STUDIA I PRACE WYDZIAŁU NAUK EKONOMICZNYCH I ZARZĄDZANIA NR 10

ELIZA BUSZKOWSKA

PORÓWNYWANIE ZDOLNOŚCI PROGNOSTYCZNEJ MODELI ZMIENNOŚCI INDEKSU WIG20 ZA POMOCĄ TESTU SPA

WPROWADZENIE

Badania nad nadrzędną zdolnością prognostyczną modeli zmienności, z wykorzystaniem testu SPA (Suprior Predictive Ability), wprowadzonego przez Hansena w roku 2001, były prowadzone szeroko, pod różnym kątem. Metoda SPA została przedstawiona w pracy [5]. Jednym z zagadnień z tego obszaru była warunkowa zdolność prognostyczna pewnych modeli. Podejście to zakłada warunkowe oczekiwania, co do prognoz i błędów. Zastosowano test warunkowej zdolności prognostycznej do oceny dokładności prognostycznej różnych metod prognozowania. Badanie to jest przedstawione w pracy [4]. W pracy [6] porównano przy użyciu testu SPA i innej klasycznej metody RC (Reality Check) 330 modeli prognostycznych typu ARCH pod względem ich zdolności do opisywania warunkowej wariancji. Pokazano, że dla IBM model GARCH(1,1) nie ma przewagi nad innymi. Dokonano analizy wartości różnych estymatorów *p*-wartości dla SPA.

Celem niniejszej pracy jest wykorzystanie testu SPA do porównania zdolności prognostycznej modeli typu GARCH(1,1) w przypadku prognoz wariancji warunkowej (zmienności) szeregu notowań indeksu WIG20. Prosty sposób budowania modeli, łatwość ich estymacji oraz istniejąca naturalna interpretacja spowodowały, że modele z rodziny GARCH są najpowszechniej stosowanymi modelami zmienności instrumentów finansowych. Prognozy zmienności porównujemy z dzienną zmiennością zrealizowaną, obliczaną jako suma kwadratów zwrotów śróddziennych. Badamy, czy wyniki otrzymane za pomocą testu SPA zmieniają się dla różnych miar zmienności zrealizowanej. Następnie porównujemy wskazania testu SPA dla zmienności zrealizowanej, wyliczonej na podstawie danych śróddziennych o różnej częstotliwości. Odnosimy się do pracy [13], w której podano, że dobrym oszacowaniem zmienności zrealizowanej jest zmienność wyznaczona na podstawie zwrotów śróddziennych 30 minutowych, podczas gdy użycie danych 5 minutowych nie eliminuje efektu mikrostruktury rynku. Następnie porównujemy wskazania testu SPA dla różnych typów funkcji błędu. Odnosimy się również do tezy pracy [15], że do oceny prognoz zmienności ważny jest wybór funkcji błędu. Jako wzorcowe, przyjmujemy modele najprostsze: podstawowy GARCH(1,1) i RiskMetrics.

SPECYFIKACJA MODELI WARIANCJI WARUNKOWEJ

W badaniu wykorzystano przedstawione poniżej typy modeli GARCH. Rozważano również inne modele, ale nie uzyskano zbieżności przy ich estymacji dla badanego szeregu.

Model GARCH(p,q) opisany jest równaniami:

$$y_t = \sigma_t \varepsilon_t, \tag{1}$$

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

$$\varepsilon_t \sim \mathrm{idd}(0,1),\ \omega > 0,\ \beta_j \geq 0,\ \alpha_i \geq 0. \tag{3}$$

EGARCH

EGARCH jest to historycznie pierwszy wariant modelu GARCH, uwzględniający efekt asymetrii, wprowadzony przez Nelsona w roku 1991:

$$\log(\sigma_t^2) = \omega + \sum_{i=1}^{q} \left[\alpha_i \varepsilon_{t-i} + \gamma_i \left(|\varepsilon_{t-i}| - E|\varepsilon_{t-i}|\right)\right] + \sum_{j=1}^{p} \beta_j \cdot \log(\sigma_{t-j}^2). \tag{4}$$

