Bank i Kredyt 45(2), 2014, 105–132

BANK I KREDYT
BANK & CREDIT

Model GARCH – wykorzystanie dodatkowych informacji o cenach minimalnych i maksymalnych

Grzegorz Perczak*, Piotr Fiszeder#

Nadesłany: 19 marca 2013 r. Zaakceptowany: 12 listopada 2013 r.

Streszczenie

W pracy zaprezentowano modele GARCH wprowadzone przez Lildholdta (2002) oraz Ventera, de Jongha i Griebenowa (2005), które zostały skonstruowane na podstawie cen minimalnych, maksymalnych i zamknięcia. Zakładając, że śróddzienne stopy zwrotu mogą być opisane przez arytmetyczny ruch Browna z rozkładem NIG, przedstawiono łączne rozkłady wektorów losowych. Ich współrzędnymi są zmienne losowe wartości minimalnej, maksymalnej i końcowej logarytmicznych stóp zwrotu. Rozkłady te zostały następnie wykorzystane do skonstruowania funkcji wiarygodności, służących do estymacji parametrów modeli.

Ponadto w pracy zaproponowano rozszerzenie modeli Lildholdta (2002) oraz Ventera, de Jongha, Griebenowa (2005). Polega ono na zastosowaniu bardziej efektywnych estymatorów dziennej wariancji, konstruowanych na podstawie cen minimalnych i maksymalnych, zamiast estymatora wyznaczanego wyłącznie na podstawie cen zamknięcia. Dokonano również pewnych uproszczeń wspomnianych parametryzacji modeli. Na podstawie szeregów stóp zwrotu z indeksu WIG20 i kursu walutowego EUR/PLN pokazano, że wykorzystanie informacji o cenach minimalnych i maksymalnych do oszacowania parametrów modelu GARCH, bez zwiększenia jego parametryzacji, poprawia jakość modelu mierzoną wartością funkcji wiarygodności.

Słowa kluczowe: model GARCH, estymacja zmienności, rozkład NIG, ruch Browna, ceny minimalne i maksymalne

JEL: C13, C22, C51

* Uniwersytet Mikołaja Kopernika w Toruniu; e-mail: grzegorz.perczak@gmail.com.
# Uniwersytet Mikołaja Kopernika w Toruniu; e-mail: piotr.fiszeder@umk.pl.

1. Wstęp

Zmienność cen instrumentów finansowych jest jednym z ważniejszych pojęć współczesnych finansów. Znaczenie zmienności wynika zarówno z teorii finansów, jak i z licznych zastosowań praktycznych (Schwert 1989; Fleming, Kirby, Ostdiek 2001; Poon, Granger 2003; Andersen i in. 2006). Wyniki badań empirycznych wskazują, że rozszerzanie istniejących już modeli zmienności stóp zwrotu i tworzenie kolejnych parametryzacji w celu coraz lepszego dopasowania modelu do danych empirycznych jest mniej efektywne niż wykorzystanie informacji powszechnie dostępnych na rynku (Fiszeder 2009). Do niedawna modele parametryczne stosowane do opisu zmienności stóp zwrotu były konstruowane wyłącznie na podstawie cen zamknięcia. Dodatkowe dane rynkowe, które można uwzględnić przy budowie modeli zmienności, dzieli się na dwie grupy: 1) dane o częstotliwości wyższej niż dzienna, tzw. dane *intraday*, które są zbiorem uporządkowanych par zawierających wartość ceny i czas jej odnotowania, 2) informacje o minimalnych i maksymalnych cenach. W tym przypadku nie jest jednak uwzględniany moment ani nawet kolejność ich wystąpienia.

Obecnie znanych jest wiele parametryzacji modeli służących do opisu zmienności stóp zwrotu na rynkach finansowych; wśród nich najliczniejszą klasę stanowią modele GARCH. Rozważania prowadzone w niniejszej pracy będą dotyczyły właśnie tej grupy modeli, w przyszłości mogą być jednak rozszerzone na inne ich rodzaje, np. modele SV (ang. *stochastic volatility*).

Zastosowanie notowań *intraday* dostarcza więcej informacji o badanym instrumencie finansowym, wiąże się jednak z koniecznością pokonania szeregu problemów:

1. Dane o częstotliwości wyższej niż dzienna nie zawsze są dostępne dla wszystkich aktywów, w przeciwieństwie do cen zamknięcia oraz cen minimalnych i maksymalnych.

2. Pozyskanie danych *intraday* wymaga z reguły poniesienia dodatkowych kosztów, ponieważ nie są ogólnie dostępne.

3. Zastosowanie danych *intraday* z dłuższych okresów wymaga bardzo dużych baz danych, liczących setki tysięcy obserwacji, co znacznie wydłuża czas prowadzenia analiz.

4. Czynniki związane z mikrostrukturą rynku (Doman 2011), szczególnie w przypadku mało płynnych aktywów, powodują znaczne obciążenie estymatorów konstruowanych na podstawie danych śróddziennych.

5. Dane o wysokiej częstotliwości mają wiele cech utrudniających ich bezpośrednią analizę. Przykładowo występują silne wahania cykliczne w ciągu dnia i silna autokorelacja, informacje makroekonomiczne mają duży wpływ na notowania cen instrumentów finansowych; zmienność obliczona na ich podstawie może być znacznie przeszacowana lub niedoszacowana; często wymagają też podjęcia dodatkowych działań, np. filtrowania.

6. Informacje o cenach w ciągu dnia nie są na ogół wykorzystywane do budowy modelu GARCH i funkcji wiarygodności służącej do estymacji parametrów tego modelu dla okresów dziennych, a pojawiające się w literaturze propozycje wymagają znacznego rozbudowania postaci modelu i zwiększenia liczby jego parametrów. Do nielicznych wyjątków można zaliczyć prace Lildholdta (2002) oraz Ventera, de Jongha i Griebenowa (2005), w których dodatkowo wykorzystano wartości minimalne i maksymalne w ciągu dnia.

Korzystanie w modelowaniu finansowym z informacji o wartościach minimalnych i maksymalnych jest uzasadnione. Estymator Parkinsona (1980), najprostszy i najmniej efektywny spośród estymatorów dziennej wariancji (termin ten określa w dalszej części pracy wariancję stopy zwrotu zaobserwowaną między zamknięciami notowań w kolejnych dniach), konstruowanych na podstawie cen minimalnych, maksymalnych i zamknięcia, daje podobne wyniki jak zmienność zrealizowana, szacowana na podstawie czterech, pięciu lub sześciu obserwacji w ciągu dnia (Parkinson 1980; Andersen, Bollerslev 1998). Ponadto Parkinson pokazał, że w przypadku arytmetycznego ruchu Browna zaproponowany przez niego estymator wariancji jest ponadpięciokrotnie efektywniejszy od estymatora konstruowanego na podstawie cen zamknięcia.

Powstanie modeli klasy GARCH umożliwiło opisanie zjawiska tzw. grupowania zmienności stóp zwrotu. Początkowo przyjmowano założenie o warunkowej normalności składnika losowego, które w przypadku większości procesów finansowych nie jest spełnione z powodu występowania wysokiej kurtozy oraz asymetrii warunkowych rozkładów stóp zwrotów. Problem ten rozwiązywano przez wprowadzenie asymetrycznych rozkładów zawierających grube ogony, najczęściej skośnego rozkładu t Studenta. Możliwości wykorzystania tego rozkładu w przypadku dodatkowej analizy danych śróddziennych są jednak ograniczone. W literaturze trudno znaleźć opisy procesów stochastycznych z czasem ciągłym, których przyrosty byłyby zmiennymi losowymi o tym rozkładzie. Zastosowanie rozkładów t Studenta do procesów GARCH znacznie komplikuje także wycenę instrumentów pochodnych. W przypadku modelowania logarytmicznych stóp zwrotu oczekiwana stopa zwrotu takiego procesu jest nieograniczona, w związku z tym nie istnieje drugi moment funkcji wypłaty. W konsekwencji cena opcji opisanej za pomocą takiego modelu jest nieskończona, co jest nonsensem. Z opisanych powodów pożądane jest stosowanie w modelowaniu rozkładu o grubych ogonach, który równocześnie miałby wszystkie momenty skończone. Warunek ten spełnia przedstawiony w niniejszej pracy rozkład NIG, którego wykorzystanie w modelowaniu stochastycznej zmienności zaproponowali Barndorff-Nielsen (1997) oraz Andersson (2001). Z kolei Jensen i Lunde (2001) opracowali tzw. model NIG-S&ARCH. Część rezultatów przedstawionych w tej ostatniej pracy zostanie wykorzystana w niniejszym opracowaniu. Analizy Ventera i de Jongha (2004) pokazują, że dzięki zastosowaniu modelu GARCH z warunkowym rozkładem NIG składnika losowego uzyskiwano trafniejsze oszacowania zmienności stóp zwrotu i ryzyka niż na podstawie modelu ze skośnym rozkładem t Studenta.

W wyniku przeprowadzonych badań Forsberg i Bollerslev (2002) stwierdzili, że dla kursu walutowego ECU/USD:

1) zmienność zrealizowana stóp zwrotu (obliczana jako suma kwadratów danych śróddziennych) ma rozkład odwrotny gaussowski,

2) dzienna stopa zwrotu standaryzowana przez pierwiastek zmienności zrealizowanej ma w przybliżeniu rozkład normalny.

Analizy te potwierdziły dobre dopasowanie modelu GARCH-NIG, zaproponowanego wcześniej przez Forsberga (2002). Skonstruowany model nie jest jednak w stanie opisać asymetrii rozkładów stóp zwrotu – własności często występującej na przykład na rynkach akcji.

Omawiane dotąd badania wykorzystywały szeregi czasowe zawierające tylko dzienne stopy zwrotu konstruowane na podstawie cen zamknięcia. W pracach Lildholdta (2002) oraz Ventera, de Jongha i Griebenowa (2005) zaproponowano parametryzacje modeli GARCH, w których dodatkowo wykorzystano wartości cen minimalnych i maksymalnych. Wyznaczono funkcje gęstości rozkładu łącznego minimalnej, maksymalnej i końcowej wartości logarytmicznych stóp zwrotu w danym dniu. Założono, że śróddzienny proces logarytmicznych stóp zwrotu instrumentu jest procesem Levy'ego. W opracowaniu Lildholdta (2002) jest to arytmetyczny ruch Browna o ustalonej na dany dzień wariancji, a składnik losowy w zaproponowanym modelu GARCH dla obserwacji dziennych ma warunkowy rozkład normalny.

Ze względu na występowanie leptokurtozy oraz asymetrii rozkładów stóp zwrotów założenie to nie jest na ogół spełnione w przypadku empirycznych szeregów finansowych. Z kolei w pracy Ventera, de Jongha i Griebenowa (2005) dynamika śróddziennych stóp zwrotu jest bardziej złożona. Wykorzystując wnioski z pracy Forsberga i Bollersleva (2002), zaproponowano parametryzację modelu GARCH-NIG, w którym warunkowy rozkład składnika losowego jest rozkładem normalnym, odwrotnym, gaussowskim NIG. Przedstawione funkcje gęstości zostały następnie wykorzystane do określenia funkcji wiarygodności dla modeli GARCH.

