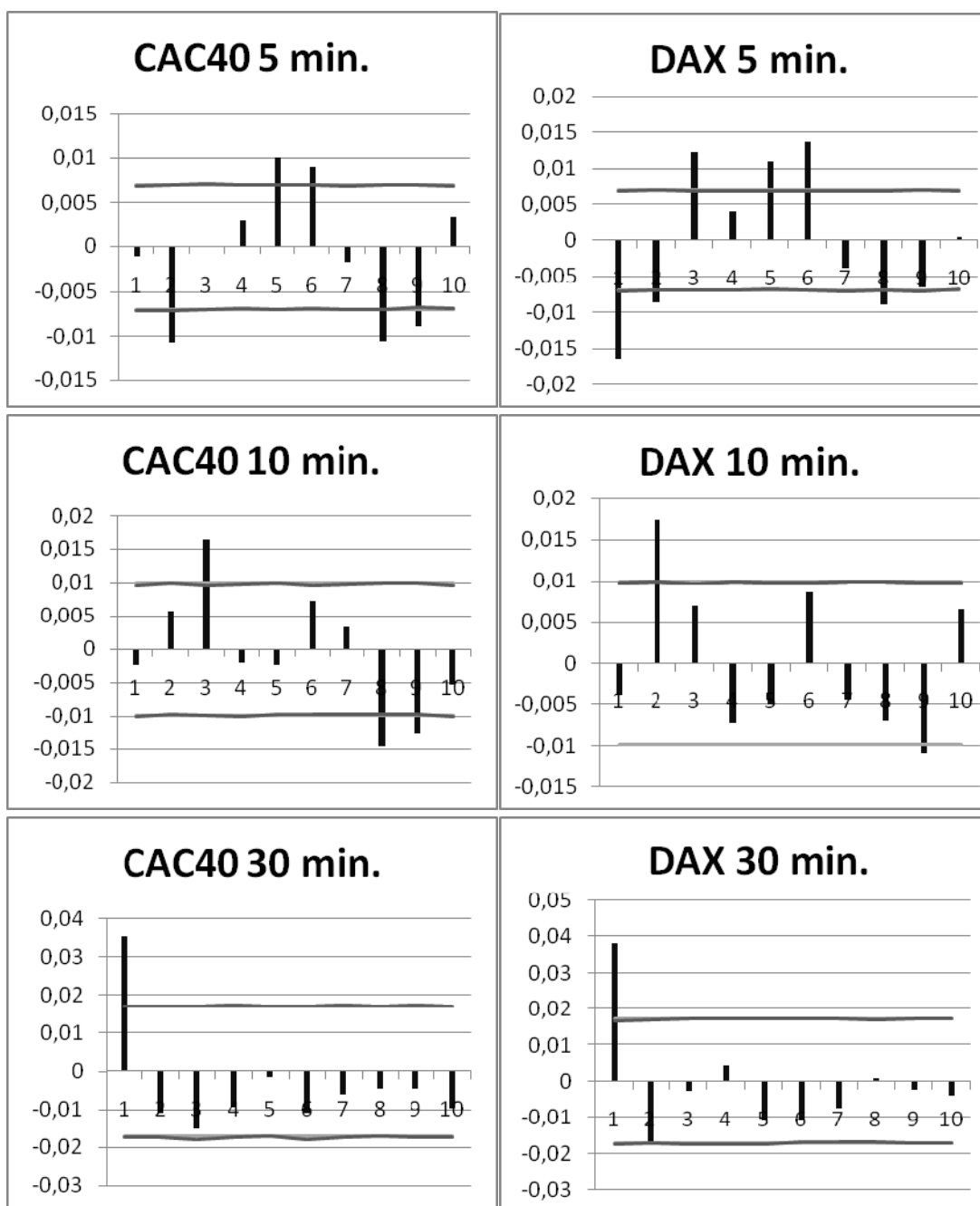
**Twierdzenie 7.2.1.** Niech  $y_t$  będzie białym szumem. Wówczas, dla odpowiednio dużego  $n$  i dowolnego  $h$ , współczynnik autokorelacji z próby,  $\hat{\rho}(h)$ , ma w przybliżeniu rozkład normalny ze średnią 0 i odchyleniem standardowym

$$\sigma_{\hat{\rho}(h)} = \frac{1}{\sqrt{n}}.$$

Na mocy tego twierdzenia łatwo zauważyć, że hipoteza zerowa o autokorelacji rzędu  $h$  będzie odrzucona na poziomie istotności  $\alpha$ , o ile

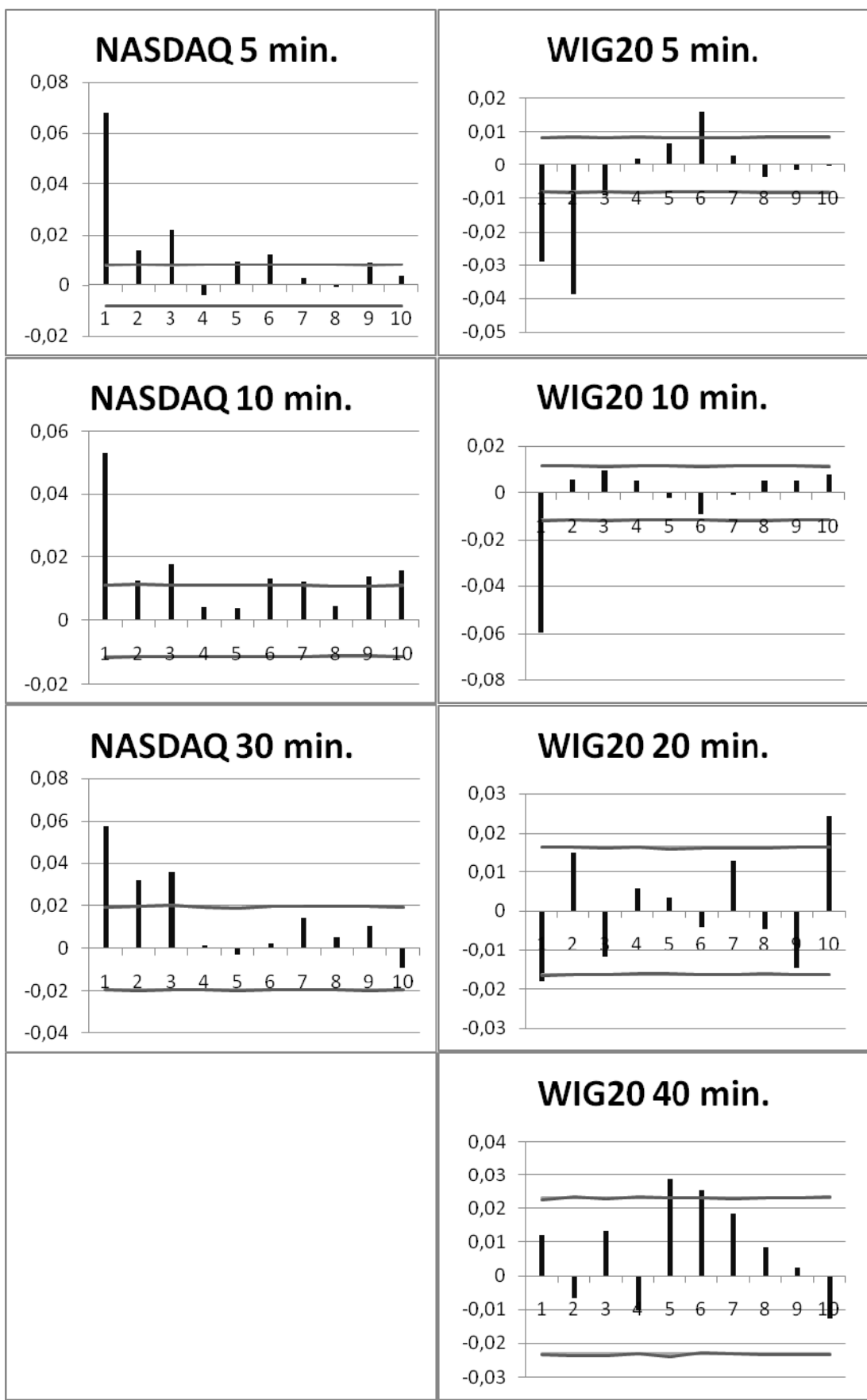
$$|\hat{\rho}(h)| > \Phi^{-1}(1-\alpha/2)/\sqrt{n}, \quad (7.1)$$

gdzie  $\Phi(x)$  jest dystrybuantą standardowego rozkładu normalnego. Progi istotności wyznaczamy niezależnie z rozkładów bootstrapowych otrzymanych z 10000 prób. Na rysunkach 7.1-7.3 przedstawione są wykresy empirycznej funkcji autokorelacji zwrotów śróddziennych różnej częstotliwości dla indeksów CAC40, DAX, NASDAQ i WIG20 wraz z progami istotności wyznaczonymi dla  $\alpha = 0,05$  z wzoru (7.1) oraz rozkładów bootstrapowych.