RiskMetrics

Jest to model IGARCH, dla którego współczynniki ARCH i GARCH są ustalone.

$$\sigma_t^2 = \omega + (1-\lambda) y_{t-1}^{\ 2} + \lambda \sigma_{t-1}^2. \tag{5}$$

*FIGARCH(**p,d,q**) w postaci BBM*

Jest to zintegrowany ułamkowo model GARCH, wprowadzony przez Bailie'go, Bolersleva i Mikelsena w roku 1996. Służy do modelowania długiej pamięci w procesie zmienności. Warunkowa wariancja w modelu FIGARCH(p,d,q) jest dana wzorem:

$$\sigma_t^2 = \omega(1-\beta(L))^{-1} + \left\{1 - [1-\beta(L)]^{-1} \phi(L)(1-L)^d \right\} y_t^2, \tag{6}$$

$$\phi(L) = 1 - \phi_1 L - \ldots - \phi_q L^q,\ \beta(L) = \beta_1 L + \beta_2 L^2 + \ldots + \beta_p L^p,\ 0 \leq d \leq 1 \tag{7}$$

lub $$\sigma_t^2 = \omega^* + \sum_{i=1}^{\infty} \lambda_i L^i y_t^2 = \omega^* + \lambda(L) y_t^2 . \tag{8}$$

FIGARCH Chunga

Model FIGARCH w postaci Chunga ma następującą postać:

$$\Phi(L)(1-L)^d \left(y_t^2 - \sigma^2\right) = [1-\beta(L)]\left(y_t^2 - \sigma_t^2\right) \tag{9}$$

HYGARCH

Hiperboliczny GARCH jest postaci:

$$\sigma_t^2 = \omega[1-\beta(L)]^{-1} + \left\{1 - [1-\beta(L)]^{-1} \varphi(L)\left\{1 + \alpha(1-L)^d\right\}\right\} y_t^2 . \tag{10}$$

GJR

GJR jest to model Glostena, Jagannathana i Runkle'a z roku 1993:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q}\left(\alpha_i y_{t-i}^2 + \gamma_i S_t^- y_{t-i}^2\right) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 , \tag{11}$$

gdzie S_t^- jest zmienną pomocniczą, która przyjmuje wartość 1 gdy wartość y_t jest ujemna a 0 kiedy jest dodatnia.

TEST NADRZĘDNEJ ZDOLNOŚCI PROGNOSTYCZNEJ (SPA)

Test nadrzędnej zdolności prognostycznej (Test of Superior Predictive Ability, SPA) pozwala orzec, czy pewien wzorcowy model prognostyczny nie jest gorszy niż inne alternatywne modele prognostyczne. Test SPA może być również użyty do sprawdzenia zdolności prognostycznych pewnych zmiennych losowych, a także do porównywania symulacji, miar zmienności i miar błędu.

Jest to nowe i konkurencyjne podejście w stosunku do testu RC po raz pierwszy wprowadzonego przez White'a w roku 2000 pod nazwą „The reality check for data snooping", RC. Test SPA jest mocniejszy od testu RC.

Niech $\left\{\hat{\sigma}_{k,t}^2 , k = 0,1,\ldots,m\right\}$ będzie skończonym zbiorem zmiennych losowych, w pracy prognoz zmienności wyznaczonych modelami $k = 0,1,\ldots,m$. Niech $\sigma_{l,t}^2$ będzie zmienną losową, w pracy zmiennością zrealizowaną typu k, $l = 0,1,3$. Wprowadźmy miarę zwaną funkcją straty:

$$L\left(\sigma_{l,t}^2, \hat{\sigma}_{k,t}^2\right). \tag{12}$$

Przykładem funkcji straty jest błąd średniokwadratowy:

$$N^{-1}\sum_{t=1}^{N}\left(\sigma_{l,t}^2 - \hat{\sigma}_{k,t}^2\right)^2 . \tag{13}$$