W porównaniu ze wspomnianymi pracami w modelach GARCH przedstawionych poniżej do estymacji dziennej wariancji stóp zwrotu zastosowano także ceny minimalne i maksymalne. Omówione propozycje mają również pewne wady. Przedstawiona w pracy Lildholdta funkcja gęstości rozkładu łącznego wartości minimalnej, maksymalnej i końcowej ma złożoną postać, co dodatkowo komplikuje model. W opracowaniu Ventera, de Jongha i Griebenowa przyjęto nietypowe parametryzacje wprowadzonych tam zmiennych losowych oraz procesów stochastycznych. Ponadto składnik losowy nie jest standaryzowany (tzn. jego wartość oczekiwana i wariancja nie są równe, odpowiednio, 0 i 1). Utrudnia to porównanie budowy modelu z innymi, opisanymi w literaturze. Modele przedstawione w niniejszym opracowaniu będą wolne od tych wad.

Niniejsze opracowanie ma trzy podstawowe cele. Pierwszym jest zaproponowanie nowej parametryzacji modelu GARCH, do której estymacji będą wykorzystywane jednocześnie informacje o cenach minimalnych i maksymalnych: do konstrukcji estymatorów dziennej wariancji i do określenia funkcji wiarygodności służącej do estymacji parametrów modeli zmienności. Drugi cel to przedstawienie w bardziej czytelnej i uporządkowanej formie dotychczas istniejących modeli opisanych w pracach Lildholdta (2002) oraz Ventera, de Jongha i Griebenowa (2005), tak aby można było bezpośrednio porównać je z modelami przedstawionymi w innych publikacjach naukowych. Trzecim celem jest pokazanie, na podstawie szeregów stóp zwrotu z indeksu WIG20 i kursu walutowego EUR/PLN, że wykorzystanie informacji o cenach minimalnych i maksymalnych do estymacji parametrów modelu GARCH, bez zwiększenia jego parametryzacji, poprawia jakość modelu mierzoną wartością funkcji wiarygodności. Tematyka ta jest praktycznie nieobecna w literaturze światowej.

Układ artykułu jest następujący. W części drugiej przedstawiona zostanie ogólna postać zaproponowanych parametryzacji modelu GARCH. Część trzecia zawiera propozycje wykorzystania dodatkowych informacji na temat dziennego minimum i maksimum przy założeniu, że warunkowy rozkład składnika losowego jest normalny oraz śróddzienne stopy zwrotu można opisać arytmetycznym ruchem Browna. W części czwartej, z myślą o uogólnieniu założeń modelowych, przedstawiono rozkład NIG oraz rozszerzoną formułę gęstości rozkładu łącznego maksimum, minimum i wartość końcową. Część piąta zawiera opis parametryzacji modelu GARCH z wykorzystaniem rozkładu NIG oraz wspomnianej funkcji gęstości. Wyniki badania empirycznego dotyczącego indeksu WIG20 notowanego na GPW w Warszawie oraz kursu walutowego EUR/PLN zamieszczono w części szóstej. Niniejszy artykuł zawiera znaczne rozszerzenie wyników przedstawionych w pracy Perczaka (2013).

2. Podstawowe specyfikacje modelu GARCH dla rozważanych rodzajów danych rynkowych

2.1. Model GARCH konstruowany na podstawie cen zamknięcia

Przez $D(\mu, \sigma^2)$ oznaczana będzie ciągła w swojej dziedzinie funkcja rozkładu prawdopodobieństwa zmiennej losowej X o wartości oczekiwanej μ i wariancji σ^2. Rozpatrywane są procesy z czasem dyskretnym t. Niech S_t oznacza cenę zamknięcia instrumentu finansowego odnotowaną w dniu t. Zdefiniowano także dzienną stopę zwrotu jako $X_t = \ln(S_t / S_{t-1})$. Symbol $\Im_t$ oznacza zbiór wszystkich informacji dostępnych w chwili t.

Niech X_t oznacza proces GARCH(p, q) w postaci:

$$X_t \mid \Im_{t-1} \sim D(\mu, h_t) \tag{1}$$

$$\varepsilon_t = X_t - \mu \tag{2}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{3}$$

Dodatnia wartość h_t jest zapewniona, gdy $\omega_0 > 0$, $\omega_i \geq 0$ dla $1 \leq i \leq q$ oraz $\xi_i \geq 0$ dla $1 \leq i \leq p$. Przy większych rzędach opóźnień q i p założenia te są zbyt restrykcyjne i często dzięki ich złagodzeniu można uzyskać lepszy opis wariancji warunkowej. Kowariancyjna stacjonarność procesu zachodzi natomiast wtedy, gdy spełniony jest warunek:

$$\sum_{i=1}^{q} \omega_i + \sum_{i=1}^{p} \xi_i < 1$$

2.2. Opis rozszerzonego zbioru danych rynkowych

Wycena wartości instrumentów finansowych zazwyczaj dokonywana jest na rynku finansowym wielokrotnie w ciągu dnia. Z dziennych notowań rynkowych można pozyskać znacznie więcej informacji niż pojedyncze notowania cen zamknięcia. Przedstawiony w poprzednim podpunkcie model GARCH konstruowany jest na podstawie tylko cen zamknięcia. Poniżej zostanie zaprezentowana propozycja modyfikacji modelu, dzięki której do estymacji parametrów będzie można wykorzystać szerszy zbiór dziennych notowań.

Przez $S_{t,\tau}$ oznaczana będzie cena instrumentu finansowego odnotowana w dniu t ($t \in N$, $0 < t$) po upływie czasu τ $(0 \leq \tau \leq 1)$ od ostatniego notowania poprzedniego dnia. Zachodzi zatem tożsamość: $S_{t-1,1} = S_{t,0}$. Dzienne (dobowe) minimum i maksimum zdefiniowano odpowiednio jako $L_t = \min_{0 \leq \tau \leq 1} S_{t,\tau}$ i $H_t = \max_{0 \leq \tau \leq 1} S_{t,\tau}$). Dodatkowo przyjęto definicję minimalnej i maksymalnej dziennej stopy zwrotu: $A_t = \ln(L_t / S_{t,0})$, $C_t = \ln(H_t / S_{t,0})$. Zredefiniowano jednodniową stopę zwrotu: $X_t = \ln(S_{t,1} / S_{t,0})$, która będzie też nazywana wartością końcową dziennej stopy zwrotu.

Wykorzystywane informacje dotyczące dziennych notowań instrumentu finansowego reprezentowane będą teraz przez uporządkowaną trójkę liczb (a_t, c_t, x_t), zamiast pojedynczej wartości x_t, jak ma to miejsce w klasycznym modelu GARCH wprowadzonym przez Bollersleva (1986).

2.3. Ogólna postać proponowanych modeli GARCH dla danych rozszerzonych o cenę minimalną i cenę maksymalną

Przez $LHC(\mu, \sigma^2)$ oznaczana będzie ciągła w swojej dziedzinie łączna funkcja rozkładu prawdopodobieństwa wektora zmiennych losowych (A, C, X), dla których spełnione są następujące warunki: $A \le 0 \le C$, $A \le X \le C$, $E[X] = \mu$, $Var[X] = \sigma^2$.

Dla ustalonego dnia t $(t \in N, 0 < t)$ rozpatrywany jest wektor losowy (A_t, C_t, X_t) o rozkładzie $LHC(\mu_t, h_t)$, dla którego zdefiniowano także funkcję $\hat{Z}(A_t, C_t, X_t)$, będącą nieobciążonym estymatorem dziennej wariancji stóp zwrotu. Funkcja ta spełnia również warunek $E\left[\hat{Z}(A_t, C_t, X_t) \mid \Im_{t-1}\right] = h_t$. Kilka wybranych postaci tej funkcji przedstawiono w dalszej części pracy.

Ogólna postać zaproponowanych parametryzacji w niniejszym opracowaniu wygląda następująco:

$$(A_t, C_t, X_t) \mid \Im_{t-1} \sim LHC(\mu, h_t) \tag{4}$$

$$\varepsilon_t^2 = \hat{Z}(A_t, C_t, X_t) \tag{5}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{6}$$

gdzie ε_t^2 określono na podstawie estymatora $\hat{Z}$.

3. Model GARCH z warunkowym rozkładem normalnym składnika losowego konstruowany na podstawie rozszerzonego zbioru danych

Zaprezentowana ogólna postać parametryzacji stanowi punkt wyjścia do prezentacji bardziej szczegółowych postaci modeli. Poniżej przyjęto, że warunkowy rozkład procesu X_t jest rozkładem normalnym. Ponadto założono, że stop zwrotu w ciągu dnia (doby) można opisać arytmetycznym ruchem Browna, co umożliwiło zastosowanie rozkładu łącznego stóp zwrotu dziennego minimum, maksimum i wartości końcowej. Do skonstruowania modeli wykorzystano również bardziej efektywne estymatory dziennej wariancji stopy zwrotu niż w klasycznym modelu AR(*r*)-GARCH(*p, q*).

3.1. Trójwymiarowy wektor losowy stóp zwrotu minimum, maksimum i wartości końcowej arytmetycznego ruchu Browna

Dla standardowego procesu Wienera B_τ, $\tau \ge 0$ zdefiniowany zostaje ruch Browna $X_\tau = \mu\tau + \sigma^2 B_\tau$ oraz $A_s = \min\limits_{0<\tau\le s} X_\tau$ i $C_s = \max\limits_{0<\tau\le s} X_\tau$. Dla dowolnej, ustalonej wartości $s > 0$ zachodzi $X_s \sim N(\mu s, \sigma^2 s)$. Jeśli proces X_τ osiąga dla pewnego τ_1 wartość $a < 0$, a następnie dla wszystkich $\tau > \tau_1$ zachodzi $X_\tau \le a$, wówczas liczbę a nazywamy dolną barierą absorbującą. Analogicznie, jeśli proces X_τ osiąga dla pewnego τ_2 wartość $c > 0$, a następnie dla wszystkich $\tau > \tau_2$ zachodzi $X_\tau \geq c$, wówczas liczbę c nazywamy górną barierą absorbującą.

Gęstość rozkładu X_s, z dolną barierą absorbującą równą a i górną barierą równą c, opisano za pomocą następującego wzoru (porównaj Cox, Miller 1965, s. 222, formuła 78, a także Li 1999):

$$f_N(A_s > a, C_s \leq c, x; \mu s, \sigma^2 s) = \\ = \frac{1}{\sqrt{2\pi\sigma s^{\frac{1}{2}}}} \sum_{k=-\infty}^{k=\infty} e^{\frac{2k(c-a)\mu}{\sigma^2}} \left(e^{-\frac{(x-2k(c-a)-\mu s)^2}{2\sigma^2 s}} - e^{\frac{2c\mu s-(x-2c-2k(c-a)-\mu s)^2}{2\sigma^2 s}} \right) \tag{7}$$

gdzie: $a \leq 0 \leq c$, $a \leq x \leq c$, średnik oddziela argumenty funkcji od jej parametrów.

