

Ewa Kwiatkowska

ANALIZA AUTOKORELACJI SKŁADNIKÓW LOSOWYCH RÓWNAŃ AUTOREGRESJI WSKAŹNIKA INFLACJI W POLSCE

Wprowadzenie

W opracowaniu podjęto problem autokorelacji wyższych rzędów reszt modeli autoregresji. Za punkt wyjścia wzięto model pętli inflacyjnej. Zbadano, jaką postać może przyjąć równanie końcowe modelu ze względu na podstawową jego zmienną – przyrost cen. Jego postaci poszukiwano pośród modeli autoregresji. Dużą uwagę skupiono na rzędzie autokorelacji reszt otrzymanych modeli, do badania którego zastosowano test *J* Pawłowskiego oraz metodę wykorzystującą oceny parametrów, wyznaczone dla możliwie minimalnej liczby danych, z uwzględnieniem wyznaczonych punktów zmiany struktury modelu. Tę samą metodę, opartą na ocenach parametrów, zastosowano do wyznaczenia rzędu autokorelacji reszt.

Model pętli inflacyjnej rozważa się najczęściej w kontekście analizy kointegracji ze względu na jego wielowymiarową strukturę. Model kointegracji pozwala wyznaczać krótko- i długookresowe zależności pomiędzy zadanymi zmiennymi. Idea modeli wektorowej autoregresji (VAR) jest ściśle związana z modelowaniem, w którym nie zakłada się *a priori* związków zachodzących między zmiennymi, lecz wyznacza się je w sposób analityczny, wykorzystując jedynie dane zawarte w próbie statystycznej. Jak pokazała praktyka, całkowite pomijanie informacji z teorii ekonomii, które pierwotnie postulowało się w tego typu modelach, prowadzi czasami do wyników, które są niezgodne z teorią. Trudno wówczas stwierdzić, czy model jest niedoskonały, czy też teoria nie jest wystarczająco ogólna, aby wyjaśnić – być może tylko pozorne – sprzeczności. W badaniach rozważono szczególny model VAR – model pętli inflacyjnej (por. Box, Jenkins, 1983; Majsterek, 1998; Welfe, Majsterek, Florczak, 1994).

Tradycyjny model pętli inflacyjnej jest modelem liniowym o dwóch równaniach. Liniowość zależności jest pewnym uproszczeniem (por. Osiewalski, Welfe, 1999) i tylko przybliżeniem prawdziwych związków. Biorąc pod uwagę to, że przy wystarczająco dużej próbie model taki ma opisywać związki długookresowe, podczas gdy przybliżenie liniowe jest na ogół dobre dla mało licznej próby, założenie liniowości może poważnie zniekształcić długookresowe wyniki. Z drugiej strony, estymatory wyznaczone na podstawie mało licznej próby mogą znacznie się różnić od rzeczywistych wielkości parametrów, szczególnie gdy model cechuje duża wariancja. Ponadto w równaniach modelu nie uwzględnia się opóźnień poszczególnych zmiennych i zakłada się niezależność składników losowych w poszczególnych równaniach oraz stałość parametrów. Założenia te ograniczają model i, jeżeli są błędne, podają nieprawdziwe związki.

W badaniach przeprowadzono analizę zachowania się zmiennej przedstawiającej inflację (obok płac), podstawowej zmiennej modelu spirali inflacyjnej. Nie rozważono modelu jako całości.

1. Model spirali inflacyjnej

Zakładając liniowość związków, klasyczny model sprzężenia cenowo-płacowego można zapisać w następujący sposób (por. *Gospodarka Polski w okresie transformacji*, 2000; Majsterek, 1998; Osiewalski, Welfe, 1999):

$$\Delta p_t = \gamma_1 \Delta w_t + \beta_1 \Delta z_t + \beta_2 \Delta m_t + \beta_3 + \xi_{P_t} \quad (1a)$$

$$\Delta w_t = \gamma_2 \Delta p_t + \beta_4 \Delta z_t + \beta_5 U_t + \beta_6 + \xi_{W_t} \quad (1b)$$

gdzie:

P – indeks dóbr konsumpcyjnych (indeks kosztów utrzymania),

W – przeciętne płace w sektorze przedsiębiorstw (ceny bieżące),

Z – wydajność pracy (ceny stałe),

M – indeks cen importu,

U – stopa bezrobocia.

Przy dodatkowym założeniu, że składniki losowe ξ_{P_t} , ξ_{W_t} mają rozkład sferyczny, zmienne Z , M i U można traktować jako zmienne egzogeniczne. Zgodnie z teorią ekonomii (Welfe, Majsterek, Florczak, 1994) w równaniach modelu (1) uwzględnia się przyrosty logarytmów zmiennych P , W , Z , M , natomiast U uwzględnia się w postaci poziomu.

Pierwsze równanie jest standardową funkcją płać (por. *Gospodarka Polski w okresie transformacji*, 2000). Drugie równanie wynika z przyjęcia kosztowej teorii inflacji i odpowiada hipotezie, że ceny zmieniają się wraz z kosztami produkcji, które są zwiększane o stały w czasie narzut*.

Wskaźnik inflacji jest jednym z najważniejszych wskaźników polityki gospodarczej. Polityka finansowa jest bardzo często podporządkowana osiągnięciu zadanego poziomu inflacji w danym okresie. Wówczas i pozostałe parametry gospodarcze są podporządkowane temu celowi. Wartości jednej grupy tych parametrów są wynikiem zmian inflacji, z kolei inne przyczyniają się do zmian inflacji. Nie można jednak w sposób jednoznaczny podzielić tych parametrów na przyczyny i skutki zmian poziomu wskaźnika cen. Do tychże parametrów zaliczyć można m.in.: poziom zatrudnienia, podaż pieniądza, kurs walutowy, politykę wynagrodzeń czy strukturę wydatków budżetowych. Niestety sposób obliczania wskaźnika powoduje jego niejednorodność i brak ekonomicznej precyzyjności, dlatego też badanie zachowania się szeregu czasowego inflacji jest utrudnione. Wyniki badań mogą być zatem mylące (por. Łyko, 2000).

2. Metody określania rzędu autokorelacji

2.1. Autokorelacja wyższych rzędów – test *J* Pawłowskiego

W celu zbadania występowania autokorelacji reszt modeli rozważonych w badaniach, zostanie zastosowany test *J* zaproponowany przez prof. Pawłowskiego. Najbardziej znanym testem na autokorelację reszt modelu ekonometrycznego jest test Durbina-Watsona (Theil, 1979; Welfe, 2003). Niestety nie są dokładnie znane kwantyle rozkładu jego statystyki testowej. Ponadto sam test, jak i jego modyfikacje wymagają założenia normalności rozkładu badanych składników losowych. Zasadniczą wadą tego testu jest jednak to, że odnosi się on jedynie do badania autokorelacji rzędu 1. W rzeczywistości związki autoregresyjne mogą być bardziej skomplikowane, co może się objawiać występowaniem autokorelacji wyższego rzędu. W szczególności współliniowość zmiennych objaśniających w modelach ekonometrycznych albo pominięcie

* Szerzej na temat budowy równań (1a) i (1b) zob. w pracy *Gospodarka Polski w okresie transformacji* (2000).

pewnych zmiennych objaśniających, co może się zdarzyć w przypadku braku informacji pochodzących z teorii ekonomicznej, może prowadzić do autokorelacji składników losowych wyższych rzędów.

