ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 177, 2004

*Małgorzata Doman**

PROGNOZOWANIE ZMIENNOŚCI POLSKICH INDEKSÓW GIEŁDOWYCH ZA POMOCĄ MODELI GARCH PRZY UŻYCIU DANYCH WYSOKIEJ CZĘSTOTLIWOŚCI

Streszczenie. Wprowadzone przez Andersena i Bollersleva pojęcie dziennej zmienności zrealizowanej dało nowy impuls badaniom poświęconym modelowaniu i prognozowaniu zmienności cen instrumentów finansowych przy użyciu modeli GARCH. Dzienna zmienność zrealizowana jest określona jako suma kwadratów zwrotów śróddziennych. Odnoszenie dziennych prognoz modeli GARCH do tak rozumianej zmienności zwykle znacznie poprawia jakość prognozy. Praca poświęcona jest prognozowaniu dziennej zmienności zrealizowanej indeksów Warszawskiej Giełdy Papierów Wartościowych za pomocą modeli z rodziny GARCH, w których opóźniona dzienna zmienność zrealizowana została również wprowadzona jako dodatkowa zmienna objaśniająca.

Słowa kluczowe: prognozowanie, zmienność zrealizowana, GARCH, dane wysokiej częstotliwości.

1. WSTĘP

Zmienność ceny jest jednym z najważniejszych parametrów charakteryzujących instrument finansowy. Modele typu GARCH ze względu na swą prostotę są najbardziej popularnym narzędziem estymowania tej nieobserwowalnej wielkości. Przez długi czas głównym zarzutem pod adresem modeli GARCH była ich słaba zdolność do prognozowania zmienności na okres poza próbą. Zapoczątkowane w 1998 r. przez Andersena i Bollersleva (1998) badania tzw. zmienności zrealizowanej doprowadziły do dość radykalnej zmiany opinii ekonometryków finansowych na temat możliwości prognostycznych modeli GARCH w tej dziedzinie. Dzienna zmienność zrealizowana, zaproponowana przez wymienionych autorów jako bardziej adekwatna miara „prawdziwej" zmienności ceny instrumentu finansowego, jest, mówiąc

*Dr, Katedra Matematyki Stosowanej, Akademia Ekonomiczna w Poznaniu.

w skrócie, sumą kwadratów zwrotów śróddziennych. Uzasadnienie takiego określenia tej wielkości związane jest z założeniem, że proces ceny instrumentu jest ciągłym procesem dyfuzji. Zastosowanie zmienności zrealizowanej jako punktu odniesienia dla prognoz uzyskiwanych za pomocą modeli GARCH skutkuje radykalną poprawą jakości prognoz. Jeżeli przyjmiemy, że zmienność zrealizowana jest dobrym estymatorem zmienności dziennej, to powstaje pytanie, czy można polepszyć jakość prognoz, wprowadzając opóźnione wartości zmienności zrealizowanej jako dodatkową zmienną objaśniającą do klasycznego modelu typu GARCH. Ostatnio pojawiły się prace (Hol i Koopman, 2002; Martens, 2002), które potwierdzają tezę, że podejście takie może istotnie poprawić jakość uzyskiwanych prognoz.

Szeregi danych dotyczących notowań ciągłych na Giełdzie Papierów Wartościowych w Warszawie są jeszcze stosunkowo krótkie. Dlatego dokonywane obecnie próby zastosowania zmienności zrealizowanej w prognozowaniu zmienności instrumentów finansowych muszą być uznane za wstępne. W niniejszym opracowaniu podejmujemy próbę wykorzystania zmienności zrealizowanej jako zmiennej objaśniającej w modelach GARCH, opisujących dzienną zmienność indeksów WIG20 i TechWIG. Otrzymane w ten sposób prognozy są porównywane z prognozami otrzymywanymi z klasycznych modeli GARCH na podstawie prób o różnych długościach. Jakość prognoz jest oceniana w stosunku do zmienności zrealizowanej, wprowadzonej przez Andersena i Bollersleva oraz dwóch modyfikacji tego pojęcia, zaproponowanych przez Martensa. Poszukujemy jednocześnie odpowiedzi na pytanie, jakiego typu modele najlepiej prognozują zmienność na polskim rynku finansowym oraz jaka miara zmienności dziennej jest najlepszym estymatorem zmienności instrumentów finansowych, rozumianej jako wariancja warunkowa stopy zwrotu tego instrumentu.

2. MODELE GARCH

Zmienność (*volatility*) ceny instrumentu finansowego (np. akcji, indeksu, kursu walutowego) jest ogólnie określana jako miara niepewności co do przyszłych zmian tej ceny. Jeśli R_t jest dziennym zwrotem z rozważanego instrumentu finansowego, a Ω_{t-1} oznacza informacje na temat procesu R_t dostępne do dnia $t-1$ włącznie, to zmienność σ_t^2 w t-tym dniu sesyjnym jest zazwyczaj definiowana jako wariancja warunkowa:

$$\sigma_t^2 = Var(R_t|\Omega_{t-1}) \tag{1}$$

Najpopularniejszym narzędziem modelowania tak określonej zmienności są uogólnione modele autoregresyjne heteroskedastyczności warunkowej

(Generalised AutoRegressive Conditional Heteroskedasticity) GARCH(p, q). Model taki został zdefiniowany w roku 1986 przez Bollersleva (1986) jako rozszerzenie modelu zaproponowanego przez Engle'a. Jeśli przez y_t oznaczymy zwrot ceny z usuniętą średnią warunkową, tzn. $y_t = R_t - E(R_t|\Omega_{t-1})$, to model GARCH($p$, q) opisuje się zależnościami:

$$y_t = \sigma_t z_t,$$
$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2, \tag{2}$$

gdzie z_t jest szeregiem niezależnych standaryzowanych zmiennych losowych o tym samym rozkładzie. Dla $p = 0$ otrzymujemy model ARCH(q) zdefiniowany przez Engle'a. Często przyjmuje się, że innowacje z_t mają rozkład normalny, co jest równoważne z tym, że rozkład warunkowy zmiennej y_t jest normalny ze średnią zero i wariancją σ_t^2. Niemniej jednak jej rozkład bezwarunkowy nie jest w tym przypadku normalny, a w szczególności kurtoza dla y_t jest większa od 3 i rozkład bezwarunkowy ma grubsze ogony niż rozkład normalny.