Prognozy są oceniane w terminach oczekiwanej straty, zdefiniowanej następująco:

$$X_{k,t} = L(\sigma^2_{1,t}, \hat{\sigma}^2_{0,t}) - L(\sigma^2_{1,t}, \hat{\sigma}^2_{k,t}) \tag{14}$$

$$k = 1, \ldots, m, \; t = 1, \ldots, n.$$

Statystyka testowa testu SPA określona jest wzorem:

$$T^{SPA}_n = \max_k \frac{n^{1/2}\overline{X}_k}{\hat{\omega}_k}, \quad k = 1, 2, \ldots, m, \tag{15}$$

gdzie

$$\overline{X}_k = \frac{1}{n}\sum_{t=1}^{n} X_{k,t}, \tag{16}$$

jest średnią z próby, a

$$\hat{\omega}^2_k = \hat{V}ar(n^{1/2}\overline{X}_k) \tag{17}$$

jest zgodnym estymatorem wariancji asymptotycznej $\omega_k^2 = \lim_{n\to\infty} Var(n^{1/2}\overline{X}_k)$, wyznaczanym za pomocą metod bootstrapowych.

Hipotezą testu SPA jest, że model obrany jako wzorcowy nie jest gorszy od żadnego z modeli alternatywnych w danym zbiorze. Ta hipoteza może być zapisana następująco:

$$\mu_k = E[X_{k,t}] \le 0, \quad k = 1, \ldots, m. \tag{18}$$

W teście SPA mamy do czynienia ze złożoną hipotezą zerową, co oznacza, że trzeba rozważyć kilka rozkładów zerowych. Rozkład statystyki testu SPA przy założeniu hipotezy zerowej, $\mu_k \le 0$ dla wszystkich *k*, nie jest standardowy. Do wyznaczenia rozkładu statystyki T_n^{SPA} stosowany jest bootstrap stacjonarny Politisa i Romano (1994).

W badaniu używamy trzech szeregów zmienności zrealizowanej, wyznaczonej na trzy różne sposoby. Estymujemy modele zmienności na podstawie zwrotów dziennych. Rozważamy 37 różnych modeli. Za pomocą każdego z nich generujemy 265 prognoz. Tworzymy rankingi dla siedmiu różnych modeli wzorcowych.

ZMIENNOŚĆ ZREALIZOWANA

Zmienność zrealizowana jest obliczana poprzez sumowanie kwadratów zwrotów śróddziennych. Za pomocą wzoru, który bierze pod uwagę zwrot nocny definiuje się ją jako:

$$\sigma^2_{2,t} = \sum_{i=0}^{N} r^2_{t,i}, \tag{19}$$

gdzie zwrot śróddzienny w dniu n i w chwili d i zwrot nocny dane są następująco:

$$r_{n,d} = 100\left(\ln P_{n,d} - \ln P_{n,d-1}\right),\ r_{n,0} = 100\left(\ln P_{n,0} - \ln P_{n-1,N}\right) \tag{20}$$

N jest liczbą obserwacji w ciągu dnia, a $P_{n,0}$ oznacza kurs otwarcia w dniu n.

Alternatywne podejście zostało zaproponowane przez Andersena i Bollersleva w roku 1997. Sugerowali oni reprezentowanie dziennej zmienności jako sumy zwrotów wyłącznie śróddziennych.

$$\sigma^2_{1,t} = \sum_{i=1}^{N} r^2_{t,i}. \tag{21}$$

Zaproponowano pomnożenie wielkości $\sigma^2_{1,t}$ przez $(1+c)$, gdzie c jest dodatnią stałą, Martens (2002). Obrano $\left(\sigma^2_{co} + \sigma^2_{oc}\right)/\sigma^2_{oc}$ jako stałą c, gdzie $\sigma^2_{co} = Var(r_{t,0})$ i $\sigma^2_{oc} = Var(\sum_{t=1}^{N} r_{t,n})$, Koopman i in. (2005). Wobec powyższego zmienność zrealizowana wyraża się wzorem:

$$\sigma^2_{3,t} = \frac{\sigma^2_{oc} + \sigma^2_{co}}{\sigma^2_{oc}} \sum_{i=1}^{N} r^2_{t,i}. \tag{22}$$

W naszej pracy, MSE i MAE oznaczają błąd średniokwadratowy i średni błąd bezwzględny, gdzie N to liczba prognoz.