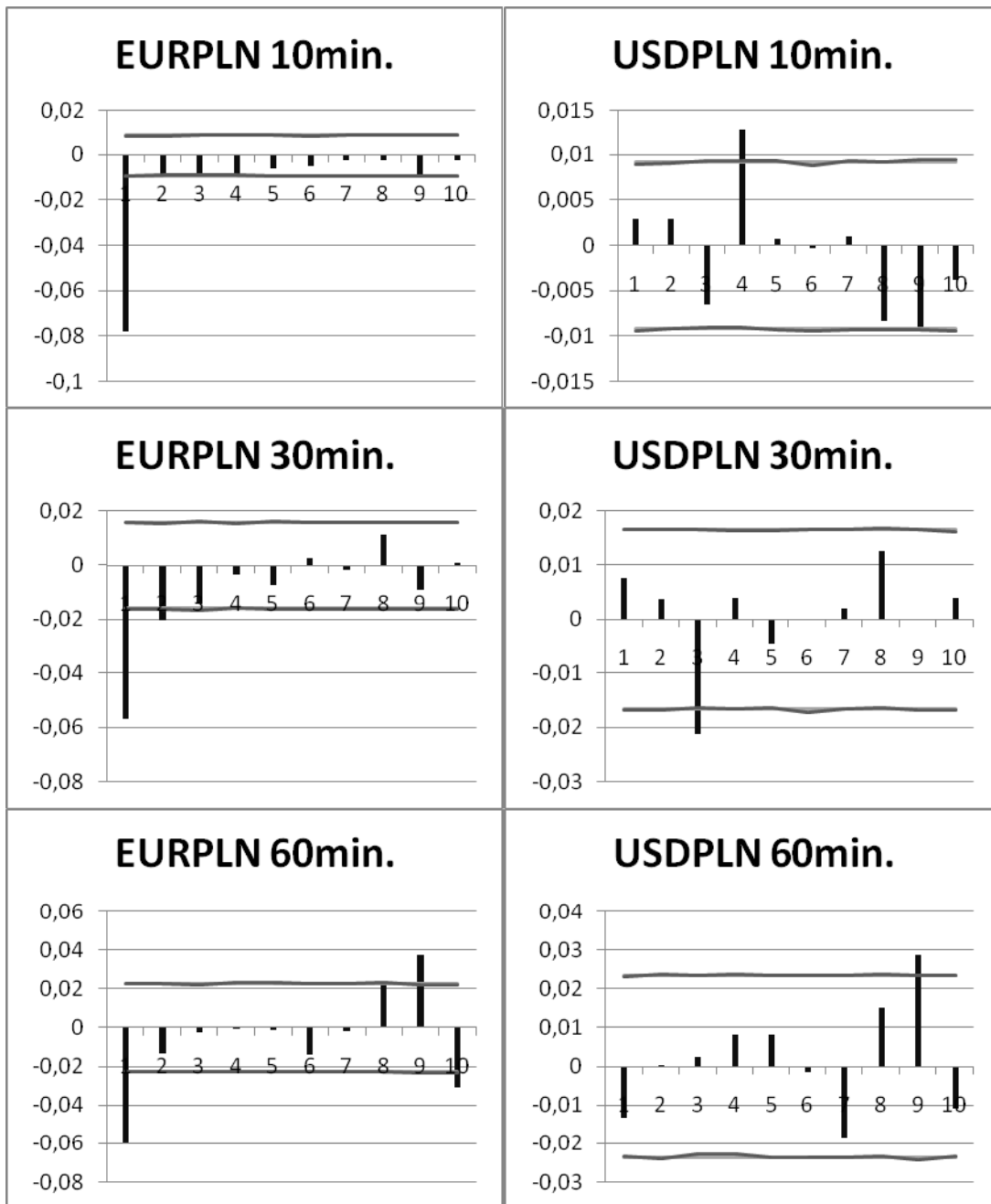


Rys. 7.1. Autokorelacja zwrotów śróddziennych różnej częstotliwości dla indeksów CAC40 i DAX, z progami istotności wyznaczonymi ze wzoru (7.1) oraz z rozkładów bootstrapowych.





Rys. 7.2. Autokorelacja zwrotów śróddziennych różnej częstotliwości dla indeksów NASDAQ i WIG20, z progami istotności wyznaczonymi ze wzoru (7.1) oraz z rozkładów bootstrapowych.



**Rys. 7.3. Autokorelacja zwrotów śróddziennych różnej częstotliwości dla indeksów kursu EURPLN i USDPLN, z progami istotności wyznaczonymi ze wzoru (7.1) oraz z rozkładów bootstrapowych.**

Jak łatwo zauważyć, progi istotności otrzymane ze wzoru (7.1) oraz te otrzymane na podstawie rozkładu bootstrapowego są niemal identyczne, co dodatkowo uwiarygodnia przeprowadzone badania. Należy pamiętać, że autokorelacja w szeregach zwrotów śróddziennych może mieć także inne uzasadnienie, jak chociażby dyskontowanie informacji z opóźnieniem, czy krótkoterminowa sezonowość. Wobec tego samo istnienie istotnej autokorelacji nie musi być następstwem efektów mikrostruktury, zwłaszcza

jeżeli pojawia się ona dla stosunkowo dużych opóźnień. Konsekwencją efektów mikrostruktury mogą być natomiast silne autokorelacje niskich rzędów. Zjawisko takie nie jest obserwowalne w pięcio- i dziesięciominutowych szeregach zwrotów CAC40 oraz dziesięciominutowych zwrotach DAX. Dla zwrotów pięciominutowych indeksu DAX można zauważyć istotną autokorelację rzędu 1 przewyższającą tylko nieznacznie autokorelacje wyższych rzędów. Natomiast dla zwrotów 30 minutowych obydwu indeksów obserwujemy silne autokorelacje rzędu 1, mogące być następstwem efektów mikrostruktury. Dla zwrotów indeksu NASDAQ obserwujemy silną autokorelację rzędu 1 i słabe autokorelacje rzędu 2 i 3. W zwrotach 30 minutowych różnice pomiędzy wartościami autokorelacji rzędu 1 oraz 2 i 3 są znacznie mniejsze, niż w innych rozważanych przypadkach. Dlatego do wyznaczania statystyk testowych stosujemy wariacje dwu- i trójpotęgowe przesunięte zarówno o 1, jak i o 3 obserwacje. W przypadku indeksu WIG20 obserwujemy silne autokorelacje rzędu 1 i 2 w zwrotach 5-minutowych i rzędu 1 w zwrotach 10-minutowych, co uzasadnia zastosowanie do wyznaczenia statystyk testowych dla tych szeregów, odpowiednio przesuniętych o 2 i o 1 obserwację wariacji dwu- i trójpotęgowych. Dla zwrotów dwudziesto i czterdziestominutowych, korelogramy sugerują brak efektów mikrostruktury w badanym szeregu. Dla zwrotów wszystkich częstotliwości kursu EURPLN obserwujemy silną autokorelację rzędu 1. Ponadto w zwrotach wyznaczonych na podstawie danych o częstotliwości 10 minut i 30 minut zauważamy autokorelację rzędu 2 nieznacznie przekraczającą próg istotności, co sugeruje stosowanie jako estymatorów zmienności scałkowanej i kwartyczności, przesuniętych o 1 obserwację wariacji dwupotęgowych i trójpotęgowych. Korelogramy wyznaczone dla zwrotów kursów USDPLN nie wskazują na efekty mikrostruktury w szeregach danych. Silna autokorelacja rzędu 9 w szeregach zwrotów 60 minutowych obydwu kursów świadczy o dziennej sezonowości w szeregach zwrotów.

Na podstawie powyższej analizy przeprowadziliśmy testy na występowanie skoków z wykorzystaniem przesuniętych wariacji wielopotęgowych jako estymatorów zmienności scałkowanej i kwartyczności. Wartości statystyk zostały wyznaczone na podstawie danych różnej częstotliwości. W tabelach 7.11 i 7.12 przedstawione są podstawowe statystyki przesuniętych wariacji wielopotęgowych wykorzystanych do wyznaczenia statystyk testowych. Dla ułatwienia przyjęliśmy oznaczenia

$$BV^{*k} = \{X_{\delta} \}_{i,[k]}^{[1,1]} \text{ i } TQ^{*k} = \{X_{\delta} \}_{i,[k]}^{[4/3,4/3,4/3]}.$$

**Tabela 7.11. Średnie ( $\mu$ ) i wariancje ( $\sigma^2$ ) zrealizowanych przesuniętych wariacji wielopotęgowych dla szeregów zwrotów intraday indeksu WIG20 oraz kursów EURPLN i USDPLN różnych częstotliwości.**

Szereg zwrotów	$k$	BV** $k$		TQ** $k$	
		$\mu$	$\sigma^2$	$\mu$	$\sigma^2$
CAC40 5 min.	1	0,3825	0,1814	0,5860	10,111
CAC40 10 min.	1	0,3781	0,1178	0,3249	1,6614
CAC40 30 min.	1	0,3412	0,1139	0,2312	4,0882
DAX 5 min	1	0,4939	0,1467	0,6075	4,0385
DAX 10 min.	1	0,4775	0,1623	0,4704	3,4167
DAX 30 min	1	0,4305	0,1403	0,2746	0,8000
NASDAQ 5 min.	1	0,4410	0,0859	0,3438	0,4483
NASDAQ 10 min.	1	0,4665	0,0983	0,3392	0,4193
NASDAQ 30 min.	1	0,4501	0,1354	0,2404	0,4732
WIG20 5 min.	2	1,0773	0,8238	2,3149	114,817
WIG20 10 min.	1	1,0477	1,1087	2,2125	107,821
WIG20 20 min.	1	0,9352	0,8906	1,6674	59,0937
WIG20 40 min.	1	0,7561	0,6773	1,0036	34,6291

**Tabela 7.12. Średnie ( $\mu$ ) i wariancje ( $\sigma^2$ ) zrealizowanych przesuniętych wariacji wielopotęgowych dla zwrotów kursów intraday EURPLN i USDPLN różnych częstotliwości.**

Szereg zwrotów	$k$	BV** $k$		TQ** $k$	
		$\mu$	$\sigma^2$	$\mu$	$\sigma^2$
EURPLN 10 min.	1	0,2784	0,0449	0,1766	0,2491
EURPLN 30 min.	1	0,2527	0,0546	0,1066	0,1157
EURPLN 60 min.	1	0,2237	0,0543	0,0776	0,1054
USDPLN 10 min.	1	0,3723	0,0506	0,2440	0,2686
USDPLN 30 min.	1	0,3733	0,0766	0,1876	0,2212
USDPLN 60 min.	1	0,3619	0,1042	0,1565	0,1500

W tabelach 7.13 i 7.14 przedstawiona jest liczba dni ze skokami wykrytymi za pomocą statystyk Barndorff-Nielsen i Sheparda z przesuniętymi wariacjami wielopotęgowymi jako estymatorami zmienności scałkowanej i kwartyczności.

Porównując tabele 7.13 i 7.14 z tabelami 7.4-7.6 widzimy, że test ilorazowy z przesuniętymi wariacjami wielopotęgowymi zastosowanymi jako estymatorami zmienności scałkowanej i kwartyczności, wykrywa nieco mniej skoków. Jednocześnie obserwujemy, że liczba dni ze skokami wykrytymi za pomocą dwóch pozostałych testów z przesuniętymi wariacjami wielopotęgowymi jest nieco większa niż liczba skoków wykrytych za pomocą testów wykorzystujących dwa pozostałe estymatory wariacji wielopotęgowych. Z technicznego punktu widzenia jest to następstwem znacznie mniejszych wartości oszacowań kwartyczności, wobec nieznacznie niższych wartości oszacowań wariancji scałkowanej.