W szeregach danych finansowych często bezwarunkowa kurtoza teoretyczna dla y_t jest jednak dużo mniejsza niż kurtoza estymowanych reszt standaryzowanych $\hat{z}_t = y_t/\sigma_t$. Ponieważ kurtoza bezwarunkowa zmiennej y_t jest funkcją rosnącą kurtozy zmiennej z_t, więc można ją zwiększyć przyjmując, że rozkład zmiennej z_t jest leptokurtyczny. W związku z tym w niniejszym artykule rozważamy również modele, w których z_t ma rozkład t-Studenta lub uogólniony rozkład błędu GED.

Model typu GARCH można uogólnić przez wprowadzenie dodatkowej zmiennej objaśniającej x_t w równaniu określającym zmienność. Przyjmuje ono wtedy postać

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \gamma x_{t-k} \tag{3}$$

gdzie k oznacza liczbę opóźnień zmiennej x_t. Tak określony model będziemy oznaczać przez GX(p, q, k).

3. OZNACZENIA I DANE

Dane wykorzystywane w tym opracowaniu dotyczą notowań indeksów WIG20 oraz TechWIG i są dwojakiego rodzaju. Rozważamy dzienne kursy zamknięcia dla indeksu WIG20 od 29.12.1994 r. do 14.03.2003 r., a dla

TechWIGu od 29.12.2000 r. oraz dla obu indeksów notowania 20-minutowe z okresu od 17.11.2000 r. do 14.03.2003 r.. Pierwsza rozważana przez nas w dniu t wartość indeksu, $P_{t,0}$, pochodzi z godziny 10.05, a ostatnia, $P_{t,D}$, z 16.00. Jeśli w notowaniach indeksów występują luki, związane np. z krótszym czasem działania giełdy w danym dniu, to przyjmujemy, że przez cały okres zawieszenia notowań wartość indeksu jest stała i równa ostatniemu notowaniu przed zawieszeniem. Zwroty śróddzienne wyliczamy według wzoru

$$R_{t,d} = 100(\ln P_{t,d} - \ln P_{t,d-1}) \tag{4}$$

gdzie $d \geqslant 1$, $t = 1, 2, ..., T$, a T jest liczbą dni, z których pochodzą obserwacje. Zwrot nocny określamy następująco

$$R_{t,N} = 100(\ln P_{t,0} - \ln P_{t-1,D}), \tag{5}$$

gdzie w przypadku notowań 20-minutowych $D = 18$.

Szeregi zwrotów dziennych wyliczano na podstawie kursów zamknięcia P_t według wzoru

$$R_t = 100(\ln P_t - \ln P_{t-1}) \tag{6}$$

4. ZMIENNOŚĆ ZREALIZOWANA

Jeżeli traktujemy kwadrat zwrotu dziennego jako estymator dziennej zmienności, to jakość prognoz zmienności jest słaba. Wynika to z samych własności modeli GARCH. W swojej fundamentalnej pracy Andersen i Bollerslev (1998) proponują nowe podejście do pojęcia zmienności instrumentu finansowego. Zakładając, że proces p_t logarytmów cen instrumentu finansowego jest ciągłym procesem dyfuzji, danym przez stochastyczne równanie różniczkowe

$$dp_t = \sigma_t \, dW_t \tag{7}$$

gdzie $t \geqslant 0$, W_t oznacza standardowy ruch Browna, a σ_t jest procesem ściśle stacjonarnym, wspomniani autorzy wykazują, że poprawna miara dziennej zmienności ceny w dniu $t+1$ wyraża się wzorem

$$\sigma^2_{t+1,1} = \int_0^1 \sigma^2_{t+\tau} \, d\tau \tag{8}$$

Nieobserwowalną wielkość $\sigma^2_{t+1,1}$ można przybliżać przez dzienne sumy kwadratów zwrotów śróddziennych. W tym celu Andersen i Bollerslev wprowadzili pojęcie tzw. dziennej zmienności zrealizowanej, określonej jako suma kwadratów zwrotów śróddziennych. W niniejszym artykule wielkość tę określamy przez

$$Real1_t = R^2_{t,N} + \sum_{d=1}^{D} R^2_{t,d} \tag{9}$$

W pracy Andersena i Bollersleva (1998) rozważane były zwroty 5-minutowe. Późniejsze analizy przeprowadzane przez różnych autorów (Oomen, 2001; Martens, 2002; Doman i Doman, 2003) wykazały, że ze względu na tzw. efekty mikrostruktury rynków finansowych w niektórych przypadkach lepszym oszacowaniem zmienności jest zmienność zrealizowana, oparta na obserwacjach dokonywanych w większych odstępach czasu. W niniejszym opracowaniu, w związku ze stosunkowo niską płynnością występującą na GPW w Warszawie, przyjmujemy 20-minutowy odstęp pomiędzy kolejnymi notowaniami.

Stosując metody ciągłych procesów stochastycznych do analizy danych giełdowych, napotykamy pewną dodatkową trudność związaną z tym, że giełda nie działa przez całą dobę. Odstęp czasowy między ostatnim notowaniem danego dnia i pierwszym następnego jest długi, co powoduje, że zwrot nocny ma inną naturę niż zwroty śróddzienne. W przypadku polskiego rynku finansowego efekt ten potęguje ogromny napływ informacji gospodarczych, pojawiających się po zamknięciu giełdy (np. po 16.00 publikowane są przez GUS i NBP dane makroekonomiczne, również spółki akcyjne zgodnie z uregulowaniami KPWiG są zobowiązane do podawania istotnych informacji dopiero po 16.00).