Inną metodą badania rzędu autokorelacji jest metoda wykorzystująca funkcję autokorelacji cząstkowej, którą wyznacza się na podstawie pewnych charakterystyk rozwiązań równań Yule'a-Walkera (Box, Jenkins, 1983; Nerlove, Grether, Carvalho, 2000). Jednakże metoda ta wymaga długiego szeregu czasowego o stosunkowo stałych związkach autokorelacyjnych dla całego ciągu obserwacji. Przyczyną takiego uwarunkowania dobrych wyników zastosowania funkcji autokorelacji cząstkowej do wyznaczania rzędu autokorelacji jest to, że podstawowymi wielkościami wykorzystywanymi w tej metodzie są współczynniki autokorelacji, do których dobrej estymacji są wymagane szeregi czasowe o dużej liczebności.

Wymienione dwa warunki często nie są spełnione w badaniach ekonomicznych szeregów czasowych. Nawet wystarczająco liczna próba nie zawsze jest realizacją procesu o stałych w czasie związkach korelacyjnych. Znajomość rzędu autokorelacji w szeregu czasowym może być cenną informacją, może inspirować do poszukiwania przyczyn takiej autokorelacji oraz z pewnością przyczynia się do lepszego opisu zjawisk ekonomicznych, ponieważ jest formą opisu zachowania się nieobserwowanych zmiennych mających wpływ na przebieg procesu.

Niech dana będzie n -elementowa próba z_1, z_2, \dots, z_n w postaci szeregu czasowego. Idea testu J jest następująca: Jeżeli w ciągu obserwacji występuje autokorelacja dodatnia rzędu l , to wartości bezwzględne różnic $v_{t,t-1} = |z_t - z_{t-1}|$ muszą mieć wartości bliskie zeru. W przypadku występowania ujemnej autokorelacji rzędu l , wartości $v_{t,t-1}$ będą należeć do największych spośród wszystkich możliwych wartości bezwzględnych różnic $|z_t - z_k|$, $t, k \in \{1, 2, \dots, n\}$ i $t \neq k$. Należy zatem zdefiniować test wykrywający tę prawidłowość. Załóżmy, że jest weryfikowana hipoteza $H_0: \rho_1 = 0$ przeciw hipotezie $H_1: \rho_1 > 0$. Niech dany będzie zbiór $J = \{|z_t - z_k|, t, k \in \{1, 2, \dots, n\} \text{ i } t \neq k\}$ wszystkich możliwych różnic między dowolnymi, ale różnymi elementami danego szeregu. Moc zbioru J wynosi

$N := |J| = \binom{n}{2}$. Podzielmy zbiór J na c podzbiorów J_1, J_2, \dots, J_c takich, że w J_1 znajduje się $\frac{N}{c}$ najmniejszych elementów zbioru J , w J_2 kolejne $\frac{N}{c}$ większych elementów zbioru J , itd., aż w J_c znajduje się $\frac{N}{c}$ największych elemen-

tów wyjściowego zbioru J . (Wartość c jest wybrana arbitralnie, autor testu nie podaje procedury wyznaczania c). Liczba c podzbiorów zbioru J musi być odpowiednia do liczby elementów próby. Proponuje się, aby dla $n < 8$, $c = 2$, a dla nieco większych prób $c = 3$ lub $c = 4$. Następnie niech dany będzie zbiór $L_l := \{|z_t - z_{t-1}|, t = l + 1, l + 2, \dots, n\}$. Liczba elementów tego zbioru wynosi $n - l$. Rozważane są zbiory $M_{(l),i} := L_l \cap J_i$, $i = 1, 2, \dots, c$ elementów odległych o l okresów i należących do zbioru o określonych granicach wartości definiujących różnicę wartości elementów wyjściowego szeregu. Oczywiście zachodzą następujące związki:

$$\sum_{i=1}^c |M_{(l),i}| = n - l$$

oraz:

$$0 \leq |M_{(l),i}| \leq \frac{N}{c}, \quad i=1, 2, \dots, c$$

Rozważa się wektor losowy w postaci $(|M_{(l),1}|, |M_{(l),2}|, \dots, |M_{(l),c}|)$ oraz prawdopodobieństwo, że $|M_{(l),i}| = m_i$ dla każdego $i=1, 2, \dots, c$ ma rozkład hipergeometryczny wyrażający się wzorem:

$$P\{|M_{(l),1}| = m_1, |M_{(l),2}| = m_2, \dots, |M_{(l),c}| = m_c\} = \frac{\binom{M}{m_1} \binom{M}{m_2} \dots \binom{M}{m_c}}{\binom{N}{n-1}} \quad (2)$$

przy czym:

$$N = \binom{n}{2}$$

oraz:

$$M = \frac{N}{c}$$

Jeżeli hipoteza H_1 jest prawdziwa, to zdecydowana większość elementów zbioru L_l znajduje się w początkowych elementach ciągu zbiorów J_1, J_2, \dots, J_c . Za prawdziwością hipotezy H_1 będą przemawiać względnie duże wartości początkowych elementów wektora $(|M_{(l),1}|, |M_{(l),2}|, \dots, |M_{(l),c}|)$ w stosunku do wartości „środkowych” i „końcowych”. Natomiast w przypadku zajścia hipo-

tezy H_0 , współrzędne tegoż wektora powinny mieć zbliżone wartości, co oznacza, że odległe o l jednostek czasu wartości obserwowanego szeregu nie charakteryzują się niską korelacją. Statystyka testowa przyjmuje postać, która zależy od przyjętej wartości c . Niech m_1, m_2, \dots, m_c oznaczają zaobserwowane wartości wektora $(|M_{(l),1}|, |M_{(l),2}|, \dots, |M_{(l),c}|)$. Dla $c = 2$ rozważa się prawdopodobieństwo:

$$P\{|M_{(l),1}| \geq m_1\} = \sum_{x=m_1}^{n-l} \frac{\binom{M}{x} \binom{M}{n-l-x}}{\binom{N}{n-l}} \quad (3)$$

Przy ustalonym poziomie istotności α , jeżeli $P\{|M_{(l),1}| \geq m_1\} \leq \alpha$, to znaczy, że zaszło zdarzenie, którego prawdopodobieństwo jest względnie małe, odrzucamy hipotezę H_0 na korzyść hipotezy H_1 mówiącej, że wystąpiła autokorelacja rzędu l . W przeciwnym razie, jeżeli $P\{|M_{(l),1}| \geq m_1\} > \alpha$, nie ma podstaw do odrzucenia hipotezy H_0 .

Podobnie poszukiwane jest prawdopodobieństwo dla $c = 3$:

$$P\{|M_{(l),1}| \geq m_1 \wedge |M_{(l),2}| = m_2\} = \sum_{x=m_1}^{n-l-m_2} \frac{\binom{M}{x} \binom{M}{m_2} \binom{M}{n-l-x-m_2}}{\binom{N}{n-l}} \quad (4)$$

Różnica między wzorem (3) i (4) wynika z postulatu utrzymania na stałym poziomie liczby obojętnych elementów $v_{t,k}$, należących do zbioru J_2 . Hipotezę H_0 przyjmuje się bądź odrzuca przy zadanym poziomie istotności w sposób analogiczny jak dla $c = 2$.