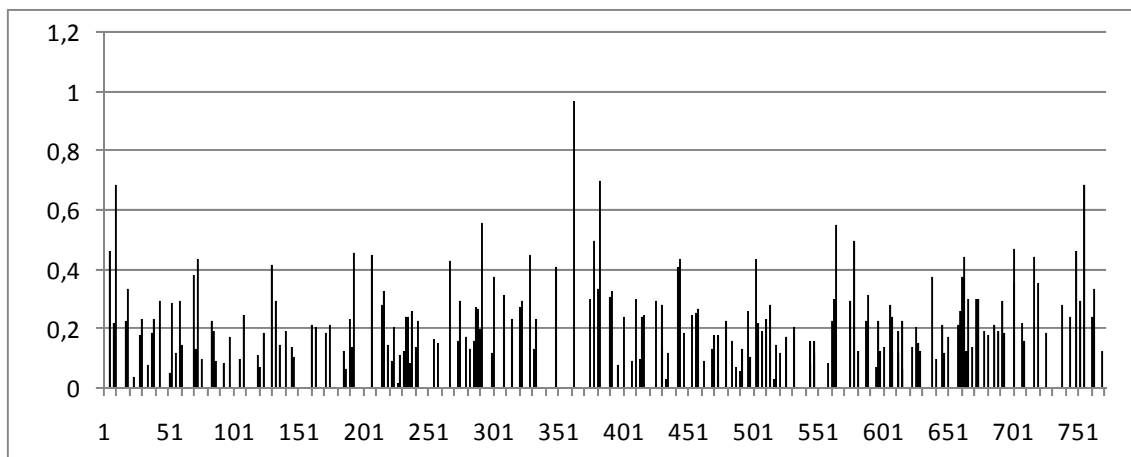
**Tabela 7.13. Liczby wykrytych dni ze skokami w procesach logarymicznych poziomów indeksów CAC40, DAX, NASDAQ i WIG20 w zależności od poziomu ufności, wyznaczone za pomocą statystyk Barndorff-Nielsen i Shepharda z wykorzystaniem przesuniętych wariacji wielopiętgowych jako estymatorów zmienności scalkowanej i kwartyczności.**

Szereg notowań (opóźnienie wariacji wielo- piętgowych)	statystyka	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
CAC40 5 min. (1)	liniowa	239	162	138	109	31,08%	21,07%	17,95%	14,17%
	logarytm.	229	139	124	84	29,78%	18,08%	16,13%	10,92%
	ilorazowa	195	139	124	84	25,36%	18,08%	16,12%	10,92%
CAC40 10 min. (1)	liniowa	267	197	176	139	34,72%	25,62%	22,89%	18,08%
	logarytm.	254	169	147	103	33,03%	21,98%	19,12%	13,40%
	ilorazowa	215	169	147	103	27,96%	21,98%	19,12%	13,39%
CAC40 30 min. (1)	liniowa	298	243	232	197	38,75%	31,60%	30,17%	25,62%
	logarytm.	278	208	193	148	36,15%	27,05%	25,10%	19,25%
	ilorazowa	195	208	193	148	25,36%	27,05%	25,10%	19,25%
DAX 5 min. (1)	liniowa	414	331	296	239	53,70%	42,93%	38,39%	31,00%
	logarytm.	396	299	267	197	51,37%	38,78%	34,63%	25,55%
	ilorazowa	380	299	267	197	49,29%	38,78%	34,63%	25,55%
DAX 10 min. (1)	liniowa	356	273	250	205	46,17%	35,41%	32,43%	26,59%
	logarytm.	326	244	211	161	42,28%	31,65%	27,37%	20,88%
	ilorazowa	291	244	211	161	37,74%	31,65%	27,37%	20,88%
DAX 30 min. (1)	liniowa	346	278	260	230	44,88%	36,06%	33,72%	29,83%
	logarytm.	317	245	222	178	41,12%	31,78%	28,79%	23,09%
	ilorazowa	234	245	222	178	30,35%	31,78%	28,79%	23,09%
NASDAQ 5 min. (1)	liniowa	246	166	135	99	31,74%	21,42%	17,42%	12,77%
	logarytm.	231	131	108	72	29,80%	16,90%	13,94%	9,29%
	ilorazowa	206	131	108	72	26,58%	16,90%	13,94%	9,29%
NASDAQ 5 min. (3)	liniowa	281	196	165	132	37,22%	25,96%	21,85%	17,48%
	logarytm.	267	164	142	99	35,36%	21,72%	18,81%	13,11%
	ilorazowa	233	164	142	99	30,86%	21,72%	18,81%	13,11%
NASDAQ 10 min. (1)	liniowa	216	157	127	98	27,87%	20,26%	16,39%	12,65%
	logarytm.	198	123	104	74	25,55%	15,87%	13,42%	9,55%
	ilorazowa	158	123	104	74	20,39%	15,87%	13,42%	9,55%
NASDAQ 30 min. (1)	liniowa	266	216	196	157	34,32%	27,87%	25,29%	20,26%
	logarytm.	242	175	149	117	31,22%	22,58%	19,23%	15,10%
	ilorazowa	144	175	149	117	18,58%	22,58%	19,23%	15,10%
WIG20 5 min. (2)	liniowa	346	276	253	203	45,71%	36,46%	33,42%	26,82%
	logarytm.	331	246	217	163	43,73%	32,50%	28,67%	21,53%
	ilorazowa	299	246	217	163	39,50%	32,50%	28,67%	21,53%
WIG20 10 min. (1)	liniowa	265	191	163	124	35,01%	25,23%	21,53%	16,38%
	logarytm.	240	146	120	81	31,70%	19,29%	15,85%	10,70%
	ilorazowa	188	146	120	81	24,83%	19,29%	15,85%	10,70%
WIG20 20 min. (1)	liniowa	234	166	139	109	30,91%	21,93%	18,36%	14,40%
	logarytm.	204	108	85	60	26,95%	14,27%	11,23%	7,93%
	ilorazowa	113	108	85	60	14,93%	14,27%	11,23%	7,93%
WIG20 40 min. (1)	liniowa	228	164	142	115	30,12%	21,66%	18,76%	15,19%
	logarytm.	161	80	58	33	21,27%	10,57%	7,66%	4,36%
	ilorazowa	52	80	58	33	6,87%	10,57%	7,66%	4,36%

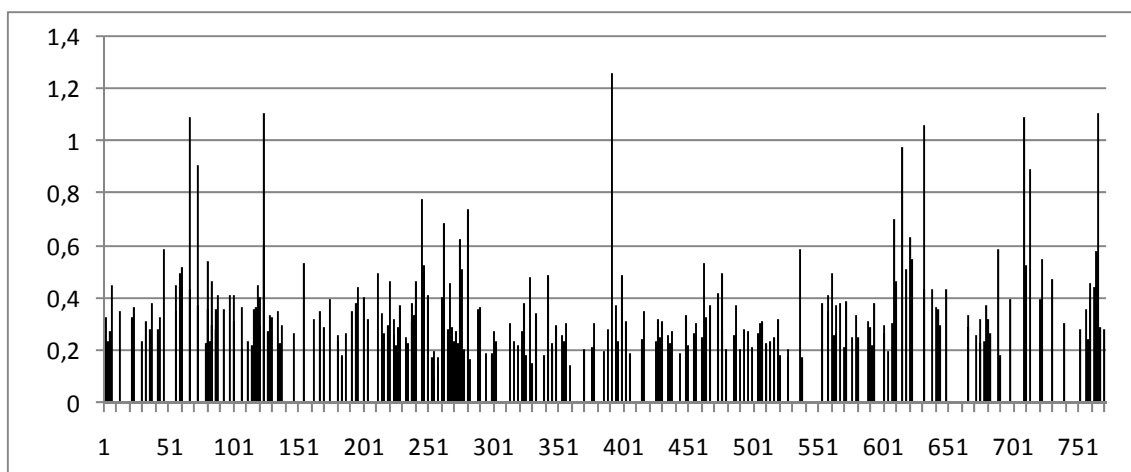
**Tabela 7.14. Liczby wykrytych dni ze skokami w procesach logarytmicznych kursów EURPLN i USDPLN w zależności od poziomu ufności, wyznaczone za pomocą statystyk Barndorff-Nielsen i Shepharda z wykorzystaniem przesuniętych wariacji wielopotęgowych jako estymatorów zmienności scałkowanej i kwartyczności.**

Szereg notowań (opóźnienie wariacji wielo- potęgowych)	statystyka	Liczba wykrytych dni ze skokami				Procentowy udział dni ze skokami			
		0,95	0,99	0,995	0,999	0,95	0,99	0,995	0,999
EURPLN 10 min. (1)	liniowa	302	224	204	164	39,17%	29,05%	26,46%	21,27%
	logarytm.	281	199	172	131	36,45%	25,81%	22,31%	16,99%
	ilorazowa	247	199	172	131	32,04%	25,81%	22,31%	16,99%
EURPLN 30 min. (1)	liniowa	351	268	250	214	45,53%	34,76%	32,43%	27,76%
	logarytm.	317	230	212	163	41,12%	29,83%	27,50%	21,14%
	ilorazowa	227	180	212	163	29,44%	29,83%	27,50%	21,14%
EURPLN 60 min. (1)	liniowa	281	229	214	188	36,45%	29,70%	27,76%	24,38%
	logarytm.	255	194	178	141	33,07%	25,16%	23,09%	18,29%
	ilorazowa	155	57	32	8	20,10%	7,39%	4,15%	1,04%
USDPLN 10 min. (1)	liniowa	304	227	214	173	39,58%	29,56%	27,86%	22,53%
	logarytm.	258	198	176	135	33,59%	25,78%	22,92%	17,58%
	ilorazowa	196	198	176	135	25,52%	25,78%	22,92%	17,58%
USDPLN 30 min. (1)	liniowa	241	180	165	138	31,38%	23,44%	21,48%	17,97%
	logarytm.	221	152	136	104	28,78%	19,79%	17,71%	13,54%
	ilorazowa	149	152	136	104	19,40%	19,79%	17,71%	13,54%
USDPLN 60 min. (1)	liniowa	277	234	222	197	36,07%	30,47%	28,91%	25,65%
	logarytm.	255	202	180	149	33,20%	26,30%	23,44%	19,40%
	ilorazowa	147	202	180	149	19,14%	26,30%	23,44%	19,40%