Wszystko to powoduje, że choć *Real*1 jest znacznie lepszym estymatorem zmienności niż kwadrat zwrotu dziennego, to jednak uważa się, iż nadal jest to estymator bardzo zanieczyszczony szumem. Alternatywna miara zmienności zrealizowanej *Real*2, zaproponowana w pracy Andersena i Bollersleva, Diebolda i Ebensa (2001) jest równa sumie kwadratów zwrotów środdziennych bez uwzględniania zwrotu nocnego

$$Real2_t = \sum_{d=1}^{D} R^2_{t,d} \tag{10}$$

Podobne podejście zostało zastosowane już wcześniej w pracy Andersena i Bollersleva (1997). Badania empiryczne wykazują jednak, że miara ta daje zaniżone estymacje zmienności zrealizowanej. W związku z tym Martens

(2002) proponuje przeskalowanie *Real*2 za pomocą pewnego współczynnika i określa zmienność zrealizowaną *Real*3 jako

$$Real3_t = (1+c)\sum_{d=1}^{D} R_{t,d}^2 \tag{11}$$

Pomysł jest prosty: chodzi o uwzględnienie wpływu zwrotu nocnego na zmienność bez wprowadzania do miary szumu nieodłącznie związanego z tym zwrotem. Opinie na temat wartości stałej c są różne. W tym opracowaniu, naśladując Koopmana i Hol (2002), przyjmujemy $c = \frac{\sigma_{co}^2}{\sigma_{oc}^2}$, gdzie σ_{co}^2 oznacza wariancję zwrotu między kursem zamknięcia a kursem otwarcia następnego dnia, a σ_{oc}^2 jest wariancją zwrotu dziennego. Zatem miara zmienności zrealizowanej *Real*3 ma postać

$$Real3_t = \left(1 + \frac{\sigma_{co}^2}{\sigma_{oc}^2}\right)\sum_{d=1}^{D} R_{t,d}^2 \tag{12}$$

5. METODY OCENY JAKOŚCI PROGNOZ

W przypadku modelu GARCH(1,1) optymalna w sensie błędu średniokwadratowego prognoza $E(\sigma_{t+h|t}^2)$ wariancji warunkowej σ_{t+h}^2 h kroków naprzód jest wyznaczona wzorem

$$E(\sigma_{t+h|t}^2) = \omega + \alpha_1 E(y_{t+h-1|t}^2) + \beta_1 E(\sigma_{t+h-1|t}^2) \tag{13}$$

gdzie $E(y_{t+i|t}^2 = E(\sigma_{t+i|t}^2)$ dla $i > 0$ oraz $E(y_{t+i|t}^2) = y_{t+i}^2$ i $E(\sigma_{t+i|t}^2) = \sigma_{t+i}^2$ dla $i \leqslant 0$.

Rozważane w tym opracowaniu miary jakości prognozy są następujące

a) współczynnik R^2 regresji Minzera-Zarnowitza:

$$RealS_{T+h+j} = a + bE(\sigma_{T+h+j|T+j}^2) + u_{T+h+j}, \quad j = 0, \ldots, N-1, \quad S = 1, 2, 3 \tag{14}$$

b) średni błąd bezwzględny (*MAE*):

$$MAE(h) = \frac{1}{N}\sum_{j=0}^{N-1} |E(\sigma_{T+h+j|T+j}^2) - RealS_{T+h+j}|, \quad S = 1, 2, 3 \tag{15}$$

c) błąd średni (ME):

$$ME(h) = \frac{1}{N}\sum_{j=0}^{N-1}(E(\sigma^2_{T+h+j|T+j}) - RealS_{T+h+j}), \quad S = 1, 2, 3 \qquad (16)$$

d) skorygowany błąd bezwzględny ($AMAPE$):

$$AMAPE(h) = \frac{1}{N}\sum_{j=0}^{N-1}\left|\frac{E(\sigma^2_{T+h+j|T+j}) - RealS_{T+h+j}}{E(\sigma^2_{T+h+j|T+j}) + RealS_{T+h+j}}\right|, \quad S = 1, 2, 3 \qquad (17)$$

e) współczynnik rozbieżności Theila (TIC):

$$TIC(h) = \frac{\sqrt{\frac{1}{N}\sum_{j=0}^{N-1}(E(\sigma^2_{T+h+j|T+j}) - RealS_{T+h+j})^2}}{\sqrt{\frac{1}{N}\sum_{j=0}^{N-1}(E(\sigma^2_{T+h+j|T+j}))^2} + \sqrt{\frac{1}{N}\sum_{j=0}^{N-1}(RealS_{T+h+j})^2}}, \quad S = 1, 2, 3 \qquad (18)$$

gdzie: T oznacza liczbę obserwacji w próbie, N jest liczbą prognoz na okres poza próbą, a h jest długością horyzontu prognozy. Ostatnie dwie miary są znormalizowane, tzn. przyjmują wartości pomiędzy zero a jeden.

6. ANALIZA EMPIRYCZNA

W artykule podejmujemy próbę oceny jakości uzyskiwanych za pomocą różnych modeli typu GARCH prognoz dziennej zmienności indeksów WIG20 i TechWIG. Jakość prognozy jest w przypadku każdego modelu odnoszona do trzech różnych estymatorów zmienności dziennej, tzw. zmienności zrealizowanych *Real*1, *Real*2, *Real*3. Tabela 1 zawiera zestawienie analizowanych szeregów zwrotów i ich oznaczenia. W tabeli 2 zawarte są statystyki opisowe szeregów zwrotów.