W przypadku testowania hipotezy $H_0: \rho_l = 0$ przeciwko hipotezie $H_1: \rho_l < 0$, postaci testowanych wyrażeń się zmieniają, ponieważ wystąpienie ujemnej autokorelacji między zmiennymi będzie się charakteryzowało tym, że większość elementów zbioru L_l będzie znajdować się w $M_{(l),c}$, ewentualnie $M_{(l),c-1}$. Dla $c = 2$ prowadzi to do rozpatrywania wartości wyrażenia:

$$P\{|M_{(l),2}| \geq m_2\} = \sum_{x=m_2}^{n-l} \frac{\binom{M}{n-l-x} \binom{M}{x}}{\binom{N}{n-l}} \quad (5)$$

natomiast dla $c = 3$ rozważa się analogicznie wartość wyrażenia:

$$P\{|M_{(l),3}| \geq m_3 \wedge |M_{(l),2}| = m_2\} = \sum_{x=m_3}^{n-l-m_2} \frac{\binom{M}{n-l-x-m_2} \binom{M}{m_2} \binom{M}{x}}{\binom{N}{n-l}} \quad (6)$$

Hipotezę H_0 odrzucamy na korzyść hipotezy alternatywnej H_1 , gdy otrzymana wartość prawdopodobieństwa nie jest większa od założonego poziomu istotności. W przeciwnym razie nie ma podstaw do odrzucenia hipotezy H_0 .

Jeżeli $\frac{N}{c}$ nie jest liczbą całkowitą, to za licznosc zbiorów J_i należy przyjmować $\left\lfloor \frac{N}{c} \right\rfloor$. Jeżeli przez Q oznaczy się liczbę „nadwyżkowych” elementów $v_{t,k}$, to elementy te zostaną zaliczone do jednego ze zbiorów J_i . Jeżeli będzie to zbiór J_c , to funkcja rozkładu prawdopodobieństwa rozkładu wektora $(|M_{(l),1}|, |M_{(l),2}|, \dots, |M_{(l),c}|)$ przyjmie następującą postać:

$$P\{M_{(l),1} = m_1 \wedge M_{(l),2} = m_2 \wedge \dots \wedge M_{(l),c} = m_c\} = \frac{\binom{M}{m_1} \binom{M}{m_2} \dots \binom{M+Q}{m_c}}{\binom{N}{n-l}} \quad (7)$$

2.2. Metoda wyznaczania rzędu opóźnień i zmian struktury w liniowym modelu autoregresji

Do opisu dynamicznych własności zmiennej ekonomicznej jest często stosowany model autoregresji:

$$y_{t+p} = a_1 y_{t+p-1} + a_2 y_{t+p-2} + \dots + a_p y_t + c + \varepsilon_{t+p}, \quad t = 0, 1, 2, \dots, N-p \quad (8)$$

w którym zakłada się, że ε_{t+p} , $t = 0, 1, \dots, N-p$ są niezależnymi zmiennymi. Ponadto a_1, a_2, \dots, a_p, c są z założenia stałymi nieznanymi współczynnikami. Zakłada się także, w celu umożliwienia obliczeń, że obserwowany szereg ma co najmniej $N = 2p+1$ elementów.

Przyjmujemy następujące oznaczenia:

$$y^t = \begin{bmatrix} y_{t+p} \\ y_{t+p+1} \\ \vdots \\ y_{t+2p} \end{bmatrix}_{p+1 \times 1}, \quad Y^t = \begin{bmatrix} y_{t+p-1} & y_{t+p-2} & \dots & y_t & 1 \\ y_{t+p} & y_{t+p-1} & \dots & y_{t+1} & 1 \\ \dots & \dots & \ddots & \vdots & \vdots \\ \dots & \dots & & \vdots & \vdots \\ y_{t+2p-1} & y_{t+2p-2} & \dots & y_{t+p} & 1 \end{bmatrix}_{p+1 \times p+1} \quad (9)$$

$$a^t = \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_p^t \\ c^t \end{bmatrix}_{p+1 \times 1}, \quad \varepsilon^t = \begin{bmatrix} \varepsilon_{t+p} \\ \varepsilon_{t+p+1} \\ \vdots \\ \varepsilon_{t+2p} \end{bmatrix}_{p+1 \times 1}$$

przy czym elementy y_{t+k} poszczególnych macierzy spełniają równanie autoregresji (8).

W celu wyznaczenia parametrów równania różnicowego odpowiadającego równaniu autoregresji rozważymy ciąg układów równań:

$$\begin{cases} y_{t+p} = a_1^t y_{t+p-1} + a_2^t y_{t+p-2} + \dots + a_p^t y_t + c^t \\ y_{t+p+1} = a_1^t y_{t+p} + a_2^t y_{t+p-1} + \dots + a_p^t y_{t+1} + c^t \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ y_{t+2p} = a_1^t y_{t+2p-1} + a_2^t y_{t+2p-2} + \dots + a_p^t y_{t+p} + c^t \end{cases}, \quad t = 0, 1, 2, \dots, N-2p$$

(10)

Układ taki można zapisać w postaci macierzowej za pomocą przyjętych oznaczeń:

$$y^t = Y^t a^t, \quad t = 0, 1, 2, \dots, N-2p \quad (11)$$

Warto zauważyć, że zadany układ równań ma tyle równań, ile niewiadomych współczynników ma równanie różnicowe rzędu p , tzn. $p+1$ równań. Równania tego układu nie zawierają składnika losowego, w przeciwieństwie do równania wyjściowego (8). W opracowaniu rozpatruje się przypadek deterministycznego równania różnicowego. Oczywiście składnik losowy istotnie wpływa na postać wyniku oraz ma szczególne znaczenie zwłaszcza dla niestacjonarnych szeregów czasowych.

Niech Y^t będzie macierzą kwadratową stopnia $p+1$ (zdefiniowaną jak w (9)), odpowiadającą liniowemu równaniu różnicowemu rzędu p postaci (8). Zachodzą wówczas następujące fakty (Kwiatkowska, 2002b):

Twierdzenie 1

$\det Y^t = 0$ wtedy i tylko wtedy, gdy rząd równania różnicowego generującego obserwowany szereg czasowy ($y_t: t=0, 1, \dots, N$) jest mniejszy od p .

Twierdzenie 2

Rząd równania (8) wynosi p ($p < r$) wtedy i tylko wtedy, gdy $\text{rz } Y^t = p+1$.

Twierdzenie 3

Założmy, że rząd równania generującego szereg czasowy ($y_t: t = 0, 1, 2, \dots, N$) wynosi p . Rozważmy ciąg układów równań w postaci (10). Wówczas dla każdego $t = 0, 1, 2, \dots, N-p$:

$$y_{t+p} = a_1 y_{t+p-1} + a_2 y_{t+p-2} + \dots + a_p y_t + c$$

wtedy i tylko wtedy, gdy dla każdego $t = 0, 1, 2, \dots, N-2p-1$:

$$y^t = Y^t a \quad (12)$$

gdzie:

$$a = [a_1, a_2, \dots, a_p, c]^T = \text{const.}$$

(tzn. dla każdego $t = 0, 1, 2, \dots, N-2p-1$ rozwiązanie układu (10) jest takie samo). Łatwo wykazać, że otrzymane oceny to estymatory MNK.

Biorąc za podstawę przytoczone twierdzenia (przy założeniu liniowości modelu), można sformułować następujący algorytm pozwalający wyznaczyć rząd opóźnień w modelu $AR(p)$ oraz ewentualne zmiany wartości parametrów.