$$MSE_l \equiv N^{-1} \sum_{t=1}^{N} \left(\sigma_{l,t} - \hat{\sigma}_{k,t}\right)^2,\ MSE \equiv N^{-1} \sum_{t=1}^{N} \left(\sigma^2_{l,t} - \hat{\sigma}^2_{k,t}\right)^2, \tag{23}$$

$$MAE_l \equiv N^{-1} \sum_{t=1}^{N} \left|\sigma_{l,t} - \hat{\sigma}_{k,t}\right|,\ MAE \equiv N^{-1} \sum_{n=1}^{N} \left|\sigma^2_{l,t} - \hat{\sigma}^2_{k,t}\right|,$$

gdzie $l \in \{1,2,3\}$, $k \in \{1,\ldots,m\}$, a m jest liczbą modeli z rozważanego zbioru. W powyższych wzorach, $\hat{\sigma}^2_{k,t}$ jest prognozą zmienności z modelu k na moment t, a $\sigma^2_{l,t}$ to wartość zmienności zrealizowanej typu l w chwili t.

DANE

W przeprowadzonym badaniu, do estymacji modeli używamy 1739 dziennych obserwacji indeksu WIG20 od 12.10.2000 do 09.14.2007. Na ostatnie 265 dni, tj. od 28.08.2006 do 14.09.2007 obliczamy jednodniowe prognozy.

W celu oceny jakości prognoz porównujemy je z dzienną zmiennością zrealizowaną obliczoną dla 10-minutowych, 5-minutowych i 30-minutowych zwrotów śróddziennych. Zmienności zrealizowane obliczone zostały ze wzorów (19), (21) i (22). Podstawowe statystyki opisowe dla szeregu zwrotów dziennych są przedstawione w tabeli 1.

Tabela 1. Podstawowe statystyki opisowe dla szeregu zwrotów dziennych od 12.10.2000 do 14.09.2007

maksimum	minimum	średnia	wariancja	skośność	kurtoza
5,4829	-5,7305	0,0447	2,2437	0,0792	4,0311

Źródło: obliczenia własne.

Rozważamy następujące typy modeli GARCH(1,1) z różnymi rozkładami błędu: RiskMetrics, GARCH, GJR, EGARCH, APARCH, IGARCH, FIGARCH estymowany metodą Chunga i FIGARCH estymowany metodą BBM, FIAPARCH estymowany metodą Chunga i FIAPARCH estymowany metodą BBM, HYGARCH z rozkładami Gaussa, GED, *t*-Studenta, rozkładem skośnym *t*- Studenta.

WYNIKI EMPIRYCZNE

Porównywano 37 modeli typu GARCH. Za pomocą każdego z modeli wyznaczono 265 jednodniowych prognoz wariancji warunkowej (zmienności) indeksu WIG20. W celu oszacowania *p*-wartości statystyki testowej testu SPA wyznaczano po 1000 prób bootstrapowych. W stosowanym programie komputerowym [7] wykorzystuje się dwie funkcje straty MSE i MAE. W badaniu dodatkowo stosowano również pierwiastki z wariancji warunkowej i pierwiastki ze zmienności zrealizowanej. Wybór funkcji straty MSE_1 wynika między innymi z wniosków zamieszczonych w pracy [15]. Niska *p*-wartość (0,05 – 0,1) informuje, że model wzorcowy jest gorszy od jednego z modeli alternatywnych. Wysoka *p*-wartość oznacza, że model wzorcowy generuje prognozy nie nie gorsze niż modele alternatywne. Wykorzystywany program komputerowy [7] podaje wartości trzech estymatorów *p*-wartości testu SPA oznaczonych: SPA_u, SPA_c i SPA_l. Aproksymuje się rozkład statystyki testowej T_n^{SPA} przy założeniu hipotezy zerowej. Zmienne bootstrapowe są scentrowane na trzy różne sposoby. Modyfikacja ta zmniejsza wpływ złych modeli na rozkład statystyki testowej. Dla każdego ze sposobów centrowania otrzymujemy jeden z estymatorów *p*-wartości: SPA_u, SPA_c i SPA_l, przy czym SPA_c jest estymatorem zgodnym. Szczegółowy opis postaci tych estymatorów można znaleźć w pracy [5].