Wyznaczona na podstawie tożsamości (7) gęstość rozkładu łącznego wektora zmiennych losowych (A_s, C_s, X_s) dana jest następującą formułą (Fiszeder, Perczak 2013):

$$f_{ACN}(a,c,x;\mu s, \sigma^2 s) = -\frac{\partial^2 f_N(x; A_s > a, C_s \leq c; \mu s, \sigma^2 s)}{\partial a \partial c} = \\ = \frac{1}{\sqrt{2\pi\sigma^5 s^{\frac{5}{2}}}} \sum_{k=-\infty}^{k=\infty} (g(a,c,x;k,k,\mu s,\sigma^2 s) - g(a,c,x;k,k+1,\mu s,\sigma^2 s)) \tag{8}$$

gdzie funkcja g opisana jest zależnością:

$$g(a,c,x;k_1,k_2,\mu s,\sigma^2 s) = 4k_1 k_2 \left((x-2(ck_2 - ak_1))^2 - \sigma^2 s\right) e^{\frac{2(ck_2-ak_1)\mu s-(x-2(ck_2-ak_1)-\mu s)^2}{2\sigma^2 s}} \tag{9}$$

Argumenty a, c, x można interpretować jako wartości, odpowiednio: minimalną, maksymalną i końcową.

Wektor losowy (A_s, C_s, X_s) ma rozkład łączny $ACN(\mu s, \sigma^2 s)$, jeśli jego gęstość opisana jest wzorem (8).

Jeżeli $(A_s, C_s, X_s) \sim ACN(\mu s, \sigma^2 s)$, wówczas statystyka:

$$\hat{V}_1(s) = 0{,}86 C_s (C_s - X_s) + 0{,}86 A_s (A_s - X_s) + 0{,}14 (X_s^2 - \mu^2 s^2) \tag{10}$$

jest nieobciążonym estymatorem wariancji $\sigma^2 s$ procesu X_s (Perczak, Fiszeder 2013).

Ponadto wariancja tego estymatora dla $|\mu| << \sigma$ jest mniejsza od wariancji nieobciążonego estymatora Rogersa-Satchella (1991). Wartości 0,86 i 0,14 są wartościami przybliżonymi.

3.2. Estymacja parametrów modelu GARCH z warunkowym rozkładem normalnym z wykorzystaniem cen minimalnych, maksymalnych i zamknięcia

W przedstawionym w punkcie 2.1 modelu GARCH(p, q) w postaci (1–3) przyjmuje się bardzo często w miejsce założenia (1), że $X_t \mid \Im_{t-1} \sim N(\mu, h_t)$. Model taki oznaczany w dalszej części pracy jako N_{11}:

$$X_t \mid \Im_{t-1} \sim N(\mu, h_t) \tag{11}$$

$$\varepsilon_t = X_t - \mu \tag{12}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{13}$$

Parametry takiego modelu estymuje się najczęściej metodą najmniejszej wiarygodności (MNW) na podstawie funkcji wiarygodności, wykorzystując wyłącznie ceny zamknięcia (podejście tradycyjne):

$$\left\{\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\varphi}}\right\} = \underset{\{\boldsymbol{\omega}, \boldsymbol{\xi}, \boldsymbol{\varphi}\}}{\arg\max} \ln L_N(\boldsymbol{\omega}, \boldsymbol{\xi}, \boldsymbol{\varphi}) = \underset{\{\boldsymbol{\omega}, \boldsymbol{\xi}, \boldsymbol{\varphi}\}}{\arg\max} \sum_{t=1}^{n} \ln f_N(x_t;\ \mu_t(\boldsymbol{\varphi}), h_t(\boldsymbol{\omega}, \boldsymbol{\xi})) \tag{14}$$

gdzie $\boldsymbol{\omega} = (\omega_0, \omega_1, \ldots, \omega_q)$, $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_p)$, n jest liczebnością próby, natomiast f_N jest funkcją gęstości rozkładu normalnego.

Wprowadźmy bardziej restrykcyjne założenia modelu N_{11}. Przyjęto, że

$$(A_t, C_t, X_t) \mid \Im_{t-1} \sim ACN(\mu_t, h_t)$$

Zgodnie z definicją rozkładu ACN prawdą jest, że

$$((A_t, C_t, X_t) \mid \Im_{t-1} \sim ACN(\mu_t, h_t)) \Rightarrow (X_t \mid \Im_{t-1} \sim N(\mu_t, h_t))$$

co jednak nie zachodzi w drugą stronę. Znajomość funkcji gęstości rozkładu łącznego wektora (A_t, C_t, X_t) umożliwia sformułowanie modelu GARCH(p, q) o odmiennej postaci, którą nazwano GARCH-HL z HLC. HL oznacza, że do budowy funkcji wiarygodności wykorzystano ceny maksymalne i minimalne, HLC oznacza zastosowanie jednego z estymatorów wariancji, konstruowanego na podstawie cen: maksimum, minimum i zamknięcia. Zbiór ten, oznaczony jako N_{22}, może być rozszerzony o cenę otwarcia:

$$(A_t, C_t, X_t) \mid \Im_{t-1} \sim ACN(\mu_t, h_t) \tag{15}$$

$$\varepsilon_t^2 = 0{,}86 C_t(C_t - X_t) + 0{,}86 A_t(A_t - X_t) + 0{,}14(X_t^2 - \mu^2) \tag{16}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{17}$$

Parametry tego modelu mogą być estymowane za pomocą MNW na postawie funkcji wiarygodności konstruowanej z wykorzystaniem cen minimalnych, maksymalnych i zamknięcia oraz nowego estymatora dziennej wariancji, do którego budowy wykorzystano również ten sam rodzaj informacji (równanie 16):

$$\left\{\hat{\omega}, \hat{\xi}\right\} = \underset{\{\omega,\xi\}}{\arg\max} \ln L_{ACN}(\omega, \xi) = \underset{\{\omega,\xi\}}{\arg\max} \sum_{i=1}^{n} \ln f_{ACN}(a_t, c_t, x_t;\, \mu, h_t(\omega, \xi)) \tag{18}$$

Kolejna propozycja modelu GARCH(p, q), nazwana GARCH-HL i oznaczana w dalszej części pracy jako N_{21}, została zaprezentowana poniżej:

$$(A_t, C_t, X_t) | \Im_{t-1} \sim ACN(\mu, h_t) \tag{19}$$

$$\varepsilon_t^2 = (X_t - \mu)^2 \tag{20}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{21}$$

Parametry tego modelu mogą być estymowane za pomocą MNW na postawie funkcji wiarygodności konstruowanej z wykorzystaniem cen minimalnych, maksymalnych i zamknięcia oraz tradycyjnego estymatora dziennej wariancji opartego na podstawie cen zamknięcia (równanie 20):

$$\left\{\hat{\omega}, \hat{\xi}\right\} = \underset{\{\omega,\xi\}}{\arg\max} \ln L_{ACN}(\omega, \xi) = \underset{\{\omega,\xi\}}{\arg\max} \sum_{t=1}^{n} \ln f_{ACN}(a_t, c_t, x_t;\, \mu, h_t(\omega, \xi)) \tag{22}$$

Równania (23–25) przedstawiają trzecią propozycję modelu GARCH(p, q). Postać tę nazwano GARCH z HLC i oznaczono w dalszej części pracy jako N_{12}:

$$(A_t, C_t, X_t) | \Im_{t-1} \sim ACN(\mu, h_t) \tag{23}$$

$$\varepsilon_t^2 = 0{,}86 C_t (C_t - X_t) + 0{,}86 A_t (A_t - X_t) + 0{,}14 (X_t^2 - \mu) \tag{24}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{25}$$

Parametry tego modelu można estymować na podstawie klasycznej funkcji wiarygodności, wykorzystując wyłącznie ceny zamknięcia, oraz nowego estymatora dziennej wariancji konstruowanego na podstawie cen minimalnych, maksymalnych i zamknięcia (równanie 24):

$$\left\{\hat{\omega}, \hat{\xi}\right\} = \underset{\{\omega,\xi\}}{\arg\max} \ln L_N(\omega, \xi) = \underset{\{\omega,\xi\}}{\arg\max} \sum_{t=1}^{n} \ln f_N(x_t;\, \mu, h_t(\omega, \xi)) \tag{26}$$

Zauważmy, że parametryzacje N_{12} i N_{22} są identyczne, a różnica między nimi wynika z zupełnie innej konstrukcji funkcji wiarygodności. W tabeli 1 zestawiono wszystkie zaprezentowane modele.

Parametryzacja N_{22} różni się od klasycznej parametryzacji modelu GARCH N_{11} tym, że przy jej konstrukcji zastosowano informacje o cenach minimalnych i maksymalnych zarówno przy estymacji dziennej wariancji, jak i do określenia funkcji wiarygodności służącej do estymacji parametrów modelu GARCH. Zasadniczą przyczyną tego, że w równaniach (16) i (24) zastosowano estymator dziennej wariancji określony w równaniu (10), jest jego ponadsiedmiokrotnie większa efektywność niż efektywność estymatora o postaci $(X_t-\mu_t)^2$ (Perczak, Fiszeder 2013). Można przypuszczać, że pozwoli to na uzyskanie dokładniejszego oszacowania wariancji i jej trafniejszych prognoz.

Parametryzacja N_{21} zasadniczo nie różni się od propozycji zawartej w pracy Lildholdta (2002), choć zawiera inną propozycję postaci funkcji f_{ACN}. Wynika to z faktu, że Lildholdt jako punkt wyjścia do dalszych rozważań zastosował funkcję przedstawioną przez Coxa i Millera (1965, formuły 80–82, s. 222–223). W niniejszym opracowaniu postać funkcji przedstawionej w równaniu (7) została natomiast wyznaczona przez przekształcenie innej funkcji zawartej w tej pracy (formuła 78, s. 222). Uzyskana w ten sposób postać funkcji gęstości rozkładu łącznego wektora (A_s, C_s, X_s) jest dużo prostsza. Jak zaznaczono, założenia przyjęte w modelach N_{22}, N_{21} i N_{12} są silniejsze od tych, które przyjęto dla modelu N_{11}. Poza warunkową normalnością rozkładu składnika losowego X_t (przyjętą również w modelu N_{11}) wymagane jest, aby (A_t, C_t, X_t) miał rozkład ACN. Nie oznacza to jednak, że proces śróddziennych stóp zwrotu musi być arytmetycznym ruchem Browna, mimo że założenie to było punktem wyjścia do wyprowadzenia formuły gęstości funkcji f_{ACN}.

Ponadto modele N_{22}, N_{21} i N_{12} są oszczędnie sparametryzowane. Nie zostały do nich wprowadzone żadne dodatkowe parametry wymagające oszacowania w stosunku do modelu N_{11}.

