A C T A   U N I V E R S I T A T I S   N I C O L A I   C O P E R N I C I
EKONOMIA XXXIX – NAUKI HUMANISTYCZNO-SPOŁECZNE – ZESZYT 389 – TORUŃ 2009

*Uniwersytet Mikołaja Kopernika w Toruniu*
*Katedra Ekonometrii i Statystyki*

*Witold Orzeszko*

# WSPÓŁCZYNNIK INFORMACJI WZAJEMNEJ JAKO MIARA ZALEŻNOŚCI NIELINIOWYCH W SZEREGACH CZASOWYCH

Z a r y s   t r e ś c i. W artykule scharakteryzowano konstrukcję, estymację oraz możliwości zastosowania współczynnika informacji wzajemnej. Przedstawiono wyniki symulacji, prowadzących do weryfikacji jego przydatności w procesie identyfikacji zależności nieliniowych w szeregach czasowych. Ponadto zaprezentowano wyniki zastosowania tego współczynnika do analizy indeksów Giełdy Papierów Wartościowych w Warszawie.

S ł o w a   k l u c z o w e: nieliniowość, współczynnik informacji wzajemnej, mutual information, identyfikacja zależności.

## 1. WSTĘP

Pomiar zależności pomiędzy zmiennymi jest niezwykle istotnym obszarem badań w ekonometrii. W tym celu najczęściej wykorzystuje się współczynnik korelacji Pearsona. Współczynnik ten nie jest jednak właściwym narzędziem pomiaru zależności nieliniowych. Do analizy nieliniowości należy zastosować inne miary, wśród których, jedną z najważniejszych jest współczynnik informacji wzajemnej. Współczynnik ten wywodzi się z teorii informacji i w swej konstrukcji opiera się na pojęciu entropii. Można go wykorzystać do pomiaru zależności pomiędzy dwoma szeregami czasowymi lub autozależności w pojedynczym szeregu.

## 2. POMIAR ZALEŻNOŚCI NIELINIOWYCH W SZEREGACH CZASOWYCH

Istnieje szereg metod, które można zastosować do pomiaru zależności o charakterze nieliniowym (por. Granger, Terasvirta, 1993; Maasoumi, Racine, 2002; Bruzda, 2004). Jedną z najważniejszych z nich jest miara informacji wzajemnej (*ang*. Mutual Information – MI), określona wzorem:

$$I(X,Y)=\iint p(x,y)\log\left(\frac{p(x,y)}{p_1(x)p_2(y)}\right)dxdy\,, \tag{1}$$

gdzie $p(x,y)$ jest funkcją gęstości rozkładu łącznego, natomiast $p_1(x)$ oraz $p_2(y)$ są gęstościami brzegowymi.

Miara informacji wzajemnej identyfikując zależności dowolnego typu (zarówno liniowe, jak i nieliniowe), informuje o możliwości wykorzystania zmiennej $X$ do prognozowania $Y$.

W celu unormowania, miarę tę można przekształcić do współczynnika informacji wzajemnej – $R(X,Y)$, według wzoru:

$$R(X,Y)=\sqrt{1-e^{-2I(X,Y)}}\,. \tag{2}$$

Współczynnik $R(X,Y)$ posiada następujące własności (por. Granger, Terasvirta, 1993; Granger, Lin, 1994):

1. $0\le R(X,Y)\le 1$,
2. $R(X,Y)=0 \Leftrightarrow X$ i $Y$ są niezależne,
3. $R(X,Y)=1 \Leftrightarrow Y=f(X)$, gdzie $f$ jest pewną odwracalną funkcją,
4. jest niezmienniczy względem transformacji zmiennych, tzn. $R(X,Y)=R(h_1(X),h_2(Y))$, gdzie $h_1$ oraz $h_2$ są dowolnymi funkcjami różnowartościowymi,
5. jeśli $(X,Y)$ (lub $(h_1(X),h_2(Y))$, gdzie $h_1$ oraz $h_2$ są różnowartościowe) ma dwuwymiarowy rozkład normalny ze współczynnikiem korelacji $\rho(X,Y)$, to $R(X,Y)=|\rho(X,Y)|$.