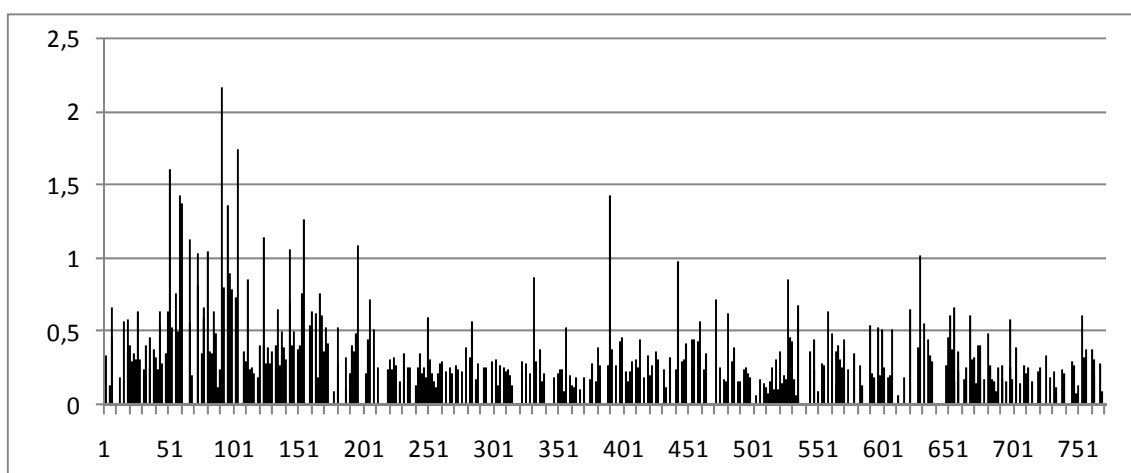
Jak wspomnieliśmy już w rozdziale 4, przybliżone wartości skoków można wyznaczyć na podstawie różnic wariacji zrealizowanej i wariacji dwupotęgowej (wzór (4.10)) oraz wariacji zrealizowanej i wariacji swapowej (wzór (4.14)). Na rysunkach 7.14 i 7.15 przedstawione są wartości skoków wyznaczone z powyższych wzorów pozytywnie zweryfikowane, odpowiednio, przez logarytmiczną statystykę Barndorff-Nielsen i Shepharda z zastosowanymi przesuniętymi wariacjami wielopotęgowymi jako estymatorami wariacji scałkowanej i kwartyczności oraz logarytmiczną statystykę Jianga i Oomena z przeskalowaną przeskalowaną wariacją sześciopotęgową jako estymatorem sekstyczności. Badania empiryczne wykonane w rozdziale 4 wykazały, że im wyższa jest częstotliwość danych, tym moc testu jest większa, a jego rozmiar mniejszy. Dlatego we wszystkich przypadkach wykorzystaliśmy dane o najwyższej dostępnej częstotliwości.



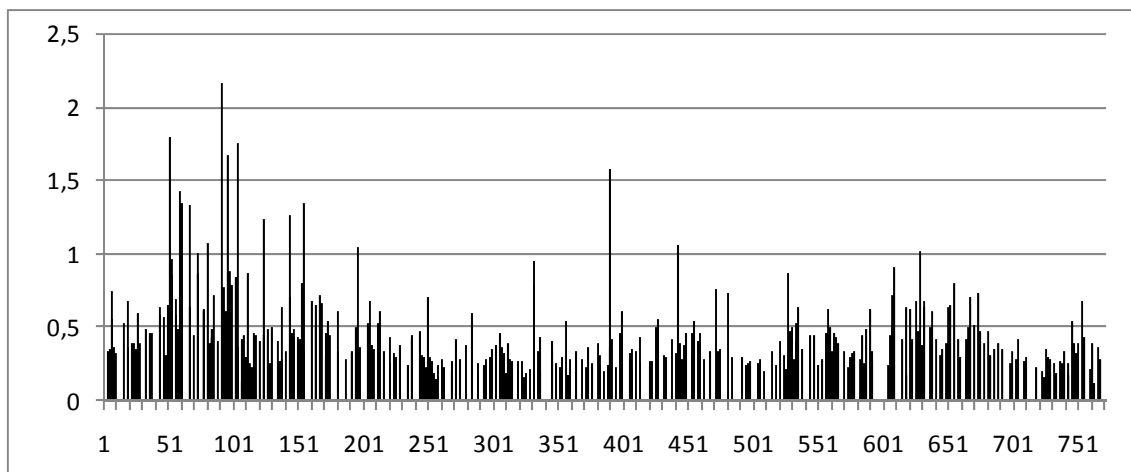
**7.4. Wartości skoków w procesie cen logarytmicznych CAC40 wyznaczonych na podstawie zależności (4.10), pozytywnie zweryfikowanych przez logarytmiczny test Barndorff-Nielsen i Shepharda.**



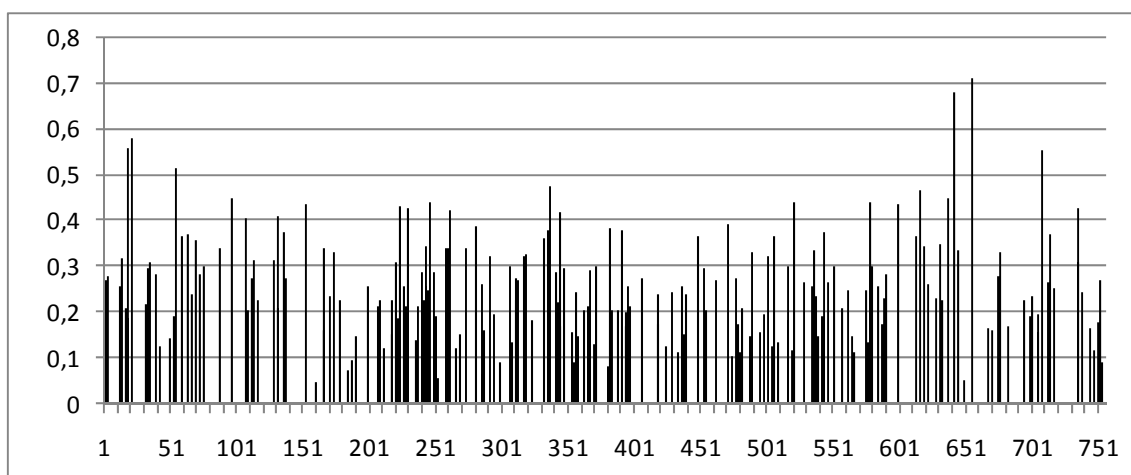
**7.5. Wartości skoków w procesie cen logarytmicznych CAC40 wyznaczonych na podstawie zależności (4.13), pozytywnie zweryfikowanych przez logarytmiczny test Jianga i Oomena.**



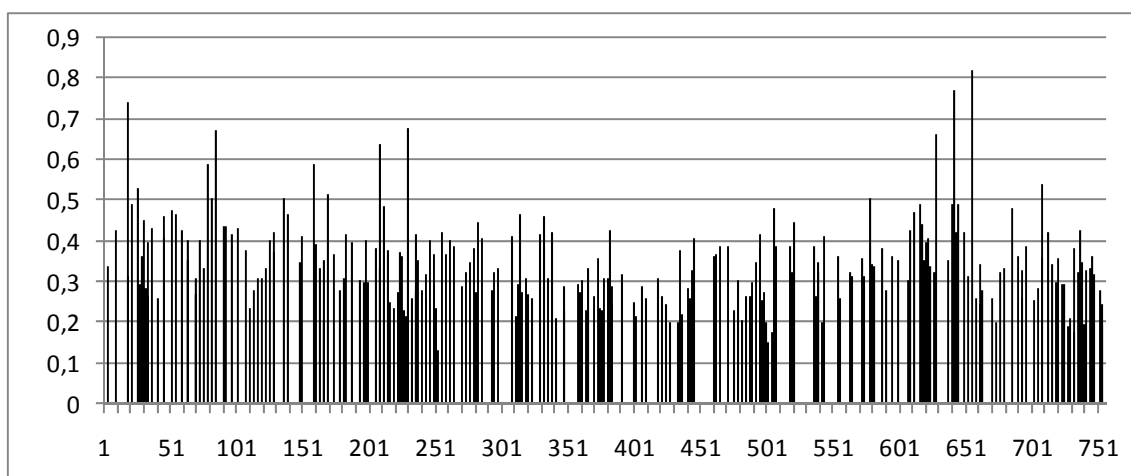
**7.6 Wartości skoków w procesie cen logarytmicznych DAX wyznaczonych na podstawie zależności (4.10), pozytywnie zweryfikowanych przez logarytmiczny test Barndorff-Nielsen i Shepharda.**



**7.7. Wartości skoków w procesie cen logarytmicznych DAX wyznaczonych na podstawie zależności (4.13), pozytywnie zweryfikowanych przez logarytmiczny test Jianga i Oomena.**

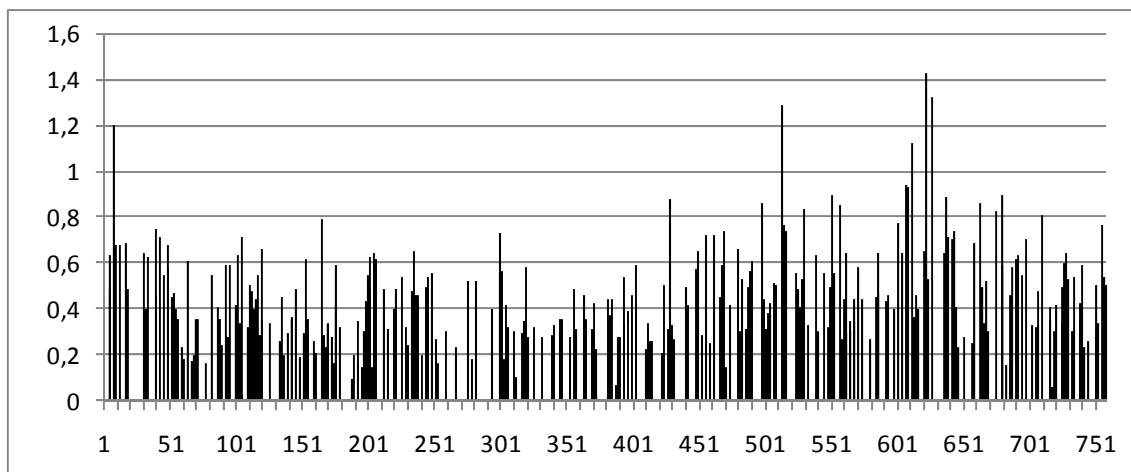


**7.8. Wartości skoków w procesie cen logarytmicznych NASDAQ wyznaczonych na podstawie zależności (4.10), pozytywnie zweryfikowanych przez logarytmiczny test Barndorff-Nielsen i Shepharda.**