Tabela 1. Analizowane szeregi zwrotów i ich oznaczenia

Indeks	Nazwa szeregu zwrotów	Okres	Liczba obserwacji
WIG20	WIG20K	17.11.2000–14.03.2003	528 + 52 = 580
WIG20	WIG20D	02.01.1995–14.03.2003	1996 + 52 = 2048
TechWIG	TechWIGK	17.11.2000–14.03.2003	528 + 52 = 580
TechWIG	TechWIGD	03.01.2000–14.03.2003	749 + 52 = 801

Tabela 2. Statystyki opisowe analizowanych szeregów zwrotów

Szereg	Średnia	Min	Max	Odch. stand.	Skośność	Kurtoza
WIG20K	–0,0670	–5,0446	5,4830	1,6619	0,3405	3,5880
WIG20D	0,0201	-14,161	13,709	2,0931	–0,0751	6,4367
TechWIGK	–0,2160	–11,358	8,8510	2,3850	0,1596	4,5189
TechWIGD	–0,1255	–11,612	11,200	2,8395	0,2743	4,9635

Do każdego z rozważanych szeregów dziennych zwrotów indeksów został dopasowany model typu GARCH(p, q). Szeregi WIG20K oraz TechWIGK zawierają zwroty dzienne z okresu, kiedy dostępne były już informacje na temat notowań ciągłych. Dzięki temu można było wyestymować dla nich modele GX(p, q, k). Jako dodatkowa zmienna objaśniająca została wybrana opóźniona zmienność zrealizowana *Real*2. Interesującym faktem jest to, że właściwe (w sensie występowania zależności) opóźnienie wynosi 3. Jest to specyficzna cecha badanych przez nas szeregów. W znanych nam pracach poświęconych modelowaniu zmienności zwykle wprowadza się do modelu zmienność zrealizowaną z dnia poprzedniego. Wszystkie wyliczenia zostały wykonane przy użyciu pakietu G@RCH2.3 (Laurent i Peters 2002). Modele wybrano w oparciu o kryteria informacyjne Akaike'go i Schwarza. Tabele 3 i 4 zawierają informacje dotyczące rodzajów i parametrów dopasowanych modeli. Symbol DF oznacza liczbę stopni swobody w rozkładzie t-Studenta lub GED zmiennej z_t. Współczynniki a_0 i a_1 modelu AR(1) dla zwrotu oraz DF były estymowane jednocześnie z parametrami modeli GARCH.

Za pomocą przedstawionych modeli wyliczyliśmy dla każdego szeregu zwrotów 52 prognozy 1-dniowe (na okres od 2.01.2003 r. do 14.03.2003 r.). Oceny jakości prognozy dokonujemy w odniesieniu do trzech różnych miar zmienności zrealizowanej *Real*1, *Real*2 oraz *Real*3. Występujący w określeniu *Real*3 współczynnik $1 + c = 1 + \frac{\sigma_{co}^2}{\sigma_{oc}^2}$ dla zwrotów indeksu WIG20 ma wartość **1,28676**, a dla TechWIGu **1,284833**.

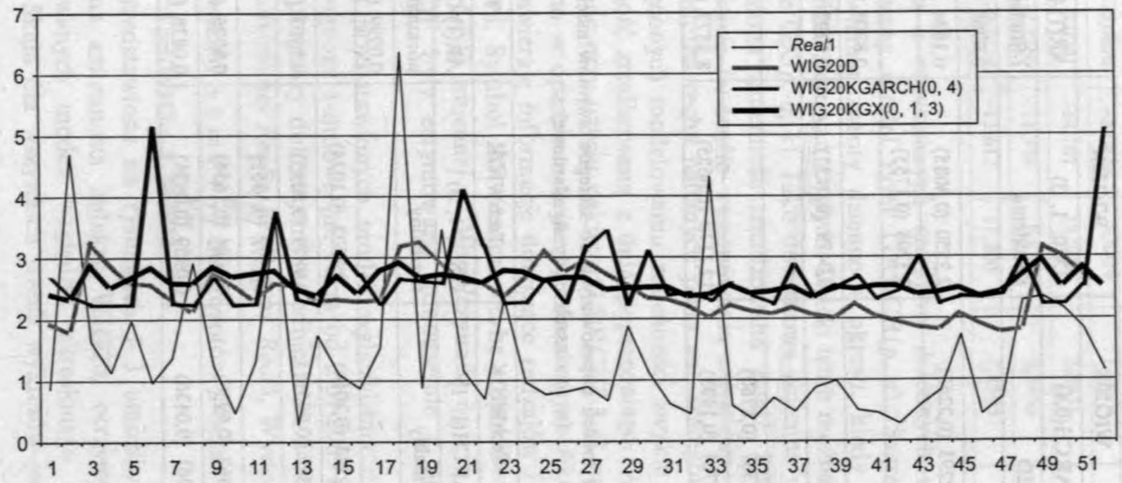
Wykresy przedstawione na rysunkach 1–3 umożliwiają wstępną ocenę jakości prognoz zmienności indeksu WIG20, otrzymanych przy użyciu trzech dopasowanych modeli względem określonego rodzaju zmienności zrealizowanej. Skala dla wszystkich trzech wykresów jest taka sama.

Tabela 3. Parametry modeli dopasowanych do szeregów zwrotów indeksu WIG20 (w nawiasach błędy standardowe)