Model najbardziej znaczący, podawany przez program SPA, [7] jest to taki model, dla którego wartość statystyki testowej jest najwyższa. Modelami wskazywanymi jako najbardziej znaczące były HYGARCH z rozkładem GED i FIGARCH-BBM ze rozkładem skośnym *t*-Studenta.

Modelem najlepszym, podawanym przez program SPA jest model, dla którego obrana funkcja straty ma wartość najmniejszą. Można mówić o modelu najlepszym dla danego typu zmienności zrealizowanej i dla danej funkcji błędu. Sześć różnych modeli zostało wskazanych jako najlepsze dla różnych rodzajów zmienności zrealizowanej. Są to modele: RiskMetrics(1,1) z rozkładem Gaussa, GED, *t*- Studenta oraz z rozkładem skośnym *t*-Studenta, GARCH(1,1) z rozkładem Gaussa (Model GJR ze względu na wartości oszacowań parametrów może być traktowany jak GARCH(1,1)). W przeanalizowanych rankingach najgorszymi były dwa modele FIGARCH(1,1) postaci Chunga ze rozkładem skośnym *t*-Studenta i EGARCH(1,1) z rozkładem skośnym *t* Studenta.

Wszystkie modele z najmniejszymi funkcjami straty zostały obrane jako wzorcowe i przeprowadzono dla nich test SPA. Test SPA potwierdził przewagę tych modeli nad pozostałymi modelami w rozpatrywanym zbiorze (*p*-wartości testu SPA były znacznie wyższe od 0.05)

Są to GARCH(1,1) z rozkładem Gaussa

$$\mu = 0.0759(0.0350),\ \omega = 0.015317(0.0066),$$
$$\alpha_1 = 0.0403(0.0067), \beta_1 = 0.953(0.0078).$$

Są to GJR(1,1) z rozkładem Gaussa

$$\mu = 0{,}0762(0{,}0354),\ \ \omega = 0.0152(0.0067),$$
$$\gamma = -0.0005(0{,}0107), \alpha_1 = 0.0405(0.0087), \beta_1 = 0.9531(0.0078).$$

RiskMetrics z rozkładem skośnym *t*-Studenta

$$\mu = 0.0883(0.0307), \lambda = 0.94,$$
$$W.\ Ogona{=}9.4524\,(2.0108),\ W.\ Asymetrii{=}\,0.04322(0.0357).$$

RiskMetrics z rozkładem GED,

$$\mu = 0.0762(0.0297), \lambda = 0.94\,,\ GED{=}\ 1.4367\,(0.0728).$$

RiskMetrics z rozkładem *t*-Studenta

$$\mu = 0.079(0.0307), \lambda = 0.94\,,\ \text{Student DF} = 9.4884\ \ (2.0336).$$

RiskMetrics z rozkładem Gaussa

$$\mu = 0.0822(0.029), \lambda = 0.94.$$

Tabela 2. Oszacowania p-wartości SPA dla zmienności zrealizowanych $\sigma_{1,t}^2$ i dla częstotliwości 5, 10 i 30 minut dla wybranych dwóch modeli

Model wzorcowy	RiskMetrics z rozkładem skośnym t-Studenta			GARCH z rozkładem Gaussa		
Błąd	SPA_l	SPA_c	SPA_u	SPA_l	SPA_c	SPA_u
mse	0.166	0.21	0.352	0.905	1	1
mae	0.148	0.249	0.498	0.077	0.206	0.323
mse_1	0.188	0.276	0.50	0.2	0.515	0.685
mae_1	0.323	0.323	0.79	0.01	0.06	0.12

Źródło: obliczenia własne.