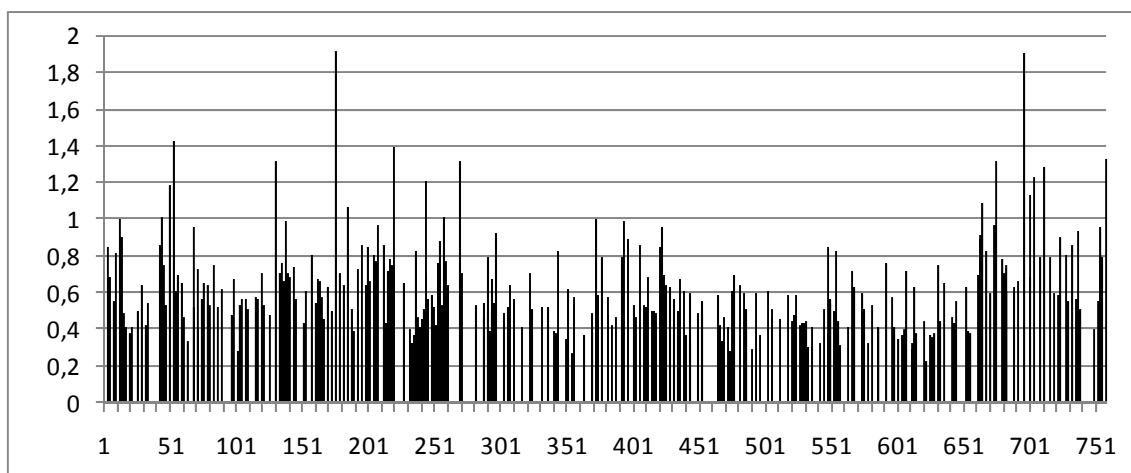


**7.9. Wartości skoków w procesie cen logarytmicznych NASDAQ wyznaczonych na podstawie zależności (4.13), pozytywnie zweryfikowanych przez logarytmiczny test Jianga i Oomena.**

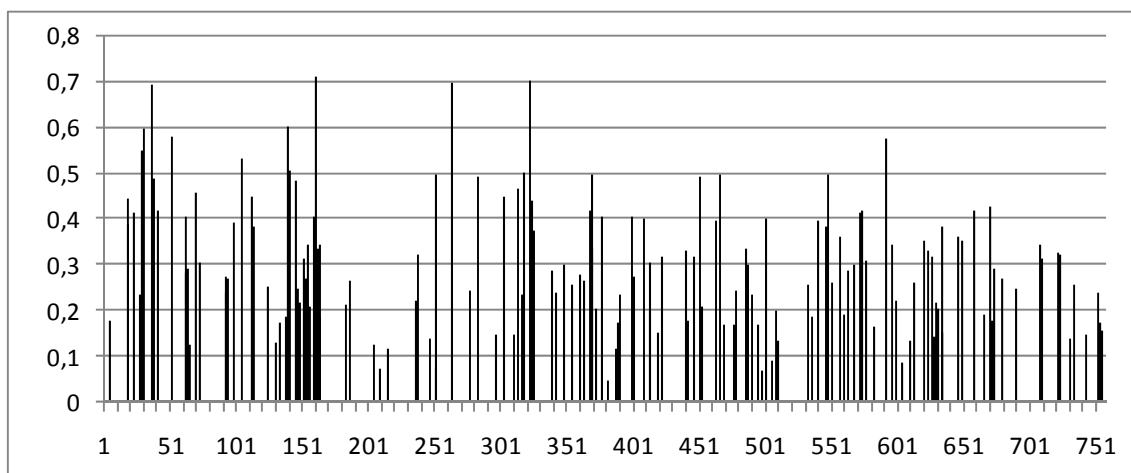




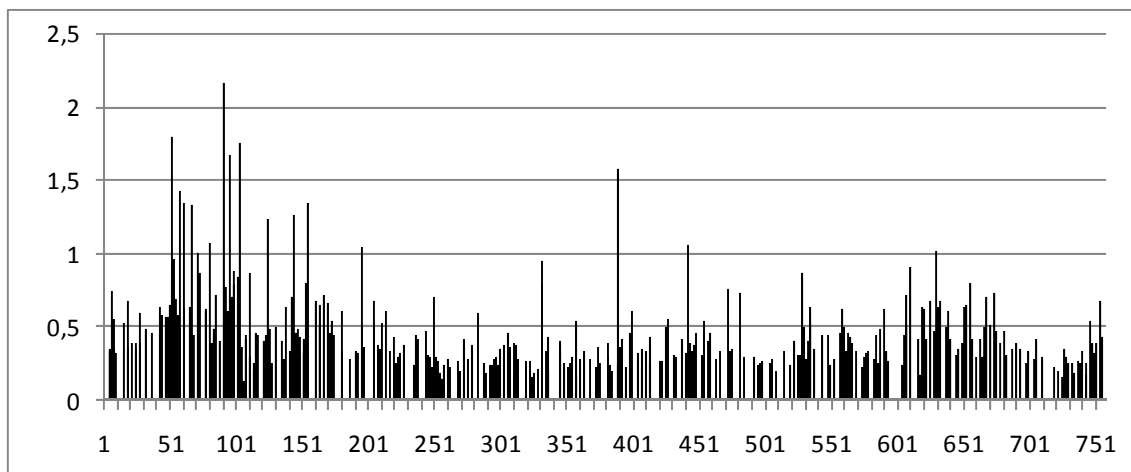
**7.10. Wartości skoków w procesie cen logarytmicznych WIG20 wyznaczonych na podstawie zależności (4.10), pozytywnie zweryfikowanych przez logarytmiczny test Barndorff-Nielsen i Shepharda.**



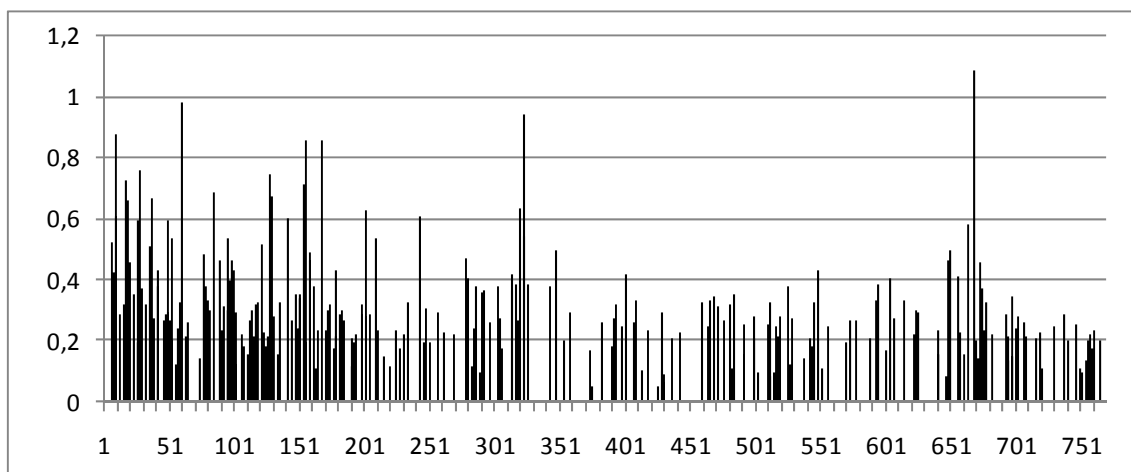
**7.11. Wartości skoków w procesie cen logarytmicznych WIG20 wyznaczonych na podstawie zależności (4.13), pozytywnie zweryfikowanych przez logarytmiczny test Jianga i Oomena.**



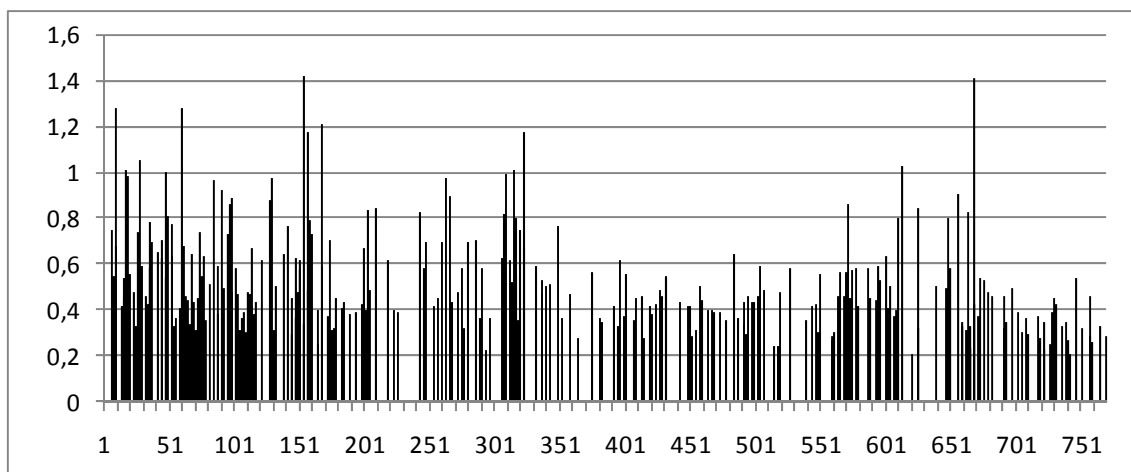
**7.12. Wartości skoków w procesie kursów logarytmicznych EURPLN wyznaczonych na podstawie zależności (4.10), pozytywnie zweryfikowanych przez logarytmiczny test Barndorff-Nielsen i Shepharda.**



**7.13. Wartości skoków w procesie kursów logarytmicznych EUR/PLN wyznaczonych na podstawie zależności (4.13), pozytywnie zweryfikowanych przez logarytmiczny test Jianga i Oomena.**



**7.14. Wartości skoków w procesie kursów logarytmicznych USD/PLN wyznaczonych na podstawie zależności (4.10), pozytywnie zweryfikowanych przez logarytmiczny test Barndorff-Nielsen i Shepharda.**



**7.15. Wartości skoków w procesie kursów logarytmicznych USD/PLN wyznaczonych na podstawie zależności (4.13), pozytywnie zweryfikowanych przez logarytmiczny test Jianga i Oomena.**

Wartości wykrytych skoków, podobnie jak ich liczba, są dosyć zbliżone, chociaż skoki wyznaczone na podstawie różnicy wariancji zrealizowanej i wariacji swapowej przyjmują nieco wyższe wartości. Największe rozbieżności widoczne są w procesie logarytmicznych kursów EURPLN. Natomiast wartości skoków są sobie najbliższe w przypadku logarytmicznych poziomów indeksu DAX. Zbliżone wielkości skoków otrzymane za pomocą dwóch zupełnie różnych podejść, świadczą o tym, że jakość oszacowań skoków jest satysfakcjonująca.