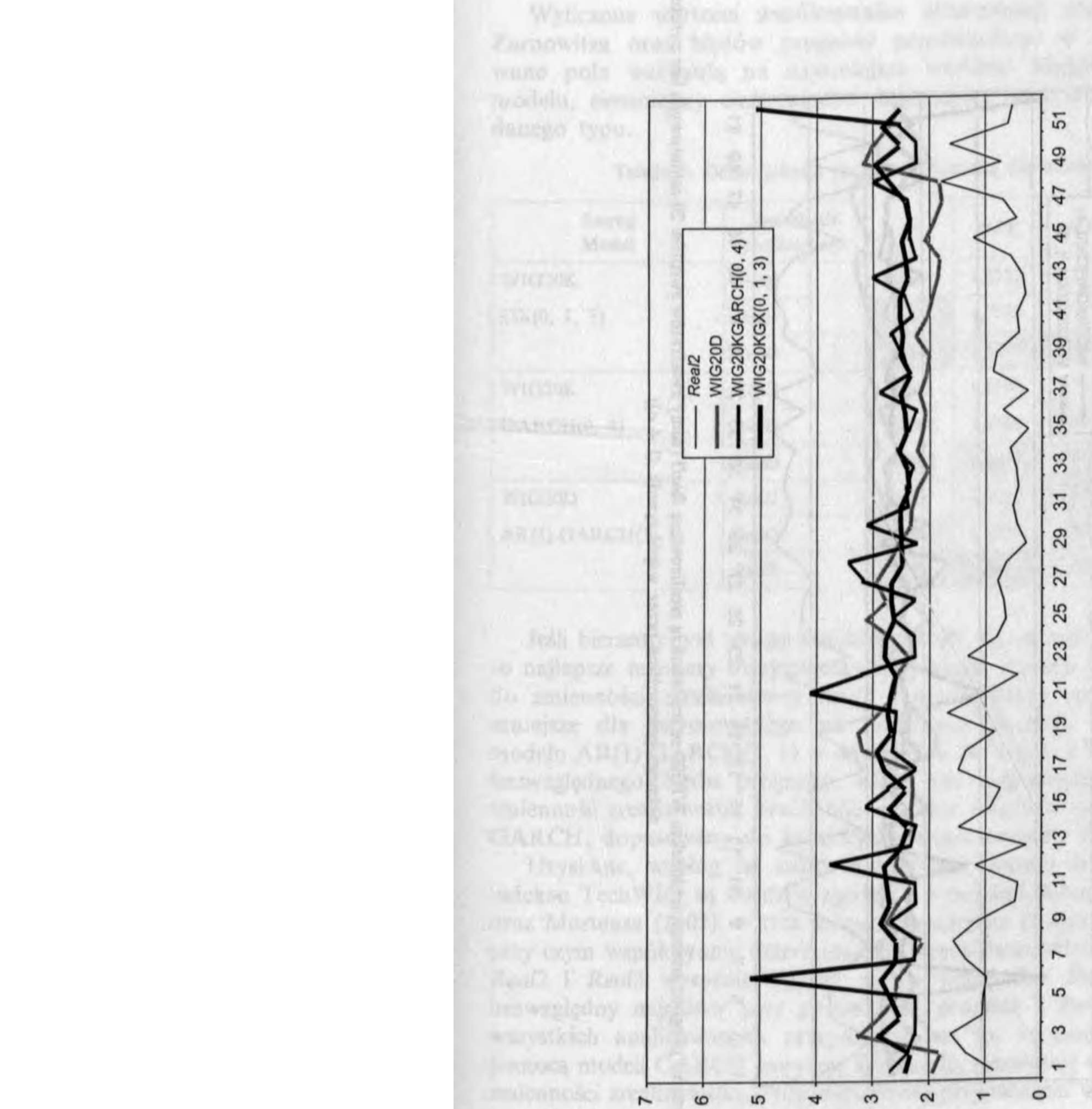
Szereg	WIG20K	WIG20K	WIG20D
Typ modelu	GARCH(0,4)	GX(0, 1, 3)	AR(1)-GARCH(1, 1)
Rozkład błędu	GED	*t*-Studenta	*t*-Studenta
a_1			0,0708 (0,0235)
ω	2,2293 (0,2268)	2,2550 (0,3085)	0,1834 (0,0548)
γ		0,3208 (0,1755)	
β_1	0		0,8306 (0,0289)
α_1	0	0,0479 (0,0453)	0,1292 (0,0220)
$\alpha_2 = \alpha_3$	0		
α_4	0,2367 (0,8760)		
DF	1,5764 (0,1493)	12,1170 (6,6625)	8,5573 (1,5589)

Tabela 4. Parametry modeli dopasowanych do szeregów zwrotów indeksu TechWIG (w nawiasach błędy standardowe)

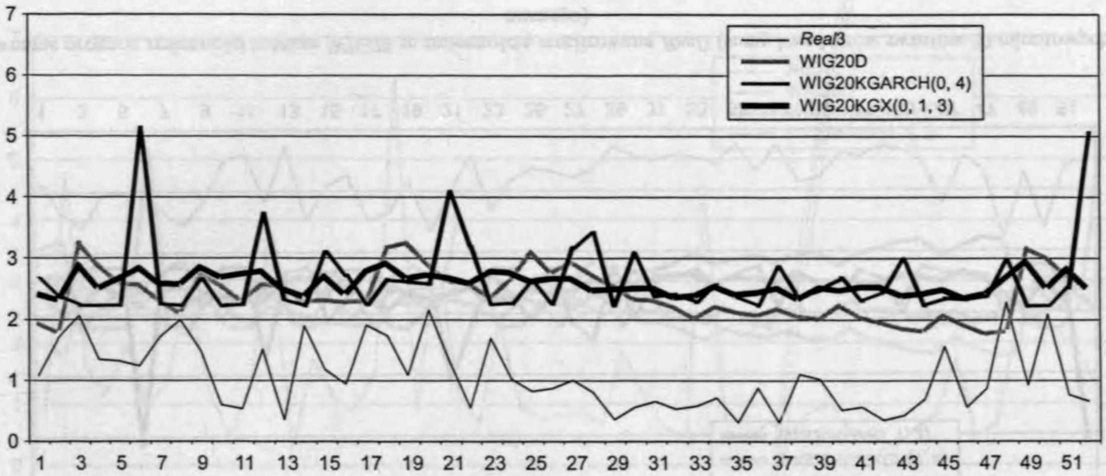
Szereg	TechWIGK	TechWIGK	TechWIGD
Typ modelu	GARCH(1, 1)	GX(1, 1, 3)	AR(1)-GARCH(1, 1)
Rozkład błędu	normalny	normalny	normalny
a_1			0,0999 (0,0378)
a_0	–0,2243 (0,1014)	–0,2303 (0,1024)	
ω	0,6585 (0,3951)	0,9879 (0,5695)	0,1907 (0,0970)
γ		0,2372 (0,1352)	
β_1	0,7862 (0,0950)	0,6084 (0,1654)	0,8926 (0,0299)
α_1	0,1061 (0,0436)	0,0959 (0,0454)	0,0838 (0,0231)



Rys. 1. Porównanie prognoz zmienności indeksu WIG20 ze zmiennością zrealizowaną *Real*1 (sumy kwadratów zwrotów 20-minutowych)



Rys. 2. Porównanie prognoz zmienności indeksu WIG20 ze zmiennością zrealizowaną *Real2* (sumy kwadratów zwrotów 20-minutowych bez zwrotu nocnego)