WNIOSKI

Modelami najbardziej znaczącymi w zbiorze modeli typu GARCH okazały się HYGARCH(1,1) z rozkładem GED i FIGARCH(1,1) z rozkładem skośnym t-Studenta. Zauważamy, że ten sam model najbardziej znaczący pojawia się w rankingach dla danych o różnej częstotliwości, różnych rodzajów modeli wzorcowych i różnie wyznaczanej zmienności zrealizowanej. Najgorszymi okazały się FIGARCH(1,1) wyznaczony metodą Chunga z rozkładem skośnym t-Studenta oraz EGARCH z rozkładem skośnym t-Studenta.

Tabela 3. Oszacowania p-wartości SPA dla zmienności zrealizowanych $\sigma_{2,t}^2$ i dla częstotliwości 5, 10 i 30 minut dla wybranych dwóch modeli

Model wzorcowy	RiskMetrics z rozkładem skośnym t-Studenta			GARCH z rozkładem Gaussa		
Błąd	SPA_l	SPA_c	SPA_u	SPA_l	SPA_c	SPA_u
mse	0.55	0.81	0.84	0.571	0.741	0.932
mae	0.687	0.955	0.982	0.449	0.649	0.951
mse_1	0,57	0.882	0.994	0.663	0.916	0.998
mae_1	0.777	0.988	0.995	0.27	0.5	0.777

Źródło: obliczenia własne.

Wybór modelu najbardziej znaczącego zależy od rodzaju błędu spośród MSE, MSE_1, MAE, MAE_1. Modelami najlepszymi, wyznaczonymi w różnych rankingach dla naszych danych były modele RiskMetrics(1,1) z rozkładami Gaussa, GED, t-Studenta oraz z rozkładem skośnym t-Studenta i GARCH(1,1) z rozkładem Gaussa, Modele najbardziej znaczące nie należą do zbioru modeli najlepszych. Modele podawane jako najlepsze zostały obrane jako wzorcowe i przeprowadzono dla nich test SPA za pomocą programu SPA [7]. Wyniki testu potwierdziły brak przewagi pozostałych modeli nad tymi modelami pod względem zdolności prognostycznej. Wynik ten odczytano na podstawie *p*-wartości programu SPA. Oszacowane *p*-wartości były znacznie wyższe od założonego poziomu istotności 0,05. Zauważono, że w celu wyznaczenia wszystkich modeli najlepszych należy powtarzać test SPA dla różnych modeli wzorcowych, nie wyłącznie tych podawanych jako najlepsze dla ustalonego modelu benchmarkowego.

Tabela 4. Oszacowania *p*-wartości SPA dla zmienności zrealizowanych $\sigma_{3,t}^2$ i dla częstotliwości 5, 10 i 30 minut dla wybranych dwóch modeli

Model wzorcowy	RiskMetrics z rozkładem skośnym *t*-Studenta			GARCH z rozkładem Gaussa		
Błąd	SPA_l	SPA_c	SPA_u	SPA_l	SPA_c	SPA_u
mse	0.173	0.217	0.363	0.9	1	1
mae	0.148	0.249	0.5	0.079	0.216	0.332
mse_1	0.589	0.892	0.946	0.198	0.51	0.68
mae_1	0.777	0.988	0.998	0	0	0.014

Źródło: obliczenia własne.

Metoda SPA wymaga zatem wielokrotnego powtarzania testu dla wszystkich modeli obieranych kolejno jako benchmarkowe. Jej algorytm jest dłuższy niż algorytm konkurencyjnej metody zbioru ufności modeli, MCS, [8], która nie wymaga powtórzeń testu dla różnych modeli wzorcowych. Jest zatem gorszy. Zauważyliśmy, że *p*-wartości powtarzają się dla różnych częstotliwości, dla tych samych pozostałych parametrów. Wystarczy, zatem analiza *p*-wartości dla jednej obranej częstotliwości. W tabeli jest to częstotliwość 10- minutowa. Wskazany model RiskMetrics jest to model najczęściej używany w praktyce, [11]. Wynik badania może być przydatny do podejmowania decyzji inwestycyjnych na podstawie metod statystycznych, [12]. W literaturze ([3], [9]) potwierdzona jest przewaga strategii portfelowej opartej na prognozach zmienności. W pracy wytypowano pewne modele najlepiej prognozujące zmienność indeksu WIG20, co przy założeniu istnienia podobnej dynamiki pozwala na odrzucenie pozostałych modeli typu GARCH bez konieczności ponownego przeprowadzenia pracochłonnej procedury eliminacji.