### **7.3. Wnioski końcowe**

Przeprowadzone badania empiryczne mogą sugerować, że w procesach logarytmicznych poziomów indeksów giełdowych skoki, które interpretujemy jako reakcję ceny na informacje makroekonomiczne lub następstwo zamykania pozycji przez inwestorów pod koniec dnia sesji giełdowej, są dość powszechnym zjawiskiem. Zjawisko to obserwujemy zarówno na rynkach dojrzałych, takich jak rynki zachodnioeuropejskie czy amerykański, jak i rynkach mniej dojrzałych, których przykładem jest rynek polski. Nieco rzadziej występuje ono w przypadku indeksu CAC40, natomiast dla procesów poziomów logarytmicznych WIG20 i NASDAQ, przeprowadzone testy wykryły skok w około 30% dni sesyjnych. Wniosek ten jest również potwierdzony przez testy Lee-Mykland (2007). Masowe występowanie skoków, jakie zaobserwowaliśmy w modelowanych procesach, silnie przemawia za stosowaniem do ich opisu modeli dyfuzji ze skokami. Silna leptokurtoza w szeregach zwrotów logarytmicznych, którą zaobserwowaliśmy we wszystkich szeregach danych, jest dodatkowym uzasadnieniem stosowania tych modeli do opisu badanych procesów. Częstotliwość danych ma dość istotny wpływ na otrzymane wyniki. Rozbieżności w liczbie wykrytych skoków oscylują w granicach 10-20% i są uzależnione głównie od precyzji estymacji wariancji zrealizowanej oraz zrealizowanych wariacji wielopotęgowych.

Wykryty w większości szeregów czasowych szum mikrostruktury uzasadnia stosowanie przesuniętych wariacji wielopotęgowych jako oszacowań wariancji scałkowanej i kwartyczności przy wyznaczaniu wartości statystyk Barndorff-Nielsen i Shepharda. Jednocześnie obserwujemy, że zastosowanie przesuniętych wariacji wielopotęgowych jako estymatorów wariacji scałkowanej i kwartyczności znacząco wpłynęło na liczbę wykrytych skoków.

Możliwość szacowania wartości skoków w oparciu o zależności (4.10) i (4.13) daje podstawę do opisu procesów generujących dane dzienne za pomocą modeli dyfuzji ze skokami. Estymacji parametrów można dokonać poprzez pierwotne odseparowanie skoków z szeregu notowań dziennych, dopasowanie do szeregu rozmiarów skoków odpowiedniego rozkładu, np.  $N(0, \sigma_j^2)$ , jak przyjęliśmy w specyfikacji modelu (4.18), i opis występowania skoków za pomocą procesu Poissona. Następnie proces z odseparowanymi skokami można modelować za pomocą metody Phillipsa i Yu.

## Zakończenie

Niniejsza praca dotyczy problematyki modelowania i prognozowania procesów stochastycznych z czasem ciągłym, generujących ceny instrumentów finansowych. Zastosowana do estymacji parametrów modeli dyfuzji dwukrokowa metoda Phillipsa i Yu dała zadowalające rezultaty, na co wskazały przeprowadzone badania symulacyjne. Istotnym wnioskiem płynącym z przedstawionych badań jest konstatacja, iż jakość oszacowań otrzymanych za pomocą danych intraday jest znacznie lepsza niż przy wykorzystaniu danych dziennych. Ponadto porównanie wyników estymacji uzyskanych na podstawie danych różnej częstotliwości może sugerować, że zastosowanie danych półgodzinnych daje wyniki bardzo zbliżone do tych, które otrzymujemy dla danych wyższych częstotliwości. Wniosek ten pozwala na dużą oszczędność czasu nie tylko w samym procesie estymacji, którego czasochłonność przy obecnie wykorzystywanych komputerach nie stanowi zasadniczego problemu, ale przede wszystkim w bardzo pracochłonnym procesie opracowywania danych do estymacji. Jednocześnie dla danych półgodzinnych marginalizuje się problem wpływu efektów mikrostruktury na wyniki estymacji. Kolejny istotny wniosek wypływa z przeprowadzenia estymacji parametrów modeli dyfuzji dla tych samych szeregów za pomocą podejścia zaproponowanego przez Phillipsa i Yu oraz bardziej klasycznie, uogólnioną metodą momentów. Otrzymane wyniki sugerują, że metoda Phillipsa i Yu daje dokładniejsze oszacowania. Do innego ważnego wniosku dochodzimy zestawiając prognozy średniej warunkowej uzyskane różnymi sposobami. Porównując standardowe miary błędów prognoz, uzyskanych za pomocą modeli dyfuzji z parametrami estymowanymi obydwoima metodami z błędami otrzymanymi za pomocą modeli ARIMA-GARCH, zauważamy, że te pierwsze są w wielu przypadkach mniejsze. Wobec tego możemy obronić hipotezę, że własności prognostyczne modeli dyfuzji są dobre. Kolejny interesujący wniosek można wyciągnąć na podstawie badań symulacyjnych opisanych w rozdziale 4. Badania te wykazały znaczny wpływ częstotliwości notowań, wykorzystywanych do wyznaczania wariancji zrealizowanej oraz zrealizowanych wariacji wielopotęgowych, na wyniki testów na występowanie skoków. Wyniki uzyskane za pomocą trzech różnych podejść do testowania skoków jednoznacznie wskazują, że skoki w procesach logarytmicznych poziomów indeksów giełdowych oraz logarytmicznych kursach walutowych są zjawiskiem powszech-

nym, co jest silnym argumentem przemawiającym za modelowaniem procesów cen logarytmicznych za pomocą modeli dyfuzji ze skokami.

Przeprowadzone badania pozytywnie weryfikują postawione w pracy cztery hipotezy badawcze. Należy jednak zaznaczyć, że tematyka modelowania finansowych procesów za pomocą modeli z czasem ciągłym jest obecnie bardzo intensywnie rozwijana. Ukazujące się w ostatnim czasie nowe prace dają inspirację do dalszych badań empirycznych mających na celu weryfikację jakości wyników otrzymanych za pomocą nowych teoretycznych modeli oraz podjęcia kroków służących zmniejszeniu obciążenia uzyskanych oszacowań parametrów.

Ciekawym obszarem badań pozostają modele zmienności stochastycznej z czasem ciągłym, chociaż ich szybka, a zarazem dokładna, estymacja nadal jest dużym wyzwaniem. Trudności z jej przeprowadzeniem wiążą się między innymi z koniecznością modelowania nieobserwowalnego procesu wariancji chwilowej charakteryzującego się zazwyczaj burzliwą dynamiką. Jedno z podejść do estymacji modeli zmienności stochastycznej zostało wprowadzone w niniejszej pracy, niemniej w najnowszej literaturze znajdujemy wiele innych ciekawych metod szacowania ich parametrów. Gallant i Tauchen (1996) wykorzystują do estymacji modelu zmienności stochastycznej efektywną metodę momentów. Hansen i Scheinkman (1995) oraz Conley i in. (1997) proponują podejście bazujące na uogólnionej metodzie momentów. Aït-Sahalia i Kimmel (2007) szacują parametry modelu Hestona, ciągłego modelu GARCH oraz modelu CEV-SV metodą największej wiarygodności, estymując gęstości przejścia za pomocą rozwinięcia w wielomian Hermite'a, podobnie jak było to robione w pracach Aït-Sahalii (1999, 2001) dla modeli dyfuzji. Ciekawe podejście prezentuje również Renò (2006), proponując stosunkowo łatwą w implementacji metodę estymacji funkcji dryfu i funkcji dyfuzji w modelu zmienności stochastycznej.

Powszechne występowanie skoków w procesach cen logarytmicznych, które wykazaliśmy w ostatnim rozdziale pracy, uzasadnia stosowanie modeli dyfuzji ze skokami do modelowania tych procesów. Należy również pamiętać o tym, że istnieją już modele wyceny bardzo złożonych instrumentów pochodnych oparte na procesach dyfuzji ze skokami. Modelowanie tych procesów jest na razie bardzo słabo rozwinięte. Aït-Sahalia (2004) proponuje metodę pozwalającą na precyzyjne oddzielenie skoków w procesie dyfuzji ze skokami od zaburzeń generowanych przez ruch Browna. Jiang i Oomen (2007) estymują parametry modeli dyfuzji ze skokami za pomocą uogólnionej

metody momentów, natomiast Mancini i Renò (2008) wykorzystują podejście bazujące na wprowadzonych przez Mancini (2004) progach skoków.

Wymienione powyżej metody w znacznej części nie zostały jeszcze zbadane pod względem wykorzystania ich do rzeczywistych szeregów czasowych, a więc w dalszym ciągu jest to interesujący obszar badawczy. Ciekawe może również okazać się porównanie wartości instrumentów pochodnych uzyskanych za pomocą podejść klasycznych oraz nowoczesnych modeli wyceny bazujących na modelach zmienności stochastycznej czy modelach dyfuzji ze skokami.

## Dodatek A

W dodatku przedstawione są standardowe miary błędów wykorzystane w niniejszej pracy do oceny jakości prognoz (por. Welfe 1998, Doman, Doman 2004).

1. Błąd średniokwadratowy

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (X_{T+i} - \hat{X}_{T+i})^2.$$

2. Średni błąd medianowy

$$\text{MedSE} = \text{mediana}\{(X_{T+i} - \hat{X}_{T+i})^2\}_1^N.$$

3. Błąd średni

$$\text{ME} = \frac{1}{N} \sum_{i=1}^N (X_{T+i} - \hat{X}_{T+i}).$$

4. Średni błąd bezwzględny

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |X_{T+i} - \hat{X}_{T+i}|.$$

5. Pierwiastek błędu średniokwadratowego

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_{T+i} - \hat{X}_{T+i})^2}.$$

6. Średni bezwzględny błąd procentowy

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{X_{T+i} - \hat{X}_{T+i}}{X_{T+i}} \right|,$$

7. Skorygowany bezwzględny błąd procentowy

$$\text{AMAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{X_{T+i} - \hat{X}_{T+i}}{X_{T+i} + \hat{X}_{T+i}} \right|,$$

8. Logarytmiczna funkcja straty oznaczona jest przez LL i dana jest wzorem

$$\text{LL} = \frac{1}{N} \sum_{i=1}^N \left( \ln \frac{\hat{X}_{T+i}}{X_{T+i}} \right)^2.$$



W powyższych formułach  $N$  oznacza liczbę prognoz natomiast  $T$  liczbę wyrazów w próbie.