W literaturze przedmiotu istnieją różne propozycje szacowania wartości miary informacji wzajemnej w oparciu o szeregi czasowe $\{x_t\}$ oraz $\{y_t\}$. Zasadniczo, sprowadzają się one do oszacowania funkcji gęstości $p(x,y)$, $p_1(x)$ oraz $p_2(y)$ (por. wzór (1)). Istniejące metody można podzielić na trzy główne grupy (por. Dionisio, Menezes, Mendes, 2003):

- odwołujące się do histogramu,
- oparte na estymatorach jądrowych,
- metody parametryczne.

Jak wykazały przeprowadzone symulacje, estymatory z pierwszej grupy są mniej odporne na małą liczbę obserwacji niż metody oparte na estymatorach jądrowych. Jednak z drugiej strony, estymatory jądrowe są wrażliwe na wartości parametrów, przez co mogą prowadzić do bardzo zróżnicowanych wyników.

Z kolei istotnym ograniczeniem możliwości wykorzystania trzeciej grupy metod jest konieczność posiadania informacji o procesach generujących.

Oczywiście miarę informacji wzajemnej można wykorzystać również do pomiaru autozależności w pojedynczym szeregu czasowym $(x_t)$. W tym celu za $(y_t)$ należy przyjąć szereg opóźnionych wartości $(x_t)$. Ponadto miarę tę można wykorzystać również do określenia opóźnień czasowych w nieliniowych metodach prognozowania a także do analizy niestacjonarności (por. Dionisio, Menezes, Mendes, 2003).

## 3. ZASTOSOWANIE MIARY INFORMACJI WZAJEMNEJ DO IDENTYFIKACJI AUTOZALEŻNOŚCI W SZEREGACH CZASOWYCH

### 3.1. DANE GENEROWANE

W pierwszej kolejności badaniu poddano szeregi wygenerowane na potrzeby dokonanej przez Barnetta i innych (1998) analizy porównawczej metod identyfikacji nieliniowości. Było to pięć szeregów złożonych z 2000 obserwacji (ozn. M1, M2, M3, M4, M5) oraz pięć szeregów skróconych, tzn. złożonych z pierwszych 380 obserwacji (ozn. M1s, M2s, M3s, M4s, M5s)[1].

Szeregi te zostały wygenerowane z następujących modeli[2]:

I) M1 – odwzorowanie logistyczne:

$$x_t = 3{,}57 x_{t-1}(1 - x_{t-1}), \tag{3}$$

II) M2 – proces GARCH(1,1):

$$x_t = \sqrt{h_t} u_t, \tag{4a}$$

$$h_t = 1 + 0{,}1 x_{t-1}^2 + 0{,}8 h_{t-1}, \tag{4b}$$

gdzie $h_0 = 1$ i $x_0 = 0$.

III) M3 – proces nieliniowej średniej ruchomej (*ang*. nonlinear moving average process – NLMA):

$$x_t = u_t + 0{,}8 u_{t-1} u_{t-2}, \tag{5}$$

IV) M4 – proces ARCH(1):

$$x_t = \sqrt{1 + 0{,}5 x_{t-1}^2} u_t, \tag{6}$$

V) M5 – proces ARMA(2,1):

$$x_t = 0{,}8 x_{t-1} + 0{,}15 x_{t-2} + u_t + 0{,}3 u_{t-1}, \tag{7}$$

[1] Szeregi zostały pobrane ze strony http://econ.tepper.cmu.edu/barnett/Papers.html.

[2] W każdym z modeli, $u_t$ jest procesem białoszumowym o rozkładzie $N(0;1)$.

gdzie $x_0 = 1$ i $x_1 = 0{,}7$.

W każdym przypadku wyznaczono miarę MI dla szeregu oryginalnego oraz szeregu reszt z modelu ARMA. Przefiltrowanie modelem ARMA daje możliwość stwierdzenia, czy ewentualnie wykryte zależności w analizowanym szeregu mają charakter nieliniowy.

W pierwszej kolejności, przy pomocy testu ADF, zbadano zintegrowanie procesów. Test wykazał brak podstaw do odrzucenia hipotezy o istnieniu pierwiastka jednostkowego dla szeregu M5s. Z tego względu w kolejnych etapach badania przedmiotem analizy był szereg pierwszych różnic, ozn. M5s_rozn.

W tabeli 1 przedstawiono modele ARMA dla analizowanych szeregów[3].

Tabela 1. Modele ARMA dla szeregów symulowanych

| Szereg | model ARMA | Szereg | model ARMA |
|---|---|---|---|
| M1 | Biały szum (EX=0,648) | M1s | Biały szum (EX=0,649) |
| M2 | Biały szum (EX=0,034) | M2s | Biały szum (EX=0,067) |
| M3 | Biały szum (EX= 0,007) | M3s | Biały szum (EX= 0,033) |
| M4 | Biały szum (EX= 0,011) | M4s | Biały szum (EX= 0,018) |
| M5 | ARMA(1,1) | M5s_rozn | MA(1) |

Źródło: obliczenia własne.

W oparciu o test Boxa-Ljunga stwierdzono, że oszacowane modele skutecznie wyeliminowały autozależności liniowe z badanych szeregów.

Do oszacowania miary informacji wzajemnej wykorzystano metodę zaproponowaną przez Frasera, Swinneya (1986)[4]. Jest to metoda oparta na analizie dwuwymiarowego histogramu. Mówiąc w uproszczeniu, polega ona na podziale przestrzeni dwuwymiarowej zawierającej pary $(x_t, y_t)$ na prostokątne komórki (według określonej zasady) i obliczeniu, jaki odsetek naniesionych punktów znajduje się w każdej z komórek. Następnie stosowany jest wzór 1, przy czym obliczone odsetki są oszacowaniem funkcji gęstości, natomiast całkowanie odbywa się numerycznie.

Niech $i_k$ oznacza oszacowaną wartość miary informacji wzajemnej między zmiennymi $X_t$ i $X_{t-k}$. W tabelach 2–6 przedstawiono obliczone wartości $i_k$ ($k = 1,2,...,10$) dla analizowanych szeregów. Ze względu na cel badania, jakim była identyfikacja autozależności, kluczową kwestią jest określenie istotności oszacowanych współczynników. W celu zweryfikowania hipotezy zerowej o nieistotności współczynnika (inaczej mówiąc: o braku autozależności) zastosowano symulacje Monte Carlo (10 000 replikacji – dla każdego z szeregów). Tak otrzymane prawdopodobieństwa empiryczne (*p*-values) umieszczono

[3] Przy konstrukcji modelu kierowano się kryterium Schwarza oraz istotnością parametrów ($\alpha$=0,05).

[4] Obliczenia wykonano w programie Matlab 6.5 w oparciu o procedurę napisaną przez A. Leontitsisa.

w tabelach pod każdą z wyznaczonych wartości $i_k$. Wytłuszczonym drukiem zaznaczono te z nich, które są nie większe niż 0,005.[5]

Tabela 2. Wartości $i_k$ dla szeregów M1s oraz M1

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| M1s | 1,6927 | 1,6963 | 1,6123 | 1,7148 | 1,5919 | 1,6849 | 1,5412 | 1,6381 | 1,5379 | 1,6560 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| M1 | 2,0139 | 2,0090 | 2,0064 | 2,2520 | 1,9981 | 1,9991 | 1,9940 | 2,2737 | 1,9891 | 1,9891 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |

Źródło: obliczenia własne.

Tabela 3. Wartości $i_k$ dla szeregów M2s oraz M2

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| M2s | 0,0848 | 0,1538 | 0,1191 | 0,1308 | 0,1231 | 0,1616 | 0,1701 | 0,1162 | 0,1281 | 0,1228 |
| | 0,9616 | 0,0201 | 0,3802 | 0,1786 | 0,3052 | 0,0081 | 0,0029 | 0,4412 | 0,2187 | 0,3090 |
| M2 | 0,0541 | 0,0562 | 0,0477 | 0,0488 | 0,0492 | 0,0509 | 0,0541 | 0,0461 | 0,0449 | 0,0334 |
| | 0,0053 | 0,0025 | 0,0808 | 0,0536 | 0,0451 | 0,0227 | 0,0052 | 0,1303 | 0,1868 | 0,9315 |

Źródło: obliczenia własne.

Tabela 4. Wartości $i_k$ dla szeregów M3s oraz M3

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| M3s | 0,1857 | 0,1586 | 0,1425 | 0,1469 | 0,1323 | 0,1028 | 0,1897 | 0,1600 | 0,1525 | 0,1606 |
| | 0,0492 | 0,3316 | 0,6241 | 0,5429 | 0,8032 | 0,9927 | 0,0353 | 0,3096 | 0,4389 | 0,2987 |
| M3 | 0,0725 | 0,0658 | 0,0307 | 0,0429 | 0,0309 | 0,0383 | 0,0372 | 0,0404 | 0,0389 | 0,0456 |
| | 0,0000 | 0,0001 | 0,9634 | 0,2065 | 0,9599 | 0,5426 | 0,6274 | 0,3724 | 0,4868 | 0,0976 |

Źródło: obliczenia własne.

Tabela 5. Wartości $i_k$ dla szeregów M4s oraz M4

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| M4s | 0,1365 | 0,1667 | 0,1442 | 0,1349 | 0,1198 | 0,1367 | 0,1347 | 0,1327 | 0,1435 | 0,1464 |
| | 0,2663 | 0,0205 | 0,1562 | 0,2940 | 0,6104 | 0,2613 | 0,2959 | 0,3361 | 0,1641 | 0,1303 |
| M4 | 0,1053 | 0,0472 | 0,0363 | 0,0379 | 0,0286 | 0,0344 | 0,0370 | 0,0475 | 0,0368 | 0,0344 |
| | 0,0000 | 0,0051 | 0,3383 | 0,2324 | 0,9261 | 0,5058 | 0,2866 | 0,0039 | 0,3074 | 0,5059 |

Źródło: obliczenia własne.

[5] Zauważmy, że dla każdego z analizowanego szeregów, znalezienie się w obszarze krytycznym choćby jednej z oszacowanych wartości $i_k$ ($k$=1,2,...,10), oznacza odrzucenie hipotezy o braku autozależności dla tego szeregu. Z tego względu przyjęcie jako wartości krytycznej 0,005 przy analizie pojedynczej wartości $i_k$ (dla ustalonego $k$) implikuje, że prawdopodobieństwo popełnienia błędu pierwszego rodzaju wynosi w przybliżeniu 0,05.

Tabela 6. Wartości $i_k$ dla szeregów M5s oraz M5

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| M5s | 1,4787 | 1,1206 | 0,9817 | 0,8640 | 0,7505 | 0,6895 | 0,6344 | 0,6310 | 0,6173 | 0,6070 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| M5s_rozn | 0,1390 | 0,1658 | 0,1288 | 0,1438 | 0,1496 | 0,2012 | 0,1642 | 0,1297 | 0,1161 | 0,1387 |
| | 0,5519 | 0,1199 | 0,7509 | 0,4542 | 0,3452 | 0,0039 | 0,1351 | 0,7340 | 0,9125 | 0,5560 |
| M5s_roznMA | 0,1224 | 0,1584 | 0,1225 | 0,1242 | 0,1444 | 0,1391 | 0,1624 | 0,1510 | 0,1495 | 0,1474 |
| | 0,7971 | 0,1595 | 0,7942 | 0,7668 | 0,3745 | 0,4816 | 0,1193 | 0,2584 | 0,2821 | 0,3179 |
| M5 | 1,7145 | 1,3154 | 1,0949 | 0,9504 | 0,8414 | 0,7597 | 0,6958 | 0,6449 | 0,5917 | 0,5584 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| M5ARMA | 0,0422 | 0,0375 | 0,0417 | 0,0412 | 0,0355 | 0,0396 | 0,0419 | 0,0486 | 0,0434 | 0,0397 |
| | 0,2714 | 0,6530 | 0,3103 | 0,3438 | 0,8012 | 0,4685 | 0,2963 | 0,0398 | 0,2030 | 0,4640 |

Źródło: obliczenia własne.

W tabelach 7–8 dokonano podsumowania wyników przeprowadzonej identyfikacji autozależności dla szeregów generowanych.

Tabela 7. Wyniki identyfikacji autozależności dla szeregów długich

| Szereg | Autozależności | Nieliniowość |
|---|---|---|
| M1 | TAK | TAK |
| M2 | TAK | TAK |
| M3 | TAK | TAK |
| M4 | TAK | TAK |
| M5 | TAK | NIE |

Źródło: opracowanie własne na podstawie tabel 2–6.

Tabela 8. Wyniki identyfikacji autozależności dla szeregów krótkich

| Szereg | Autozależności | Nieliniowość |
|---|---|---|
| M1s | TAK | TAK |
| M2s | TAK | TAK |
| M3s | NIE | NIE |
| M4s | NIE | NIE |
| M5s_rozn | TAK | NIE |

Źródło: opracowanie własne na podstawie tabel 2–6.

Jak widać, miara informacji wzajemnej poprawnie zidentyfikowała każdy z analizowanych długich szeregów. W zastosowaniu do szeregów skróconych zastosowana metoda doprowadziła do błędnych wniosków w przypadku szeregów M3s oraz M4s. Otrzymany rezultat jest zgodny z badaniami innych autorów, wskazującymi, że w przypadku małej liczby obserwacji estymatory MI oparte na analizie histogramu mogą być zawodne (np. Dionisio, Menezes, Mendes, 2003).

### 3.2. INDEKSY GIEŁDOWE

Badaniu poddano indeksy Giełdy Papierów Wartościowych w Warszawie z okresu 2.01.2001–15.04.2009 (2078 obserwacji). Dla każdego z indeksów testowaniu poddano trzy szeregi: logarytmicznych dziennych stóp zmian, otrzymanych dla nich reszt z modeli ARMA oraz ARMA-GARCH. Zbadanie reszt z modeli ARMA pozwala stwierdzić, czy ewentualnie wykryte zależności w szeregu stóp zmian mają charakter liniowy. W przypadku odpowiedzi negatywnej, przebadanie standaryzowanych reszt z modelu ARMA-GARCH umożliwia stwierdzenie, czy ta klasa modeli dobrze opisuje zidentyfikowaną nieliniowość[6].

Tabela 9. Wartości $i_k$ dla indeksu WIG

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0458 | 0,0444 | 0,0605 | 0,0612 | 0,0486 | 0,0518 | 0,0350 | 0,0365 | 0,0522 | 0,0559 |
| | 0,0000 | 0,0003 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0338 | 0,0153 | 0,0000 | 0,0000 |
| MA(1) | 0,0412 | 0,0455 | 0,0549 | 0,0632 | 0,0427 | 0,0500 | 0,0379 | 0,0313 | 0,0552 | 0,0566 |
| | 0,0010 | 0,0000 | 0,0000 | 0,0000 | 0,0002 | 0,0000 | 0,0074 | 0,1530 | 0,0000 | 0,0000 |
| MA(1)- | 0,0458 | 0,0498 | 0,0336 | 0,0395 | 0,0359 | 0,0306 | 0,0302 | 0,0352 | 0,0328 | 0,0309 |
| GARCH(3,1) | 0,0225 | 0,0033 | 0,7074 | 0,2254 | 0,5142 | 0,8964 | 0,9124 | 0,5738 | 0,7700 | 0,8823 |

Źródło: obliczenia własne.

Tabela 10. Wartości $i_k$ dla indeksu WIG20

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0514 | 0,0415 | 0,0577 | 0,0690 | 0,0489 | 0,0509 | 0,0388 | 0,0388 | 0,0438 | 0,0537 |
| | 0,0000 | 0,0106 | 0,0000 | 0,0000 | 0,0002 | 0,0000 | 0,0381 | 0,0373 | 0,0029 | 0,0000 |
| MA(1) | 0,0456 | 0,0471 | 0,0579 | 0,0687 | 0,0506 | 0,0510 | 0,0402 | 0,0458 | 0,0439 | 0,0545 |
| | 0,0011 | 0,0006 | 0,0000 | 0,0000 | 0,0001 | 0,0001 | 0,0187 | 0,0009 | 0,0028 | 0,0000 |
| MA(1)- | 0,0441 | 0,0457 | 0,0337 | 0,0382 | 0,0311 | 0,0333 | 0,0307 | 0,0384 | 0,0303 | 0,0272 |
| GARCH(3,1) | 0,0410 | 0,0222 | 0,6723 | 0,2962 | 0,8499 | 0,6978 | 0,8683 | 0,2806 | 0,8867 | 0,9756 |

Źródło: obliczenia własne.

Tabela 11. Wartości $i_k$ dla indeksu mWIG40

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0728 | 0,0508 | 0,0630 | 0,0603 | 0,0545 | 0,0660 | 0,0508 | 0,0343 | 0,0397 | 0,0428 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0058 | 0,0002 | 0,0000 |
| AR(3) | 0,0511 | 0,0458 | 0,0539 | 0,0569 | 0,0508 | 0,0462 | 0,0465 | 0,0376 | 0,0379 | 0,0460 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0002 | 0,0002 | 0,0000 |
| AR(3)- | 0,0340 | 0,0301 | 0,0278 | 0,0404 | 0,0264 | 0,0377 | 0,0250 | 0,0309 | 0,0283 | 0,0295 |
| GARCH(1,2) | 0,0964 | 0,3434 | 0,5657 | 0,0039 | 0,6980 | 0,0182 | 0,8188 | 0,2750 | 0,5131 | 0,3955 |

Źródło: obliczenia własne.

[6] Jakość oszacowanych modeli w kontekście eliminacji wspomnianych zależności została pozytywnie zweryfikowana przy zastosowaniu testów Boxa-Ljunga oraz Engle'a.

Tabela 12. Wartości $i_k$ dla indeksu sWIG80

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0911 | 0,0551 | 0,0680 | 0,0579 | 0,0597 | 0,0546 | 0,0498 | 0,0416 | 0,0440 | 0,0426 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0014 | 0,0003 | 0,0006 |
| ARMA(1,2) | 0,0478 | 0,0386 | 0,0538 | 0,0502 | 0,0397 | 0,0479 | 0,0371 | 0,0340 | 0,0369 | 0,0376 |
| | 0,0000 | 0,0031 | 0,0000 | 0,0000 | 0,0020 | 0,0000 | 0,0074 | 0,0349 | 0,0083 | 0,0056 |
| ARMA(1,2)- | 0,0268 | 0,0300 | 0,0367 | 0,0282 | 0,0258 | 0,0345 | 0,0255 | 0,0295 | 0,0278 | 0,0309 |
| GARCH(1,1) | 0,7878 | 0,5014 | 0,0616 | 0,6646 | 0,8545 | 0,1451 | 0,8722 | 0,5471 | 0,7014 | 0,4072 |

Źródło: obliczenia własne.

Tabela 13. Wartości $i_k$ dla indeksu WIG-Banki

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0429 | 0,0439 | 0,0628 | 0,0556 | 0,0485 | 0,0518 | 0,0476 | 0,0516 | 0,0602 | 0,0443 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| MA(1) | 0,0421 | 0,0469 | 0,0566 | 0,0542 | 0,0609 | 0,0530 | 0,0421 | 0,0544 | 0,0565 | 0,0496 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| MA(1)- | 0,0387 | 0,0346 | 0,0347 | 0,0278 | 0,0308 | 0,0320 | 0,0276 | 0,0302 | 0,0306 | 0,0354 |
| GARCH(1,2) | 0,0534 | 0,2279 | 0,2242 | 0,8131 | 0,5525 | 0,4328 | 0,8250 | 0,6136 | 0,5702 | 0,1790 |

Źródło: obliczenia własne.

Tabela 14. Wartości $i_k$ dla indeksu WIG-Budow

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0525 | 0,0301 | 0,0415 | 0,0400 | 0,0365 | 0,0451 | 0,0460 | 0,0326 | 0,0321 | 0,0451 |
| | 0,0000 | 0,2070 | 0,0009 | 0,0016 | 0,0119 | 0,0001 | 0,0001 | 0,0823 | 0,1004 | 0,0001 |
| ARMA(2,1) | 0,0336 | 0,0386 | 0,0428 | 0,0387 | 0,0350 | 0,0365 | 0,0391 | 0,0221 | 0,0320 | 0,0481 |
| | 0,0145 | 0,0003 | 0,0000 | 0,0003 | 0,0064 | 0,0022 | 0,0002 | 0,7270 | 0,0311 | 0,0000 |
| ARMA(2,1)- | 0,0286 | 0,0289 | 0,0321 | 0,0293 | 0,0239 | 0,0305 | 0,0263 | 0,0251 | 0,0301 | 0,0338 |
| GARCH(1,1) | 0,5966 | 0,5661 | 0,2637 | 0,5281 | 0,9422 | 0,4099 | 0,8061 | 0,8875 | 0,4452 | 0,1569 |

Źródło: obliczenia własne.

Tabela 15. Wartości $i_k$ dla indeksu WIG-Dewel

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,1392 | 0,1477 | 0,1290 | 0,1154 | 0,1859 | 0,1292 | 0,1255 | 0,1370 | 0,1699 | 0,1353 |
| | 0,0063 | 0,0013 | 0,0246 | 0,1188 | 0,0000 | 0,0240 | 0,0392 | 0,0091 | 0,0000 | 0,0118 |
| ARMA(1,1) | 0,1479 | 0,1562 | 0,1466 | 0,1144 | 0,1506 | 0,1515 | 0,1488 | 0,1258 | 0,1412 | 0,1484 |
| | 0,0022 | 0,0006 | 0,0028 | 0,1664 | 0,0017 | 0,0014 | 0,0021 | 0,0531 | 0,0079 | 0,0022 |
| ARMA(1,1)- | 0,0928 | 0,1147 | 0,0999 | 0,0929 | 0,1195 | 0,1153 | 0,1189 | 0,1199 | 0,0869 | 0,1241 |
| GARCH(1,2) | 0,9250 | 0,5124 | 0,8324 | 0,9245 | 0,3976 | 0,4980 | 0,4135 | 0,3861 | 0,9708 | 0,3014 |

Źródło: obliczenia własne.

Tabela 16. Wartości $i_k$ dla indeksu WIG-Info

| szereg \ $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0486 | 0,0388 | 0,0466 | 0,0573 | 0,0454 | 0,0443 | 0,0335 | 0,0513 | 0,0469 | 0,0434 |
| | 0,0000 | 0,0038 | 0,0000 | 0,0000 | 0,0000 | 0,0001 | 0,0580 | 0,0000 | 0,0000 | 0,0002 |
| AR(1) | 0,0585 | 0,0359 | 0,0476 | 0,0619 | 0,0553 | 0,0488 | 0,0314 | 0,0499 | 0,0543 | 0,0409 |
| | 0,0000 | 0,0449 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,2556 | 0,0000 | 0,0000 | 0,0032 |
| AR(1)-GARCH(1,1) | 0,0362 | 0,0251 | 0,0270 | 0,0339 | 0,0222 | 0,0260 | 0,0282 | 0,0244 | 0,0303 | 0,0293 |
| | 0,0778 | 0,8876 | 0,7622 | 0,1810 | 0,9796 | 0,8343 | 0,6611 | 0,9176 | 0,4641 | 0,5646 |

Źródło: obliczenia własne.

Tabela 17. Wartości $i_k$ dla indeksu WIG-Media

| szereg \ $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0481 | 0,0560 | 0,0448 | 0,0562 | 0,0475 | 0,0456 | 0,0350 | 0,0422 | 0,0304 | 0,0393 |
| | 0,0555 | 0,0049 | 0,1229 | 0,0047 | 0,0642 | 0,1020 | 0,6139 | 0,2144 | 0,8539 | 0,3531 |
| MA(1) | 0,0484 | 0,0571 | 0,0516 | 0,0529 | 0,0519 | 0,0398 | 0,0464 | 0,0450 | 0,0446 | 0,0426 |
| | 0,1063 | 0,0097 | 0,0473 | 0,0333 | 0,0432 | 0,4922 | 0,1644 | 0,2159 | 0,2322 | 0,3298 |
| MA(1)-GARCH(1,1) | 0,0484 | 0,0414 | 0,0427 | 0,0481 | 0,0510 | 0,0352 | 0,0397 | 0,0363 | 0,0465 | 0,0370 |
| | 0,1380 | 0,4673 | 0,3925 | 0,1451 | 0,0735 | 0,8260 | 0,5785 | 0,7691 | 0,2020 | 0,7320 |

Źródło: obliczenia własne.

Tabela 18. Wartości $i_k$ dla indeksu WIG-Paliwa

| szereg \ $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0825 | 0,0780 | 0,0761 | 0,0711 | 0,0862 | 0,0658 | 0,0562 | 0,0685 | 0,0829 | 0,0667 |
| | 0,0185 | 0,0503 | 0,0733 | 0,1660 | 0,0076 | 0,3335 | 0,7510 | 0,2422 | 0,0166 | 0,3009 |
| AR(2) | 0,0816 | 0,0652 | 0,0619 | 0,0824 | 0,0820 | 0,0863 | 0,0573 | 0,0823 | 0,0878 | 0,0771 |
| | 0,0203 | 0,3279 | 0,4671 | 0,0183 | 0,0195 | 0,0062 | 0,6711 | 0,0184 | 0,0043 | 0,0493 |
| AR(2)-GARCH(1,1) | 0,0451 | 0,0611 | 0,0493 | 0,0573 | 0,0716 | 0,0573 | 0,0524 | 0,0652 | 0,0592 | 0,0472 |
| | 0,8837 | 0,2362 | 0,7406 | 0,3809 | 0,0368 | 0,3787 | 0,6043 | 0,1240 | 0,3055 | 0,8178 |

Źródło: obliczenia własne.

Tabela 19. Wartości $i_k$ dla indeksu WIG-Spozy

| szereg \ $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0544 | 0,0361 | 0,0449 | 0,0473 | 0,0446 | 0,0311 | 0,0314 | 0,0337 | 0,0305 | 0,0342 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0037 | 0,0028 | 0,0003 | 0,0052 | 0,0002 |
| ARMA(1,1) | 0,0371 | 0,0365 | 0,0418 | 0,0433 | 0,0358 | 0,0270 | 0,0310 | 0,0409 | 0,0230 | 0,0298 |
| | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0055 | 0,0003 | 0,0000 | 0,0601 | 0,0007 |
| ARMA(1,1)-GARCH(1,1) | 0,0311 | 0,0340 | 0,0338 | 0,0347 | 0,0239 | 0,0309 | 0,0281 | 0,0281 | 0,0311 | 0,0362 |
| | 0,5233 | 0,2766 | 0,2873 | 0,2198 | 0,9738 | 0,5448 | 0,7944 | 0,7938 | 0,5252 | 0,1404 |

Źródło: obliczenia własne.

Tabela 20. Wartości $i_k$ dla indeksu WIG-Telkom

| szereg \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| stopy zmian | 0,0467 | 0,0395 | 0,0429 | 0,0687 | 0,0440 | 0,0393 | 0,0417 | 0,0469 | 0,0419 | 0,0514 |
| | 0,0072 | 0,1307 | 0,0369 | 0,0000 | 0,0234 | 0,1405 | 0,0579 | 0,0062 | 0,0518 | 0,0007 |
| GARCH(1,3) | 0,0311 | 0,0340 | 0,0338 | 0,0347 | 0,0239 | 0,0309 | 0,0281 | 0,0281 | 0,0311 | 0,0362 |
| | 0,5186 | 0,2752 | 0,2834 | 0,2234 | 0,9693 | 0,5411 | 0,7826 | 0,7821 | 0,5207 | 0,1448 |

Źródło: obliczenia własne.

Jak wynika z tabel 9–20 w analizowanych indeksach powszechnie występują autozależności[7]. Ponadto, zależności te mają charakter nieliniowy (za wyjątkiem indeksu WIG-Media). W większości wypadków zidentyfikowaną zależność dobrze opisuje model typu ARMA-GARCH. Jedynie w przypadku indeksów WIG oraz mWIG40 istnieją przesłanki, aby sądzić, że zidentyfikowana nieliniowość nie jest wywołana efektem ARCH.

## LITERATURA

Barnett W. A., Gallant A. R., Hinich M. J., Jungeilges J. A., Kaplan D., Jensen M. J. (1998), *A single-blind controlled competition among tests for nonlinearity and chaos*, „Journal of Econometrics”, 82.1, 157–192.

Bruzda J. (2004), *Miary zależności nieliniowej w identyfikacji nieliniowych procesów ekonomicznych*, „Acta Universitatis Nicolai Copernici”, 34, 183–203.

Dionisio A., Menezes R., Mendes D.A. (2003), Mutual Information: a dependence measure for nonlinear time series, working paper.

Fraser A.M., Swinney H.L. (1986), *Independent coordinates for strange attractors from mutual information*, „Physical Review A”, nr 33.2, 1134–1140.

Granger C. W. J., Terasvirta T. (1993), *Modelling nonlinear economic relationship*, Oxford University Press.

Granger C. W. J., Lin J-L. (1994), *Using the mutual information coefficient to identify lags in nonlinear models*, „Journal of Time Series Analysis”, 15, 371–384.

Maasoumi E., Racine J. (2002), *Entropy and predictability of stock market returns, „Journal of Econometrics*”, 107, 291–312.

Orzeszko W., *Detection of nonlinear autodependencies using Hiemstra-Jones test*, w druku.

## THE MUTUAL INFORMATION COEFFICIENT AS A MEASURE OF NONLINEAR SERIAL DEPENDENCIES

A b s t r a c t. The idea, estimation and applicability of the mutual information coefficient are presented. Simulations are provided to verify its usefulness to identify nonlinear serial dependencies. Moreover, the mutual information coefficient is applied to analyze the indices from The Warsaw Stock Exchange.

K e y w o r d s: nonlinearity, mutual information, identification of dependencies.

[7] Wyjątkiem jest indeks WIG-Paliwa. W tym przypadku otrzymano efekt dość nietypowy: przefiltrowanie danych modelem ARMA spowodowało pojawienie się istotności miary MI dla $k$=9.