Rys. 3. Porównanie prognoz zmienności indeksu WIG20 ze zmiennością zrealizowaną *Real*3 (sumy kwadratów zwrotów 20-minutowych bez zwrotu nocnego pomnożone przez współczynnik $(1 + c)$)

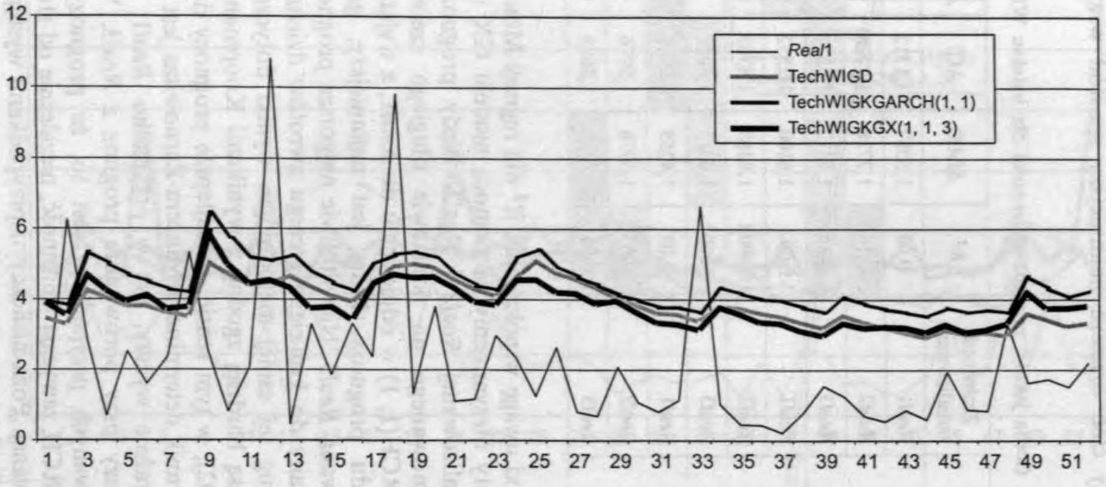
Wyliczone wartości współczynnika determinacji dla regresji Mincera-Zarnowitza oraz błędów prognozy przedstawiono w tabeli 5. Zacieniowane pola wskazują na najmniejsze wartości błędów przy ustalonym modelu, ciemniejszy cień oznacza najmniejszy spośród wszystkich błędów danego typu.

Tabela 5. Ocena jakości prognoz zmienności dla indeksu WIG20

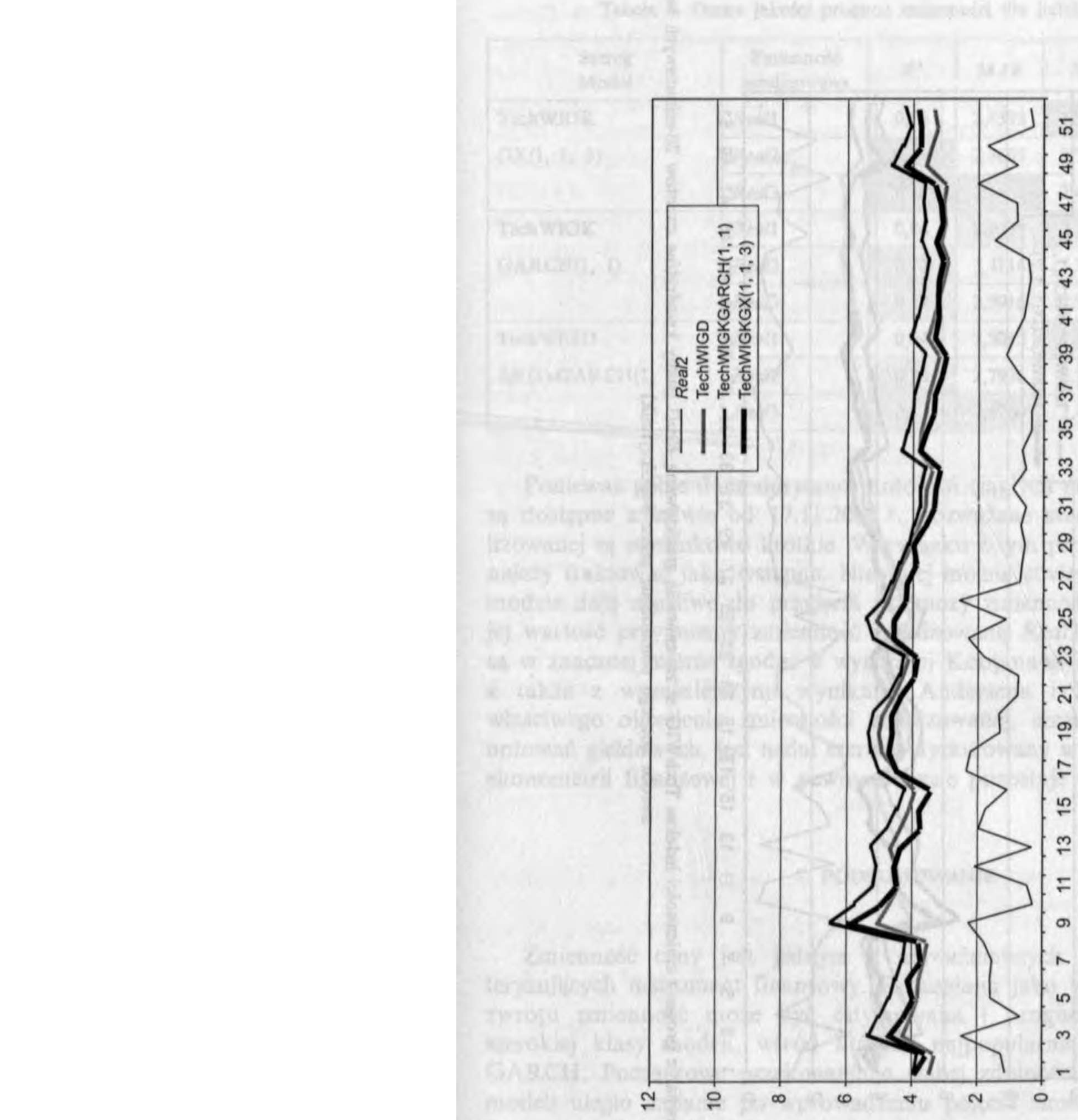
Szereg Model	Zmienność zrealizowana	R^2	*MAE*	*ME*	*AMAPE*	*TIC*
WIG20K	*Real*1	0,09	1,5387	1,1712	0,4273	0,3707
GX(0, 1, 3)	*Real*2	0,14	1,7728	1,7728	0,5441	0,5156
	*Real*3	0,14	1,5377	1,5377	0,4542	0,4279
WIG20K	*Real*1	0,002	1,6100	1,2352	0,4263	0,4052
GARCH(0, 4)	*Real*2	0,0002	1,8369	1,8369	0,5400	0,5461
	*Real*3	0,0002	1,6017	1,6017	0,4508	0,4638
WIG20D	*Real*1	0,03	1,4253	0,9959	0,4098	0,3615
AR(1)-GARCH(1, 1)	*Real*2	0,05	1,5976	1,5976	0,5158	0,4974
	*Real*3	0,05	1,3801	1,3625	0,4272	0,4095