4. Rozkład NIG oraz proces BIG

4.1. Charakterystyka rozkładu IG

Zmienna losowa X ma rozkład odwrotny gaussowski oznaczany jako $IG(\delta,\gamma)$ (ang. *inverse Gaussian*) z parametrami $\delta\in R_+$ i $\gamma\in R_+$, jeżeli jej gęstość rozkładu prawdopodobieństwa określona dla $x>0$ ma następującą postać (por. Barndorff-Nielsen 1997, s. 2, równanie 2.5):

$$f_{IG}(x;\delta,\gamma)=\frac{\delta}{\sqrt{2\pi}}x^{-\frac{3}{2}}\exp\left(\delta\gamma-\frac{\delta^2x^{-1}+\gamma^2x}{2}\right) \tag{27}$$

Funkcja charakterystyczna tej zmiennej ma postać: $\varphi_{IG}(u;\delta,\gamma)=E[e^{iuX}]=e^{\delta\left(\gamma-\sqrt{\gamma^2-2iu}\right)}$.
Wartość oczekiwana i wariancja są zatem równe odpowiednio:

$$E[X]=\frac{\delta}{\gamma},\ Var[X]=\frac{\delta}{\gamma^3} \tag{28}$$

Łatwo także sprawdzić, że dla dowolnej dodatniej stałej s zachodzą równoważności:

$$X\sim IG(\delta,\gamma)\Leftrightarrow sX\sim IG(\delta\sqrt{s},\gamma/\sqrt{s})\quad \text{oraz}\quad sX\sim IG(\delta,\gamma)\Leftrightarrow X\sim IG(\delta/\sqrt{s},\gamma\sqrt{s}) \tag{29}$$

4.2. Charakterystyka rozkładu NIG

Zmienna losowa X ma rozkład $NIG(\alpha, \beta, \delta, \mu)$ z parametrami $\alpha \in R_+$, $-\alpha < \beta < \alpha$, $\delta \in R_+$ i $\mu \in R$, jeżeli $X \mid W \sim N(\mu + \beta W, W)$ gdzie $W \sim IG(\delta, \gamma)$ oraz $\gamma = \sqrt{\alpha^2 - \beta^2}$. Gęstość rozkładu prawdopodobieństwa tej zmiennej ma zatem postać:

$$f_{NIG}(x; \alpha, \beta, \delta, \mu) = \int_0^\infty f_N(x; \mu + \beta w, w)\, f_{IG}(w; \delta, \gamma)\, dw = \tag{30}$$
$$= \frac{\alpha \delta e^{\beta(x-\mu)+\delta\gamma}}{\pi\sqrt{\delta^2 + (x-\mu)^2}} K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)$$

gdzie $K_\lambda(z) = \frac{1}{2}\int_0^\infty y^{\lambda-1} \exp\left(-\frac{1}{2} z\left(y + \frac{1}{y}\right)\right) dy$

gdzie α i β są parametrami kształtu, a δ i μ, to, odpowiednio, parametry skali i położenia (por. Jensen, Lunde 2001, s. 324).

Rozkład NIG został zaproponowany przez Barndorffa-Nielsena (1995). Jest to szczególny przypadek tzw. uogólnionego rozkładu hiperbolicznego (Barndorff-Nielsen 1977). Więcej informacji na temat rozkładu NIG można znaleźć np. w pracy Schlössera (2011, s. 129–135).

Funkcja charakterystyczna zmiennej X o rozkładzie NIG opisana jest wzorem:

$$\varphi_{IG}(u; \alpha, \beta, \delta, \mu) = E[e^{iuX}] = e^{i\mu u + \delta\left(\gamma - \sqrt{\alpha^2 - (\beta + iu)^2}\right)}$$

Wartość oczekiwana i wariancja równe są odpowiednio:

$$E[X] = \mu + \frac{\beta\delta}{\gamma}, \; Var[X] = \frac{\alpha^2\delta}{\gamma^3} \tag{31}$$

Współczynniki asymetrii i koncentracji dane są natomiast następującymi formułami:

$$A[X] = \frac{E[X - E[X]]^3}{Var[X]^{3/2}} = \frac{3\beta}{\alpha\sqrt{\delta\gamma}}, \; K[X] = \frac{E[X - E[X]]^4}{Var[X]^2} = 3 + \frac{3}{\delta\gamma}\left(1 + 4\left(\frac{\beta}{\alpha}\right)^2\right) \tag{32}$$

Dodatnia wartość współczynnika asymetrii wskazuje na prawostronną skośność rozkładu, natomiast wartości współczynnika koncentracji większe od 3 informują o podwyższonej, w stosunku do rozkładu normalnego, kurtozie rozkładu.

Przez odpowiednie przekształcenie funkcji charakterystycznej rozkładu NIG można pokazać, że dla stałego $b > 0$ zachodzi:

$$X \sim NIG(\alpha, \beta, \delta, \mu) \Leftrightarrow bX \sim NIG\left(\frac{\alpha}{b}, \frac{\beta}{b}, b\delta, b\mu\right) \tag{33}$$

W niniejszej pracy wygodniej będzie zastosować nieco inną parametryzację zaprezentowaną np. w artykułach Barndorffa-Nielsena (1997, s. 2) oraz Jensena i Lunde'a (2001, s. 324). Podstawiając: $\overline{\alpha} = \alpha\delta$, $\overline{\beta} = \beta\delta$, $\overline{\gamma} = \sqrt{\overline{\alpha}^2 - \overline{\beta}^2}$, uzyskuje się parametry kształtu, które są niezmiennicze przy zmianie parametrów położenia i skali (ang. *invariant under location-scale changes*). Zmienna losowa X ma rozkład $NIG(\overline{\alpha}, \overline{\beta}, \delta, \mu)$, jeśli jej funkcja gęstości rozkładu prawdopodobieństwa dana jest formułą:

$$f_{NIG}(x; \overline{\alpha}, \overline{\beta}, \delta, \mu) = \frac{\overline{\alpha} e^{\overline{\beta}\left(\frac{x-\mu}{\delta}\right)+\overline{\gamma}}}{\pi\delta\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}} K_1\left(\overline{\alpha}\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}\right) \tag{34}$$

Wartość oczekiwana i wariancja tej zmiennej dane są, odpowiednio, formułami:

$$E[X] = \mu + \frac{\overline{\beta}\delta}{\overline{\gamma}}, \; Var[X] = \frac{\overline{\alpha}^2 \delta^2}{\overline{\gamma}^3} \tag{35}$$

W niniejszej pracy zaletą tej parametryzacji jest fakt, że tylko dwa parametry: $\overline{\alpha}$ i $\overline{\beta}$, jednoznacznie określają wartości współczynników asymetrii i koncentracji:

$$A[X] = \frac{3\overline{\beta}}{\overline{\alpha}\sqrt{\overline{\gamma}}}, \; K[X] = 3 + \frac{3}{\overline{\gamma}}\left(1 + 4\left(\frac{\overline{\beta}}{\overline{\alpha}}\right)^2\right) \tag{36}$$

4.3. Proces BIG

Jeżeli $\beta \in R_+$, $\gamma \in R_+$, $\delta \in R_+$, $\mu \in R$, $Z \sim N(0,1)$ i $W \sim IG(\delta, \gamma)$, wówczas zmienna losowa $X \mid W = \mu + \beta W + \sqrt{W}Z$ ma rozkład $N(\mu + \beta W, W)$. Z tego oraz z równania (30) wynika, że $X \sim NIG(x; \alpha, \beta, \delta, \mu)$ dla $\alpha = \sqrt{\beta^2 + \gamma^2}$ (Barndorff-Nielsen, Shephard 2001, s. 15).

Niech $s \in R_+$ będzie ustalonym parametrem, zmienna losowa Y ma rozkład $IG(\delta\sqrt{s}, \gamma\sqrt{s})$ natomiast B_τ będzie standardowym procesem Wienera. Dla procesu opisanego równaniem:

$$X_\tau^{(s)} \mid Y = \mu\tau + \beta Y\tau + \sqrt{Y}B_\tau, \qquad Y \sim IG(\delta\sqrt{s}, \gamma\sqrt{s})$$

zachodzi: $X_\tau^{(s)} \mid Y \sim N(\mu\tau + \beta Y\tau, Y\tau)$. Ponadto z relacji opisanych w formule (29) wynika, że $sY \sim IG(\delta s, \gamma)$. Przyjmując dodatkowo, że $W = sY$, łatwo pokazać, że w punkcie τ równym s zachodzi $X_s^{(s)} \sim NIG(\alpha, \beta, \delta s, \mu s)$.

Proces $X_\tau^{(s)}$ zostanie nazwany procesem BIG (ang. *Brownian inverse Gaussian*). Jego specyfikacja jest modyfikacją procesu opisanego przez Ventera, de Jongha i Griebenowa (2006, s. 102, wzór 6).

4.4. Gęstość rozkładu łącznego stóp zwrotu minimum, maksimum i wartości końcowej procesu BIG

Niech $X_\tau^{(s)}$ będzie procesem BIG zdefiniowanym w równaniu (37) oraz $A_s = \min_{0<\tau\le s} X_\tau^{(s)}$ i $C_s = \max_{0<\tau\le s} X_\tau^{(s)}$. Dla uproszczenia oznaczeń przyjęto $X_s := X_s^{(s)}$. Tak jak w punkcie 3.1 wyznaczono gęstość rozkładu X_s z barierami absorbującymi dolną i górną, równymi a i c (Perczak 2013; Venter, de Jongh, Griebenow 2005, równanie 4.5):

$$f_{NIG}(x;\, A_s > a, C_s \le c;\, \alpha,\ \beta, \delta s, \mu s) =$$

$$= \int_0^\infty f_N(A_s > a, C_s \le c, x;\, \mu s + \beta ys, ys) f_{IG}(ys; \delta s, \gamma)\, d(ys) =$$

$$= \frac{\alpha \delta s e^{\beta(x-\mu s)+\delta s\gamma}}{\pi} \sum_{k=-\infty}^{k=\infty} \left(\frac{K_1(\alpha\sqrt{\theta_1 + \vartheta})}{\sqrt{\theta_1 + \vartheta}} - \frac{K_1(\alpha\sqrt{\theta_2 + \vartheta})}{\sqrt{\theta_2 + \vartheta}} \right)$$

gdzie: $\theta_1 = (2k(c-a) - x)^2,\ \theta_2 = (2c + 2k(c-a) - x)^2,\ \vartheta = \delta^2 s^2 + (x - \mu s)^2 - x^2$

Korzystając ze wzoru (38) lub (8), można wyznaczyć gęstość rozkładu łącznego wektora zmiennych losowych (A_s, C_s, X_s):

$$f_{ACNIG}(a, c, x;\, \alpha,\, \beta,\, \delta s,\, \mu s) = -\frac{\partial^2 f_{NIG}(x;\, A_s > a, C_s \le c;\, \alpha,\, \beta,\, \delta s,\, \mu s)}{\partial a \partial c} =$$

$$= \int_0^\infty f_{ACN}(a, c, x;\, \mu s + \beta ys, ys) f_{IG}(ys;\, \delta s, \gamma)\, d(ys) =$$

$$= \sum_{k=-\infty}^{k=\infty} \frac{4\alpha^2 \delta s k^2 e^{\beta(x-\mu s)+\delta s\gamma}}{\pi} \left(\alpha\theta_1 \frac{K_1(\alpha\sqrt{\theta_1 + \vartheta})}{(\theta_1 + \vartheta)^{\frac{3}{2}}} + (3\theta_1 - \vartheta)^2 \frac{K_2(\alpha\sqrt{\theta_1 + \vartheta})}{(\theta_1 + \vartheta)^2} \right) - \tag{39}$$

$$- \sum_{k=-\infty}^{k=\infty} \frac{4\alpha^2 \delta s k\,(k+1)\, e^{\beta(x-\mu s)+\delta s\gamma}}{\pi} \left(\alpha\theta_2 \frac{K_1(\alpha\sqrt{\theta_2 + \vartheta})}{(\theta_2 + \vartheta)^{\frac{3}{2}}} + (3\theta_2 - \vartheta)^2 \frac{K_2(\alpha\sqrt{\theta_2 + \vartheta})}{(\theta_2 + \vartheta)^2} \right)$$

Parametry $\theta_1, \theta_2, \vartheta$ wykorzystuje się w równaniu (38).