Algorytm:

1. Dla kolejnych $p = 0, 1, \dots, p_0$, gdzie p_0 jest dowolnie ustalone, rozważa się ciągi układów równań w postaci (10). Jest ich $N-p-1$.
2. Wyznacza się oceny parametrów i bada się stałość ciągów tych ocen.
3. Na ich podstawie określa się przedziały czasowe, w których oceny były stałe dla ustalonego p . Średnia arytmetyczna ocen parametrów z przedziału ich stałości stanowi ocenę parametru dla tegoż przedziału czasowego.
4. Na podstawie uzyskanych wyników określa się rzędy opóźnień zmiennej endogenicznej w poszczególnych przedziałach oraz punkty zmiany rzędów opóźnień lub szerzej – zmiany wartości parametrów.

Powyższa metoda ma mankamenty i zawiera kilka niedociągnięć nadających się do uzupełnienia. Po pierwsze, oceny parametrów dla kolejnych podciągów danych są wyznaczane MNK. Ponieważ jest to model autoregresji, otrzymane estymatory są obciążone (por. Milo, 1990). W literaturze mówi się często o zgodności estymatorów jako o pożądanej ich właściwości. W rozważanym przypadku nie ma sensu badać zgodności, gdyż kolejne oceny parametrów uzyskuje się z wykorzystaniem możliwie najmniejszej liczby danych. Po drugie, nie zaproponowano metody weryfikacji hipotezy o stałości parametrów. Temat jest problematyczny ze względu na własności estymatorów i predyktorów estymatorów MNK zastosowanych do równań autoregresji. Proponuje się weryfikację postawionej hipotezy polegającą na analizie wykresów kolejnych ocen parametrów, a następnie porównaniu dopasowania wartości teoretycznych do danych empirycznych.

Wydaje się, że metoda ma również zalety. Wiadomo, że na ogół nie można zakładać stałości parametrów modelu. Dlatego duża „część” obciążenia estymatorów może wynikać z niedopasowania modelu ze stałymi parametrami do danego szeregu czasowego, skąd otrzymuje się składnik losowy modelu o dużej wariancji warunkującej wartość obciążenia estymatora. Przy założeniu, że dobrze wyspecyfikowany model ma składnik losowy o względnie małej wariancji, wyznaczone równania generujące dla podszeregów, dla których zaobserwowano stałe ciągi ocen parametrów, są równaniami „prawie” deterministycznymi. Wówczas obciążenie estymatorów jest względnie małe. W ten sposób zbliżamy się do pewnych praw rządzących przebiegiem danego szeregu. Metoda może być wykorzystana jednocześnie do:

- 1) wyznaczania rzędu opóźnień,
- 2) określania punktów zwrotnych,
- 3) estymacji parametrów.

Pozytywną cechą metody jest również możliwość jej zastosowania w próbach o niewielkiej liczbie obserwacji, co jest zaletą w badaniach makroekonomicznych. Do przeprowadzenia obliczeń potrzebna jest oczywiście pewna minimalna liczba obserwacji. Badania empiryczne były prowadzone dla rzeczywistych szeregów, które posiadały trendy, a więc nie były stacjonarne. Wyniki pokazały, że uzyskiwane modele były na ogół dobrze dopasowane. W porównaniu z metodą Yule’a-Walkera wyniki dopasowania szeregu teoretycznego do danych są lepsze (por. Kwiatkowska, 2002a). Może to być spowodowane tym, że pomiędzy punktami zwrotnymi wpływ składnika losowego jest niewielki.

Dalsze badania nad modelami autoregresji ze zmiennymi parametrami mogą prowadzić do pewnej klasy modeli nieliniowych.

3. Wyniki badań

3.1. Dane statystyczne

W badaniach rozważono jedną z podstawowych zmiennych modelu pętli inflacyjnej opisanego w rozdziale 1. Jest nią wskaźnik cen towarów i usług konsumpcyjnych. Uwzględniono przedział czasowy 1993(3)-2002(2) zawierający dane kwartalne. Dane pochodzące ze źródeł (GUS) były łańcuchami o stałej podstawie, przy czym podstawa ta w uwzględnionym przedziale czasu zmieniała się kilkakrotnie. Ujednolicono podstawę przyjmując za okres podstawowy dane z 1 kwietnia 1990 r., następnie wyznaczono szereg łańcuchowy, po czym

obliczono logarytmy wartości szeregu łańcuchowego, uzyskując w ten sposób przyrosty logarytmów badanej zmiennej, czyli wielkość przedstawiającą dynamikę.

3.2. Rząd autokorelacji składnika losowego w modelach autoregresji cen i płac

Celem rozdziału jest przedstawienie wyników określenia występowania autokorelacji składnika losowego w modelu autoregresji cen (P). Rozważany jest ogólny model w postaci (8). Dla wyznaczenia reszt równań autoregresji ustalono najpierw (stosując metodę przedstawioną w rozdziale 2.2) punkty zmiany struktury równania generującego dla omawianego szeregu czasowego. Określono, tam gdzie było to możliwe, rząd autoregresji oraz obliczono oceny parametrów. Dla wyznaczonych w ten sposób szeregów teoretycznych obliczono reszty obrazujące odchylenie modeli od przebiegu szeregu czasowego. Z kolei dla szeregu reszt zastosowano taką samą procedurę, jak dla oryginalnych szeregów danych. Założono mianowicie, że reszty ε_t , zdefiniowane jako:

$$\varepsilon_t = y_t - \bar{y}_t \quad (13)$$

spełniają zależność wyrażającą się równaniem:

$$\varepsilon_{t+p} = a_1 \varepsilon_{t+p-1} + a_2 \varepsilon_{t+p-2} + \dots + a_p \varepsilon_t + c + \eta_{t+p}, \quad t=k, k+1, k+2, \dots, K-p \quad (14)$$

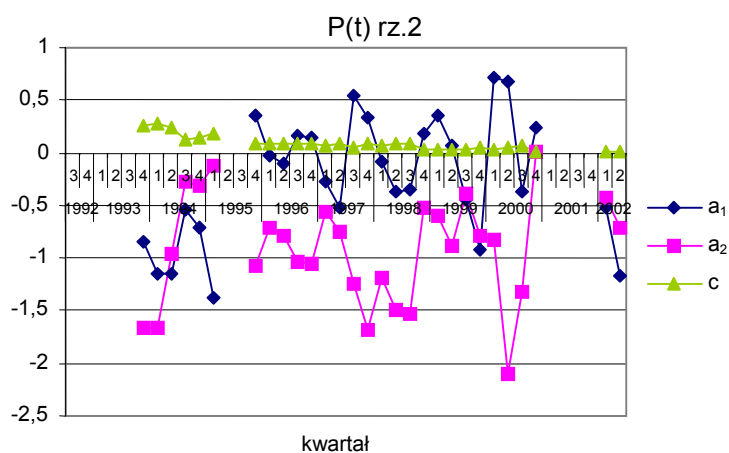
gdzie η_{t+p} jest składnikiem losowym. Zastosowano procedurę wyznaczania rzędu (i ewentualnie punktów zwrotnych) oraz wyznaczania ocen parametrów taką samą, jak w przypadku przekształconego szeregu dla zmiennej P .

Stałość ciągu ocen parametrów dla zadanego p świadczy o tym, że rząd autoregresji składnika losowego wynosi p , co pociąga za sobą autokorelację rzędu p składnika losowego. Oczywiście liniowość modelu jest tylko przybliżeniem związków, jakie mogą zachodzić w rzeczywistości. Niestety w trakcie wykonywania obliczeń pojawił się problem małej próby, co uniemożliwiło wiarygodną ocenę związków autokorelacji rzędów wyższych niż 2. Już nawet dla rzędu 2 autokorelacji, długość ciągu ocen wektora parametrów wynosiła w niektórych przypadkach 2 albo 3. Otrzymane wyniki porównano z wynikami testu J Pawłowskiego na autokorelację wyższych rzędów. Wszystkie wyniki przedstawiono w tabeli zbiorczej.