# Literatura

- Aït-Sahalia Y. (1999) *Transition Densities for Interest Rate and Other Nonlinear Diffusions*, Journal of Finance 54, 1361–1395.
- Aït-Sahalia Y. (2002) *Telling From Discrete Data Whether the Underlying Continuous-Time Model is a Diffusion*, The Journal of Finance 57, 2075–2112.
- Aït-Sahalia Y. (2004) *Disentangling diffusion from jumps*, Journal of Financial Economics 74, 487–528.
- Aït-Sahalia Y. (2008) *Closed-Form Likelihood Expansions for Multivariate Diffusions*, Annuals of Statistics.
- Aït-Sahalia Y., Kimmel R. (2007) *Maximum Likelihood Estimation of Stochastic Volatility Models*, Journal of Financial Economics 83 413–452.
- Aït-Sahalia Y., Mykland P.A., Zhang L. (2005) *How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise*, Review of Financial Studies 18, 351–416.
- Andersen T.G., Benzoni L., Lund J. (2002) *An Empirical Investigation of Continuous-Time Equity Return Models*, The Journal of Finance 57 (3) 1239–1284.
- Andersen T.G., Bollerslev T. (1998a) *Answering The Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts*, International Economic Review 39, 885–905.
- Andersen T.G., Bollerslev T. (1998b) *Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies*, Journal of Finance 53, 219-265.
- Andersen T.G., Bollerslev T., Diebold F.X. (2003a) *Some Like It Smooth, and Some Like It Rough: Untangling Continuous and Jump Components in Measuring, Modeling and Forecasting Asset Return Volatility*, Manuscript University of Pennsylvania.
- Andersen T.G., Bollerslev T., Diebold F.X. (2007a) *Roughing It Up Including Jumps Componentes in Measurement, Modeling and Forecasting of Return Volatility*, The Review of Economics and Statistics 89 (4) 701–720.
- Andersen T.G., Bollerslev T., Diebold F.X., Ebens H. (2001) *The Distribution of Realised Stock Return Volatility*. Journal of Financial Economics 61, 43–76.
- Andersen T.G., Bollerslev T., Diebold F.X. Labys P. (2000) *Exchange Rate Returns Standardised by Realized Volatility Are (Nearly) Gaussian*, Multinatl. Financial Journal 4, 159–179.
- Andersen T.G., Bollerslev T., Diebold F.X., Vega C. (2003b) *Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange*, American Economic Review 93, 38–62.
- Andersen T.G., Bollerslev T., Diebold F.X., Vega C. (2007b) *Real-Time Price Discovery in Stock, Bond and Foreign Exchange Markets*, Journal of International Economics 73 (2) 251–277
- Andrews D.W.K. (1991) *Heteroskedasticity and autocorrelation consistent covariance matrix estimation*, Econometrica 49, 817–858.
- Andrews, D.W.K., Monahan J. C. (1992) *An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator*, Econometrica 60, 953–966.
- Areal N.M.P.C., Taylor S.J. (2002) *The Realised Volatility of FTSE-100 Future Prices*. Journal of Future Markets 22, 627–648.

- Ashby F.G. (2000) *A Stochastic Version of General Recognition Theory*. Journal of Mathematical Psychology 44, 310–29
- Bachelier L., (1900) *Théorie de la spéculation*, *Annales Scientifiques de l'Ecole Normale Supérieure*, III-17, 21–86.
- Bain A. (2008) *Stochastic Calculus*, <http://www.chiark.greenend.org.uk/~alanb/>
- Bandi F.M., Phillips P. C. B. (2007) *A Simple Approach to the Parametric Estimation of Potentially Non-stationary Diffusions*, Journal of Econometrics 137, 354–395.
- Bandi F.M., Russell J. R. (2006) *Separating Microstructure Noise from Volatility* Manuscript, Journal of Financial Econometrics 79, 655–692.
- Barndorff-Nielsen O. E., Shephard N. (2002) *Econometric Analysis of Realised Covariation: High Frequency Covariance, Regression and Correlation in Financial Economics*. Unpublished paper: Nuffield College, Oxford, Economics working paper 2002–W13.
- Barndorff-Nielsen O. E., Shephard N. (2003) *Realised Power Variation and Stochastic Volatility*, Bernoulli 9, 243–265.
- Barndorff-Nielsen, O. E., Shephard N. (2004a) *Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models*, Journal of the Royal Statistical Society 64, Series B, 253–280.
- Barndorff-Nielsen O. E., Shephard N. (2004b) *Multipower Variation and Stochastic Volatility*, Oxford Financial Research Centre series OFRC Working Papers Series 2004fe22.
- Barndorff-Nielsen O. E., Shephard N. (2004c) *Power and bipower variation with stochastic volatility and Jumps*. Journal of Financial Econometrics 2 (1) 1–37.
- Barndorff-Nielsen O. E., Shephard N. (2005) *How Accurate is the Asymptotic Approximation to the Distribution of Realized Volatility?. In Identification and Inference for Econometric Models*, eds by D.W.K. Andrews, J. Powell, P. Ruud and J. Stock, Cambridge University Press
- Barndorff-Nielsen O. E., Shephard N. (2006) *Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation*, Journal of Financial Econometrics 4, 1–30.
- Barndorff-Nielsen O. E., Shephard N. Winkel M. (2006) *Limit Theorems for Multipower Variation in the Presence of Jumps*, Stochastic Processes and their Applications 116 796–806.
- Barndorff-Nielsen O.E., Shephard N. (2005) *How Accurate is the Asymptotic Approximation to the Distribution of Realized Volatility?. In Identification and Inference for Econometric Models*, eds by D.W.K. Andrews, J. Powell, P. Ruud and J. Stock, Cambridge University Press
- Barndorff-Nielsen O.E., Shephard N. (2006) *Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation*, Journal of Financial Econometrics 4, 1–30.
- Barndorff-Nielsen O.E., Shephard N., Winkel M. (2006) *Limit Theorems for Multipower Variation in the Presence of Jumps*, Stochastic Processes and their Applications 116 796–806.
- Black F. (1976) *The Pricing of Commodity Contracts*, Journal of Financial Economics, 3, 167–179.
- Black F., Scholes M. (1973) *The Pricing of Option and Corporate Liabilities*, Journal of Political Economy 81, 637–659.
- Brandt M.W., Santa-Clara P. (2002) *Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets*. Journal of Financial Economics 30, 161–210.

- Broadie M., Jain A. (2008) *Pricing and Hedging Volatility Derivatives*, Journal of Derivatives 15 (3) 7–24.
- Broyden C.G. (1970) *The Convergence of a Class of Double-rank Minimization Algorithm*, Journal of the Institute of Mathematics and Its Applications 6, 76–90.
- Burnside A.C., Eichenbaum M. (1994) *Small Sample Properties of Generalized Method of Moments Based Wald Tests*, NBER Working Paper No. T0155. Available at SSRN: <http://ssrn.com/abstract=225114>
- Busemeyer J.R., Townsend J.T. (1993) *Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment*, Psychological Review 100, 432–59.
- Carr P., Wu L. (2003) *What Type of Process Underlies Options? A Simple Robust Test*, The Journal of Finance 58 (6) 2581–2610.
- Chan K.C., Karolyi G.A., Longstaff F.A., Sanders A.B. (1992) *An Empirical Comparison of Alternative Models of Short Term Interest Rates*, Journal of Finance 47, 1209–1227.
- Chernov M., Gallant A.R., Ghysels E., Tauchen G. (2003) *Alternative Models for Stock Price Dynamics*, Journal of Econometrics 116, 225–257.
- Charemza W.W., Deadman D. (1997) *Nowa Ekonomia*, PWE, Warszawa.
- Chen L. (1996) *Stochastic Mean and Stochastic Volatility – A Three-Factor Model of the Term Structure of Interest Rates and Its Application to the Pricing of Interest Rate Derivatives*. Blackwell Publishers.
- Cliff M.T. (2003) *GMM and MINZ Program Libraries for Matlab*, Krannert Graduate School of Management Purdue University.
- Conley T., Hansen L.P., Luttmer E., Scheinkman J. (1997) *Short term interest rates as subordinated diffusions*, Review of Financial Studies 10, 525–578.
- Courtadon G. (1982) *The Pricing of Options on Default-Free Bonds*, Journal of Financial and Quantitative Analysis 17, 75–100.
- Cox J.C. (1975) *Notes On Option Pricing I: Constant Elasticity of Variance Diffusions*, Working paper, Stanford University.
- Cox J.C., Ingersoll J., Ross S. (1980) *An analysis of variable rate loan contracts*, Journal of Finance 35, 389–403.
- Cox J.C., Ingersoll J., Ross S. (1985) *A Theory of the Term Structure of Interest Rates*, Econometrica, 53, 385–407.
- Cziraky D., Kucherenko S. (2008) *Monte Carlo Forecasting from CIR Square Root Diffusion Models*, BRODA Ltd. <http://www.broda.co.uk>
- Davidon W.C. (1991) *Variable Metric Method for Minimization*, SIAM Journal on Optimization 1, 1-17
- Davidson J. (2008) *Time Series Modelling Version 4.26*, <http://www.timeseriesmodelling.com/>
- Deen Haan W.J., Levin A. (1996) *A practitioner's Guide to Robust Covariance Matrix Estimation*, University of California, Discussion paper 96-17
- Detemple J., Osakwe S. (1999) *The Valuation of Volatility Options*, CIRANO Paper, Scientific Series.
- Dickey D.A. Fuller W.A. (1979) *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*, Journal of the American Statistical Association 74, 427–431.

- Doman M., Doman R. (2004) *Ekonomiczne modelowanie dynamiki polskiego rynku finansowego*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- Dothan U.L. (1978) *On the term structure of interest rates*, Journal of Financial Economics 6, 59–69.
- Elliott R., Siu T., Chan L. (2007) *Pricing Volatility Swaps Under Heston's Stochastic Volatility Model with Regime Switching*, Applied Mathematical Finance 14 (1) 41–62.
- Engle R.F., (1982) *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*, Econometrica 50, 987–1007.
- Fletcher R. (1970) *A New Approach to Variable Metric Algorithms*, Computer Journal 13, 317–322.
- Florens-Zmirou D., (1989) *Approximate Discrete-time Schemes for Statistics of Diffusion Processes*, Statistics 20, 547–557.
- Fuller W.A. (1976) *Introduction to Statistical Time Series*, John Wiley, New York.
- Goldfarb D. (1970) *A Family of Variable Metric Updates Derived by Variational Means*, Mathematics of Computation 24, 23–26
- Gallant A.R., Tauchen, G.E. (1996) *Which Moment to Match?* Econometric Theory 12, 657–681.
- Gallant A.R. (1987) *Nonlinear Statistical Models* (John Wiley & Sons: New York).
- Gourieroux C., Jasiak J. (2001) *Financial Econometrics*, Princeton University Press, Princeton.
- Hagan P. S., Kumar D., Lesniewski A.S., Woodward. D.E., *Managing smile risk*, Wilmott Magazine, 84–108, July 2002.
- Hamilton J.D. (1994) *Time Series Analysis*, Princeton University Press, Princeton.
- Hansen L.P. (1982) *Large Sample Properties of Generalized Method Of Moments Estimators*, Econometrica 50, 1029–1054.
- Hansen L.P., Hodrick R.J. (1980) *Forward exchange rates as optimal predictors of future spot rates: An empirical analysis*, Journal of Political Economy 88, 829–853.
- Hansen P.R., Lunde A. (2002) *Volatility Estimation Using High Frequency Data With Partial Availability*, Discussion Paper. Brown University Working Paper.
- Hansen P.R., Lunde A. (2005) *A Forecast Comparison of Volatility Models: does anything beat a GARCH(1,1)?* Journal of Applied Econometrics 20, 873–889.
- Hansen P.R., Lunde A. (2006) *Realized Variance and Market Microstructure Noise*, Journal of Business and Economic Statistics 24, 127–218.
- Hansen L.P., Scheinkman, J.A. (1995) *Back to the future: generalized moment implications for continuous time Markov process*, Econometrica 63, 767–804.
- Heston S.L. (1993) *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies 6 (2) 327–343
- Howison S., Rafailidis A., Rasmussen H. (2004) *On the Pricing and Hedging of Volatility Derivatives*, Applied Mathematical Finance 11 (4) 317–346.
- Huang X., Tauchen G. (2005) *The Relative Contribution of Jumps to Total Price Variation*, Journal of Financial Econometrics 3, 456–499.
- Hull J., White A. (1987) *The Pricing of Options on Assets with Stochastic Volatilities*, Journal of Finance 42, 281–300.
- Itô K. (1951) *On Stochastic Differential Equations*, Memoirs, American Mathematical Society 4, 1–51.