Wektor losowy (A_s, C_s, X_s) będzie miał rozkład oznaczony jako ACNIG, tj. $(A_s, C_s, X_s) \sim ACNIG(\alpha,\, \beta,\, \delta s,\, \mu s)$, jeżeli jego gęstość będzie określona równaniem (39).

Podobnie jak w przypadku rozkładu NIG można określić dla rozkładu ACNIG alternatywną parametryzację $ACNIG(\overline{\alpha},\, \overline{\beta},\, \delta s,\, \mu s)$ z oznaczeniami niezmienniczych wielkości $\overline{\alpha} = \alpha\delta s,\ \overline{\beta} = \beta\delta s, \overline{\gamma} = \sqrt{\overline{\alpha}^2 - \overline{\beta}^2}$. Wymaga to dokonania stosunkowo prostych przekształceń wzoru (39).

Wyżej pokazano, że jeżeli $X_\tau^{(s)}$ jest procesem BIG zdefiniowanym na podstawie formuły (37), to $(A_s, C_s, X_s) \sim ACNIG(\alpha, \beta, \delta s, \mu s)$. Zależność ta nie musi jednak zachodzić w drugą stronę. Z faktu, że $(A_s, C_s, X_s) \sim ACNIG(\alpha, \beta, \delta s, \mu s)$, nie wynika, iż zmienne losowe A_s, C_s muszą być zdefiniowane na podstawie procesu $X_\tau^{(s)}$.

4.5. Estymacja wariancji procesu BIG na podstawie stóp zwrotu minimum, maksimum i wartości końcowej

Na podstawie dotychczasowych ustaleń:

$$\begin{gathered} W = sY \quad \wedge \quad Y \sim IG(\delta\sqrt{s}, \gamma\sqrt{s}\,) \quad \wedge \quad (A_s, C_s, X_s) \,|\, W \sim ACN(\mu s + \beta W, W) \\ \Downarrow \\ (A_s, C_s, X_s) \sim ACNIG(\alpha, \beta, \delta s, \mu s) \end{gathered} \tag{40}$$

Estymator Rogersa-Satchella (1991) jest nieobciążonym estymatorem wariancji ruchu Browna: $E[C_s(C_s - X_s) + A_s(A_s - X_s) \,|\, sY] = sY$, czyli:

$$E[C_s(C_s - X_s) + A_s(A_s - X_s)] = E[E[C_s(C_s - X_s) + A_s(A_s - X_s) \,|\, sY]] = E[sY] = \frac{\delta s}{\gamma} \tag{41}$$

Estymator o postaci:

$$\hat{V}_2(s) = \frac{\alpha^2}{\gamma^2}(C_s(C_s - X_s) + A_s(A_s - X_s)) = \frac{\overline{\alpha}^2}{\overline{\gamma}^2}(C_s(C_s - X_s) + A_s(A_s - X_s)) \tag{42}$$

ma zatem wartość oczekiwaną równą:

$$E[\hat{V}_2(s)] = \frac{\alpha^2}{\gamma^2}\frac{\delta s}{\gamma} = \frac{\alpha^2 \delta s}{\gamma^3} = \frac{\overline{\alpha}^2 \delta^2 s^2}{\overline{\gamma}^3} = Var[X_s] \tag{43}$$

W punkcie 3.1 przedstawiono estymator wariancji ruchu Browna $\hat{V}_1(s)$, który zastosowano w parametryzacjach modeli N_{12} i N_{22}. Estymator $\hat{V}_2(s)$ jest estymatorem wariancji procesu śróddziennych stóp zwrotu X_τ $\tau \le s$. Zostanie on wykorzystany do konstrukcji nowych parametryzacji modeli w piątej części pracy.

5. Model GARCH z warunkowym rozkładem NIG składnika losowego konstruowany na podstawie rozszerzonego zbioru danych

5.1. Model NIG-S&ARCH

Model NIG-S&ARCH (ang. NIG *stochastic and autoregressive conditional heteroskedasticity*) został wprowadzony przez Jensena i Lundego, (2001). Jego postać uzupełnioną równaniem autoregresyjnym oraz ogólną specyfikację modelu GARCH można przedstawić w formie następujących równań:

$$X_t \mid \Im_{t-1} \sim NIG\left(\bar{\alpha}, \bar{\beta}, \frac{\bar{\gamma}^{3/2}}{\bar{\alpha}}\sqrt{h_t}, \mu\right) \tag{44}$$

$$\varepsilon_t = X_t - \mu - \frac{\bar{\beta}\sqrt{\bar{\gamma}}}{\bar{\alpha}}\sqrt{h_t} \tag{45}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{46}$$

Parametry tego modelu mogą być estymowane za pomocą MNW na podstawie funkcji wiarygodności z wykorzystaniem wyłącznie cen zamknięcia:

$$\begin{aligned} \left\{\hat{\bar{\alpha}}, \hat{\bar{\beta}}, \hat{\omega}, \hat{\xi}\right\} &= \underset{\{\bar{\alpha}, \bar{\beta}, \omega, \xi\}}{\arg\max} \ln L_{NIG}(\bar{\alpha}, \bar{\beta}, \omega, \xi) = \\ &= \underset{\{\bar{\alpha}, \bar{\beta}, \omega, \xi\}}{\arg\max} \sum_{t=1}^{n} \ln f_{NIG}\left(x_t; \bar{\alpha}, \bar{\beta}, \frac{\bar{\gamma}^{3/2}}{\bar{\alpha}}\sqrt{h_t(\omega, \xi)}, \mu\right) \end{aligned} \tag{47}$$

gdzie $\boldsymbol{\omega} = (\omega_0, \omega_1, ..., \omega_q)$, $\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_p)$.

Model ten w dalszej części pracy będzie określany jako S&GARCH-NIG i oznaczany NIG_{11}.

5.2. Propozycje procedur estymacji parametrów modelu S&GARCH-NIG

Przedstawione w tym punkcie parametryzacje modeli zostały zaproponowane w pracy Perczak (2013). Tak jak w przypadku przedstawionych w części trzeciej parametryzacji modeli M_{ij}, do nowych parametryzacji modelu S&GARCH-NIG zostaną wykorzystane dodatkowe informacje o cenach minimalnych i maksymalnych. W pierwszej kolejności przyjęto, że:

$$(A_t, C_t, X_t) \mid \Im_{t-1} \sim ACNIG\left(\overline{\alpha}, \overline{\beta}, \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t}, \mu\right) \text{ dla } 1 \leq t \leq n$$

Wykorzystanie dodatkowych informacji będzie przebiegać w dwóch niezależnych kierunkach:

1) zostanie zastosowany estymator dziennej wariancji przedstawiony w równaniu (42), który jest efektywniejszy niż kwadrat dziennej stopy zwrotu obliczonej na podstawie cen zamknięcia,

2) parametry modelu będą estymowane za pomocą MNW na podstawie funkcji wiarygodności konstruowanej z wykorzystaniem cen minimalnych, maksymalnych i zamknięcia.

Pierwszą propozycją jest model, który zostanie nazwany S&GARCH-HL-NIG z HLC. Będzie on oznaczany w dalszej części pracy jako NIG_{22}:

$$(A_t, C_t, X_t) \mid \Im_{t-1} \sim ACNIG\left(\overline{\alpha}, \overline{\beta}, \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t}, \mu\right) \tag{48}$$

$$\varepsilon_t^2 = \frac{\overline{\alpha}^2}{\overline{\gamma}^2} (C_t(C_t - X_t) + A_t(A_t - X_t)) \tag{49}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{50}$$

Parametry tego modelu mogą być estymowane za pomocą MNW na podstawie funkcji wiarygodności konstruowanej z wykorzystaniem cen minimalnych, maksymalnych i zamknięcia oraz nowego estymatora dziennej wariancji przedstawionego w równaniu (42):

$$\begin{aligned} \left\{\hat{\overline{\alpha}}, \hat{\overline{\beta}}, \hat{\omega}, \hat{\xi}\right\} &= \underset{\{\overline{\alpha}, \overline{\beta}, \omega, \xi\}}{\arg\max} \ln L_{ACNIG}(\overline{\alpha}, \overline{\beta}, \omega, \xi) = \\ &= \underset{\{\overline{\alpha}, \overline{\beta}, \omega, \xi\}}{\arg\max} \sum_{i=1}^{n} \ln f_{ACNIG}\left(a_t, c_t, x_t; \overline{\alpha}\ \overline{\beta}, \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t(\omega, \xi)}, \mu\right) \end{aligned} \tag{51}$$

Poniżej zaprezentowano kolejną parametryzację modelu, nazwaną S&GARCH-HL-NIG i oznaczoną w dalszej części pracy jako NIG_{21}:

$$(A_t, C_t, X_t) \mid \Im_{t-1} \sim ACNIG\left(\overline{\alpha}, \overline{\beta}\ \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t}, \mu\right) \tag{52}$$

$$\varepsilon_t^2 = \left(X_t - \mu_t - \frac{\overline{\beta}\sqrt{\overline{\gamma}}}{\overline{\alpha}} \sqrt{h_t}\right)^2 \tag{53}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{54}$$

Parametry tego modelu mogą być estymowane na podstawie funkcji wiarygodności:

$$\left\{ \hat{\overline{\alpha}}, \hat{\overline{\beta}}, \hat{\omega}, \hat{\xi} \right\} = \underset{\{\overline{\alpha},\, \overline{\beta},\, \omega,\, \xi,\, \varphi\}}{\arg\max} \ln L_{ACNIG}(\overline{\alpha}, \overline{\beta}, \omega, \xi) =$$
$$= \underset{\overline{\alpha},\, \overline{\beta},\, \omega,\, \xi\}}{\arg\max} \sum_{i=1}^{n} \ln f_{ACNIG}\left(a_t, c_t, x_t; \overline{\alpha}, \overline{\beta}, \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t(\omega, \xi)}, \mu \right) \tag{55}$$

W tym modelu parametry estymuje się na podstawie funkcji wiarygodności, do której skonstruowania wykorzystano ceny minimalne, maksymalne i zamknięcia. Z kolei składnik losowy wyznaczono tak jak w pracy Jensena i Lundego (2001).

Czwartą propozycję modelu, nazwaną S&GARCH-NIG z HLC i oznaczoną w dalszej części pracy jako NIG_{12}, przedstawiają poniższe równania:

$$(A_t, C_t, X_t) | \Im_{t-1} \sim ACNIG\left(\overline{\alpha}, \overline{\beta}, \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t}, \mu \right) \tag{56}$$

$$\varepsilon_t^2 = \frac{\overline{\alpha}^2}{\overline{\gamma}^2} (C_t(C_t - X_t) + A_t(A_t - X_t)) \tag{57}$$

$$h_t = \omega_0 + \sum_{i=1}^{q} \omega_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \xi_i h_{t-i} \tag{58}$$

Parametry tego modelu mogą być estymowane za pomocą MNW na podstawie funkcji wiarygodności z wykorzystaniem cen zamknięcia oraz nowego estymatora dziennej wariancji przedstawionego w równaniu (42):

$$\left\{ \hat{\overline{\alpha}}, \hat{\overline{\beta}}, \hat{\omega}, \hat{\xi} \right\} = \underset{\{\overline{\alpha},\, \overline{\beta},\, \omega,\, \xi,\}}{\arg\max} \ln L_{NIG}(\overline{\alpha}, \overline{\beta}, \omega, \xi) =$$
$$= \underset{\{\overline{\alpha},\, \overline{\beta},\, \omega,\, \xi,\}}{\arg\max} \sum_{i=1}^{n} \ln f_{NIG}\left(x_t; \overline{\alpha}, \overline{\beta}, \frac{\overline{\gamma}^{3/2}}{\overline{\alpha}} \sqrt{h_t(\omega, \xi)}, \mu \right) \tag{59}$$

Zauważmy, że parametryzacje NIG_{12} i NIG_{22} są identyczne, a różnica między nimi wynika z zupełnie innej konstrukcji funkcji wiarygodności.