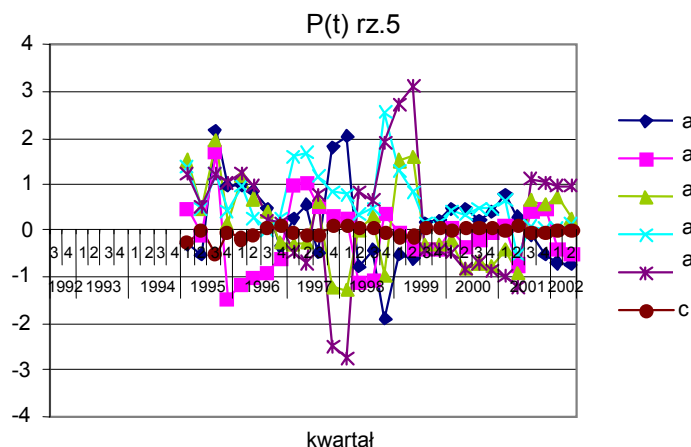
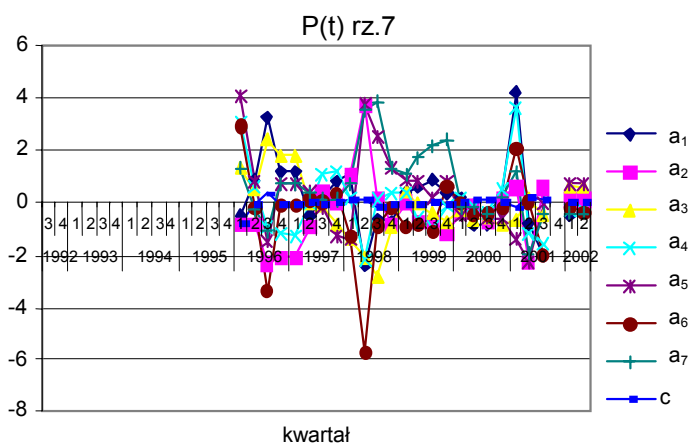
Wykorzystano kilka dodatkowych parametrów, które w sposób opisowy mogą określać zachodzenie związków autokorelacyjnych i mieć wpływ na wnioskowanie dotyczące autokorelacji. Wielkości te zostaną przedstawione w dalszej części opracowania.

3.3. Wybór modeli autoregresji

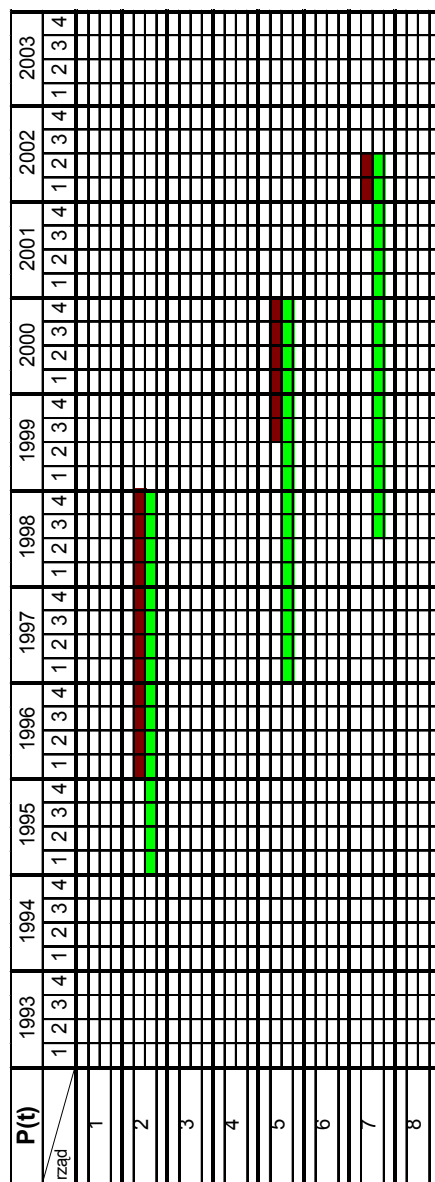
Rozważono szereg czasowy będący szeregiem różnic logarytmów naturalnych kolejnych wartości szeregu inflacji, który z kolei był szeregiem czasowym o stałej podstawie z 1 kwietnia 1990 r. W wyniku zastosowania algorytmu podanego w rozdziale 2.2 dla $p = 1, 2, \dots, 8$ otrzymano ciągi wektorów ocen parametrów modelu typu (14). Wybrano te z nich, które co najmniej przedziałami miały stosunkowo małe wahania. Na rys. 1, 2 i 3 przedstawiono ilustrację graficzną wybranych ciągów ocen parametrów.



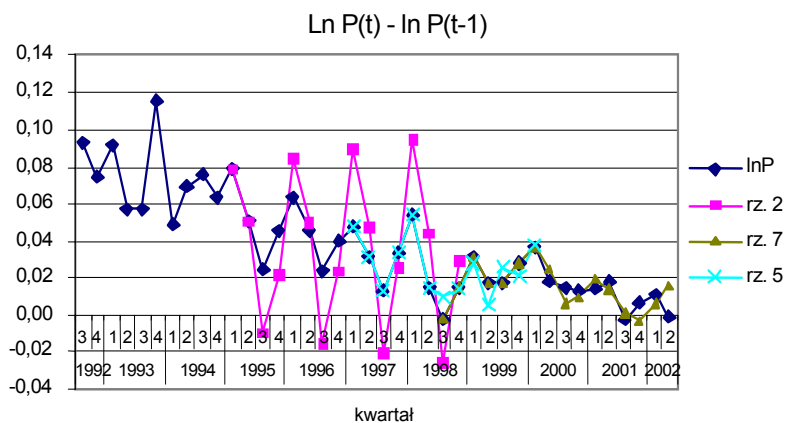
Rys. 1. Oceny parametrów równania autoregresji rzędu 2 dla $P(t)$

Rys. 2. Oceny parametrów równania autoregresji rzędu 5 dla $P(t)$ Rys. 3. Oceny parametrów równania autoregresji rzędu 7 dla $P(t)$

Dla modeli pozostałych rzędów oceny parametrów były dużo bardziej niestabilne, dlatego zostały pominięte w rozważaniach. Wybrano modele rzędu 2, 5 i 7. Na rys. 4 przedstawiono wybrane podokresy stosunkowo stałych podciągów parametrów (zaznaczone ciemniejszym odcieniem szarości) i okresu podciągów szeregu czasowego, dla których je wyznaczono (zaznaczone jaśniejszym odcieniem szarości).

Rys. 4. Okresy stałości struktury równań generujących dla wybranych rządów opóźnień $p - P(t)$

Z rys. 4 można odczytać, że poszczególne modele odpowiadają podszeregom mającym części wspólne. Na rys. 5 przedstawiono przebieg szeregów teoretycznych i szeregu danych.