- Jacod J. (1994) *Limit of random measures associated with the increments of a Brownian semimartingale*, Working paper, Laboratoire de Probabilités, Université Pierre et Marie Curie, Paris.
- Jagannathan R., Kaplin A., Sun S. G. (2004) *An Evaluation of Multi-Factor CIR Models Using Libor, Swap Rates, and Cap and Swaption Prices*, Journal of Econometrics 116, 113–146.
- Jakubowski J., Palczewski A., Rutkowski M., Stettner Ł. (2003) *Matematyka finansowa – instrumenty pochodne*, Wydawnictwo Naukowo Techniczne, Warszawa.
- Javaheri A., Wilmott P., Haug E. (2002) *GARCH and Volatility Swaps*, Wilmott Magazine, January 1–17.
- Jiang G.J., Oomen, R.C.A. (2007) *Estimating Latent Variables and Jump Diffusion Models Using High Frequency Data*, Journal of Financial Econometrics 5 (1) 1–30.
- Jiang G.J., Oomen, R.C.A. (2008) *Testing for Jumps When Asset Prices are Observed with Noise - A Swap Variance Approach*, Journal of Econometrics 144 (2) 352–370.
- Johannes M., Polson N., Stroud J. (2004a) *Nonlinear Filtering of Stochastic Differential Equations with Jumps*, manuscript GSB Columbia University.
- Johannes M., Polson N., Stroud J. (2004b) *Sequential Parameter Estimation in Stochastic Volatility Models with Jumps*, manuscript GSB Columbia University.
- Kloeden P.E., Platen E. (1992) *Numerical Solution of Stochastic Differential Equations*, Springer Verlag, Berlin.
- Koopman S.J., Jungbacker B., Hol E. (2005) *Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realized and Implied Volatility Measurements*, Journal of Empirical Finance 12, 445–475.
- Kostrzewski M. (2006a) *Bayesowska analiza finansowych szeregów czasowych modelowanych procesami dyfuzji*, Wydawnictwo AGH, Kraków.
- Kostrzewski M. (2006b) *Bayesowska estymacja, prognoza i porównywanie procesów Ito modelujących finansowe szeregi czasowe*, w: *Metody Ilościowe w Naukach Ekonomicznych*, SGH, Warszawa.
- Kou S.G. (2000) *A jump diffusion model for option pricing with three properties: leptokurtic feature, volatility smile, and analytical tractability*, Computational Intelligence for Financial Engineering, Proceedings of the IEEE/IAFE/INFORMS 2000.
- Kou S.G. (2002) *A Jump-Diffusion Model for Option Pricing*, Management Science 48 (8) 1086–1101.
- Kou S.G., Wang H. (2004) *Option Pricing Under a Double Exponential Jump Diffusion Model*, Management Science 50 (9) 1178–1192
- Lánska V. (1979) *Minimum Contrast Estimation in Diffusion Processes*, Journal of Applied Probability, 16, 65–75.
- Lahaye J., Laurent S., Neely C. J. (2007) *Jumps, Cojumps and Macro Announcements*, Working Paper 2007–032A Federal Reserve Bank of St. Louis Research Division P.O. Box 442 St. Louis, MO 63166.
- Lee S.S., Mykland P.A. (2007) *Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics*, Review of Financial Studies, doi:10.1093/rfs/hhm056.
- Ljung G.M., Box G.E.P. (1978) *On a measure of lack of fit in time series models*, Biometrika 65, 297–303.

- Newey W., West K. (1987) *A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix*, *Econometrica* 55, 703–708.
- MacKinnon J.G. (1991) *Critical values for cointegration tests*, chapter 13 in *Long-run Economic Relationships: Readings in Cointegration*, eds. R. F. Engle and C. W. J. Granger, Oxford, Oxford University Press.
- Matychak Ya.S., Pavlyna V. S., Fedirko V. M. (1998) *Diffusion Processes and Mechanics of Materials*, *Materials Science* 34 (3), 13–23.
- McLeod A.I., Li W.K. (1983) *Diagnostic checking ARMA time series models using squared residual autocorrelations*, *Journal of Time Series Analysis* 4, 269–273.
- Mancini C. (2004) *Estimation of the characteristics of jump of a general Poisson-diffusion process*, *Scandinavian Actuarial Journal* 2004 (1), 42–52.
- Mancini C., Renò R. (2008) *Threshold Estimation of Jump-Diffusion Models and Interest Rate Modeling*, SSRN Working paper series.
- Mannolini A., Mari C., Renò R. (2008) *Pricing caps and floors with the extended CIR model*, *International Journal of Finance & Economics*, Volume 13 (4) 386–400
- Maruyama G. (1955) *Continuous Markov Processes and Stochastic Equations*, *Rend. Circ. Math. Palermo* 4, 1–43.
- Martens M. (2002) *Measuring and Forecasting S&P 500 Index-Futures Volatility Using High-Frequency Data*, *Journal of Futures Markets* 22, 497–518
- Merton R.C. (1973) *Theory of rational option pricing*, *Bell Journal of Economics and Management Science* 4 (1) 141–183.
- Merton R.C. (1974) *On The Pricing of Corporate Debt: The Risk Structure of Interest Rates*, *Journal of Finance* 29, 449–470.
- Milstein G.N. (1974) *Approximate Integration of Stochastic Differential Equations*, *Theory of Probability and Its Applications* 19, 557–562.
- Milstein G.N. (1978) *A Method of Second-Order Accuracy Integration of Stochastic Differential Equations*, *Theory of Probability and Its Applications* 23, 396–401.
- Nelson D.B. (1990) *ARCH Models as Diffusion Approximations*, *Journal of Econometrics* 45, 7–38.
- Neuberger A. (1994) *The Log Contract: A New Instrument to Hedge Volatility*, *The Journal of Portfolio Management*, Winter, 74–80.
- Newey W.K., West K.D. (1987) *A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*, *Econometrica* 59, 347–370.
- Nosofsky R.M., Palmeri T.J. (1997) *An Exemplar Based Random Walk Model of Speeded Classification*, *Psychological Review* 104, 266–300.
- Osińska M. (2006) *Ekonometria Finansowa*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Perron P. (1988) *Trends and Random Walks in Macroeconomic Time Series*, *Journal of Economic Dynamics in Control* 12, 297–332.
- Phillips P.C.B., Perron P. (1988) *Testing for a Unit Root in Time Series Regressions*, *Biometrika* 75, 335–346.

- Phillips P.C.B., Yu J. (2007) *A Two-Stage Realized Volatility Approach to Estimation of Diffusion Processes with Discrete Data*, Journal of Econometrics, ukaze się w 2009.
- Platen E. (1999) *An Introduction to Numerical Methods for Stochastic Differential Equations*, Acta Numerica, 197–246.
- Protter P. (2005) *Stochastic Integration and Differential Equation*, 2<sup>nd</sup> edition, version 2.1, Springer Verlag, Berlin.
- Psychoyios D., Skiadopoulos G. (2006) *Volatility Options: Hedging Effectiveness, Pricing, and Model Error*, Journal of Futures Markets 26, 1–31.
- Ratcliff R. (1978) *A Theory of Memory Retrieval*, Psychological Review 85, 59–108.
- Ratcliff R., Van Zandt T., McKoon G. (1999) *Connectionist and Diffusion Models of Reaction Time*, Psychological Review 106, 261–300
- Renò R. (2006) *Nonparametric Estimation of Stochastic Volatility Models*, Economics Letters 90 (3) 390–395.
- Saito Y., Mitsui T. (1996) *Stability Analysis of Numerical Schemes for Stochastic Differential Equations*, SIAM Journal of Numerical Analysis 88 (6) 2254–2267.
- Said E., Dickey D.A. (1984) *Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order*, Biometrika 71, 599–607.
- Shumway R.H., Stoffer D.S. (2006) *Time Series Analysis and its Applications*, Springer, New York.
- Shreve S. (2004) *Stochastic Calculus and Finance*, Springer Verlag, New York.
- Sepp A. (2008a) *Pricing Options on Realized Variance in Heston Model with Jumps in Returns and Volatility*, Journal of Computational Finance 11 (4) 33–70.
- Sepp A. (2008b) *VIX option pricing in a jump-diffusion model*, Risk, April, 84–89.
- Shanno D.F. (1970) *Conditioning of Quasi-Newton Methods for Function Minimization*, Mathematics of Computation 24, 647–656.
- Vašiček O. (1977) *An Equilibrium Characterization of the Term Structure*, Journal of Financial Economics 5, 177–188.
- Verbeek M. (2005) *A Guide to Modern Econometrics*, John Willey and Sons Ltd., West Sussex.
- Welfe A. (1998) *Ekonometria*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Weron A., Weron R. (1999) *Inżynieria finansowa*, Wydawnictwo Naukowo-Techniczne, Warszawa.
- Yoshida N. (1992) *Estimation for Diffusion Processes From Discrete Observation*, Journal of Multivariate Analysis 41, 220–242.
- Zhang L., Aït-Sahalia Y., Mykland P. A. (2005) *A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data*, Journal of the American Statistical Association 100, 1394–1411.