Jak wspomniano, modele NIG_{12} i NIG_{22} zostały zbudowane na podstawie estymatora, który jest modyfikacją estymatora Rogersa i Satchella. W modelach NIG_{12} i NIG_{22} wykorzystano natomiast estymator zaprezentowany w równaniu (10), będący efektywniejszym estymatorem wariancji arytmetycznego ruchu Browna. Gdyby estymator Rogersa i Satchella został zastosowany w modelach N_{ij} wówczas z własności rozkładu NIG wynika, że dla $\alpha \to \infty$, $\delta = \alpha\sigma^2$, $\beta = 0$ wszystkie modele NIG_{ij} byłyby równoważne modelom N_{ij}. Oznacza to, że modele NIG_{ij} nie stanowią odrębnej klasy, lecz są uogólnieniami odpowiednich modeli N_{ij}.

Podobnie jak w przypadku analizy ruchu Browna przeprowadzonej w części trzeciej, parametryzacja NIG_{22} różni się od parametryzacji modelu NIG_{11} tym, że ceny minimalne i maksymalne zostały wykorzystane do konstrukcji zarówno estymatora dziennej wariancji, jak i funkcji wiarygodności służącej do estymacji parametrów modelu GARCH. Zestawienie wszystkich modeli przedstawionych w opracowaniu zawiera tabela 1.

W modelach NIG_{22}, NIG_{21} i NIG_{12} przyjęto, że stopy zwrotu minimum, maksimum i wartości końcowej mają łączny rozkład ACNIG. Nie wprowadzono żadnych dodatkowych założeń dotyczących śróddziennych zmian stóp zwrotu. W szczególności śróddzienna stopa zwrotu nie musi być zmienną losową, której rozkład jest mieszaniną dwóch rozkładów. Jest to osłabienie założeń w stosunku do tych, które przyjęto w pracy Ventera, de Jongha i Griebenowa (2005).

Ponadto modele NIG_{22}, NIG_{21} i NIG_{12} są oszczędnie sparametryzowane i nie zostały do nich wprowadzone żadne dodatkowe parametry w stosunku do modelu NIG_{11}.

6. Analiza zmienności zwrotów z indeksu WIG20 i kursu EUR/PLN

6.1. Zarys badania

Przedstawione w pracy propozycje parametryzacji modeli GARCH zastosowano do szacowania zmienności zwrotów z indeksu rynku akcji WIG20 notowanego na GPW w Warszawie oraz kursu walutowego EUR/PLN notowanego na międzynarodowym rynku walutowym FOREX (dane pochodziły z agencji Bloomberg). Analizę przeprowadzono dla okresu od 30 września 2002 r. do 28 września 2012 r., w którym obserwowano zarówno hossę, jak i bessę, a także, co istotne, kryzys finansowy. Długości badanych szeregów czasowych stóp zwrotu wynosiły 2513 dla indeksu WIG20 oraz 2571 dla kursu EUR/PLN.

Powszechnie dostępne dane rynkowe zawierające szeregi czasowe z cenami: otwarcia, minimum, maksimum i zamknięcia, są nieco inaczej określone, niż przyjęto w punkcie 2.2. W szczególności cena otwarcia w dniu bieżącym jest na ogół różna od ceny zamknięcia z poprzedniego dnia (występują tzw. nocne stopy zwrotu). Z tego powodu dokonano redefinicji zmiennych: $A_t = \ln(\min(S_{t-1}, L_t)/S_{t-1})$, $C_t = \ln(\max(S_{t-1}, H_t)/S_{t-1})$ i $X_t = \ln(S_t/S_{t-1})$. Następnie wyznaczono wektory dziennych stóp zwrotu: $(a_1, c_1, x_1), (a_2, c_2, x_2), \ldots, (a_n, c_n, x_n)$.

W badaniu zastosowano osiem parametryzacji modeli GARCH przedstawionych w trzeciej i piątej części: N_{11}, N_{12}, N_{21}, N_{22}, NIG_{11}, NIG_{12}, NIG_{21}, NIG_{22}. Do estymacji parametrów modeli zastosowano metodę największej wiarygodności. Logarytmy funkcji wiarygodności $\ln L_N$ i $\ln L_{NIG}$ zbudowane wyłącznie na podstawie cen zamknięcia zostały oznaczone jako $\ln L_1$, a funkcje $\ln L_{ACN}$ i $\ln L_{ACNIG}$, które były konstruowane dla wektorów losowych, oznaczono jako $\ln L_2$. Parametry modeli N_{11}, N_{12}, NIG_{11}, NIG_{12}, estymowano, maksymalizując logarytm funkcji wiarygodności $\ln L_1$. Do celów informacyjnych podano jednak także wartości $\ln L_2$ (tj. $\ln L_{ACN}$ w przypadku N_{11} i N_{12} oraz $\ln L_{ACNIG}$ dla modeli NIG_{11} i NIG_{12}). Również parametry modeli N_{21}, N_{22}, NIG_{21}, NIG_{22} szacowano, maksymalizując wartości $\ln L_2$. Podano również wartości $\ln L_1$: $\ln L_N$ w przypadku N_{21} i N_{22} oraz $\ln L_{NIG}$ dla modeli NIG_{21} i NIG_{22}). Miara $\ln L_2$ zawiera więcej istotnych informacji na temat kształtowania się cen instrumentu finansowego, gdyż ceny minimalne i maksymalne są z pewnością ważnymi informacjami z punktu widzenia pomiaru zmienności (por. Parkinson 1980; Rogers, Satchell 1991). Można zatem przypuszczać, że jest bardziej wiarygodną miarą oceny jakości danego modelu.

Dodatkowo w przypadku każdego modelu badano również wartości kryterium informacyjnego Schwarza, obliczone dla obu funkcji wiarygodności, oraz wartości statystyki Riversa i Vuonga (2002), oznaczanej dalej jako RV. Test ten pozwala na weryfikację hipotezy o asymptotycznej równości wartości funkcji wiarygodności dwóch niezagnieżdżonych modeli. Jest on rozszerzeniem testu Vuonga (1989) i może być stosowany między innymi w przypadku modeli szeregów czasowych. Dokonano również oceny jakości przedstawionych modeli dla cen zamknięcia, stosując wybrane testy statystyczne: test autokorelacji Ljunga i Boxa, test efektu ARCH Engle'a oraz test zgodności Andersona i Darlinga. Zarówno dla indeksu WIG20, jak i dla kursu walutowego nie zaobserwowano istotnej statystycznie autokorelacji składnika losowego (wyjątkiem jest model NIG_{22} dla kursu EUR/PLN). W przypadku obu szeregów czasowych przyjęto rzędy opóźnień w modelach GARCH równe jeden. Dla indeksu WIG20 we wszystkich modelach poza N_{22} występował istotny statystycznie efekt ARCH składnika losowego. Z kolei w przypadku kursu EUR/PLN efekt taki występował w modelach, w których zastosowano nowe estymatory dziennych wariancji. W obu przypadkach modele o wyższych rzędach opóźnień były zdecydowanie odrzucane ze względu na mniejsze wartości funkcji wiarygodności ln L_2 oraz większe wartości kryterium informacyjnego Schwarza i dlatego zostały pominięte w dalszej prezentacji. Wyniki estymacji zaprezentowano w tabelach 2 i 3, odpowiednio dla indeksu WIG20 oraz kursu EUR/PLN.

Jak można było przypuszczać, hipoteza zakładająca normalność rozkładów warunkowych została odrzucona na podstawie wyników testu Andersona i Darlinga dla wszystkich modeli N_{ij}. Z kolei hipoteza zakładająca warunkowy rozkład NIG nie została odrzucona tylko dla modeli N_{11} (oba szeregi) oraz N_{12} (indeks WIG20).

Wszystkie obliczenia wykonano za pomocą samodzielnie napisanych kodów źródłowych w języku C++, wykorzystując bibliotekę obliczeń numerycznych GSL (http://www.gnu.org/software/gsl/).

6.2. Wnioski z badania empirycznego

Wykorzystanie informacji o cenach minimalnych i maksymalnych w procesie estymacji parametrów modelu GARCH, bez zwiększenia jego parametryzacji, poprawia jakość modelu mierzoną wartością funkcji wiarygodności. Według testu RV zarówno dla indeksu WIG20, jak i kursu EUR/PLN wartości funkcji wiarygodności ln L_2 modeli N_{12}, N_{21}, N_{22} są istotnie większe od wartości dla modelu N_{11}. Takie same zależności występują w modelach klasy NIG_{ij}, z wyjątkiem modelu NIG_{12} dla indeksu WIG20. We wszystkich przypadkach najbardziej zaawansowane modele, wykorzystujące dane o cenach minimalnych i maksymalnych zarówno do konstrukcji funkcji wiarygodności, jak i do estymacji dziennej wariancji, czyli modele N_{22} i NIG_{22} mają dużą przewagę nad modelami N_{11} i NIG_{11}.

Zastosowanie estymatorów dziennej wariancji (równanie (10) dla modeli N_{ij} oraz (42) dla modeli NIG_{ij}), konstruowanych na podstawie danych o cenach minimalnych i maksymalnych dla kursu EUR/PLN, zwiększa wartości ln L_2. Znaczy to, że wartość funkcji wiarygodności ln L_2 jest wyższa w modelu N_{12} niż w modelu N_{11}, w N_{22} jest wyższa niż N_{21}, w NIG_{12} jest wyższa niż NIG_{11}, a w NIG_{22} jest wyższa niż w NIG_{21}. Wzrost funkcji wiarygodności dla każdej z przedstawionych czterech relacji jest istotny statystycznie, na co wskazują oceny statystyki RV, wynoszące odpowiednio: -5,9560; -6,8120; -5,9864 i -8,0456. W przypadku indeksu WIG20 tylko model N_{12} ma istotnie wyższą wartość funkcji wiarygodności ln L_2 niż w modelu N_{11} (wartość testu RV jest równa -2,0713). Zmiany ln L_2 dla pozostałych trzech par modeli nie są istotne statystycznie (oceny statystyki RV wynoszą odpowiednio: -0,9486; -0,0506 i 1,1548).

Poprawa jakości modelu wynikająca z zastosowania rozkładu NIG, czyli rozkładu o grubszych ogonach niż w rozkładzie normalnym, jako warunkowego rozkładu składnika losowego jest większa niż w przypadku miary ln L_2. Oszacowano również parametry modelu GARCH z warunkowym skośnym rozkładem t Studenta składnika losowego na podstawie szeregu czasowego zawierającego wyłącznie ceny zamknięcia. W przypadku indeksu WIG20 dla modelu GARCH(1,1) uzyskano wartość funkcji wiarygodności ln L_1 równą 7128,16, czyli niższą niż w przypadku modelu NIG_{11}. Z kolei w modelu GARCH(1,1) dla kursu EUR/PLN wynosiła ona 9607,033, czyli nieznacznie więcej niż w modelu NIG_{11}. Szczegółowe wyniki estymacji zostały pominięte w tabelach.