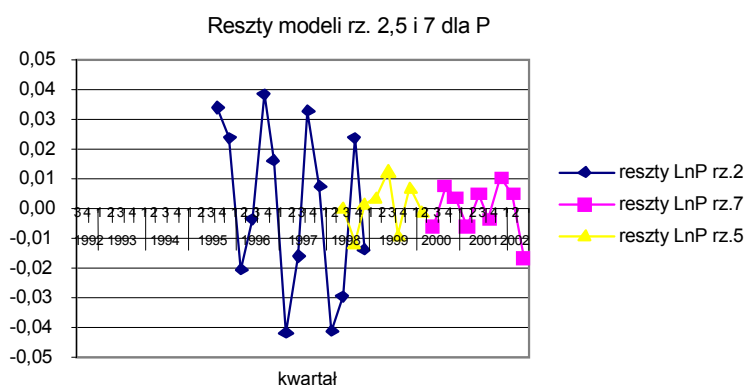
Rys. 5. Szeregi teoretyczne i szereg obserwacji P

Na podstawie rys. 5 można stwierdzić słabe dopasowanie do danych modelu równania rzędu 2, obejmującego przedział czasowy od 1995(1) do 1998(4). Nie należy jednak zbyt pochopnie odrzucać takiego modelu. W kolejnym kroku oszacujemy rząd autoregresji reszt, co pozwoli stwierdzić, czy w przebiegu odchyień szeregu teoretycznego od szeregu obserwacji występuje regularność. Drugi krok procedury ma za zadanie określić rząd autokorelacji składnika losowego.

Na podstawie względnie stałych szeregów ocen parametrów wyznaczono ich średnie arytmetyczne, jako oszacowania parametrów. Tabela 1 przedstawia wybrane szeregi ocen parametrów dla poszczególnych rzędów autoregresji. Na ciemniejszych polach podano oszacowania parametrów w postaci średnich arytmetycznych względnie stałych ciągów ocen znajdujących się powyżej średnich.

Rysunek 6 przedstawia reszty modeli wyróżnionych dla zmiennej P .

W wyniku zastosowania tej samej procedury, jak dla szeregów danych dotyczących zmiennej P , dla $p = 1,2$ wyznaczono ciągi ocen parametrów dla trzech wymienionych powyżej szeregów reszt. Wyniki przedstawiono poniżej. Rysunki 7a i 7b przedstawiają przebiegi szeregów ocen parametrów autoregresji rzędu 1 i 2 reszt modelu rzędu 2 dla okresu 1998(1)-1998(4).



Rys. 6. Przebieg reszt modeli autoregresji dla szeregu cen P

Łatwo zauważyć, że szereg ocen parametru a_1 dla modelu rzędu 1 jest niestabilny. Przemawia to za brakiem autoregresji rzędu 1 szeregu reszt. Jednakże średnia wartość ocen parametrów jest większa od zera, co może sugerować dodatnią korelację między resztami dla następujących po sobie okresów. Na rys. 7b oceny parametrów odpowiadających poszczególnym opóźnieniom charakteryzują się większą stabilnością niż dla modelu rzędu 1 (skala o mniejszej podziałce i mniejsze wychylenia ocen parametrów).

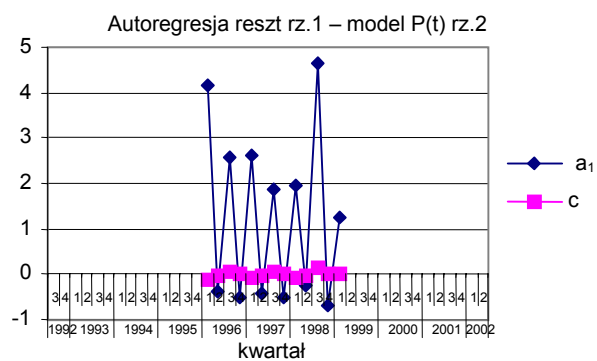
Na rys. 7b oceny parametru a_2 oscylują wokół wartości -1 . Wynik ten sugeruje, że występuje ujemna korelacja rzędu 2 w rozważanym szeregu reszt. Oceny parametru a_1 wahają się wokół zera. Cechą charakterystyczną i oczekiwaną jest w obydwu przypadkach zerowa wartość ocen wyrazu wolnego c . Wyklucza to dryf szeregu reszt.

W wyniku szacowania kolejnych wartości parametrów dla reszt modelu rzędu 5 otrzymano przebiegi ocen przedstawione na rys. 8a i 8b.

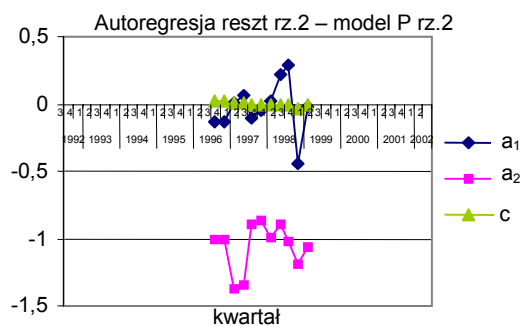
Żaden z szeregu ocen parametrów nie wskazuje na stałość równania liniowego wiążącego zależności między resztami modelu rzędu 5. Szeregi ocen odpowiadające wyrazowi wolnemu c modelu równania (14) są stałe, tzn. równe zero.

Dla modelu rzędu 7 otrzymano następujące przebiegi ocen parametrów dla modeli rzędu 1 i 2.

a)

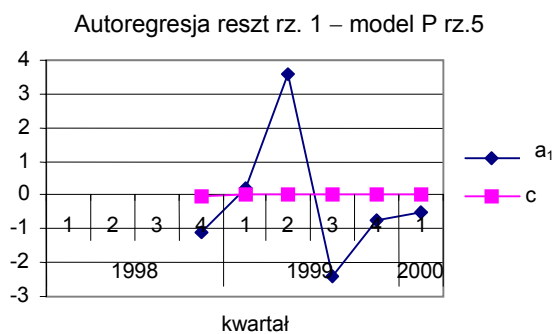


b)

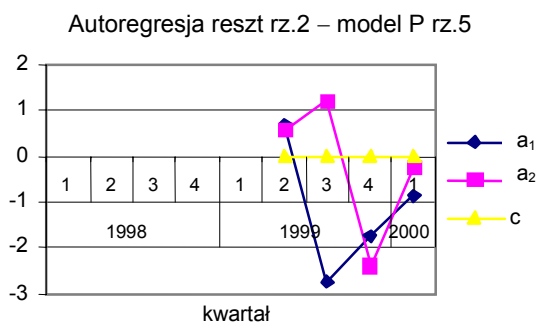


Rys. 7. Przebieg ocen parametrów równań autoregresji rzędu 1 i 2 dla reszt modelu rz.2 – P

a)

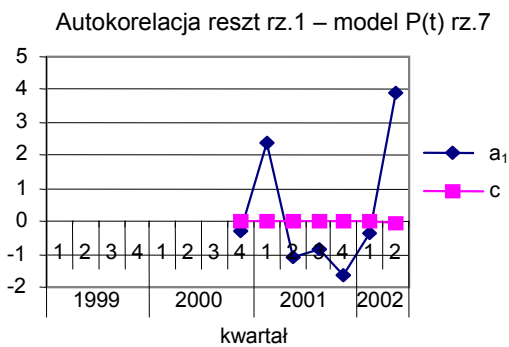


b)

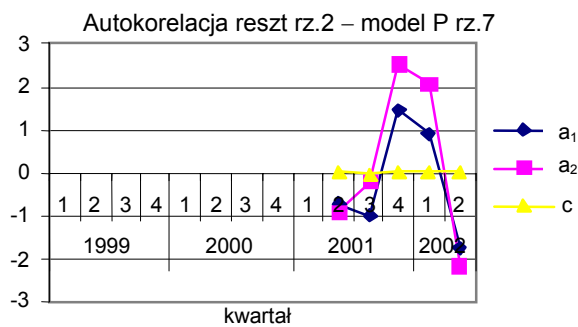


Rys. 8. Przebieg ocen parametrów równań autoregresji rzędu 1 i 2 dla reszt modelu rz.5 – P

a)



b)



Rys. 9. Przebieg ocen parametrów równań autoregresji rzędu 1 i 2 dla reszt modelu rz.7 – P

Podobnie jak w przypadku modelu rzędu 5 reszt, nie można stwierdzić stałości ocen parametrów. W związku z tym nie ma podstaw, aby twierdzić, że jest zależność autoregresyjna między resztami modelu rzędu 7 przekształconego wskaźnika cen.

W tab. 2 przedstawiono wyniki testu J na autokorelację wszystkich rzędów zastosowaną do reszt wyznaczonych poprzednich modeli. Poniżej omówiono poszczególne kolumny tabeli.

Pierwsze cztery kolumny dotyczą określonych powyżej modeli autoregresji oszacowanych dla różnic logarytmów szeregu cen. Kolumny zatytułowane „miary dopasowania” dotyczą wartości następujących wielkości:

$$D_B = \frac{\sqrt{\sum_{t=t_0+p-1}^{t=t_0+k-1} (y_t - \hat{y}_{pt})^2}}{p+k} \quad (15)$$

oraz:

$$D_W = \frac{\sqrt{\sum_{t=t_0+p-1}^{t=t_0+k-1} \left(\frac{2(y_t - \hat{y}_{pt})}{|y_{t+1} - y_t| + |y_t - y_{t-1}|} \right)^2}}{p+k} \quad (16)$$

Wzór (15) określa średnią bezwzględną odległość szeregu teoretycznego od szeregu obserwacji. D_W jest to średnia wartość błędu w stosunku do średniej arytmetycznej wartości bezwzględnych przyrostów szeregu obserwacji bezpośrednio przed i za t -tą obserwacją, gdzie $(\hat{y}_t : t_0 \leq t \leq t_0 + k + 2p)$ będzie ciągiem teoretycznym, przy założeniu, że p początkowych wyrazów szeregu teoretycznego jest równe odpowiednim wyrazom szeregu obserwowanego. Oczywiście, im mniejsze wartości obydwu wielkości, tym lepiej dopasowany jest szereg teoretyczny do szeregu danych (Kwiatkowska, 2002c).

Następny blok tabeli zawiera dane i wyniki dotyczące szacunków na podstawie chwilowych ocen parametrów równania autoregresji. Wyróżniono kolumnę „przedział czasowy”, ponieważ może się zdarzyć, że trzeba będzie uwzględnić węższy przedział, ze stałym szeregiem ocen, niż zadany w pierwszej kolumnie tablicy.

Kolumna zatytułowana „ a_p ” zawiera parametr stojący przy opóźnieniu o numerze p , oszacowany jako średnia arytmetyczna ocen parametrów odpowiedniego modelu. Jego wartość bezwzględna sugeruje siłę zależności liniowej, a zatem i korelacji, między resztami odległymi o p jednostek czasu. Znak a_p wskazuje, czy autokorelacja rzędu p jest dodatnia, czy też ujemna.

Podstawą wniosku jest stałość ciągu ocen parametrów. Stąd do tabeli wprowadzono wielkość zatytułowaną „Wahania ocen a_p ”. Oblicza się ją według wzoru:

$$W = \frac{\sum_{t=t_0+p}^{t_0+p+k-1} |a_{p,t+1} - a_{p,t}|}{k} \quad (17)$$

gdzie t_0 oznacza numer pierwszego elementu szeregu reszt, dla którego obliczono średnią a_p , a k jest liczbą ocen parametrów uwzględnioną przy obliczaniu wartości średniej a_p .

Kolumna nazwana „Znak oceny wyst. częściej” może przyjąć następujące oznaczenia:

- „+” – w ciągu ocen parametrów wystąpiło więcej wartości dodatnich,
- „-” – w ciągu ocen parametrów wystąpiło więcej wartości ujemnych,
- „-/+” – w ciągu ocen parametrów wystąpiło tyle samo wartości ujemnych, co dodatnich.

W ostatniej kolumnie tego bloku określono, czy autokorelacja może być ujemna czy dodatnia. W celu wyznaczenia tego znaku posłużono się algorytmem:

- jeżeli „-/+”, to $\text{sgn}(a_p)$
- jeżeli „-”, i $\text{sgn}(a_p) = -1$, to -
- jeżeli „+” i $\text{sgn}(a_p) = 1$, to +
- jeżeli „-”, i $\text{sgn}(a_p) = 1$, to +
- jeżeli „+” i $\text{sgn}(a_p) = -1$, to -

Ostatnie dwie implikacje wynikają z założenia, że znak wartości średniej jest ważniejszy od większej liczby obserwacji o znaku przeciwnym niż znak średniej. Takie postępowanie może spotkać się z krytyką. Mianowicie, jeżeli wszystkie oceny, oprócz jednej, mają wartość ujemną, a wartość bezwzględna oceny o znaku różnym od znaku większości powoduje, że średnia nie przyjmuje znaku ujemnego, to postępując, jak opisano wyżej, pierwszeństwo daje się ob-

serwacji nietypowej. W dalszych badaniach podejmie się pracę nad udoskonaleniem algorytmu, co będzie wymagało nałożenia dodatkowych warunków i rozpatrzenia większej liczby przypadków.

Trzeci blok tabeli odnosi się do testu J na autokorelację. Oznaczenia kolumn są analogiczne do definicji odpowiednich wielkości w rozdziale 2.1, gdzie:

n – liczba obserwacji (tu: reszt modelu),

$N = \binom{n}{2}$ – liczba możliwych do utworzenia różnic z elementów szeregu obserwacji,

c – liczba bloków, na które jest podzielony porządkowany ciąg różnic,

M – długości bloków,

Q – pozostała liczba obserwacji, niezaliczona do żadnego bloku ($Q < \left\lceil \frac{N}{c} \right\rceil$),

p – badany rząd autokorelacji,

J – wartość statystyki testowej,

H_1 – postać hipotezy alternatywnej:

H_1 : „+” – występuje autokorelacja dodatnia,

H_1 : „-” – występuje autokorelacja ujemna,

podczas gdy hipoteza zerowa ma zawsze postać: H_0 : brak autokorelacji. Ostatnia kolumna wskazuje na wybór hipotezy.

Przeanalizujemy tablicę wierszami ze zwróceniem uwagi na test J . Za poziom istotności przyjęto $\alpha = 0,05$.

Dla modelu rzędu 2 w okresie 1995(1)-1998(4) istnieje ujemna autokorelacja składnika losowego rzędu 2, lecz jednocześnie nie istnieje autokorelacja rzędu 1.

Dla modelu rzędu 5 w okresie 1997(1)-2000(4) istnieje ujemna autokorelacja składnika losowego rzędu 1, natomiast nie istnieje autokorelacja rzędu 2.

Dla modelu rzędu 7 w okresie 1998(3)-2002(2) istnieje dodatnia autokorelacja składnika losowego rzędu 2 i obserwuje się brak autokorelacji rzędu 1.

Tabela 2

Porównanie wyników testu J na autokorelację składnika losowego z wynikami rzędu autoregresji składnika losowego, jako miary autokorelacji – modele zm. P-ceny

Model autoregresji dla LnPt-LnPt-1				Rząd autoregresji składników losowych												alfa = 0,05				
przedział czasu	rząd autoregresji	miary dopasowania		metoda chwilowych ocen parametrów – szeregi reszt										test J						wybór H0 albo H1
		DW	DB	przedział czasu	rząd autoregresji reszt p	ap	wahania ocen ap	znak oceny wyst. części	znak auto-korelacji	n	N	c	M	Q	p	statystyka J	H1			
95(1)-98(4)	2	0,4051	0,0072	95(1)-98(4)	1	1,1343	3,3107	„+”	+	14	91	3	30	1	1	0,0810	„+”	H0		
				95(1)-98(4)	2	-1,1057	0,1567	„+”	-	14	91	3	30	1	2	0,202*(e(-4))	„+”	H1		
97(1)-00(4)	5	0,3219	0,0064	97(1)-00(4)	1	-0,1658	2,5055	„+”	-	8	28	3	9	1	1	0,0327	„+”	H1		
				97(1)-00(4)	2	-0,2223	2,1459	„+”	-	8	28	3	9	1	2	0,1196	„+”	H0		
98(3)-02(2)	7	0,6930	0,0090	98(3)-02(2)	1	-0,0285	2,1430	„+”	-	9	36	3	12	0	1	0,2411	„+”	H0		
				98(3)-02(2)	2	0,1631	2,0528	„+”	+	9	36	3	12	0	2	0,0354	„+”	H1		

Dokładnie takie same wyniki, w przypadku wykrycia autokorelacji przez test J , daje procedura przeprowadzona na podstawie ciągów chwilowych ocen parametrów, przy przedstawionych wcześniej założeniach. Ponadto tam, gdzie nie ma podstaw do odrzucenia hipotezy zerowej przez test J , wyniki konkurencyjnej procedury mogą wskazywać na brak autokorelacji. Istotnie, w przypadku autoregresji składnika losowego rzędu 1 dla modelu rzędu 2 istnieje duże wahanie kolejnych ocen parametru a_p , a liczba ocen o znaku ujemnym jest równa liczbie ocen o znaku dodatnim. Podobnie jest w przypadku autokorelacji rzędu 2 dla reszt modelu rzędu 5. Poza tym przy badaniu autokorelacji reszt rzędu 1 w modelu rzędu 7 występuje duże wahanie ocen, które dają średnią arytmetyczną bliską zeru, co odpowiada brakowi związku między kolejnymi resztami.

Należy dodać, że dopasowanie szeregów teoretycznych do danych jest dość dobre biorąc pod uwagę to, że odległość szeregów teoretycznych od obserwacji waha się od 0,3 do 0,7, średnio rzecz biorąc, wartości różnicy między dwoma kolejnymi „skokami” obserwacji.

Podsumowanie

W opracowaniu podjęto próbę scharakteryzowania modelu autoregresji cen przyjmując, że składa się na niego równanie w postaci zredukowanej. Po wyznaczeniu potencjalnych modeli generujących szeregi, podjęto próbę określenia związków korelacyjnych między resztami modelu. W tym celu zastosowano dwie metody – test J Pawłowskiego i metodę wykorzystującą oceny parametrów wyznaczone dla określonej długości okien szeregu zmiennych. W niemal wszystkich przypadkach uzyskano porównywalne wyniki. Oprócz tego zaobserwowano, że w przypadku dobrego dopasowania modelu autoregresji do danych, występuje brak autokorelacji składnika losowego. Może to być wynikiem małej wariancji reszt. Model cechujący się małą wariancją reszt jest zbliżony do modelu deterministycznego. Należy tu dodać, że właśnie wariancja reszt modelu jest przyczyną obciążenia estymatorów MNK, które zostały zastosowane do wyznaczania ocen parametrów poszczególnych modeli. Stąd można sądzić, że dużego obciążenia estymatorów można uniknąć, stosując procedury wykorzystujące jak najmniejszą liczbę danych.

Literatura

- Box G.E.P., Jenkins G.M.: *Analiza szeregów czasowych*. PWN, Warszawa 1983.
- Fuller W.A.: *Introduction to Statistical Time Series*. John Willey & Sons, INC, New York.
- Gospodarka Polski w okresie transformacji*. Red. A. Welfe. PWE, Warszawa 2000.
- Kwartalnik Statystyki Międzynarodowej*. GUS, Warszawa 1994-2003, nr 1-2.
- Kwiatkowska E.: *Analiza zmian przebiegu wybranych makroekonomicznych szeregów czasowych w Czechach, Polsce i na Węgrzech*. Red. J. Mika. AE, Katowice 2002a.
- Kwiatkowska E.: *Analiza zmian strukturalnych modeli autoregresyjnych dla wybranych zmiennych ekonomicznych*. AE, Katowice 2002b.
- Kwiatkowska E.: *Zastosowanie metody określania rodzajów i sił związków między wybranymi zmiennymi ekonomicznymi w Polsce w latach 1993-2001*. Badania własne. AE, Katowice 2002c.
- Łyko J.: *Wskaźnik cen towarów i usług konsumpcyjnych jako miernik inflacji w Polsce*. *Ekonomia Matematyczna*. AE, Wrocław 2000, nr 4.
- Majsterek M.: *Zastosowanie procedury Johansena do analizy sprzężenia inflacyjnego w gospodarce polskiej*. „Przegląd Statystyczny” 1998, z. 1.
- Milo W.: *Szeregi czasowe*. PWE, Warszawa 1990.
- Nerlove M., Grether D.M., Carvalho J.L.: *Analysis of Economic Time Series*. Academic Press, 2000.
- Osiewalski J., Welfe A.: *A short-run price-wage nexus: an Application of Endogenous Switching*. „Przegląd Statystyczny” 1999, z. 4.
- Pawłowski Z.: *Prognozy ekonometryczne*. PWE, Warszawa.
- Theil H.: *Zasady ekonometrii*. PWN, Warszawa 1979.
- Tobin J.: *The Natural Rate as a New Classical Macroeconomics*. In: *The Natural Rate of Unemployment*. Ed. R. Cross. Cambridge University Press, Cambridge 1995.
- Tobin J.: *The Wage-Price Mechanism: Overview of the Conference*. In: *The Econometrics of Price Determination Conference*. Federal Reserve System, Washington D.C.
- Welfe A.: *Ekonometria*. PWE, Warszawa 2003.
- Welfe A., Majsterek M., Florczak W.: *Model pętli inflacyjnej w gospodarce polskiej – analiza kointegracyjna*. „Przegląd Statystyczny” 1994, z. 3.

**AUTOCORRELATION ANALYSIS OF THE RESIDUALS
OF THE AUTOREGRESSIVE MODELS
OF THE PRICE INDEX IN POLAND**

Summary

The paper presents application of two methods for determining the row of the residuals autocorrelation. Considered time series of the prices index, transformed into the form, which is used in the inflationary spiral model. After evaluating the breaking points and the row of autoregression of the model, the parameters' estimators have been determined. For this purposes the method based on the series of LSE estimators for the possibly shortest parts of the considered time series has been used. The residuals have been determined for the estimated model. They constitute background for further research on their autocorrelation. In this part, firstly *Pawłowski test* for autocorrelation of the row higher than one has been applied. Secondly, the same procedure as for determining the model of autoregression has been used for exploring the autocorrelation of residuals. This procedure has been applied under assumption that residuals are generated by the autoregressive equation. These two methods give similar results.