Jeśli bierzemy pod uwagę współczynnik R^2 dla regresji Mincera-Zarnovitza, to najlepsze rezultaty otrzymujemy za pomocą modelu GX i w odniesieniu do zmienności zrealizowanej *Real*2 i *Real*3. Błędy prognozy okazują się mniejsze dla estymowanego na podstawie długiego szeregu obserwacji modelu AR(1)-GARCH(1, 1) w odniesieniu do *Real*1, z wyjątkiem średniego bezwzględnego błędu prognozy, który jest najmniejszy, gdy rozważamy zmienność zrealizowaną *Real*3. Niewątpliwie najgorsze prognozy daje model GARCH, dopasowany do krótkiego szeregu zwrotów dziennych.

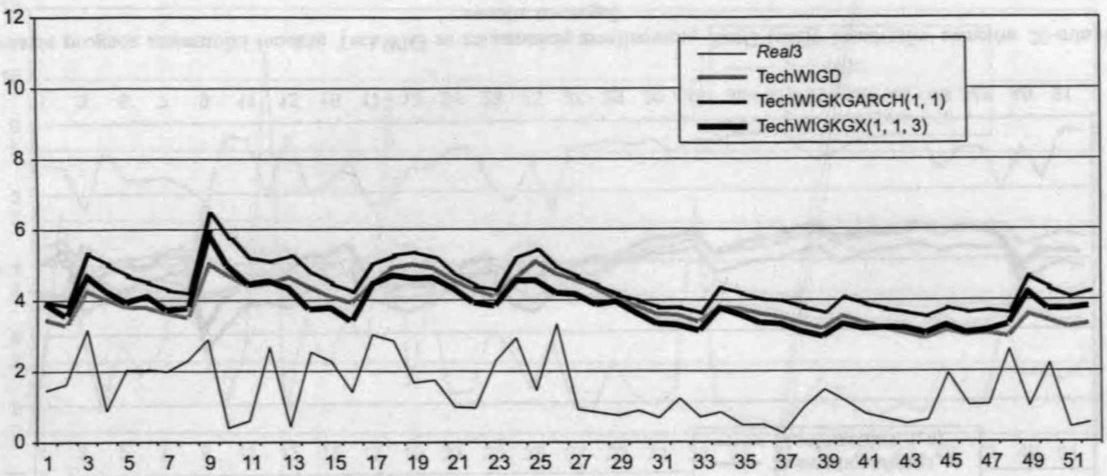
Uzyskane, według tej samej metodologii, wyniki dotyczące zmienności indeksu TechWIG są bardziej zgodne z wynikami Koopmana i Hol (2002) oraz Martensa (2002), w tym sensie, że najlepsze prognozy daje model GX, przy czym współczynnik determinacji Mincera-Zarnowitza jest dla zmienności *Real*2 i *Real*3 wyraźnie wyższy, niż w przypadku *Real*1, a średni błąd bezwzględny najniższy przy porównaniu prognoz z *Real*3. Wspólną cechą wszystkich analizowanych przypadków jest to, że prognozy uzyskane za pomocą modeli GARCH zawyżają zmienność, niezależnie od stosowanej miary zmienności zrealizowanej. Poza nielicznymi przypadkami występujących w szeregu *Real*1 obserwacji nietypowych, uzyskane przez nas prognozy dają oszacowanie zmienności z góry. Wykresy prognoz zmienności i zmienności zrealizowanych dla zwrotów indeksu TechWIG przedstawiono na rysunkach 4–6, a dane dotyczące błędów zawiera tabela 6.



Rys. 4. Porównanie prognoz zmienności indeksu TechWIG ze zmiennością zrealizowaną *Real*1 (sumy kwadratów zwrotów 20-minutowych)



Rys. 5. Porównanie prognoz zmienności indeksu TechWIG ze zmiennością zrealizowaną *Real2* (sumy kwadratów zwrotów 20-minutowych bez zwrotu nocnego)



Rys. 6. Porównanie prognoz zmienności indeksu TechWIG ze zmiennością zrealizowaną *Real3* (sumy kwadratów zwrotów 20-minutowych bez zwrotu nocnego pomnożone przez współczynnik)

Tabela 6. Ocena jakości prognoz zmienności dla indeksu TechWIG

Szereg Model	Zmienność zrealizowana	R^2	*MAE*	*ME*	*AMAPE*	*TIC*
TechWIGK	*Real*1	0,06	2,4595	1,7276	0,4510	0,3880
GX(1, 1, 3)	*Real*2	0,18	2,7607	2,7607	0,5803	0,5446
	*Real*3	0,18	2,4431	2,4431	0,4964	0,4604
TechWIGK	*Real*1	0,04	2,9239	2,2785	0,4697	0,4133
GARCH(1, 1)	*Real*2	0,17	3,3116	3,3116	0,6208	0,5866
	*Real*3	0,17	2,9918	2,9918	0,5418	0,5054
TechWIGD	*Real*1	0,04	2,5095	1,7606	0,4547	0,3917
AR(1)-GARCH(1, 1)	*Real*2	0,13	2,7936	2,7936	0,5820	0,5491
	*Real*3	0,13	2,4760	2,4760	0,4985	0,4656