Powyższe wnioski zostały również potwierdzone przez wartości kryterium informacyjnego Schwarza.

Zastosowanie danych o cenach minimalnych i maksymalnych do budowy oraz estymacji modeli najczęściej powoduje pogorszenie jakości modeli, ocenianej przez pryzmat statystycznych własności składników losowych konstruowanych na podstawie cen zamknięcia (pojawiający się efekt ARCH oraz gorsze dopasowanie warunkowych rozkładów). Takiego wyniku można się było spodziewać i nie oznacza on, że modele te mają gorszą jakość. Do ich oceny należałoby wykorzystać większy zbiór informacji o cenach minimalnych, maksymalnych i zamknięcia, jednak nie istnieją jeszcze odpowiednie procedury i testy (będzie to przedmiotem dalszych badań autorów).

Średnie błędy oszacowania parametrów są na ogół mniejsze w przypadku modeli, w których parametry były szacowane z wykorzystaniem cen minimalnych i maksymalnych.

Dla indeksu WIG20 asymetria rozkładów bezwarunkowych procesów opisanych modelami NIG_{21} i NIG_{22} jest większa od tej, która występuje w rozkładach procesów opisanych modelami NIG_{11} i NIG_{12}.

Wykorzystanie w estymacji cen minimalnych i maksymalnych powoduje zmiany ocen parametrów w modelu GARCH. Zwiększają się na ogół oceny parametru ω_1, a zmniejszają się oceny ξ_1 oraz maleje suma ocen $\omega_1 + \xi_1$ w porównaniu z modelem GARCH szacowanym na podstawie cen zamknięcia. Ma to duże znaczenie dla modelowania i prognozowania zmienności stóp zwrotu. Oznacza to bowiem, że wpływ szoków w poprzednim okresie na bieżącą wariancję jest większy, a zatem reakcja na zmieniającą się sytuację rynkową jest szybsza, zgodnie z modelem, w którym zastosowano również ceny minimalne i maksymalne. Ponadto oddziaływanie szoków na zmienność trwa krócej, niż wynikałoby to z modelu GARCH szacowanego wyłącznie na podstawie cen zamknięcia. Przeczy to wynikom uzyskiwanym na przykład na podstawie zmienności implikowanej (por. Engle, Mustafa 1992). Właśnie to zastrzeżenie było dotychczas wymieniane jako jedna z największych słabości modeli GARCH. Wydaje się zatem, że dzięki zastosowaniu danych o cenach minimalnych i maksymalnych do estymacji parametrów modelu uzyskuje się oceny, które są bliższe prawdziwym parametrom.

Wartości funkcji wiarygodności ln L_2 otrzymane dla modelu N_{22} są zbliżone do wartości uzyskanych dla modelu NIG_{11}. Wykorzystanie informacji o cenach minimalnych i maksymalnych do konstrukcji zarówno funkcji wiarygodności, jak i estymatora dziennej wariancji podobnie poprawia jakość modelu, jak zastosowanie asymetrycznego warunkowego rozkładu składnika losowego o grubych ogonach.

Warto również odnotować, że dla indeksu WIG20 w przypadku modeli N_{12} i NIG_{12} oceny parametrów nie spełniają warunku $\omega_1 + \xi_1 < 1$ co sugeruje, że wariancja stóp zwrotu procesu GARCH jest nieskończona. Zjawisko to często dotyczy także innych instrumentów notowanych na rynku finansowym w przypadku estymacji parametrów modelu wyłącznie na podstawie cen zamknięcia (Diebold 1986; Lamoureux, Lastrapes 1990). Wynik ten nie występuje dla modeli N_{21} i NIG_{21} oraz N_{22} i NIG_{22}.

Zróżnicowanie wartości funkcji ln L_1 w poszczególnych modelach jest znacznie mniejsze niż w przypadku ln L_2 i nie przekracza 3%, podczas gdy dla nowej postaci funkcji wiarygodności dochodzi nawet do 8–10%.

7. Podsumowanie

W pracy zaprezentowano modele GARCH wprowadzone przez Lildholdta (2002) oraz Ventera, de Jongha i Griebenowa (2005), skonstruowane nie tylko na podstawie cen zamknięcia, ale również na podstawie informacji o dobowych minimach i maksimach cen. Jako warunkowe rozkłady składnika losowego przyjęto rozkład normalny oraz NIG. Zakładając, że procesy śróddziennych stóp zwrotu są arytmetycznym ruchem Browna oraz procesem o warunkowym rozkładzie NIG, przedstawiono rozkłady łączne wektorów losowych, których współrzędnymi są zmienne losowe wartości minimalnej, maksymalnej i końcowej logarytmicznych stóp zwrotu. W wyniku zastosowania dodatkowych informacji o cenach minimalnych i maksymalnych skonstruowano funkcje wiarygodności, które wykorzystują znacznie szerszy zbiór danych rynkowych opisujących instrumenty finansowe. Dzięki temu uzyskano znacznie większe wartości funkcji wiarygodności.

W pracy zaproponowano rozszerzenie modeli Lildholdta (2002) oraz Ventera, de Jongha i Griebenowa (2005). Polega ono na zastosowaniu efektywniejszych estymatorów dziennej wariancji (w tym propozycji nowego estymatora), konstruowanych na podstawie cen minimalnych i maksymalnych, zamiast estymatora wyznaczanego wyłącznie na podstawie cen zamknięcia. Ponadto dokonano pewnych uproszczeń wymienionych parametryzacji modeli.

Na podstawie obserwacji stóp zwrotu z indeksu WIG20 i kursu walutowego EUR/PLN pokazano, że wykorzystanie informacji o cenach minimalnych i maksymalnych do estymacji parametrów modelu GARCH, bez zwiększenia jego parametryzacji, poprawia jakość modelu mierzoną wartością funkcji wiarygodności. Zmienność stóp zwrotu oszacowana na podstawie zaprezentowanych modeli ma własności bliskie zmienności implikowanej. Stosowanie klasycznych modeli GARCH i wyłącznie cen zamknięcia może zatem prowadzić do błędów poznawczych.

Wszystkie przedstawiane modele GARCH są oszczędnie sparametryzowane i mają tyle samo parametrów, co tradycyjny model GARCH ze składnikiem losowym opisanym warunkowymi rozkładami: normalnym oraz skośnym t Studenta. Wykorzystanie informacji z notowań śróddziennych (w postaci ceny minimalnej i maksymalnej) nie zwiększa częstotliwości analizowanego szeregu czasowego. Do budowy omawianych modeli nadal wykorzystuje się dane o częstotliwości dziennej, tzn. szeregi wektorów stóp zwrotu minimum, maksimum i wartości końcowej. Pozwala to uniknąć problemów związanych z analizą danych o bardzo wysokiej częstotliwości.

Można przypuszczać, że wykorzystanie dodatkowych danych do estymacji parametrów modelu GARCH zwiększy trafność prognozowania zmienności stóp zwrotu w porównaniu z modelami, których parametry są estymowane tylko na podstawie cen zamknięcia. Wymaga to jednak przeprowadzenia dodatkowych badań. Dostępność cen minimalnych i maksymalnych na równi z cenami zamknięcia daje nadzieję, że stosowanie zaprezentowanych modeli stanie się powszechne.

Bibliografia

Andersen T.G., Bollerslev T. (1998), Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, 39(4), 885–905.

Andersen T.G., Bollerslev T., Christoffersen P.F., Diebold F.X. (2006), Volatility and correlations forecasting, w: G. Elliott, C.W.J. Granger, A. Timmermann (red.), *Handbook o f ec onomic forecasting*, 1, North-Holland, Elsevier, Amsterdam.

Andersson J. (2001), On the normal inverse Gaussian stochastic volatility model, *Journal of Business and Economic Statistics*, 19, 44–54.

Barndorff-Nielsen O.E. (1977), Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society London, Series A*, 353, 401–419.

Barndorff-Nielsen O.E. (1997), Normal inverse Gaussian Distributions and stochastic volatility modelling, *Scandinavian Journal of Statistics*, 24(1), 1–13.

Barndorff-Nielsen O.E., Shephard N. (2001), Modelling by Lévy processes for financial econometrics, w: O.E. Barndorff-Nielsen, T. Mikosch, S. Resnick (red.), *Lévy processes – theory and applications,* Birkhauser, Boston.

Bollerslev T. (1986), Generalised autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307–327.

Cox D.R., Miller M.D. (1965), *The theory of stochastic processes*, Methuen and Co., London.

Diebold F.X. (1986), Modelling the persistence of conditional variances: a comment, *Econometric Reviews*, 5, 51–56.

Doman M. (2011), *Mikrostruktura giełd papierów wartościowych*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Engle R.F., Mustafa C. (1992), Implied ARCH models from options prices, *Journal of Econometrics*, 52, 289–311.

Fiszeder P. (2009), *Modele klasy GARCH w empirycznych badaniach finansowych*, Wydawnictwo UMK, Toruń.

Fiszeder P., Perczak G. (2013), A new look at variance estimation based on low, high and closing prices taking into account the drift, *Statistica Neerlandica*, 67(4), 456–481.

Fleming J., Kirby C., Ostdiek B. (2001), The economic value of volatility timing, *Journal of Finance*, 56, 329–352.

Forsberg L. (2002), *On the normal inverse Gaussian distribution in modeling volatility in the financial markets*, Acta Universitatis Upsaliensis, Studia Statistica Upsaliensia, 5, http://uu.diva-portal.org/smash/get/diva2:161613/FULLTEXT01.

Forsberg L., Bollerslev T. (2002), Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model, *Journal of Applied Econometrics*, 17(5), 535–548.

Jensen M.B., Lunde A. (2001), The NIG-S&ARCH model: a fat tailed, stochastic, and autoregressive conditional heteroscedastic volatility model, *Econometrics Journal*, 4, 319–342.

Li A., (1999), The pricing of double barrier options and their variations, *Advances in Futures and Options Research*, 10, 17–41.

Lildholdt P.M. (2002), *Estimation of GARCH models based on open, close, high, and low prices*, Working Paper, 128, Centre for Analytical Finance, Aarhus School of Business.

Lamoureux C.G., Lastrapes W.D. (1990), Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics*, 8, 225–234.

Parkinson M. (1980), The extreme value method for estimating the variance of the rate of return, *Journal of Business*, 53, 61–68.

Perczak G. (2014), Zastosowanie rozkładu NIG w modelowaniu danych finansowych przy wykorzystaniu dodatkowych informacji o cenach minimalnych i maksymalnych, w: T. Czerwińska, A.Z. Nowak (red.), *Rynek kapitałowy wobec wyzwań dekoniunktury*, Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego, Warszawa (przyjęta do druku).

Perczak G., Fiszeder P. (2013), Estymacja wariancji arytmetycznego ruchu Browna na podstawie znanych wartości minimum, maksimum, końcowej oraz dryfu, *Przegląd Statystyczny*, 60(1), 39–62.