Literatura

1. Andersen T.G., Bollerslev T., *Intraday Seasonality and Volatility Peresistence in Foreign Exchange and Equity Markets*. Journal of Empirical Finance, 1997, 4.
2. Andersen T.G., Bollerslev, Diebold F.X., Labys P., *The Distribution of Exchange Rate Volatility*. Journal of American Statistical Associacion, 2001, 96.
3. Busse J. A., *Volatility Timing In Mutual Funds: Evidence from Daily Returns*. Review of Financial Studies, 1999, 12.
4. Giacomini R., White H., Koopman S.J., *The Test of Conditional Predictive Ability*, Uniwersytet of California, San Diego 2003.
5. Hansen P.R., *A Test for Superior Predictive Ability*. Journal of Business and Economic Statistics, 2005, 23.
6. Hansen P. R., Lunde A., *A Forecast Comparison of Volatility Models. Does Anything Beat a GARCH(1,1)?* Journal of Applied Econometrics, 2005, 20.
7. Hansen P. R., Kim J., Lunde A., *Testing for Superior Predictive Ability using Ox*. A Manual for SPA for Ox. 2003.
8. Hansen P. R., Lunde A., Nason J. M., *Choosing the Best Volatility Models: The Model Confidence Set Approach*. Working Paper, 2003.
9. Johannes M., Polson N., Stroud J., *Sequential Optimal Portfolio Performance: Market and Volatility Timing.Wharton School*, University of Pensylvania, Working Paper, 2002.
10. Koopman S.J., Jungbacker, B., Hol, E., *Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realized and Implied Volatility Measurement*. Journal of Empirical Finance, 2005, 12.
11. Laurent S., *Estimating and Forecasting ARCH Models Using G@RCH* ™ 5. Timberlake Consultants, London 2007.
12. Nagayasu J., *Modeling and Predicting Japanese Stock Returns Based on ARFIMA-FIGARCH*, [w:] Milo W. Wdowiński P. [red.], Financial Markets. Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2006.
13. Patton A.J., Sheppard K., *Evaluating Volatility and Correlation Forecasts*. Departament of Economics and Oxford-Man Institute of Quantitative Finance, University of Oxford 2007.
14. Politis D.N., Romano J.P., *The Stationary Bootstrap*. Journal of the American Statistical Association, 1994, 89.

15. Souza L., Veiga A., Medeiros M.C., *Evaluating the Forecasting Performance of GARCH Models Using White's Reality Check*. Departamento de Economia PUC-RIO. Texto Para Discussao, 2002, 453.

STRESZCZENIE

Test nadrzędnej zdolności prognostycznej (SPA) pozwala orzec, czy pewien model prognostyczny daje nie gorsze wyniki niż inne modele. SPA jest przydatny, jeśli chcemy zbadać, czy dostępny jest lepszy model prognostyczny. Celem niniejszej pracy jest zbadanie za pomocą testu SPA, czy wybrane proste modele zmienności dopasowane do indeksu WIG20 dostarczają nie gorszych prognoz zmienności niż bardziej wyrafinowane modele typu GARCH, a także sprawdzenie wpływu różnych czynników na wyniki testu.

COMPARING THE PREDICTIVE ABILITY OF VOLATILITY MODELS FOR THE WIG20 INDEX USING THE SPA TEST

SUMMARY

The Test of Superior Predictive Ability (SPA) allows to test whether a particular forecasting procedure is outperformed by alternative forecasting models. The SPA test is useful in investigating whether any alternative forecasting model is better then the benchmark. The purpose of this paper is to use the SPA method to determine whether the volatility forecasts from some selected simple volatility models fitted to the WIG20 are outperformed by those from more sophisticated ones. Moreover, we check the impact of different factors on the results of the test.

Translated by E. Buszkowska

Mgr Eliza Buszkowska
Uniwersytet im. Adama Mickiewicza w Poznaniu
eliza_b2@o2.pl