Ponieważ pełne dane dotyczące notowań ciągłych na GPW w Warszawie są dostępne zaledwie od 17.11.2000 r., rozważane szeregi zmienności zrealizowanej są stosunkowo krótkie. W związku z tym przedstawione tu wyniki należy traktować jako wstępne. Niemniej można stwierdzić, że analizowane modele dają możliwe do przyjęcia prognozy zmienności, zwłaszcza jeśli za jej wartość przyjmiemy zmienność zrealizowaną *Real*3. Uzyskane rezultaty są w znacznej mierze zgodne z wynikami Koopmana i Hol oraz Martensa, a także z wcześniejszymi wynikami Andersena i Bollersleva. Problem właściwego określenia zmienności zrealizowanej, szczególnie w przypadku notowań giełdowych, jest nadal szeroko dyskutowany w literaturze dotyczącej ekonometrii finansowej i w pewnym sensie pozostaje otwarty.

7. PODSUMOWANIE

Zmienność ceny jest jednym z najważniejszych parametrów charakteryzujących instrument finansowy. Rozumiana jako wariancja warunkowa zwrotu zmienność może być estymowana i prognozowana za pomocą szerokiej klasy modeli, wśród których najpopularniejsze są modele typu GARCH. Początkowe przekonanie o słabej zdolności prognostycznej tych modeli uległo zmianie po wprowadzeniu pojęcia zmienności zrealizowanej. Jest ona definiowana jako suma kwadratów zwrotów śróddziennych. Częstotliwość, z jaką powinny być wyliczane zwroty, zależy od mikrostruktury rynku i na ogół waha się od 5 do 30 minut. W literaturze występują również inne miary zmienności zrealizowanej, nie uwzględniające zwrotu nocnego. Modele GARCH można rozszerzyć, wprowadzając jako dodatkową zmienną objaśniającą opóźnione wartości zmienności zrealizowanej. W niniejszym opracowaniu dokonujemy oceny jakości prognoz dziennej zmienności indeksów WIG20 i TechWIG, przyjmując kolejno za punkt odniesienia prognozy zmienność zrealizowaną, określoną na trzy różne sposoby: jako sumę kwadratów zwrotów śróddziennych z uwzględnieniem zwrotu nocnego, sumę kwadratów zwrotów śróddziennych bez uwzględniania zwrotu nocnego oraz tę ostatnią przeskalowaną za pomocą współczynnika wiążącego wariancję zwrotu dziennego z wariancją zwrotu nocnego. Dla każdego z indeksów dopasowujemy trzy modele GARCH; dwa klasyczne oparte na szeregach zwrotów dziennych różnej długości oraz model rozszerzony przez wprowadzenie opóźnionej zmienności zrealizowanej jako zmiennej objaśniającej. Otrzymane w ten sposób prognozy dają oszacowanie zmienności indeksu z góry, w zasadzie niezależnie od stosowanej miary zmienności zrealizowanej. Wydaje się, że spośród wymienionych, najbardziej adekwatną miarą zmienności jest odpowiednio przeskalowana suma kwadratów zwrotów śróddziennych bez uwzględniania zwrotu nocnego. Modele GX, czyli GARCH ze zmiennością zrealizowaną jako dodatkową zmienną objaśniającą, dają lepsze prognozy niż zwykłe modele GARCH, oparte na szeregu danych o tej samej długości. Zwiększenie liczby obserwacji w próbie skutkuje w przypadku indeksu WIG20 zmniejszeniem błędów prognoz, przy dużo mniejszym niż w przypadku modelu GX współczynniku R^2 regresji Mincera-Zarnowitza. Zmienność indeksu TechWIG najlepiej opisuje i prognozuje model GX, uwzględniający zmienność zrealizowaną sprzed trzech dni.

LITERATURA

Andersen T.G., Bollerslev T. (1997), *Intraday Periodicity and Volatility Persistence in Financial Markets*, „Journal of Empirical Finance", 4.

Andersen T.G., Bollerslev T. (1998), *Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts*, „International Economic Review", 39.

Andersen T.G., Bollerslev T., Diebold F., Ebens H. (2001), *The Distribution of Realized Stock Return Volatility*, „Journal of Financial Economics", 61.

Bollerslev T. (1986), *Generalized Autoregressive Conditional Heteroscedasticity*, „Journal of Econometrics", 31.

Doman M., Doman R. (2003), *Prognozowanie dziennej zmienności indeksu WIG określonej za pomocą danych o wyższej częstotliwości*, „Acta Universitatis Lodziensis. Folia Oeconomica", 166.

Engle R.F. (1982), *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*, „Econometrica", 50.

Hol E., Koopman S.J. (2002), *Stock Index Volatility Forecasting with High Frequency Data*, Tinbergen Instytute, „Discussion Paper", 068/4.

Laurent S., Peters J.-P. (2002), *G@RCH2.2: An Ox Package for Estimating and Forecasting Various ARCH Models*, „Journal of Economic Surveys", 16.
Martens M. (2002), *Measuring and Forecasting S&P500 Index-Futures Volatility Using High Frequency Data*, „Journal of Futures Markets", 22.
Oomen R. (2001), *Using High Frequency Stock Market Index Data to Calculate, Model & Forecast Realized Return Variance*, „Economic Working Papers", Eco16.

Małgorzata Doman

FORECASTING POLISH STOCK INDICES VOLATILITY USING GARCH MODELS AND HIGH FREQUENCY DATA

Summary

The notion of daily realized volatility introduced by Andersen and Bollerslev gave a new impulse to research connected with modeling and forecasting the volatility of financial returns using GARCH models. Daily realized volatility is a sum of squared intraday returns. Volatility forecasts obtained from GARCH models improve when instead of daily squared returns they are evaluated against the realized volatility. In this paper we calculate and investigate volatility forecasts for stock indices from the Warsaw Stock Exchange delivered by GARCH models with realized volatility as an additional explanatory variable.