Poon S.-H., Granger C. (2003), Forecasting volatility in financial markets: a review, *Journal of Economic Literature*, 41, 478–539.

Rivers D., Vuong Q. (2002), Model selection tests for nonlinear dynamic models, *Econometrics Journal*, 5, 1–39.

Rogers L.C.G., Satchell S.E. (1991), Estimating variance from high, low and closing prices, *The Annals of Applied Probability*, 1 (4), 504–512.

Schwert G.W. (1989), Why does stock market volatility change over time?, *The Journal of Finance*, 44, 1115–1153.

Schlösser A. (2011), *Pricing and risk management of synthetic CDOs*, Springer Verlag, Berlin.

Venter J.H., de Jongh P.J. (2004), Selecting an innovation distribution for Garch models to improve efficiency of risk and volatility estimation, *The Journal of Risk*, 6, 27–52.

Venter J.H., de Jongh P.J., Griebenow G. (2005), NIG-GARCH models based on open, close, high and low prices, *South African Statistical Journal*, 39(2), 79–101.

Venter J.H., de Jongh P.J., Griebenow G. (2006), GARCH-type volatility models based on Brownian inverse Gaussian intra-day return processes, *Journal of Risk*, 8(4), 97–116.

Vuong Q. (1989), Likelihood ratio tests for model selection and non-nested hypotheses, *Econometrica*, 57, 307–333.

Podziękowania

Autorzy pragną podziękować anonimowym recenzentom za uwagi, które przyczyniły się do znacznej poprawy artykułu w stosunku do wcześniejszej wersji tekstu.

Praca została sfinansowana ze środków Narodowego Centrum Nauki projekt numer 2012/05/B/HS4/00675 nt. „Modelowanie i prognozowanie zmienności – wykorzystanie dodatkowych informacji o cenach minimalnych i maksymalnych”.

Aneks

Tabela 1

Zestawienie zaprezentowanych w opracowaniu modeli

<table>
<tr><th colspan="3" rowspan="2">Zastosowany rozkład warunkowy oraz przyjęta funkcja wiarygodności</th><th colspan="2">Estymatory dziennej wariancji</th></tr>
<tr><th>klasyczny konstruowany dla cen zamknięcia</th><th>efektywniejsze konstruowane dla cen HLC</th></tr>
<tr><td rowspan="4">Rozkład normalny</td><td rowspan="2">standardowa funkcja wiarygodności</td><td>nazwa modelu</td><td>S&GARCH
Bollerslev (1986)</td><td>S&GARCH z HLC</td></tr>
<tr><td>oznaczenie</td><td>N_{11}</td><td>N_{12}</td></tr>
<tr><td rowspan="2">rozszerzona funkcja wiarygodności</td><td>nazwa modelu</td><td>S&GARCH-HL
Lildholdt (2002)</td><td>S&GARCH-HL z HLC</td></tr>
<tr><td>oznaczenie</td><td>N_{21}</td><td>N_{22}</td></tr>
<tr><td rowspan="4">Rozkład NIG</td><td rowspan="2">standardowa funkcja wiarygodności</td><td>nazwa modelu</td><td>S&GARCH-NIG
Jensen, Lunde (2001)</td><td>S&GARCH-NIG z HLC</td></tr>
<tr><td>oznaczenie</td><td>NIG_{11}</td><td>NIG_{12}</td></tr>
<tr><td rowspan="2">rozszerzona funkcja wiarygodności</td><td>nazwa modelu</td><td>S&GARCH-HL-NIG
Venter, de Jongh, Griebenow (2005)</td><td>S&GARCH-HL-NIG z HLC</td></tr>
<tr><td>oznaczenie</td><td>NIG_{21}</td><td>NIG_{22}</td></tr>
</table>

Uwaga: modele zaprezentowane w pracy nie są dokładnymi parametryzacjami przedstawionymi w cytowanych artykułach, ale ich modyfikacjami dokonanymi na potrzeby niniejszej pracy.

Tabela 2

Wyniki estymacji dla indeksu WIG20

Parametr	**Parametryzacje modelu GARCH**							
	N_{11}	N_{12}	N_{21}	N_{22}	NIG_{11}	NIG_{12}	NIG_{21}	NIG_{22}
$\bar{\alpha}$					2,95600 (0,68023)	3,20715 (0,74968)	3,00467 (0,17284)	2,84584 (0,15986)
$\bar{\beta}$					-0,02975 (0,09352)	-0,08676 (0,10159)	-0,37455 (0,09011)	-0,54468 (0,08985)
μ	0,00072 (0,00026)	0,00050 (0,00026)	0,00133 (0,00016)	0,00054 (0,00021)	0,00070 (0,00069)	0,00108 (0,00071)	0,00337 (0,00055)	0,00415 (0,00056)
ω_0	2,04e-06 (6,61e-07)	1,71e-06 (9,43e-07)	3,17e-06 (2,75e-07)	3,23e-06 (4,42e-07)	1,86e-06 (7,10e-07)	1,75e-06 (1,06e-06)	3,45e-06 (4,93e-07)	2,56e-06 (5,69e-07)
ω_1	0,06009 (0,00747)	0,11327 (0,01756)	0,07822 (0,00295)	0,16651 (0,00866)	0,05493 (0,00821)	0,11266 (0,02178)	0,08001 (0,00545)	0,11743 (0,01089)
ξ_1	0,93196 (0,00811)	0,90761 (0,01379)	0,86033 (0,00526)	0,78951 (0,01103)	0,93787 (0,00889)	0,91270 (0,01618)	0,87003 (0,00869)	0,86921 (0,01198)
$\ln L_1$	7108,59	7128,30	6918,41	6959,05	7129,86	7144,78	7041,19	7071,64
SIC L_1	-14178,03	-14217,45	-13797,67	-13878,95	-14204,92	-14234,75	-14027,58	-14088,48
$\ln L_2$	27849,75	28046,59	29273,95	29341,05	29591,04	29592,34	29949,87	29918,53
SIC L_2	-55660,36	-56054,03	-58508,76	-58642,95	-59127,27	-59129,88	-59844,93	-59782,25
RV	–	-2,0713*	-9,6472*	-9,3264*	–	-0,0506	-8,4861*	-6,9021*
LB(8)	7,4309	7,0155	6,1990	6,8660	7,6305	7,4085	6,4492	7,5991
LM(8)	25,6543*	21,3554*	44,3944*	14,8902	27,5269*	29,0166*	36,763*	19,5065*
AD	3,5865*	2,9092*	3,8528*	3,3194*	0,7279	0,2872	17,8657*	8,9140*

* Oceny statystyk istotnie różne od zera na poziomie 0,05.

Uwaga: w nawiasach podano średnie błędy szacunku parametrów, RV oznacza statystykę testu Riversa, Vuonga (2002), porównywania dokonywano dla modeli klasy N_{ij} z modelem N_{11}, a w przypadku modeli klasy NIG_{ij} z modelem NIG_{11}, LB to statystyka testu autokorelacji Ljunga-Boxa, LM statystyka testu efektu ARCH Engle'a, AD to statystyka testu zgodności Andersona-Darlinga.

Tabela 3

Wyniki estymacji dla kursu EUR/PLN

Parametr	Parametryzacje modelu GARCH							
	N_{11}	N_{12}	N_{21}	N_{22}	NIG_{11}	NIG_{12}	NIG_{21}	NIG_{22}
$\bar{\alpha}$					2,10205 (0,34480)	1,47889 (0,22286)	1,14614 (0,08023)	1,97952 (0,10041)
$\bar{\beta}$					0,22648 (0,07391)	0,15800 (0,05628)	0,14949 (0,04780)	0,15740 (0,05191)
μ	-0,00022 (0,00010)	-0,00004 (0,00011)	-0,00049 (0,00036)	-0,00005 (0,00011)	-0,00101 (0,00024)	-0,00081 (0,00023)	-0,00135 (0,00021)	-0,00071 (0,00019)
ω_0	8,725e-07 (2,05e-07)	4,667e-06 (8,39e-07)	1,769e-06 (9,53e-08)	2,774e-06 (1,97e-07)	7,797e-07 (2,29e-07)	4,991e-06 (1,11e-06)	4,412e-06 (3,28e-07)	1,107e-06 (2,20e-07)
ω_1	0,8866 (0,01139)	0,16444 (0,02536)	0,08111 (0,00361)	0,29653 (0,00949)	0,0879 (0,01367)5	0,15068 (0,02938)	0,17111 (0,02140)	0,19834 (0,01772)
ξ_1	0,89196 (0,01342)	0,68230 (0,04639)	0,87310 (0,00522)	0,59690 (0,01292)	0,89438 (0,01584)	0,73463 (0,05717)	0,78144 (0,02430)	0,78272 (0,01942)
ln L_1	9540,67	9486,97	9527,48	9449,17	9605,08	9535,15	9556,04	9493,73
SIC L_1	-19042,07	-18934,68	-19015,70	-18859,08	-19155,20	-19015,33	-19057,11	-18932,50
ln L_2	32076,39	33651,33	32205,82	34062,57	34791,01	35310,88	34980,03	35456,78
SIC L_2	-64113,53	-67263,40	-64372,37	-68085,89	-69527,05	-70566,80	-69905,10	-70858,60
RV	–	-5,9560*	-2,0561*	-6,7484*	–	-5,9864*	-3,5391*	-7,0237*
LB(8)	8,2695	11,5736	9,4457	10,6462	8,6418	13,9906	9,2419	16,6924*
LM(8)	3,0272	52,773*	6,3042	40,3654*	3,0267	102,955*	4,4186	212,359*
AD	6,4284*	8,8111*	6,7096*	10,4318*	0,5567	5,2963*	14,8275*	3,5221*

* Oceny statystyk istotnie różne od zera na poziomie 0,05.

Uwaga: w nawiasach podano średnie błędy szacunku parametrów, RV oznacza statystykę testu Riversa, Vuonga (2002), porównywania dokonywano dla modeli klasy N_{ij} z modelem N_{11}, a w przypadku modeli klasy NIG_{ij} z modelem NIG_{11}, LB to statystyka testu autokorelacji Ljunga-Boxa, LM statystyka testu efektu ARCH Engle'a, AD to statystyka testu zgodności Andersona-Darlinga.

The GARCH model – the application of additional information about low and high prices

Abstract

The paper presents the GARCH models established by Lildholdt (2002) and Venter, de Jongh, Griebenow (2005), which are formulated on the basis of low, high and closing prices. Assuming that intraday processes of returns are the arithmetic Brownian motion and then the process, for description of which the NIG distribution was used, the joint distributions of random vectors of minimum, maximum and finish values of logarithmic returns are presented. These distributions are then applied in the formulation of likelihood functions used for the estimation of parameters of considered models.

Moreover, extensions of models of Lildholdt (2002) and Venter, de Jongh, Griebenow (2005) are proposed. These modifications include more efficient estimators of daily variance based on low and high prices in place of the estimator based on only closing prices. Some simplifications of the mentioned parameterizations of models are also performed. It is demonstrated for series of returns of the stock index WIG20 and the EUR/PLN exchange rate that the application of information about low and high prices in the estimation process of the GARCH model parameters, without increasing its parameterization, improves the quality of the model measured by the value of its likelihood function.

Keywords: GARCH model, volatility estimation, NIG distribution, Brownian motion, low and high prices