

MATERIAŁY I STUDIA

Zeszyt nr 220

X-12-ARIMA i TRAMO/SEATS –
empiryczne porównanie metod wyrównania
sezonowego w kontekście długości próby

Sylwia Grudkowska
Ewa Paśnicka

Warszawa, lipiec 2007 r.

Projekt graficzny:

Oliwka s.c.

Skład i druk:

Drukarnia NBP

Wydął:

Narodowy Bank Polski
Departament Komunikacji Społecznej
00-919 Warszawa, ul. Świętokrzyska 11/21
tel. 022 653 23 35, fax 022 653 13 21

© Copyright Narodowy Bank Polski, 2007

Materiały i Studia rozprowadzane bezpłatnie.

Dostępne są również na stronie internetowej NBP: <http://www.nbp.pl>

Spis treści

Streszczenie	5
Wstęp	6
1. Charakterystyka sezonowości	8
2. Charakterystyka X-11, X-11-ARIMA, X-12-ARIMA i TRAMO/SEATS . .	11
2.1. X-11	11
2.2. X-11-ARIMA	12
2.3. X-12-ARIMA	13
2.4. TRAMO/SEATS	15
2.5. Różnice pomiędzy X-12-ARIMA i TRAMO/SEATS.	16
3. Kryteria porównawcze oceny jakości modelu	18
4. Analiza porównawcza na podstawie wyników empirycznych	22
4.1. Informacje podstawowe	22
4.2. Badanie stabilności modelu	25
4.3. Kryterium zgodności	31
4.4. Gładkość dopasowania komponentów	32
4.5. Względny błąd prognozy	32
4.6. Kryterium idempotencji	33
4.7. Analiza komponentu nieregularnego	35
5. Podsumowanie	38
6. Aneks	41
7. Bibliografia	58

Spis tabel i wykresów

Tabela 1. Informacje podstawowe	23
Tabela 2. <i>Sliding spans</i> – analiza czynnika sezonowego <i>SS</i>	26
Tabela 3. Analiza miesięcznych stóp wzrostu danych odsezonowanych <i>MM</i>	26
Tabela 4. Analiza rocznych stóp wzrostu danych odsezonowanych <i>YY</i>	27
Tabela 5. Rewizje historyczne – dane odsezonowane	28
Tabela 6. Rewizje historyczne – trend	29
Tabela 7. Różnica rocznych stóp wzrostu danych surowych i odsezonowanych	31
Tabela 8. Miary gładkości dopasowania komponentów.	32
Tabela 9. Względny błąd prognozy dla szeregu odsezonowanego i komponentu trend-cykl	33
Tabela 10. Ponowne odsezonowanie danych jednokrotnie wyrównanych sezonowo.	34
Tabela 11. Analiza komponentu nieregularnego – test na biały szum	36
Tabela 12. Analiza komponentu nieregularnego	37
Tabela 13. Podsumowanie – pełna próba vs próba skrócona.	39
Tabela 14. Podsumowanie – porównanie metod.	39
Wykres 1. Indeks produkcji sprzedanej przemysłu, przeciętna miesięczna 2000=100	22
Wykres 2. Gęstość widmowa – pełna próba	24
Wykres 3. Gęstość widmowa – próba 2001–2006	25
Wykres 4. Rewizje szeregu odsezonowanego – oszacowanie wstępne vs ostateczne	28
Wykres 5. Rewizje TRAMO/SEATS pełna próba – roczne stopy wzrostu trendu-cyklu	29
Wykres 6. Rewizje X-12-ARIMA pełna próba – roczne stopy wzrostu trendu-cyklu.	30
Wykres 7. Rewizje TRAMO/SEATS krótki szereg – roczne stopy wzrostu trendu-cyklu	30
Wykres 8. Rewizje X-12-ARIMA krótki szereg – roczne stopy wzrostu trendu-cyklu	31
Wykres 9. Komponent sezonowy jednokrotnie (<i>S</i>) i dwukrotnie wyrównanego szeregu (<i>S(S)</i>).	35
Wykres 10. Komponent nieregularny TRAMO/SEATS vs X-12-ARIMA (pełna próba vs próba okrojona).	36

Streszczenie

Niniejsze opracowanie zwięźle prezentuje najważniejsze pojęcia, zagadnienia oraz podstawowe metody wygładzania sezonowego i kryteria ich porównania. Ze względu na to, że znaczna część danych polskich zalicza się do krótkich szeregów czasowych, badanie empiryczne objęło przede wszystkim wieloaspektową analizę jakości odsezonowania w zależności od długości dostępnej próby. W celu wyboru optymalnej metody wyrównania sezonowego estymacje wykonano dwiema najbardziej rozpowszechnionymi – TRAMO/SEATS oraz X-12-ARIMA. W badaniu wykorzystano indeks produkcji sprzedanej przemysłu z odpowiednim doбором długości szeregu, przy czym długi szereg czasowy obejmuje pełną próbę, a krótkie szeregi – wyodrębnione z niej próby skrócone. Wyniki analizy porównawczej wskazują na wyższą jakość odsezonowania długich szeregów czasowych. W przypadku indeksu produkcji sprzedanej przemysłu w Polsce sezonowe wyrównanie szeregu metodą TRAMO/SEATS, nawet w przypadku znacznej redukcji długości próby, daje lepsze jakościowo wyniki niż X-12-ARIMA. Ponadto dobór okresu dla krótkiej próby nie wpływa zasadniczo na wyniki analizy jakości odsezonowania.

Słowa kluczowe: odsezonowanie, X-12-ARIMA, TRAMO/SEATS, szereg czasowy, jakość dopasowania.

Klasyfikacja JEL: C13, C22, E23

Wstęp

W niniejszym opracowaniu przedstawiono podstawowe zagadnienia dotyczące sezonowego wygładzenia szeregów czasowych oraz metody służące ocenie otrzymanych rezultatów.

Wśród wielu metod sezonowego wyrównania danych najczęściej stosowane są obecnie TRAMO/SEATS i X-12-ARIMA¹. Równoległe wykorzystywanie dwóch uznanych metod przyczynia się do powstawania szeregu analiz komparatywnych. Ze względu na odmienną konstrukcję metod oszacowanie jakości wyników estymacji otrzymanych przy zastosowaniu tych podejść staje się problemem. Przegląd literatury wskazuje na brak zgodności odnośnie do wyboru kluczowego kryterium porównawczego. Przykładowo B. Cabra² koncentruje się na minimalizacji rewizji, natomiast R. Feldpausch, C. Hood i K. Wills³ przykładają największą wagę do kryterium *sliding spans*⁴. Ze względu na trudności w wyborze decydującego wskaźnika, najpowszechniej spotykanym podejściem jest wieloaspektowe porównanie otrzymanych wyników. Szeroki zbiór kryteriów omówił m.in. B. Fisher⁵. Wielopłaszczyznową analizę jakości wyrównania danych stosują m.in. R. Astolfi, D. Ladiray i G. Mazzi⁶ oraz M. Ongan⁷.

Badania jakości sezonowego wyrównania danych dla polskich szeregów czasowych są znikome i obejmują niewielkie zbiory danych. Jednym z nielicznych jest badanie P. Woźniaka dotyczące sezonowego wyrównywania wskaźnika cen dóbr i usług konsumpcyjnych. Należy przy tym zaznaczyć, że analiza ta nie jest kompleksowa i obejmuje jedynie wybrane kryteria porównawcze.

W procesie wyrównania sezonowego istotne jest uwzględnienie charakteru analizowanych danych, takich jak stopień agregacji szeregu i długość próby. Jeśli szereg czasowy jest sumą wielu elementów składowych, tak jak w przypadku agregatów monetarnych lub PKB, wówczas istnieją dwa sposoby uzyskania wyrównania sezonowego. Jednym z nich jest wygładzenie poszczególnych składowych, a w dalszej kolejności ich zagregowanie w celu uzyskania sezonowo wyrównanego agregatu. Jest to tzw. podejście pośrednie do odsezonowania szeregów. Odmiennym sposobem jest natomiast odsezonowanie bezpośrednie, które polega na stosowaniu wyrównania sezonowego do zagregowanych danych. Każde z podejść, jak pokazują to m.in. prace C. Hood i D. Findley⁸, D. Ladiray i G. Mazzi⁹, czy D. Willoughby¹⁰, związane jest z uzyskaniem odmiennych jakościowo wyników.

Jak wskazują dotychczasowe analizy na jakość otrzymanego oszacowania poszczególnych komponentów szeregu istotnie wpływa również liczba dostępnych obserwacji. Kompleksowe badanie takiej zależności przeprowadzone przez G. Mazzi¹¹ wskazuje, że w miarę skracania długości próby wyraźnie pogarsza się merytoryczna wartość wyników.

¹ Metody te są obecnie najbardziej rozpowszechnione, jednak równoległe stosuje się wiele innych metod: X-11, X-11-ARIMA, SABL, STAMP, BV4, DAINTIES. Więcej informacji na ich temat zawartych jest m.in. w: Fischer (1995).

² Cabra, Esteban i Afonso (2003).

³ Feldpausch, Hood i Wills (2004), s. 7.

⁴ Koncepcja *sliding spans* polega na wybraniu z analizowanego szeregu określonej ilości zachodzących na siebie podprób i ich odsezonowania. Jakość wyrównania sezonowego jest oceniana na podstawie stopnia rozbieżności obserwacji danego okresu pochodzących z różnych podprób.

⁵ Fischer (1995).

⁶ Astolfi, Ladiray i Mazzi (2003).

⁷ Ongan (2002).

⁸ Hood i Findley (2003).

⁹ Ladiray i Mazzi (2003).

¹⁰ Willoughby (2003).

¹¹ Ladiray i Mazzi (2003).

Oszacowany przez tego autora spadek jakości wyrównania był znacznie większy w przypadku odsezonowania metodą TRAMO/SEATS niż X-12-ARIMA. Zbliżone rezultaty uzyskali C. Hood, J. Ashley i D. Findley, wykonując analizy na danych symulowanych.

Odsezonowanie danych przeprowadzone na pełnej i skróconej próbie ma na celu porównanie jakości dekompozycji długich i krótkich szeregów oraz dokonanie oceny wyrównania sezonowego przy przejściu od krótkiego do długiego szeregu czasowego, w miarę jak kolejne obserwacje stają się dostępne. Analiza wpływu jest szczególnie istotna w przypadku polskich szeregów czasowych, które w przeważającej części są szeregami krótkimi i średnimi. W badaniach sezonowości za krótkie szeregi czasowe powszechnie uznaje się szeregi liczące około 60 obserwacji¹² dla danych miesięcznych.

Dla danych polskich brakuje opracowań dotyczących badania jakości odsezonowania w zależności od długości próby. W niniejszym opracowaniu poddano analizie indeks produkcji sprzedanej przemysłu 2000=100. Ze względu na charakter tej zmiennej, należącej do statystyki sfery realnej, uwaga zostanie skierowana na zależności pomiędzy długością szeregu czasowego a jakością odsezonowania. Wpływ poziomu agregacji na jakość uzyskanych wyników jest odrębnym tematem, który zostanie przeanalizowany w oddzielnej pracy na przykładzie agregatów monetarnych.

Rozdział pierwszy niniejszego opracowania poświęcony jest przybliżeniu zjawiska sezonowości. Omówiono tu cechy poszczególnych komponentów szeregu czasowego.

W kolejnym rozdziale przedstawiono podstawowe metody odsezonowania. Scharakteryzowano w nim poszczególne podejścia, a także krótko je porównano. Nacisk położono na omówienie metod TRAMO/SEATS i X-12-ARIMA, jako że są one powszechnie stosowane przez europejskie urzędy statystyczne i banki centralne oraz są rekomendowane przez Eurostat.

W rozdziale trzecim przedstawiono przegląd najczęściej stosowanych kryteriów oceny jakości odsezonowania. Omówiono tu te wskaźniki porównawcze, które mogą być stosowane zarówno w przypadku TRAMO/SEATS, jak i X-12-ARIMA.

Część empiryczna pracy, zawarta w rozdziale czwartym, obejmuje porównanie jakości sezonowego wyrównania szeregu czasowego dla pełnej i skróconej próby, otrzymanego metodami TRAMO/SEATS i X-12-ARIMA.

¹² Wg Mazzi i Savio (2005), w przypadku danych o częstotliwości miesięcznej, krótki szereg czasowy obejmuje 5 lat, średni szereg – 10 lat, długi szereg – 20 lat, s. 5.

Charakterystyka sezonowości

W literaturze przedmiotu sezonowość definiowana jest jako regularnie powtarzająca się zależność pomiędzy obserwacjami oddalonymi od siebie o stałą liczbę okresów. Analizując zjawiska ekonomiczne, najczęściej rozpatruje się sezonowość w obrębie roku kalendarzowego. Występowanie wahań sezonowych w szeregu czasowym utrudnia porównania i interpretację zmian zjawiska z okresu na okres (z kwartału na kwartał, z miesiąca na miesiąc, etc.). Wygładzenie sezonowe pozwala na określenie rzeczywistego kierunku średnio- i długookresowych zmian w poszczególnych okresach poprzez usunięcie krótkookresowych wahań, które określają charakterystyczny dla danego okresu czynnik sezonowy. Ponadto możliwe staje się uwzględnienie specyfiki kraju, z którego pochodzą dane, poprzez wykorzystanie w dekompozycji informacji o układzie świąt państwowych i religijnych. W ten sposób wnioskowanie na temat znaczenia danego zjawiska oraz kierunku zmian aktywności gospodarczej jest trafniejsze.

Dekompozycja szeregów czasowych jest ważnym zagadnieniem z punktu widzenia docelowych użytkowników i decydentów. Polega ona na wyodrębnieniu z wejściowego szeregu składnika trend-cykl (T_t), składnika losowego (I_t), efektu różnej liczby dni roboczych (D_t), efektu świąt (E_t) i składnika sezonowego (S_t):

$$X_t = T_t \square S_t \square I_t \square E_t \square D_t, \quad (1)$$

przy czym \square – w zależności od rozpatrywanego modelu multiplikatywnego lub addytywnego – oznacza odpowiednio znak mnożenia lub dodawania¹³.

W celu uzyskania sezonowo wygładzonych danych z szeregu czasowego usuwa się składnik sezonowy (S_t) oraz efekty kalendarza, na które składają się efekty dni roboczych i efekt świąt ruchomych ($D_t \square E_t$). Oczyszczony w ten sposób szereg (X_{SA_t}) zawiera tylko składnik nieregularny oraz trend-cykl ($X_{SA_t} = I_t \square T_t$)¹⁴.

Do wspomnianych wyżej efektów kalendarza zalicza się: efekt liczby dni roboczych oraz efekt ruchomych i stałych świąt. Pierwszy z nich polega na tym, że liczba dni roboczych w badanym okresie (wynikającym z częstotliwości szeregu) wpływa na obserwowane zjawisko. W dni robocze aktywność ekonomiczna podmiotów gospodarczych cechuje się z reguły większą intensywnością w porównaniu do dni wolnych od pracy. W zależności od tego czy dana zmienna występuje w postaci strumienia, czy też zasobu, w ujęciu modelowym efekt liczby dni roboczych wyrażany jest za pomocą odpowiedniej kombinacji zmiennych.

Rozpatrując kategorie ekonomiczne jako stan na koniec okresu, ważny staje się dzień tygodnia dokonania pomiaru. Sytuacja ta jest modelowana za pomocą zmiennych zerojedynkowych postaci:

$$D_{i,t} = \begin{cases} 1 & \text{dla } i\text{-tego dnia tygodnia} \\ -1 & \text{dla niedzieli} \\ 0 & \text{w p.p.} \end{cases}$$

gdzie: $i = \{1, \dots, 6\}$, $i=1$ oznacza poniedziałek, $i=2$ oznacza wtorek, ...

¹³ Możliwe jest także stosowanie modeli innych typów, np. pseudoaddytywnego $X_t = T_t \times [S_t + I_t + 1]$ lub logaddytywnego $\log(X_t) = \log(T_t) + \log(S_t) + \log(I_t)$.

¹⁴ W literaturze przedmiotu komponent sezonowy bywa również definiowany jako suma czynników S_t , D_t i E_t . Podejście takie stosuje m.in. Australian Bureau of Statistics.

Ogólna zasada modelowania ekonometrycznego wskazuje na konieczność uwzględnienia w równaniu regresji o jedną zmienną mniej niż liczba analizowanych przypadków (tu: dni tygodnia) w wyniku czego nie powstaje współliniowość tych regresorów z wyrazem wolnym równania. Ostatni dzień tygodnia jest identyfikowany przez analizę wartości przyjmowanych przez pozostałe zmienne zerojedynkowe.

W przypadku zmiennych ekonomicznych o charakterze strumieni, dni robocze można traktować jako zbiór rozróżnialnych lub nierozróżnialnych elementów. Gdy wszystkie dni robocze mają to samo znaczenie, przypisywane są im równe wagi. Istotne znaczenie ma jedynie łączna liczba dni roboczych w badanym okresie. W estymacji uwzględnia się zatem tylko dwie zmienne odpowiadające za relację liczby dni roboczych do liczby dni wolnych od pracy oraz wpływ roku przestępnego. Pierwsza z nich jest modelowana jako:

$$N_t = N_{roboczy_t} - \frac{5}{2}N_{weekend_t} \quad (2)$$

gdzie:

$N_{roboczy_t}$ – liczba dni roboczych w miesiącu t ,

$N_{weekend_t}$ – liczba sobót i niedziel w miesiącu t .

Zmienna obrazująca efekt roku przestępnego jest równa różnicy długości danego miesiąca i średniej długości miesiąca w roku wynoszącej 30,4375 dni.

Jeżeli rozpatrywane zjawisko zmienia swoje natężenie w zależności od dnia tygodnia, wówczas do modelu włączane są zmienne odpowiadające każdemu z tych dni w postaci:

$$k_{i,t} = K_{i,t} - K_{7,t} \quad (3)$$

gdzie: $K_{i,t}$ – liczba i -tych dni tygodnia w miesiącu t , $i = \{1, \dots, 6\}$, $i=1$ oznacza poniedziałek, $i=2$ oznacza wtorek, ...

$K_{7,t}$ – liczba niedziel w miesiącu t .

Dodatkowo do równania może być włączany regresor uwzględniający efekt roku przestępnego.

Efekt świąt polega na modyfikacji aktywności ekonomicznej w okresie okołoswiątecznym. W przypadku świąt ruchomych, takich jak Wielkanoc i Boże Ciało, istotny staje się okres (miesiąc, kwartał), w którym wypada dane święto oraz liczba dni w danym okresie przypadających przed świętem. Aktywność ekonomiczna dotycząca świąt stałych zależy od liczby dni roboczych wolnych od pracy w badanym okresie.

W procesie dekompozycji szeregu czasowego wyodrębnia się także, wspomniany już, składnik nieregularny, w skład którego wchodzi:

1. Obserwacje nietypowe, nieujęte w estymacji (*outliers*), wśród których, ze względu na charakter zmian, wyróżnia się zjawiska:

- jednorazowe (*additive outliers*), tj. stanowiące istotne odchylenie od przewidywanej wartości badanego zjawiska tylko w jednym okresie, niewpływające na wartości szeregu w następnych okresach,
- długotrwałe (*level shift*), tj. powodujące trwałą zmianę poziomu zmiennej,
- powodujące przejściową zmianę poziomu zmiennej (*temporary change*), przy czym powrót do stanu początkowego następuje przeważnie zgodnie z funkcją wykładniczą lub liniową,
- innowacyjne (*innovation outliers*), które, w przeciwieństwie do wyżej omawianych, powodują zmianę procesu generującego dane, w szczególności zmianę posta-

ci trendu¹⁵. Zdarzeniem o charakterze innowacyjnym może być np. zastosowanie nowej technologii produkcji.

2. Efekty nieregularne, wywołane przez czynniki losowe lub niemożliwe do przewidzenia, takie jak: klęski żywiołowe, nagłe zmiany w polityce państwa, strajki.

Celem wyrównywania sezonowego jest eliminacja efektów kalendarza, składnika sezonowego oraz odpowiednie uwzględnienie w estymacji zmiennych nietypowych i brakujących obserwacji. Do najbardziej kompleksowych metod wyrównania sezonowego należą X-12-ARIMA oraz TRAMO/SEATS.

¹⁵ Muirhead (1986), s. 1-2.

Charakterystyka X-11, X-11-ARIMA, X-12-ARIMA i TRAMO/SEATS

2.1. X-11

Jedne z pierwszych metod wyrównania sezonowego polegały na dekompozycji szeregu czasowego za pomocą średnich ruchomych. Do takich metod zaliczyć należy X-11, opracowaną w 1965 r. przez Bureau of the Census¹⁶.

Metoda ta wykorzystuje w procesie estymacji symetryczne filtry¹⁷, zbudowane na podstawie średniej ruchomej Hendersona¹⁸. Ich zaletą jest satysfakcjonująca eliminacja nieregularnych zaburzeń, trwających krócej niż 6 miesięcy, przy jednoczesnym zachowaniu w wyrównanym szeregu czasowym znacznej części wahań cyklicznych. Symetryczność filtru nie zaburza położenia i amplitudy punktów zwrotnych badanego zjawiska. Co więcej, jego wyniki najczęściej nie ulegają istotnym zmianom po przeprowadzeniu rewizji¹⁹.

Cechą charakterystyczną metody X-11 jest zastosowanie w niej podwójnych średnich ruchomych $M_{m \times p}(X_t)$. Podwójna średnia ruchoma $M_{m \times p}(X_t)$ rzędu $(m + p - 1)$ dla szeregu X_t jest obliczana w zależności od tego, czy parametry m i p są nieparzyste lub parzyste, odpowiednio jako:

$$\bullet M_{(2k+1) \times (2n+1)}(X_t) = \frac{1}{2k+1} \sum_{j=-k}^k S_{t+12j}^{2n+1}, \text{ gdzie } S_t^{2n+1} = \frac{1}{2n+1} \sum_{j=-n}^n X_{t+12j}, \quad (4)$$

$$\bullet M_{2k \times 2n}(X_t) = \frac{1}{2k} \sum_{j=-k+1}^{k-1} S_{t+12j}^{2n}, \text{ gdzie } S_t^{2n} = \frac{1}{2n} \sum_{j=-n+1}^{n-1} X_{t+12j} \quad (5)$$

przy czym $m, n \in N$.

Jest to zatem średnia $\{(2k+1) \times (2n+1)\}$ lub $(2k \times 2n)$ liczona w dwóch etapach: pierwszy obejmuje kalkulację na pierwotnym szeregu czasowym średniej ruchomej rzędu $2n+1$ (odpowiednio $2n$), której współczynniki są równe $\frac{1}{2n+1}$ (odpowiednio $\frac{1}{2n}$), natomiast drugi policzenie średniej rzędu $2k+1$ (odpowiednio $2k$) na średnich z etapu pierwszego o współczynnikach $\frac{1}{2k+1}$ (odpowiednio $\frac{1}{2k}$). Stąd też, przykładowo, podwójna średnia ruchoma: $M_{3 \times 3}$ jest symetryczną średnią rzędu 5 liczoną na pierwotnym szeregu o wagach: $\left(\frac{1}{9}, \frac{2}{9}, \frac{3}{9}, \frac{2}{9}, \frac{1}{9}\right)$; $M_{3 \times 5}$ rzędu 7 – $\left(\frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{3}{15}, \frac{2}{15}, \frac{1}{15}, \frac{1}{15}\right)$; $M_{2 \times 4}$ rzędu 5 – $\left(\frac{1}{8}, \frac{2}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}\right)$. W ogólnym przypadku podwójna średnia ruchoma ma następującą postać:

¹⁶ Ladiray i Quenneville (2004).

¹⁷ Filtr jest symetryczny, jeżeli jego wagi są symetrycznie rozłożone względem środka ciągu wag.

¹⁸ Konstrukcja filtru Hendersona przedstawiona jest u Findley, Monsell, Bell, Otto i Bor-Chung Chen (1998), s. 60-61.

¹⁹ *Seasonal adjustment (2006)* w: „Seasonal and Prior Adjustment Factors for 2006, Trends for 2002-2006”, s. 16.

$$M_{n \times m}(X_t) = \sum_{i=-f}^f \theta_i X_{t+12i}, \quad (6)$$

gdzie: f jest równe $\frac{m+n}{2}-1$, a θ_i są odpowiednimi wagami.

Metoda X-11, zastosowana do wyrównania szeregu X_t , składa się z dwóch etapów:

1. Wstępnej estymacji komponentów szeregu, przy czym składnik trend-cykl jest przybliżany za pomocą podwójnej średniej ruchomej $M_{2 \times 12}$, oczyszczającej szereg z czynnika losowego i wahań sezonowych ($C_t^{(1)} = M_{2 \times 12}(X_t)$). Komponent sezonowy i nieregularny zostają oszacowane jako różnica między szeregiem X_t a pierwszym oszacowaniem trendu-cyklu $(S_t + I_t)^{(1)} = X_t - C_t^{(1)}$. W wyniku zastosowania średniej ruchomej $M_{3 \times 3}$ wyodrębnia się składnik sezonowy: $S_t^{(1)} = M_{3 \times 3}[(S_t + I_t)^{(1)}]$. Komponent nieregularny otrzymywany jest jako różnica $I_t^{(1)} = (S_t + I_t)^{(1)} - S_t^{(1)}$, a wstępnie wygładzony szereg szacowany jest jako $A_t^{(1)} = (C_t + I_t)^{(1)} = X_t - S_t^{(1)}$.
2. Estymacji wahań sezonowych $S_t^{(2)}$ i ostatecznej postaci długookresowego trendu $C_t^{(2)}$ za pomocą 13-okresowego filtra Hendersona. Trend-cykl jest szacowany przez zastosowanie filtra do wstępnie wygładzonego szeregu $C_t^{(2)} = H_{13}(A_t^{(1)})$. Natomiast łączne oszacowanie komponentu sezonowego i nieregularnego odbywa się tak jak w poprzednim etapie: $(S_t + I_t)^{(2)} = X_t - C_t^{(2)}$. Komponent sezonowy jest tu szacowany jako średnia ruchoma na wielkości oszacowanej w poprzednim równaniu $S_t^{(2)} = M_{3 \times 5}[(S_t + I_t)^{(2)}]$, a następnie poddawany jest normalizacji, tak aby suma czynników sezonowych w poszczególnych latach wynosiła w przybliżeniu zero: $\tilde{S}_t^{(2)} = S_t^{(2)} - M_{2 \times 12}(S_t^{(2)})$. Końcowym etapem jest ustalenie ostatecznych wartości odsezonowanego szeregu $A_t^{(2)} = (C_t + I_t)^{(2)} = X_t - \tilde{S}_t^{(2)}$.

Ze względu na swoją prostotę X-11 jest nadal stosowana, mimo że ma wady. Do najpoważniejszych z nich należy brak zastosowania modelu statystycznego. Należy ona do tzw. metod *ad hoc*, które zakładają, że każdy szereg może być rozłożony na trzy składniki za pomocą tej samej procedury. Ponadto stosowanie filtrów średniej ruchomej prowadzi do eliminacji wszystkich komponentów szeregu, których rozkład nie jest symetryczny.

Kolejną wadą X-11 jest to, że ze względu na brak odpowiedniej ilości obserwacji filtr średnich ruchomych nie może być stosowany na końcach próby. Do wyrównania tych obserwacji stosuje się filtry asymetryczne, które mogą zniekształcać oszacowany efekt sezonowy i wprowadzać do wyrównanych danych zmiany strukturalne, które w rzeczywistości nie wystąpiły. Ponadto, jak wskazuje Planas²⁰, metoda X-11 nie może być stosowana w przypadku szeregów czasowych o bardzo niestabilnych wahaniami sezonowych, ze względu na ograniczony zbiór dostępnych filtrów.

2.2. X-11-ARIMA

Kolejnym krokiem w ewolucji metod wyrównania sezonowego było wprowadzenie metody X-11-ARIMA, eliminującej większość wad poprzednio omawianego rozwiązania. X-11-ARIMA przeprowadza wstępną procedurę rozszerzenia szeregu za pomocą pro-

²⁰ Planas (1997), s. 10.

gnoz (*forecast*) i ekstrapolacji wstecz (*backcast*) używając w tym celu modelu ARIMA²¹. W wyniku tego symetryczne filtry średniej ruchomej mogą być stosowane na całej długości próby. Zastosowanie takiego zabiegu znacznie zmniejsza rewizję oraz poprawia jakość estymacji trendu na końcu próby w porównaniu do wyników X-11²².

2.3. X-12-ARIMA

Rozwinięciem metody X-11-ARIMA jest opracowana przez U.S. Bureau of the Census X-12-ARIMA. Idea jej funkcjonowania opiera się na tradycyjnej dekompozycji szeregów czasowych i modelowaniu szeregów przy użyciu sezonowych modeli ARIMA $(p, d, q) \times (P, D, Q)_s$ ²³. Główną zmianę stanowi dodanie wstępnej estymacji przed zasadniczym odsezonowaniem. Proces ten w X-12-ARIMA określany jest mianem REGARIMA. Polega na estymacji modelu wyjściowego, w którym uwzględnia się wszystkie możliwe zmienne objaśniające, takie jak składnik trend-cykl, wahania sezonowe, efekt kalendarza, efekt świąt czy składnik losowy. Zabieg ten wstępnie oczyszcza szereg z czynników, takich jak efekt kalendarza lub obserwacje nietypowe, które mogą powodować błędy w dekompozycji szeregu. Estymacja i identyfikacja postaci modelu ARIMA stanowi kolejny krok procedury REGARIMA. W tym celu określany jest stopień różnicowania niesezonowego i sezonowego (odpowiednio d i D), zapewniający stacjonarność szeregu. W dalszej kolejności następuje dobór parametrów p i q modelu ARMA, odpowiadających części niesezonowej modelu oraz P i Q , odpowiadających części sezonowej.

Zakładając, że wektor obserwacji ma postać: $\mathbf{X}_t = (X_1, X_2, \dots, X_m)$, X-12-ARIMA estymuje następujące równanie regresji²⁴:

$$\mathbf{X}_t = \mathbf{Z}_t^T \boldsymbol{\beta} + \mathbf{e}_t \quad (7)$$

gdzie:

$\boldsymbol{\beta}$ – wektor estymowanych parametrów strukturalnych,

\mathbf{Z}_t^T – macierz zmiennych niezależnych,

\mathbf{e}_t – składnik losowy podlegający procesowi ARIMA(p, d, q) \times (P, D, Q):

$$\phi(B) \cdot \delta(B) \cdot \mathbf{e}_t = \theta(B) \cdot \mathbf{a}_t \quad (8)$$

gdzie:

\mathbf{a}_t – proces białego szumu o zerowej wartości oczekiwanej i wariancji wynoszącej δ^2 ,

β – operator opóźnień,

²¹ Ogólny model ARIMA (*Autoregressive Integrated Moving Average*) został wprowadzony przez Boxa i Jenkinsa w 1976 r. Jest on określony przez zestaw parametrów (p, d, q) , gdzie p oznacza rząd procesu autoregresyjnego, d – rząd różnicowania, q – rząd procesu średniej ruchomej. Przykładowo model ARIMA (1,1,2) zawiera 1 parametr autoregresyjny i 2 parametry średniej ruchomej obliczone dla szeregu po jednokrotnym różnicowaniu. Metodologia ARIMA zaproponowana przez Boxa i Jenkinsa obejmuje identyfikację modelu ARIMA (tj. określenie parametrów modelu na podstawie analizy funkcji autokorelacji i częściowej autokorelacji), estymację modelu i prognozowanie.

²² *Seasonal adjustment (2006)* w: „Seasonal and Prior Adjustment Factors for 2006, Trends for 2002-2006”, s. 12.

²³ W dalszej części opracowania pojęcie modelu ARIMA obejmuje również modele w postaci sezonowej, o ile nie jest to inaczej zaznaczone.

²⁴ Fischer (1995), s. 6.

$$\begin{aligned} \delta(B) &= (1-B)^d (1-B^S)^D \\ \phi(B) &= \underbrace{\left(1 - \phi_1 B - \dots - \phi_p B^p\right)}_{w_1} \cdot \underbrace{\left(1 - \Phi_1 B^S - \dots - \Phi_p B^{pS}\right)}_{w_2} \\ \theta(B) &= \underbrace{\left(1 - \theta_1 B - \dots - \theta_q B^q\right)}_{w_3} \cdot \underbrace{\left(1 - \Theta_1 B^S - \dots - \Theta_q B^{qS}\right)}_{w_4} \end{aligned} \quad (9)$$

przy czym:

w_1 – wielomian związany z niesezonową częścią modelu, w postaci procesu AR(p)²⁵,

w_2 – wielomian związany z sezonową częścią modelu,

w_3 – wielomian wyrażający niesezonową część modelu, w postaci procesu MA(q)²⁶,

w_4 – wielomian wyrażający sezonową część modelu w postaci średniej ruchomej,
S – liczba obserwacji w roku,

d – ilość wyznaczonych pierwszych przyrostów zmiennej e_t ,

D – ilość wyznaczonych przyrostów w stosunku do analogicznego okresu roku poprzedniego.

Pierwszemu etapowi dekompozycji podlega równanie powstałe po podstawieniu za nieobserwowalny składnik losowy (8) równania (7):

$$\phi(B) \cdot \delta(B) \cdot (\mathbf{X}_t - \mathbf{Z}_t^T \beta) = \theta(B) \cdot \mathbf{a}_t \quad (10)$$

Kolejne etapy odsezonowania w X-12-ARIMA przebiegają analogicznie do całej procedury X-11, jednak wybór dostępnych opcji znacznie rozszerzono. Zwiększył się m.in. zbiór dostępnych filtrów Hendersona, wprowadzono możliwość szacowania modelu pseudoadytywnego oraz wykorzystania filtru sezonowego (3 x 15).

Końcowy etap procedury to ocena jakości i stabilności modelu przy wykorzystaniu statystyk zbudowanych na podstawie wartości poszczególnych komponentów szeregu otrzymanych w poprzednich etapach estymacji²⁷.

Jednym z podstawowych narzędzi diagnostyki odsezonowania jest analiza spektralna. Polega ona na określeniu funkcji gęstości spektralnej szeregu, czyli rozkładu jego wariancji w dziedzinie częstości. Pozwala to na identyfikację wartości przyjmowanych przez częstotliwości sezonowe i stwierdzenie, czy różnią się one istotnie od wartości sąsiednich. Analizie poddawane są również częstotliwości odpowiadające efektowi dni roboczych. W szeregu podlegającym wahaniom sezonowym wartości spektrum dla częstotliwości sezonowych mają wyraźne, dodatnie odchylenia od średniej. Odpowiednio wyrównany szereg nie powinien mieć tego typu wahań.

Aby otrzymać funkcję gęstości spektralnej (widmo) szeregu, jest on przedstawiony przy pomocy transformaty Fouriera jako suma sinusów i cosinusów różnych częstotliwości. Dekompozycja szeregu X_t ma postać:

²⁵ AR(p) to proces autoregresyjny, definiowany jako: $X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t$, gdzie ε_t – błąd będący realizacją procesu białego szumu. AR(p) jest to zatem regresja bieżących wartości X_t na opóźnione wartości X_t . W notacji uwzględniającej wprowadzenie operatora opóźnień, proces ten wyrazić można jako: $\varepsilon_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t$.

²⁶ MA(q) to proces średniej ruchomej, definiowany jako: $X_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$, gdzie ε_t – błąd w postaci procesu białego szumu. MA(q) jest zatem procesem będącym średnią ważoną bieżących i opóźnionych błędów standardowych w postaci białego szumu. Używając innej notacji, która stosowana jest w niniejszym opracowaniu, proces ten przedstawić można jako: $X_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t$.

²⁷ Szczegółowa konstrukcja statystyk została opisana w Ladiray i Quenneville (2004).

$$X_t = \frac{a_0}{2} + \sum_{k=1}^m [A_k \cos(\omega_k t) + B_k \sin(\omega_k t)] \quad (11)$$

gdzie:

m – numer kolejny częstości kołowej, obliczany jako $m = \frac{n}{2}$, gdy n jest parzyste;

$m = \frac{n-1}{2}$, gdy n jest parzyste, gdzie n oznacza liczbę obserwacji szeregu X_t ,

a_0 – podwojona średnia z wartości szeregu X_t ,

A_k i B_k – współczynniki przy sinusie i cosinusie, uzyskane metodą najmniejszych wiarygodności jako estymatory równania $X_t = A \cos 2\pi kt + B \sin 2\pi kt$,

ω_k – częstość kołowa, zdefiniowana jako $\omega_k = \frac{2\pi k}{n}$.

Współczynniki A_k i B_k są wykorzystywane do konstrukcji funkcji $J_k = \frac{n}{2}(A_k^2 + B_k^2)$ określonej jako periodogram. Jest ona estymatorem gęstości spektralnej. Ze względu na brak zgodności²⁸ tego estymatora wygładza się go przy pomocy funkcji wag, zwanej oknem spektralnym²⁹.

2.4. TRAMO/SEATS

W przeciwieństwie do omawianych wyżej metod TRAMO/SEATS, opracowana przez A. Maravall i V. Gomez, ma charakter modelowy, tzw. AMB (*ARIMA-Model-Based*). Podejście to polega na optymalnym filtrowaniu, bazującym na metodologii ARIMA. Składają się na nią dwie części: TRAMO (*Time series Regression with ARIMA noise, Missing values, and Outliers*) i SEATS (*Signal Extraction in ARIMA Time Series*).

Procedura TRAMO przeprowadza estymację i predykcję modelu, bez względu na charakterystykę reszt. Umożliwia ona również wykonanie interpolacji ewentualnych braków danych, identyfikacji i korekty obserwacji nietypowych, a także estymacji efektów dni roboczych i świąt ruchomych. Z kolei SEATS pozwala na estymację każdego z nieobserwowalnych składników zgodnie z podejściem opartym na modelu ARIMA. Podobnie jak w przypadku poprzednio omawianych metod, dekompozycja szeregu czasowego ma na celu wyodrębnienie jego poszczególnych składników.

Metoda wyrównywania sezonowego TRAMO/SEATS przebiega wieloetapowo. Początkowo TRAMO dokonuje automatycznego wyboru modelu ARIMA, zapisanego analogicznie jak dla X-12-ARIMA na stronach 14-15 niniejszego opracowania. Następnie, przy użyciu metody największej wiarygodności, wykrywa się obserwacje nietypowe, a także estymuje zmienne niezależne, odpowiedzialne m.in. za efekt liczby dni roboczych czy odzwierciedlające efekt świąt. Efekty kalendarzowe i obserwacje nietypowe zidentyfikowane na tym etapie są usuwane z szeregu. Zlinearyzowany w ten sposób szereg cza-

²⁸ Estymator jest zgodny, jeżeli jego wariancja maleje wraz ze wzrostem liczności próby.

²⁹ Więcej informacji na ten temat zawiera artykuł Findley, Monsell, Bell i Otto (1998) oraz Skrzypczyński (2006).

sowy³⁰ zostaje poddany procedurze SEATS, podczas której dokonuje się dekompozycji szeregu. SEATS estymuje funkcje gęstości spektralnej zgodnie z metodą przedstawioną w punkcie 2.3. Przy założeniu wzajemnej niezależności poszczególnych składowych szeregu czasowego, funkcja gęstości spektralnej szeregu jest dekomponowana na funkcje gęstości spektralnej komponentów nieobserwowalnych: trendu-cyklu (T_t), komponentu sezonowego (S_t) i nieregularnego (I_t). Zakłada się przy tym, że składniki te spełniają warunek ortogonalności. Otrzymane komponenty są rozszerzane wprzód (*forecast*) i wstecz (*backcast*), co umożliwia zastosowanie na nich filtru Wienera-Kołmogorowa. Jest on wykorzystywany do oszacowania estymatorów poszczególnych komponentów i obliczenia odchyłeń standardowych estymatorów i ich prognoz. W kolejnym etapie SEATS poddaje analizie rewizję, jakim podlega estymator w miarę zwiększania długości próby. Procedurę TRAMO/SEATS kończy wprowadzenie do otrzymanych komponentów uprzednio usuniętych z szeregu efektów i obserwacji nietypowych.

2.5. Różnice pomiędzy X-12-ARIMA i TRAMO/SEATS

Dwie najpowszechniej stosowane metody wyrównania sezonowego – X-12-ARIMA i TRAMO/SEATS mają zasadniczo odmienną konstrukcję. TRAMO/SEATS dobiera rodzaj stosowanego filtru do analizowanego szeregu, podczas gdy X-12-ARIMA stosuje ujednolicony zestaw filtrów, bez względu na strukturę komponentów szeregu. W TRAMO/SEATS analiza spektralna jest podstawowym narzędziem służącym do dekompozycji szeregu na poszczególne komponenty, podczas gdy w X-12-ARIMA jest ona wykorzystywana jedynie w diagnostyce jakości odsezonowania. TRAMO/SEATS dopasowuje model ARIMA do każdego z komponentów szeregu, natomiast w drugiej z metod jest on szacowany tylko dla oryginalnego szeregu, tak aby rozszerzyć długość próby na potrzeby dalszej estymacji.

W obu metodach stosowany jest również odmienny zestaw kryteriów, które umożliwiają selekcję modelu. Podczas wstępnego etapu estymacji TRAMO modeluje szereg czasowy za pomocą procesu AR (1) i ARMA (1,1), aby wyeliminować różnice sezonowe występujące z okresu na okres. Selekcji odpowiedniego modelu ARMA dokonuje się m.in. na podstawie kryteriów informacyjnych Bayesa (BIC), Akaike (AIC), testu losowości znaków reszt z modelu oraz badaniu stosunku wariancji estymatora komponentu nieregularnego do wariancji tego komponentu, otrzymanego w wyniku estymacji. W przypadku X-12-ARIMA wybór modelu następuje na podstawie kryterium Akaike (AIC), zmodyfikowanego kryterium Akaike (AICC), kryterium Hannan-Quinn oraz kryterium Schwarz-Bayes (BIC). Za optymalny uznaje się model, dla którego kryteria informacyjne osiągają wartość minimalną.

Procedury X-12-ARIMA i TRAMO/SEATS różnią się również podejściem do wstępnej fazy wygładzania sezonowego. TRAMO/SEATS dobiera filtry wygładzania sezonowego na podstawie wyników testów statystycznych. REGARIMA używa tzw. filtrów *ad hoc* wygładzania sezonowego.

TRAMO w porównaniu do X-12-ARIMA dysponuje bardziej kompleksowymi możliwościami automatycznego wyrównywania sezonowego. Program ten zapewnia, podobnie jak X-12-ARIMA, m.in. dekompozycję multiplikatywnego i addytywnego szeregu czasowego oraz kompleksową, pod względem optymalności doboru filtrów, identyfikację modelu. Procedura wykrywania *outlierów* przebiega jednak inaczej. TRAMO automatycznie wykrywa poszczególne typy obserwacji nietypowych

³⁰ Linearyzacja szeregu omówiona przez Fisher (1995), s. 18, polega na jego przedstawieniu w postaci: $Z_t = S_t + N_t$, gdzie Z_t – logarytmowany oryginalny szereg, S_t – komponent sezonowy, N_t – komponent nie-sezonowy. Komponenty S_t i N_t są od siebie niezależne. Szereg Z_t jest opisany przez sezonowy model ARIMA; S_t , N_t są opisane przez nieznanne modele ARIMA, odpowiednio $S(B)S_t = S(B)b_t$ i $N(B)N_t = N(B)c_t$, gdzie b_t i c_t są procesami białego szumu.

w postaci zmian przejściowych, natomiast pozostałe typy *outlierów*, tj. długotrwałe i jednorazowe są wykrywane podobnie jak w X-12-ARIMA. Dlatego też metoda TRAMO/SEATS może zakwalifikować daną obserwację nietypową jako zmianę przejściową, podczas gdy X-12-ARIMA jako obserwację o charakterze jednorazowym³¹.

³¹ Atuk i Ural (2002), s. 11.

3

Kryteria porównawcze oceny jakości modelu

Porównanie wyników wyrównania sezonowego otrzymanego metodami TRAMO/SEATS i X-12-ARIMA przebiegło na podstawie analizy różnorodnych statystyk, które pozwoliły na ocenę jakości dekompozycji i stabilności wygładzonych szeregów czasowych.

Ocenie, czy zastosowana metoda wyrównywania sezonowego była poprawna służy wiele kryteriów. Jednym z nich jest kryterium idempotencji. Polega ono na sprawdzeniu występowania efektów kalendarza i resztowych efektów sezonowych w wygładzonym szeregu czasowym, np. za pomocą analizy spektralnej lub sprawdzenia losowości składnika resztowego. W przypadku dobrego dopasowania sezonowego w wygładzonym szeregu i składniku losowym te efekty nie powinny zostać zdiagnozowane. Wykrycie w wyrównanym szeregu lokalnych maksimów dla częstości, odpowiadających wahaniom sezonowym (w przypadku danych miesięcznych są to częstości postaci $k/12$, przy czym $1 \leq k \leq 6$) lub efektom kalendarza, wskazuje na nieprawidłowe przeprowadzenie dekompozycji szeregu.

Szczegółowej analizie poddawany jest również komponent nieregularny. Oprócz badania obecności efektów kalendarza oraz sezonowości, zawiera ona testy na biały szum i normalność rozkładu. Hipoteza o występowaniu procesu białego szumu może być zweryfikowana na podstawie statystyk Fishera-Kappa³² lub Kołmogorowa-Smirnowa³³. W przypadku modelu multiplikatywnego hipoteza zerowa zakłada występowanie procesu o rozkładzie z wartością oczekiwaną równą 1^{34} .

W literaturze przedmiotu przedstawia się różnorodne miary, które służą ocenie stopnia wygładzenia poszczególnych składników szeregu czasowego (*roughness of components*)³⁵. Pierwsza z nich (R_1) mierzy średnią sumę kwadratów różnic pomiędzy odsezonowanymi wielkościami z okresu bieżącego a tymi samymi wielkościami z okresu poprzedniego. Badaniu podlega zatem jakość wygładzenia całego szeregu:

$$R_1 = \frac{1}{N-1} \sum_{t=2}^N (X_SA_t - X_SA_{t-1})^2 \quad (12)$$

Druga z miar (R_2), odnosząca się również do jakości wyrównania całego szeregu, bazuje na 13-okresowym filtrze Hendersona. Przy użyciu tego filtru wyrównane dane podane zostają ponownie wygładzeniu w celu oszacowania następującej wielkości³⁶:

$$R_2 = \frac{1}{N-1} \sum_{t=1}^N (X_SA_t \square H_{13} X_SA_t)^2 = \frac{1}{N-1} \sum_{t=1}^N [(I \square H_{13}) X_SA_t]^2 \quad (13)$$

Zastosowanie powyższej transformacji służy zniwelowaniu składników o małej częstości, które odpowiadają za wariancję składnika trend-cykl. Zatem statystyka ta mierzy w przybliżeniu średnie odchylenie danych odsezonowanych od trendu, czyli aproksymuje składnik nieregularny I_t (gdyż $X_SA_t = I_t \square T_t$).

Ponadto Maravall i Gomez³⁷ proponują zastosowanie miary smukłości dopasowania odniesieniu do składnika trendu. W tym celu należy się posłużyć poniższym wzorem:

$$MAR(TC) = \sum_{t=2}^N (\Delta TC_t) \quad (14)$$

³² Fuller (1976).

³³ Miller (1956), s. 1.

³⁴ Peronaci (2003), s. 99.

³⁵ Mazzi i Savio (2005), s. 4.

³⁶ W przypadku modelu addytywnego komponent nieregularny obliczany jest jako różnica danych wyrównanych sezonowo i trendu. Natomiast dla modelu multiplikatywnego jest on równy ilorazowi tych wielkości.

³⁷ Mazzi i Savio (2005), s. 6.

Miary te mogą być zastosowane zarówno w przypadku wyrównania sezonowego przy użyciu procedury TRAMO/SEATS, jak i X-12-ARIMA; dla całej próby, jak i dla podokresów, np. ostatnich trzech lat. Reguła dobrego dopasowania mówi, że wielkości te powinny być jak najmniejsze. Trzeba jednak zwrócić uwagę na to, że składnik nieregularny jest integralną częścią szeregu czasowego, odzwierciedlającą wpływ zjawisk nieprzewidywalnych, takich jak wydarzenia polityczne czy klęski żywiołowe, dlatego też nie należy dążyć do przesadnej minimalizacji wartości powyższych miar.

Kolejnym narzędziem pozwalającym na ocenę jakości wyrównania sezonowego są rewizje historyczne. Prawidłowo wyrównane szeregi powinny być stabilne, tzn. po tym, jak kolejne obserwacje wyrównanego szeregu (dla $t = 1, \dots, n$) stają się dostępne i dokonywana jest ponowna estymacja wydłużonego szeregu (dla $t=1, 2, \dots, n, n+1, n+2, \dots$), różnice pomiędzy wstępnie wyrównanymi danymi a ostatecznie wyrównanymi danymi dla danego okresu $t=1, 2, \dots, n$ nie powinny być znaczące. Rewizja historyczna polega na porównaniu różnicy pomiędzy pierwotnymi a ostatecznymi wartościami wyestymowanego szeregu dla danego okresu n . Przez pierwotną estymację rozumieć należy wyrównanie szeregu czasowego w momencie, gdy ostatnią obserwacją jest ta o numerze n . Kiedy dostępne są kolejne obserwacje, należy dokonać kolejnej dekompozycji szeregu czasowego dla obserwacji n . W analizowanym przypadku szeregu reestymuje się na podstawie ustalonej raz do roku postaci modelu. Oznacza to, że wraz z każdorazowym pojawianiem się nowych obserwacji parametry modelu nie ulegają zmianie.

Aby ocenić jakość rewizji szeregu, można posłużyć się następującą statystyką mierzącą względną różnicę poziomów wyrównanego szeregu³⁸:

$$X_SA_n\% = \left(\frac{X_SA_{n/N}}{X_SA_{n/n}} - 1 \right) \times 100\% , \quad (15)$$

definiując przez $X_SA_{n/t}$ wyrównaną wartość szeregu X_t dla okresu n na podstawie obserwacji X_1, X_2, \dots, X_t , przy czym $n \leq t \leq N$, a szereg czasowy X_t dostępny jest dla okresów $t = 1, 2, \dots, N$, to:

$X_SA_{n/n}$ oznacza pierwotną wartość wyrównanego szeregu czasowego dla okresu n ,

$X_SA_{n/N}$ oznacza wyestymowaną wartość ostateczną dla okresu n na podstawie pełnej próby.

Do analizy porównawczej wykorzystuje się klasyczne miary statystyczne, tj. średnią z rewizji $X_SA_n\%$ oraz odchyleniem standardowym rewizji $X_SA_n\%$. Rekomendowana metoda powinna mieć niższą wartość wymienionych powyżej statystyk.

Kolejna metoda oceny dopasowania – *sliding spans* – porównuje dekompozycję dokonaną na nierozłącznych podzbiórach obserwacji danego szeregu pod względem kryterium stabilności (podobnie jak w przypadku rewizji historycznej). Zasada konstrukcji tych podzbiorów mówi, że pierwszy z nich zawiera obserwacje pochodzące z analizowanego szeregu od pierwszej do $1 + k$ obserwacji, gdzie k jest długością każdego z podzbiorów. Kolejne podzbiory powinny obejmować podpróby, których pierwsza i ostatnia obserwacja są przesunięte o rok naprzód w stosunku do poprzedniej podpróby (dla drugiej podpróby będą to odpowiednio obserwacje $k + 13$ i $k + 25$; dla trzeciej: $k + 37$, $k + 49$, etc). Literatura przedmiotu zaleca podział analizowanej próby na maksymalnie 4 podzbiory obserwacji³⁹. Miarą, która pozwala na ocenę stabilności wyrównanego szeregu czasowego jest:

$$SS_t = \frac{\max_{k \in N_t} S_t(k) - \min_{k \in N_t} S_t(k)}{\min_{k \in N_t} S_t(k)} , \quad (16)$$

³⁸ Hood (2002), s. 2.

³⁹ Findley, Monsell, Shulman i Pugh (1990), s. 2.

gdzie:

$S_t(k)$ – czynnik sezonowy wyestymowany w k -tej podpróbie dla okresu t ,

$N_t = \{k: \text{okres } t\text{-ty znajduje się w } k\text{-tej podpróbie}\}$.

Alternatywną miarą, wykorzystywaną do oceny stabilności szeregu czasowego, bazującą na *sliding spans*, jest:

$$MM_t(k) = \frac{X_{-SA_t}(k)}{X_{-SA_{t-1}}(k)} - 1, \quad (17)$$

gdzie:

$X_{-SA_t}(k)$ – sezonowo wyrównana wartość z z próby k dla okresu t .

W odróżnieniu od poprzedniej statystyki, $MM_t(k)$ odnosi się do zmian z okresu na okres w wyrównanym szeregu czasowym. Następnie oblicza się maksymalną różnicę zmian miesięcznych dla odsezonowanego szeregu:

$$MM_t^{\max} = [\max_{k \in N_t} MM_t(k) - \min_{k \in N_t} MM_t(k)], \quad (18)$$

gdzie:

$N1_t = \{k: \text{okres } t\text{-ty i } t-1\text{-szy znajduje się w } k\text{-tej podpróbie}\}$.

Analogicznie można wyznaczyć zmiany procentowe odsezonowanego szeregu rok do roku:

$$YY_t(k) = \frac{X_{-SA_t}(k)}{X_{-SA_{t-2}}(k)} - 1, \quad (19)$$

które są niestabilne, jeśli zachodzi poniższy warunek:

$$YY_t^{\max} = \max YY_t(k) - \min YY_t(k) > 3\% \quad (20)$$

Niskie wartości powyższych miar wskazują, że wygładzony szereg jest stabilny. Jako wartość progową oznaczającą niestabilność dla statystyk SS_t , MM_t^{\max} , YY_t^{\max} przyjmuje się najczęściej 3%⁴⁰. Obserwacje, dla których miary SS_t , MM_t^{\max} , YY_t^{\max} przekraczają 3% uważa się za niestabilne. Jeśli łączny odsetek (odpowiednio SS , MM , YY) obserwacji niestabilnych przekracza dla poszczególnych miar odpowiednio 25%, 40% i 10%, wówczas procedurę odsezonowania należy uznać za nieprawidłową⁴¹.

Porównanie metod odsezonowania szeregów czasowych X-12-ARIMA i TRAMO/SEATS może być przeprowadzone także na podstawie miary (nie)zgodności. Bada ona stopień zgodności znaków między rocznymi stopami wzrostu sezonowo wyrównanych szeregów:

$$S_{-A_t}\% = \frac{X_{-SA_{t+12}} - X_{-SA_t}}{X_{-SA_t}}, \quad (21)$$

a odpowiednimi stopami wzrostu liczonymi na pierwotnych danych:

$$S_{-A_t}\% = \frac{X_{t+12} - X_t}{X_t}, \quad (22)$$

gdzie

t – numer obserwacji dla szeregu stóp wzrostu,

X_{-SA_t} – t -ta obserwacja szeregu wyrównanego sezonowo,

X_t – t -ta obserwacja szeregu niewyrównanego sezonowo.

Statystyka ta bada spójność między kierunkiem zmian danych surowych i odsezonowanych, rozumianą jako zgodność odpowiadających sobie stóp wzrostu tych dwóch szeregów. Miarą niezgodności znaków jest odsetek obserwacji, dla których stopy wzrostu

⁴⁰ Findley, Monsell, Shulman i Pugh (1990), s. 4-5.

⁴¹ Findley, Monsell, Shulman i Pugh (1990), s. 2-7.

wyrównanych danych (za pomocą TRAMO/SEATS lub X-12 ARIMA) miały inny kierunek niż stopy wzrostu z okresu na okres obliczone dla obserwacji pierwotnych. Metoda o niższym odsetku powinna zostać uznana za rekomendowaną do wyrównywania danych.

W przypadku analizy szeregu o silnym trendzie można spodziewać się, że ilość niezgodności stóp wzrostu obliczonych według wzorów (21) i (22) będzie stosunkowo niewielka ze względu na dominację trendu nad pozostałymi komponentami szeregu. Jednakże, jak wskazuje Peronaci⁴², nawet nieliczne przypadki braku zgodności są bardzo niepożądane. Szczególną uwagę na to zjawisko zwraca Europejski Bank Centralny, analizując dynamikę agregatów monetarnych.

W celu uzupełnienia analizy można posłużyć się także wartościami miar klasycznych:

– średnią wartość różnic pomiędzy stopą wzrostu dla danych pierwotnych a stopą wzrostu dla danych odsezonowanych jedną z metod, tj. X-12-ARIMA lub TRAMO/SEATS:

$$\bar{\Delta} = \frac{1}{N-12} \sum_{t=1}^{N-12} \left(\underbrace{S_{-}A_t \% - S_{-}X_t \%}_{a_k} \right), \quad (23)$$

– odchylenia standardowego powyższych różnic (a_k):

$$\delta = \sqrt{\frac{1}{N-12} \sum_{t=1}^{N-12} (a_t - \bar{\Delta})^2}. \quad (24)$$

Statystyki charakteryzujące się niższymi wartościami, wskazywać będą na lepszą metodę odsezonowania danych.

⁴² Peronaci (2003), s. 2.

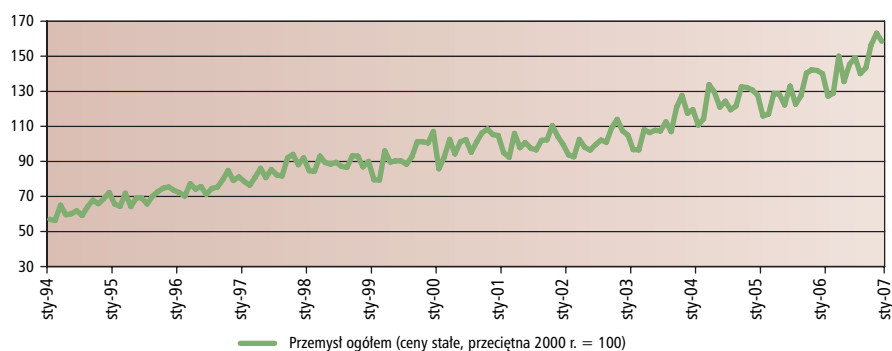
4

Analiza porównawcza na podstawie wyników empirycznych

4.1. Informacje podstawowe

Przedmiotem analizy empirycznej jest jakość odsezonowania długiego i krótkiego szeregu czasowego wyrównanych metodami TRAMO/SEATS i X-12-ARIMA. Badaniu poddano indeks produkcji sprzedanej przemysłu w Polsce; przeciętna miesięczna 2000 = 100 z okresu 1994.01 – 2006.11. Szczegółowo omówiono i zestawiono ze sobą kryteria jakości odsezonowania pełnej próby i krótkiego szeregu czasowego, obejmującego okres 2001.12 – 2006.11. Wyodrębnienie krótkiego szeregu czasowego poprzez wybranie najbardziej aktualnych obserwacji jest powszechnie spotykane w literaturze. Stosowane jest m.in. przez G. Mazzi i G. Savio. W celu dokonania bardziej kompleksowej analizy wyrównania sezonowego krótkich szeregów czasowych, z pełnej próby wyodrębniono 7 dalszych, pięcioletnich podprób.

Wykres 1. Indeks produkcji sprzedanej przemysłu, przeciętna miesięczna 2000=100



Źródło: opracowanie własne.

Poddany analizie szereg zobrazowano na wykresie 1. Wstępne badanie dynamiki produkcji przemysłowej wskazuje, że w analizowanym okresie czynniki wpływające na kształtowanie wielkości produkcji ulegały zmianie. Znacznym wahaniom podlegał popyt zarówno wewnętrzny, jak i zewnętrzny. Gospodarka doświadczyła silnego szoku, jakim był kryzys rosyjski na przełomie lat 1997 i 1998. Istotne zmiany odnotowano również w nakładach inwestycyjnych. Biorąc to wszystko pod uwagę można sformułować hipotezę, że wyniki wyrównania sezonowego uzyskane dla krótkiego szeregu czasowego mogą być determinowane przez okres, jaki obejmuje próba, a nie samą jej długość. Tym samym porównanie rezultatów odsezonowania pełnej próby ze skróconą mogą być zafałszowane wyborem okresu, z którego pochodzi krótki szereg. Spostrzeżenie to stało się przyczyną rozszerzenia niniejszej analizy o siedem dodatkowych krótkich szeregów (1994.12 – 1999.11, 1995.12 – 2000.11, 1996.12 – 2001.11, 1997.12 – 2002.11, 1998.12 – 2003.11, 1999.12 – 2004.11, 2000.12 – 2005.11). Jakość odsezonowania długiego szeregu czasowego została odniesiona przede wszystkim do wyników otrzymanych dla najbardziej aktualnego z krótkich szeregów, obejmującego okres 2001.12 – 2006.11. W tych częściach opracowania, w któ-

rych było to możliwe odniesiono się również do średniej wartości statystyk otrzymanych dla wszystkich krótkich szeregów czasowych. Pełne wyniki otrzymane dla tych szeregów zamieszczono w aneksie.

Pierwotnym źródłem danych jest Główny Urząd Statystyczny. Ze względu na zmianę przez GUS podstawy indeksu badanego szeregu z 1995=100 na 2000=100, do badania wykorzystano dane przeliczone wstecz z podstawą 2000=100. Sezonowe wyrównanie danych wykonano w pakiecie Demetra. Diagnostykę modeli przeprowadzano w programach SAS i EViews.

Podstawowe informacje dotyczące wygładzonych szeregów przedstawiono w tabeli 1. Uwzględniono w nich:

1. Typ modelu, jaki został zidentyfikowany w wyniku przeprowadzenia wygładzenia sezonowego.
2. Rodzaj transformacji dokonanej na danych wejściowych.
3. Zidentyfikowany efekt dni roboczych, przedstawiony jako ilość wykorzystanych zmiennych.
4. Efekt Wielkanocy, jako najbardziej charakterystyczny dla Polski efekt kalendarza.
5. Obecność outlierów poszczególnych typów.
6. Współczynnik skośności i kurtozy reszt estymowanego modelu.
7. Test Jarque-Berra na normalność składnika resztowego.

Tabela 1. Informacje podstawowe

Metoda		TRAMO/SEATS pełna próba	X-12-ARIMA pełna próba	TRAMO/SEATS 2001-2006	X-12-ARIMA 2001-2006
Wynik					
Model		(0,1,1)×(0,1,1)	(2,1,0)×(0,1,1)	(0,1,1)×(0,1,1)	(2,1,2)×(0,1,1)
Trans.		log	log	log	log
Efekt dni roboczych		7 regr.	7 regr.	7 regr.	1 regr.
Efekt Wielkanocy		nie	tak (6 dni)	nie	tak (6 dni)
Outliery*	TC	-	-	-	-
	LC	-	-	-	-
	AO	-	-	-	-
Skośność/Kurtoza		0,04/3,17	0,02/3,62	-0,10/2,85	-0,14/2,72
Test Jarque-Berra		0,20 [0,90]	2,54 [0,28]	0,15 [0,93]	0,38 [0,83]
Test Ljungu –Boxa		20,52** [0; 33,9]	17,78** [0; 51,2]	30,48** [0; 33,9]	23,86*** [0; 51,20]

Uwagi: wartości podane w nawiasach kwadratowych przy teście Jarque-Berra oznaczają prawdopodobieństwo P, traktowane jako najniższy poziom istotności, przy którym następuje odrzucenie testowanej hipotezy (tu: o normalności składnika losowego). Przy każdym poziomie istotności mniejszym od danego prawdopodobieństwa P, nie ma podstaw do odrzucenia hipotezy zerowej.

* TC oznacza outlier typu *temporary change*, LS – *level shift*, AO – *additive outlier*.

** oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej o braku autokorelacji na poziomie istotności 5%,

*** oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej na poziomie istotności 10%.

W nawiasach kwadratowych podano przedział ufności danego estymatora.

Źródło: opracowanie własne.

W każdym przypadku na danych pierwotnych została dokonana transformacja logarytmiczna. Dla obu szeregów wyrównanych metodą X-12-ARIMA efekt Wielkanocy okazał się istotny, a jego oddziaływanie oszacowano na 6 dni. Zaprezentowane powyżej wyniki nie wskazują na obecność outlierów⁴³. Otrzymane wartości skośności i kurtozy oraz wartość statystyki Jarque-Berra nie dają podstaw do odrzucenia hipotezy zerowej o normalnym rozkładzie reszt dla każdego z wykonanych oszacowań. Wartość statystyki

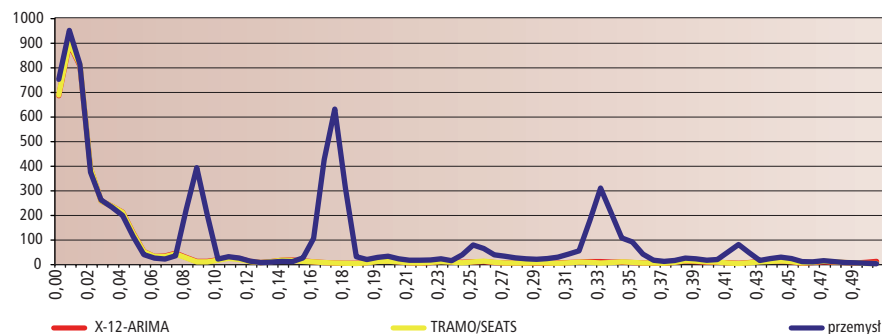
⁴³ Zamieszczone w aneksie wyniki wskazują, że w przypadku niektórych szeregów czasowych wykryto outlierzy typu addytywnego. Uznanie obserwacji za nietypową jest uzależnione m.in. od wartości przyjmowanych przez pozostałe obserwacje szeregu.

Ljunga-Boxa obliczona na resztach, nie wskazuje na obecność w nich liniowej struktury autokorelacji dla poszczególnych estymacji. We wszystkich analizowanych przypadkach w sposób satysfakcjonujący zostały wyeliminowane wahania sezonowe, co potwierdzają widma szeregów (wykres 2 i 3).

Na poniższych wykresach zamieszczono funkcje gęstości widmowych danych surowych i odsezonowanych obiema metodami⁴⁴. Lokalne maksima gęstości widmowej nieodsezonowanego szeregu odpowiadają częstotliwościom sezonowym, mającym postać $k/12$, gdzie k oznacza liczbę impulsów w roku, przy czym $1/12$ odpowiada częstotliwości rocznej, $2/12$ – półrocznej, $4/12$ – kwartalnej, $3/12$ oznacza występowanie anomalii z częstotliwością trzech razy w roku, a $6/12$ – co dwa miesiące⁴⁵. Badania empiryczne wskazują, że efekt dni roboczych odpowiada częstotliwościom $4,176/12$ oraz $5,184/12$, czyli w ciągu roku obserwowanych jest średnio odpowiednio 4,176 oraz 5,184 impulsów⁴⁶.

Na podstawie poniższego wykresu widma można przypuszczać, że w długiej próbie występuje sezonowość o częstotliwości rocznej, półrocznej i kwartalnej. Lokalne maksimum występuje również w okolicy częstotliwości dni roboczych. Porównanie widm otrzymanych po wyrównaniu szeregu metodą X-12-ARIMA i TRAMO/SEATS wskazuje, że w obu przypadkach efekty sezonowe i kalendarzowe zostały usunięte.

Wykres 2. Gęstość widmowa – pełna próba



Źródło: opracowanie własne.

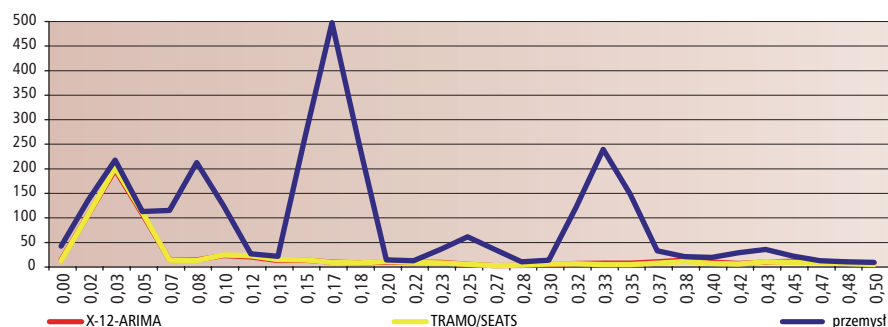
W krótkim szeregu czasowym również zostały zidentyfikowane ekstrema, wskazujące na obecność sezonowości rocznej, półrocznej oraz kwartalnej. Podobnie jak w przypadku pełnej próby najsilniejszy efekt sezonowy występuje dla częstotliwości półrocznej. Pomimo niewielkiej ilości dostępnych obserwacji analiza widma wskazuje na występowanie tych samych częstości sezonowych, które zostały wykryte dla pełnej próby. Wynik ten sugeruje, że efekty sezonowe i kalendarzowe w badanym szeregu są silne i regularne. Również w tym przypadku obie metody odsezonowania w niemal identycznym stopniu wyeliminowały występujące wahania sezonowe.

⁴⁴ Do sporządzenia wykresu gęstości i widmowej wykorzystano okno Bartletta. Ze względu na konieczność spełnienia warunku stacjonarności szeregu poddawanego analizie widmowej, z analizowanych szeregów usunięto trend oraz odjęto średnią.

⁴⁵ Obserwowana amplituda widma może wynikać z nakładania się efektów pochodzących z różnych częstotliwości. Z twierdzenia Nyquista wynika, że dla szeregu mierzonego z częstotliwością x możliwe jest obserwowanie zjawisk pojawiających się z częstotliwością co najwyżej $2x$. Z tego względu dla szeregu o częstotliwości miesięcznej maksymalną wartością k jest 6.

⁴⁶ Hood (2002), s. 2.

Wykres 3. Gęstość widmowa – próba 2001–2006



Źródło: opracowanie własne.

4.2. Badanie stabilności modelu

W tabelach 2-4 przedstawiono miary *SS*, *MM*, *YY* służące ocenie stabilności modelu zgodnie z metodą *sliding spans*, które skonstruowano na podstawie opisu zawartego w rozdziale 3 na stronach 18-19 niniejszego opracowania. Oznaczają one odpowiednio łączny odsetek niestabilnych obserwacji liczonych: wewnątrz wszystkich podprób, z miesiąca na miesiąc, z roku na rok. W analizie posłużono się konstrukcją złożoną z 4 podprób (*spanów*) w przypadku długiego szeregu oraz 2 podprób w przypadku szeregu skróconego. Pierwsza z podprób dla całego szeregu czasowego obejmowała okres od grudnia 1994 r. do listopada 2003 r., druga – od grudnia 1995 r. do listopada 2004 r., trzecia – od grudnia 1996 r. do listopada 2005 r. i ostatnia od grudnia 1997 r. do listopada 2006 r. Analogicznie przebiegała konstrukcja *spanów* dla niepełnej próby, pierwszy *span* obejmował okres od grudnia 2001 r. do listopada 2005 r., drugi od grudnia 2002 r. do listopada 2006 r. Na tej samej zasadzie utworzono podpróby dla siedmiu pozostałych krótkich szeregów.

W analizie stabilności uwzględniono, oprócz miar klasycznych, także rozbieżności w oszacowaniach czynników sezonowych pomiędzy wszystkimi podpróbami, analogicznie do kryterium (nie)zgodności pomiędzy pierwotnym a wygładzonym szeregiem. Uwzględniono przy tym wszystkie możliwe obserwacje uzyskane dla każdego ze *spanów*. Porównania stóp wzrostu czynników sezonowych miały określić rozbieżności pomiędzy szeregami odsezonowanymi poszczególnymi metodami. W tym celu wyznaczono odsetek obserwacji niezgodnych, tj. tych, dla których czynnik sezonowy w co najmniej jednym ze *spanów* przyjmował wartości większe (mniejsze) od 1, natomiast w pozostałych *spanach* przyjmował on wartości mniejsze (większe) od 1⁴⁷. Rozbieżność nie występuje wtedy i tylko wtedy, kiedy we wszystkich podpróbach komponent sezonowy przyjmował wartości wskazujące na ten sam znak stopy wzrostu.

Na podstawie danych zawartych w tabeli 2 można wywnioskować, że zdecydowanie lepsze wartości poszczególnych statystyk, tj. najniższą wartość oczekiwaną, odchylenie standardowe, łączny odsetek niestabilnych obserwacji (*SS*), liczbę i odsetek niezgodnych stóp wzrostu, uzyskano przy zastosowaniu metody X-12-ARIMA dla pełnej próby. Jednak w przypadku krótkiego szeregu czasowego lepszą metodą okazało się TRAMO/SEATS. Co prawda odsetek rozbieżności w oszacowaniach czynników sezonowych nie charakteryzuje się tak dużą przewagą względną jak w przypadku przewagi X-12-ARIMA nad TRAMO/SEATS dla długiej próby. Porównanie w obrębie metody TRAMO/SEATS, przy uwzględnieniu miar rozbieżności oszacowań oraz odchylenia standardowego, wskazuje na większą stabilność estymowanego modelu w przypadku długiego szeregu. Dodatkowo można stwierdzić, że

⁴⁷ W przypadku metody TRAMO/SEATS czynnik sezonowy oscyluje wokół 100. Na potrzeby niniejszego opracowania dokonano jego normalizacji do 1, ze względu na zapewnienie porównywalności z X-12-ARIMA. Różnica w zapisie czynnika sezonowego między X-12-ARIMA a TRAMO/SEATS jest związana tylko z konwencją przyjętą w programie *Demetra*.

średnia niestabilnych komponentów sezonowych w każdym z analizowanych przypadków nie była większa od wartości progowej 3%⁴⁸. Dopuszczalny odsetek wszystkich niestabilnych obserwacji w każdym przypadku przyjął wartość równą 0%. Pozwala to stwierdzić, że estymacja została przeprowadzona poprawnie (SS powinien być mniejszy od 25%).

Tabela 2. Sliding spans – analiza czynnika sezonowego SS

Wynik Metoda	Średnia	Maksimum	Odchylenie stand.	SS (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach czynników sezonowych	
					w miesiącach *	w %
TRAMO/SEATS pełna próba	0,009	0,029	0,006	0,0%	13/120	10,8%
X-12-ARIMA pełna próba	0,005	0,020	0,003	0,0%	3/120	2,5%
TRAMO/SEATS 2001-2006	0,006	0,016	0,004	0,0%	16/36	44,4%
X-12-ARIMA 2001-2006	0,010	0,024	0,006	0,0%	18/36	50,0%

Uwagi: * ilość okresów, dla których wykryto niezgodność znaków stóp wzrostu w stosunku do liczby wszystkich analizowanych obserwacji.

Źródło: opracowanie własne.

W tabelach 3 i 4 przedstawiono analizę komponentu sezonowego metodą *sliding spans* odpowiednio miesiąc do miesiąca (*MM*) i rok do roku (*YY*). W tych przypadkach lepszą metodą dla długiego szeregu czasowego okazuje się ponownie X-12-ARIMA, natomiast dla krótkiego szeregu czasowego TRAMO/SEATS. Co prawda, w przypadku długiego szeregu, wartości średnie nie potwierdzają tych wniosków, jednak odchylenia standardowe dla miar *MM* i *YY* oraz łączny odsetek niestabilnych obserwacji mieszczący się w dopuszczalnym przedziale (odpowiednio (0%; 40%) i (0%; 10%)) przyjmują relatywnie mniejsze wartości. Problemy z zachowaniem stabilności pojawiają się jedynie w przypadku krótkiego szeregu czasowego odsezonowywanego metodą X-12-ARIMA (*YY* = 12,5%). Analogiczne wnioski uzyskano, analizując rozbieżności w oszacowaniach miesięcznych stóp wzrostu. Dysproporcje te są odpowiednio prawie o połowę mniejsze dla krótszego szeregu odsezonowanego za pomocą TRAMO/SEATS niż X-12-ARIMA i o kilka punktów procentowych mniejsze w przypadku wyboru metody X-12-ARIMA niż TRAMO/SEATS dla długiego szeregu czasowego.

Tabela 3. Analiza miesięcznych stóp wzrostu danych odsezonowanych MM

Wynik Metoda	Średnia	Maksimum	Odchylenie stand.	MM (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach miesięcznych stóp wzrostu	
					w miesiącach	w %
TRAMO/SEATS pełna próba	0,095	6,790	1,903	2,8%	28/119	23,5%
X-12-ARIMA pełna próba	0,827	2,143	0,504	0,0%	24/119	20,2%
TRAMO/SEATS 2001-2006	0,788	2,312	0,620	0,0%	5/35	14,29%
X-12-ARIMA 2001-2006	2,103	6,491	1,763	28,6%	13/35	37,14%

Uwagi: * ilość okresów, dla których wykryto niezgodność znaków stóp wzrostu w stosunku do liczby wszystkich analizowanych obserwacji.

Źródło: opracowanie własne.

⁴⁸ Zgodnie z metodologią oceny stabilności modelu za pomocą *sliding spans*, przedstawionej w rozdziale 3, za niestabilną obserwację uznać należy taką, dla której miara $SS_t > 3\%$.

Większe dysproporcje są widoczne w oszacowaniach otrzymanych dla rocznych stóp wzrostu rozpatrywanych podprób zarówno dla pełnego szeregu czasowego, jak i szeregu okrojonego. W przypadku zastosowania metody X-12-ARIMA dla pełnej próby tylko jedna stopa wzrostu była niezgodna z pozostałymi *spanami*, natomiast w przypadku metody TRAMO/SEATS rekomendowanej dla krótkiego szeregu indeksu produkcji sprzedanej zaobserwowano dwie niezgodne obserwacje między podpróbami.

Tabela 4. Analiza rocznych stóp wzrostu danych odsezonowanych YY

Metoda	Wynik	Średnia	Maksimum	Odchylenie stand.	YY (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach rocznych stóp wzrostu	
						w miesiącach	w %
TRAMO/SEATS pełna próba		0,095	6,790	1,903	2,8%	4/108	3,70%
X-12-ARIMA pełna próba		0,170	4,670	0,663	0,9%	1/108	0,9%
TRAMO/SEATS 2001-2006		0,277	9,545	2,254	4,2%	2/24	8,33%
X-12-ARIMA 2001-2006		-1,111	41,677	12,358	12,5%	10/24	41,67%

Źródło: opracowanie własne.

Zauważyć przy tym należy, zgodnie z wynikami zaprezentowanymi w aneksie, że dobór próby okrojonej z końca analizowanego okresu nie przyczynił się do zniekształcenia uzyskanych rezultatów przy porównaniu wyników w kontekście długości próby. Średnie statystyk dla rocznych YY i miesięcznych stóp wzrostu MM danych odsezonowanych przyjmują niższe wartości w przypadku metody TRAMO/SEATS. Przewaga ta, w odniesieniu do odsetka obserwacji niestabilnych (patrz aneks), kształtuje się odpowiednio:

- dla miary MM – 4,64% TRAMO/SEATS w stosunku do 22,1% X-12-ARIMA,
- dla miary YY – 5,7% w stosunku do 7,3%.

Analiza *sliding spans* dostarcza jednoznacznych wniosków w przypadku rozpatrywania kontekstu długości próby w procesie odsezonowania. Wartości miar (SS, YY) wskazują na lepszą jakość wyrównania na długich szeregach czasowych. Odsetek niestabilnych składników sezonowych na poziomie 0%, liczony zgodnie z zaprezentowaną metodologią, wskazuje na stabilność podrób. Zdecydowaną przewagę widać w analizie niestabilnych obserwacji liczonych rok do roku, które są kilka razy mniejsze dla długich szeregów (odpowiednio 0,028 dla TRAMO/SEATS i 0,009 dla X-12-ARIMA) niż dla krótkich (odpowiednio 0,042 dla TRAMO/SEATS i 0,125 dla X-12-ARIMA).

Drugim kryterium służącym ocenie stabilności estymowanego modelu są rewizje (korekty). W tabeli 5 przedstawiono wyniki uzyskane na podstawie analizy średniej i odchylenia standardowego przeprowadzonych rewizji. Najniższe wartości miar klasycznych otrzymano dla metody TRAMO/SEATS. Dla X-12-ARIMA w przypadku pełnej próby uzyskano względnie dużą wartość odchylenia od średniej (0,32 w porównaniu z 0,20), co wskazuje na gorszą jakość dopasowania. