DYNAMICZNE MODELE EKONOMETRYCZNE
IX Ogólnopolskie Seminarium Naukowe, 6–8 września 2005 w Toruniu
Katedra Ekonometrii i Statystyki, Uniwersytet Mikołaja Kopernika w Toruniu

Magdalena Osińska, Joanna Górka
Uniwersytet Mikołaja Kopernika w Toruniu

Identyfikacja nieliniowości w ekonomicznych szeregach czasowych. Analiza symulacyjna*

1. Wstęp

Współczesna ekonometria coraz częściej zwraca się ku zastosowaniu modeli nieliniowych opisujących mechanizm generujący dane w empirycznych szeregach czasowych. Szczególnie prężna stała się pod tym względem ekonometria finansowa (por. Doman, Doman (2004)). Jedną z ostatnio powstałych specyfikacji modelowych, generujących szeregi czasowe są procesy zawierające stochastyczny pierwiastek jednostkowy STUR (por. Granger, Swanson (1997)). Są to procesy definiowane przez model z losowymi parametrami, które znajdują się pomiędzy klasą procesów zawierających dokładny pierwiastek jednostkowy (np. błądzenie przypadkowe lub ogólnie ARIMA (p,1,q)) a klasą procesów stacjonarnych (np. ARMA (p,q)). Modele te poprzez przekształcenia mogą być sprowadzone go klasy modeli nieliniowych np. dwuliniowych lub GARCH. Z drugiej strony istnieje niewiele możliwości skutecznej identyfikacji rodzaju nieliniowości w rzeczywistych szeregach czasowych.
Celem referatu jest próba określenia możliwości identyfikacji odpowiadających sobie procesów: GARCH, zmienności stochastycznej, dwuliniowych oraz STUR na podstawie eksperymentu Monte Carlo. Dla porównania pokazane zostaną charakterystyki szeregów generowanych jako biały szum, SETAR oraz błądzenie przypadkowe. Jako, że zastosowanie omawianych modeli odnosi się najczęściej do finansowych szeregów czasowych, w końcowej części pracy pokazane zostały wyniki identyfikacji empirycznych szeregów finansowych.

* Praca finansowana przez KBN w ramach grantu nr 2-H02B-015-25.

2. Charakterystyka nieliniowych szeregów czasowych

W literaturze znajduje się szereg testów pozwalających na identyfikację nieliniowości. Ogólnie możemy je podzielić na dwie grupy. Do grupy pierwszej zaliczymy wszystkie testy, które testują nieliniowość bez określenia z jakim jej rodzajem mamy do czynienia. W tej grupie najbardziej znanym jest test BDS. Zaliczyć tu można także testy oparte na reprezentacji spektralnej momentów wyższych rzędów, w tym bikowariancji i bispekrtum, np. test Hinicha. Z drugiej strony dysponujemy szeregiem testów nakierowanych na weryfikację konkretnej alternatywy, np. testy modeli STAR i SETAR, a także testy Leybourne'a i innych nakierowane na badanie alternatywy w postaci modelu stochastycznego pierwiastka jednostkowego. W niniejszej pracy wykorzystane zostały: test Hinicha (1982) oraz testy Leybourne'a, McCabe'a i Tremayne'a (1996) i Leybourne'a. McCabe'a i Millsa (1996).

Za pomocą testu Hinicha, opartego na bispektrum, badamy dane ze względu na występowanie zależności trzeciego rzędu. Załóżmy, że $\mathrm{y_t}$ jest realizacją procesu stochastycznego o średniej zero, stacjonarnego aż do trzeciego rzędu. Oznacza to, że $\mathrm{E} \mid \mathrm{y_t} \mid^3 = \mathrm{K} < \infty$ oraz momenty $E(y_1, y_2, y_3)$ nie zależą od przesunięć w czasie. Bispektrum $B_y(\omega_1\omega_2)$ szeregu y_t nazywamy transformatę Fouriera funkcji $c_3(r,s) = E(y_t y_{t+r} y_{t+s})$ postaci:

$$B_y(\omega_1\omega_2) = \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} c_3(r,s) \exp[-2\pi i(\omega_1 r + \omega_2 s)]. \tag{1}$$

Funkcja (1) jest dwuokresową funkcją dwóch zmiennych $(\omega_1\omega_2)$ o dziedzinie podstawowej $\{0 < \omega_2 < \frac{2}{3}\pi; \omega_2 < \omega_1 < -\frac{1}{2}\omega_2 + \pi\}$. Statystyka Hinicha oparta jest na współczynniku bikoherencji, którego estymator ma postać:

$$\Psi^2(\omega_1\omega_2) = \frac{|\hat{B}_y(\omega_1\omega_2)|^2}{\hat{f}_y(\omega_1)\hat{f}_y(\omega_2)\hat{f}_y(\omega_{1+2})} \tag{2}$$

gdzie $\hat{B}_y(\omega_1\omega_2)$ jest estymatorem bispektrum szeregu, zaś $\hat{f}_y(\omega_j)$ $j = \{1,2\}$ oznacza uśredniony periodogram dla częstości ω_j. Formuła (2) znajduje zastosowanie w weryfikacji hipotez dwojako: po pierwsze testuje się normalność rozkładu z którego pochodzi badany szereg, po drugie natomiast liniowość tych szeregów, dla których została odrzucona hipoteza o normalności. W pierwszym przypadku przy założeniu prawdziwości hipotezy zerowej statystyka

$$H = 2\Psi^2(\omega_1\omega_2) \tag{3}$$

ma asymptotyczny rozkład $\chi^2(2P)$, gdzie P jest liczbą par częstości $(\omega_1\omega_2)$ w dziedzinie podstawowej. Przyjęcie hipotezy zerowej kończy badanie, natomiast jej odrzucenie wymaga testowania hipotezy o liniowości badanego procesu. Oznacza to, że proces może być nie-Gaussowski ale liniowy albo nie-

Gaussowski i nieliniowy. Przy założeniu prawdziwości hipotezy o liniowości procesu, statystyka (3) ma nie scentrowany rozkład $\chi^2(2,\lambda)$, gdzie λ jest wartością średnią tego procesu. Wówczas estymator $2\Psi^2(\omega_1\omega_2)$ dostarcza P niezależnych wartości z rozkładu $\chi^2(2,\lambda)$. Jeżeli zatem badany proces jest liniowy, to dyspersja próbkowa P wartości jest zgodna z tym rozkładem. W przeciwnym wypadku statystyka $2\Psi^2(\omega_1\omega_2)$ dostarcza P niezależnych wartości pochodzących z nie scentrowanego rozkładu χ^2 ze zmienną wartością średnią, którego dyspersja jest oczywiście różna. Testowanie polega na porównaniu dyspersji obu rozkładów, na przykład poprzez rozstęp decylowy lub rozstęp kwartylowy. Ten ostatni został wykorzystany w niniejszej pracy. Hipoteza o liniowości szeregu będzie odrzucana wówczas, gdy wartości empiryczna i teoretyczna rozstępu kwartylowego będą się znacznie różnic między sobą. Hinich (1982) udowodnił, że zastosowanie powyższych testów nie wymaga uprzedniego filtrowania ewentualnych zależności liniowych.

W celu zbadania czy dany szereg jest realizacja procesu stochastycznego zawierającego STUR rozważmy następujący model autoregresyjny ze zmiennymi parametrami:

$$y_t = \alpha_t y_{t-1} + \varepsilon_t \tag{4}$$

gdzie:

$$\alpha_t = \alpha_0 + \delta_t,\ \ \delta_0 = 0,\ \ \delta_t = \rho\delta_{t-1} + \eta_t \tag{5}$$

przy czym $\alpha_0 = 1$ i $|\rho| \le 1$. Ponadto $\varepsilon_t \sim N(0,\sigma^2)$ i $\eta_t \sim N(0,\omega^2)$ są od siebie niezależne.

Hipotezy w teście LMT (Leybourne, McCabe, Tremayne) dotyczą wariancji ω^2. Jeżeli $H_0 : \omega^2 = 0$, to model (4) sprowadza się do modelu błądzenia przypadkowego (lub w uogólnionym przypadku ARIMA(p,1,0)). Hipoteza alternatywna $H_1 : \omega^2 > 0$ oznacza model STUR.

W celu uniknięcia wpływu ewentualnego trendu deterministycznego, autorzy proponują rozszerzenie modelu o trend liniowy lub kwadratowy. Ponadto możliwe jest rozszerzenie specyfikacji równania poprzez włączenie do modelu opóźnionych wartości zmiennej endogenicznej. Procedura testowania przebiega w następujący sposób:

Niech $y_t^* = \alpha_t y_{t-1}^* + \varepsilon_t$ (6)

gdzie $y_t^* = y_y - P_t - \sum_{i=1}^{p} \varphi_i y_{t-i}$ (7)

przy czym P_t jest składnikiem deterministycznym, np. trendem, natomiast część autoregresyjna w (8) jest stacjonarna. Jeżeli w H_1 $|\rho| < 1$, to statystykę Z oblicza się na podstawie następującej zależności, oszacowanej KMNK

$$\Delta y_t = \Delta P_t + \sum_{i=1}^{p} \varphi_i \Delta y_{t-1} + \varepsilon_t \tag{8}$$

Statystyka Z ma postać:

$$Z = T^{-\frac{3}{2}} \sigma^{-2} \kappa^{-1} \sum_{t=2}^{T} \left(\sum_{j=1}^{t-1} \varepsilon_j \right)^2 \left(\varepsilon_t^2 - \sigma^2 \right) \tag{9}$$

gdzie: $\sigma^2 = T^{-1} \sum_{t=1}^{T} \varepsilon_t^2$ oraz $\kappa^2 = T^{-1} \sum_{t=1}^{T} \left(\varepsilon_t^2 - \sigma^2 \right)^2$.

Zależnie od wyboru postaci trendu P_{1t} (liniowy) lub P_{2t} (kwadratowy) oznacza się ją jako Z_1 lub Z_2. Dla $\rho = 1$ w H_1 proponuje się przyjęcie następującej statystyki (por. Leybourne, McCabe, Mills (1996))

$$E = T^{-3} \sigma^{-4} \sum_{i=2}^{T} \left\{ \left[\sum_{t=i}^{T} \varepsilon_t \left(\sum_{j=1}^{t-1} \varepsilon_j \right) \right]^2 - \sigma^2 \sum_{t=i}^{T} \left(\sum_{j=1}^{t-1} \varepsilon_j \right)^2 \right\} \tag{10}$$

i, podobnie jak w poprzednim wypadku, zależnie od przyjętej funkcji trendu statystykę powyższą oznacza się przez E_1 lub E_2.
Wyniki symulacji pokazują, że moc testu E jest wyższa niż testu Z (por. Taylor, van Dijk (1999)) Wartości krytyczne testów Z i E dla wybranych poziomów istotności oraz liczebności próby zostały przedstawione w pracy Osińska (2004).

Generowane oraz empiryczne szeregi czasowe zostały także scharakteryzowane za pomocą statystycznych miar opisu. Ponadto za pomocą standardowych testów zbadano autokorelację w zakresie poziomów (test Boxa-Ljunga) oraz w zakresie zmienności (test Engle'a i McLeoda-Li). Wyniki te nie zostały zaprezentowane w celu oszczędności miejsca, są natomiast dostępne u autorów.

3. Wyniki eksperymentu Monte Carlo

Celem analizy symulacyjnej jest określenie wrażliwości testów wykrywających stochastyczny pierwiastek jednostkowy, na różne formy hipotezy alternatywnej, podczas gdy H_0 standardowo oznacza model błądzenia przypadkowego. Podobne badanie dla wybranych alternatyw przeprowadzone zostały w pracy Taylor, van Dijk (1999). Ponadto wykorzystany został test Hinicha w celu określenia na ile generowane szeregi czasowe mogą zostać uznane za Gaussowskie, a następnie za nieliniowe. Własności testów opartych na bispektrum były przedmiotem prac Bruzda (2002) i Doman, Doman (2004), jak dotąd jednak nie rozważano pod tym kątem procesów zawierających stochastyczny pierwiastek jednostkowy. Wynika to zapewne z faktu, że zastosowanie analizy spektralnej wymaga stacjonarności procesów aż do rzędu badanego momentu. Procesy STUR tej własności z definicji nie posiadają. W niniejszym opracowaniu podjęto jednak próbę zastosowania testu Hinicha w celu określenia możliwości identyfikacji liniowości szeregów generowanych przez model STUR. W analizie symulacyjnej przyjęte zostały następujące postaci modeli:

WN		$\varepsilon_t \sim N(0,1)$
RW	$y_t = y_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$
BI[1]	$y_t = \sigma_t \cdot u_t$	$\varepsilon_t \sim N(0,1)$
	$\sigma_t = \sigma_{t-1} + \varepsilon_t$	$u_t \sim N(0,1)$
BL diag	$y_t = -1 - 0.5 y_{t-1}\varepsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$
BL nad	$y_t = -1 - 0.5 y_{t-2}\varepsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$
BL pod	$y_t = -1 - 0.5 y_{t-1}\varepsilon_{t-2} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$
ARCH	$y_t = \varepsilon_t \sqrt{h_t}$	$\varepsilon_t \sim N(0,1)$
	$h_t = 0{,}1 + 0.5\varepsilon_{t-1}^2 + 0.4\varepsilon_{t-2}^2$	
GARCH	$y_t = \varepsilon_t \sqrt{h_t}$, $h_t = 0{,}1 + 0.89 h_{t-1} + 0.1\varepsilon_{t-1}^2$	$\varepsilon_t \sim N(0,1)$
SV[2]	$y_t = \varepsilon_t \sqrt{h_t}$	$\varepsilon_t \sim N(0,1)$
	$\ln h_t = 0{,}1 + 0.95 \ln h_{t-1} + \sqrt{0.009}\eta_t$	$\eta_t \sim N(0,1)$
SETAR(2,1,2)	$y_t = \begin{cases} 0.5 - 0.22 y_{t-1} + 0.1\varepsilon_t & y_{t-2} < 0.07 \\ 0.82 y_{t-1} + 0.4\varepsilon_t & y_{t-2} \geq 0.07 \end{cases}$	$\varepsilon_t \sim N(0,1)$
RCA(1,1)	$y_t = b_t \cdot \varepsilon_t$	$\varepsilon_t \sim N(0,1)$
	$b_t = 0.3 b_{t-1} + 0.6\eta_t$	$\eta_t \sim N(0,1)$
STUR1	$\alpha_0 = 1$, $\rho = 0.98$, $\omega^2 = 0.01$	$\varepsilon_t \sim N(0,1)$
STUR2	$\alpha_0 = 1$, $\rho = 0.95$, $\omega^2 = 0.01$	$\varepsilon_t \sim N(0,1)$
STUR3	$\alpha_0 = 0.98$, $\rho = 0.98$, $\omega^2 = 0.01$	$\varepsilon_t \sim N(0,1)$
STUR4	$\alpha_0 = 0.98$, $\rho = 0.95$, $\omega^2 = 0.01$	$\varepsilon_t \sim N(0,1)$
STUR5	$\alpha_0 = 1$, $\rho = 0.98$, $\omega^2 = 0.05$	$\varepsilon_t \sim N(0,1)$
STUR6	$\alpha_0 = 1$, $\rho = 0.95$, $\omega^2 = 0.05$	$\varepsilon_t \sim N(0,1)$
STUR7	$\alpha_0 = 0.98$, $\rho = 0.98$, $\omega^2 = 0.05$	$\varepsilon_t \sim N(0,1)$
STUR8	$\alpha_0 = 0.98$, $\rho = 0.95$, $\omega^2 = 0.05$	$\varepsilon_t \sim N(0,1)$
STUR9	$\alpha_0 = 1$, $\rho = 0.98$, $\omega^2 = 0.1$	$\varepsilon_t \sim N(0,1)$
STUR10	$\alpha_0 = 1$, $\rho = 0.95$, $\omega^2 = 0.1$	$\varepsilon_t \sim N(0,1)$
STUR11	$\alpha_0 = 0.98$, $\rho = 0.98$, $\omega^2 = 0.1$	$\varepsilon_t \sim N(0,1)$
STUR12	$\alpha_0 = 0.98$, $\rho = 0.95$, $\omega^2 = 0.1$	$\varepsilon_t \sim N(0,1)$

[1] Model BI został zdefiniowany w pracy Hansen (1992).
[2] Parametry modelu SV zostały zaczerpnięte z pracy Pajor (2003).

Dla każdego z powyższych modeli wygenerowano 500 obserwacji i dokonano 1000 powtórzeń. Wyniki testów Z i E oraz testu Hinicha zaprezentowane zostały w tablicy 1.

Tablica 1. Wyniki testów Z, E oraz testu Hinicha dla danych generowanych. W tabeli podane zostały odsetki odrzucenia hipotezy zerowej przy poziomie istotności 0.05.

Model	Test Z		Test E		Test Hinicha	
	Z1	Z2	E1	E2	Normalność	Liniowość
WN	0.000	0.000	0.000	0.000	0.000	--
RW	0.170	0.060	0.000	0.000	--	--
BI	0.026	0.008	0.001	0.001	--	--
BL diag	0.109	0.085	0.003	0.002	0.647	0.445
BL nad	0.006	0.002	0.000	0.000	0.189	0.328
BL pod	0.108	0.105	0.004	0.002	0.590	0.412
ARCH	0.163	0.165	0.025	0.021	0.703	0.703
GARCH	0.012	0.008	0.000	0.000	0.087	0.322
SV	0.000	0.000	0.000	0.000	0.037	0.405
SETAR	0.004	0.000	0.000	0.000	0.391	0.205
RCA(1.1)	0.042	0.036	0.002	0.002	0.145	0.538
STUR1	0.851	0.919	0.793	0.855	0.996	0.996
STUR2	0.892	0.942	0.823	0.852	1.000	0.995
STUR3	0.845	0.923	0.794	0.851	0.997	0.997
STUR4	0.895	0.938	0.839	0.863	1.000	0.993
STUR5	0.885	0.940	0.806	0.835	0.999	1.000
STUR6	0.913	0.953	0.855	0.868	1.000	0.999
STUR7	0.895	0.929	0.759	0.786	1.000	0.998
STUR8	0.916	0.953	0.842	0.848	1.000	1.000
STUR9	0.866	0.926	0.741	0.755	1.000	0.998
STUR10	0.927	0.960	0.827	0.853	1.000	1.000
STUR11	0.878	0.930	0.732	0.745	1.000	0.996
STUR12	0.922	0.957	0.839	0.841	1.000	1.000

Źródło: opracowanie własne.

Testy na istnienie stochastycznego pierwiastka jednostkowego poprzestają na akceptacji hipotezy zerowej w przypadku szeregów generowanych przez stacjonarne procesy stochastyczne, takich jak biały szum, model współzintegrowany (BI), modele dwuliniowe, ARCH, GARCH, model zmienności stochastycznej, RCA(1,1) oraz SETAR. W przypadku szeregów niestacjonarnych reprezentowanych przez model błądzenia przypadkowego RW, testy STUR wyraźnie wskazują na hipotezę zerową, która mówi, że mamy do czynienia z dokładnym pierwiastkiem jednostkowym. W przypadku szeregów generowanych przez różne typy modeli STUR wyniki testów Z i E są obiecujące. Zależnie od wartości parametrów modeli procent odrzucenia hipotezy zerowej waha się w granicach od 73% do 96%. Wyniki testu Hinicha są dość zróżnicowane. Na przykład dla procesów dwuliniowych diagonalnych i poddiagonalnych w ok. 60% odrzucamy hipotezę zerową o normalności rozkładu, natomiast w przypadku procesów naddiagonalnych tylko w 18.9%. Podobne wyniki obserwujemy w odniesieniu do modeli zmienności ARCH, GARCH i SV. Jeżeli chodzi o wyniki testu Hinicha dla szeregów STUR to niemal w 100% przypadków obserwujemy brak normalności rozkładu i nieliniowość, co może być obciążone niespełnieniem założenia o stacjonarności.

4. Wyniki analiz empirycznych

Identyfikacja empirycznych szeregów czasowych skoncentrowana została na badaniu dziennych obserwacji wybranych indeksów notowanych na GPW w Warszawie oraz ich logarytmicznych stóp zmian w okresie 2.01.01 – 13.02.04. Wyniki testów Z i E oraz testu Hinicha przedstawione zostały w tablicy 2.

Tablica 2. Wyniki testowania empirycznych szeregów czasowych.

Model	Test Z		Test E		Test Hinicha	
	Z1	Z2	E1	E2	Normalność	Liniowość
WIG	-0.901	0.088	-0.107	0.023	---	---
ln st zw WIG	0.015	0.015	0.002	0.002	5.516	---
WIG20	0.785*	0.667*	0.808*	0.627*	---	---
ln st zw WIG20	0.069	0.069	0.051	0.051	7.914	---
WIG-BANKI	-0.295	-0.022	-0.010	-0.001	---	---
ln st zw WIG-BANKI	0.024	0.024	0.001	0.001	7.229	---
WIG-BUDOW	-0.558	-0.372	-0.011	-0.010	---	---
ln st zw WIG-BUDOW	0.011	0.016	0.001	0.001	5.717	---
WIG-INFO	-0.174	0.222*	-0.048	0.000	---	---
ln st zw WIG-INFO	0.064	0.057	0.001	0.001	14.187	---
WIG-SPOZYW	-1.792	0.198*	-0.047	0.018	---	---
ln st zw WIG-SPOZYW	0.041	0.036	0.001	0.001	14.785	---
WIG-TELKOM	0.493*	0.374*	-0.133	-0.006	---	---
ln st zw WIG-TELKOM	0.035	0.036	0.001	0.001	12.929	---

* oznaczono odrzucenie H_0 o dokładnym pierwiastku jednostkowym.
Źródło: opracowanie własne.

Z przedstawionych rezultatów empirycznych wynika, że indeksy WIG20, WIG-info, WIG-spożywczy oraz WIG-telkom charakteryzują się stochastycznym pierwiastkiem jednostkowym. Test Hinicha dla indeksów, które potencjalnie są procesami I(1) nie został policzony. Dla logarytmicznych stóp zmian w żadnym przypadku nie było podstaw do odrzucenia hipotezy zerowej.

Literatura

Bruzda, J. (2002), Bispektra procesów ekonomicznych – kierunki zastosowań i analiza symulacyjna, *Acta Universitatis Nicolai Copernici*, Toruń.

Doman, M., Doman, R. (2004), *Ekonometryczne modelowanie dynamiki polskiego rynku finansowego,* Wyd. AE w Poznaniu, Poznań.

Hansen, B. E. (1992), Heteroskedastic cointegration, *Journal of Econometrics*, vol.54.

Hinich, M.J. (1982), Testing for Gaussianity and Linearity of a Stationary Time Series, *Journal of Time Series Analysis*, vol.3.

Granger, C.W.J., Swanson, N.R. (1997), An introduction to stochastic unit-root process, *Journal of Econometrics*, vol.80.

Leybourne, S.J., McCabe, B.P.M., Mills, T.C. (1996), Randomized unit root processes for modeling and forecasting financial time series: theory and applications*, Journal of Forecasting*, vol.15.

Leybourne, S.J., McCabe, B.P.M., Tremayne, A.R. (1996), Can economic time series be differenced to stationarity?, *Journal of Business and Economic Statistics*, vol.14.

Osińska, M. (2004), Stochastic unit roots processes - properties and application, w: *MACROMODELS'2003*, red. A. Welfe, W. Welfe, Wyd. Uniwersytetu Łódzkiego, Łódź.

Pajor, A. (2003), *Procesy zmienności stochastycznej SV w bayesowskiej analizie szeregów czasowych,* Wydawnictwo AE , Kraków.

Taylor, A.M.R, van Dijk, D. (1999), Testing for Stochastic Unit Roots. Some Monte Carlo Evidence*, Econometric Institute Research Report EI-9922/A*.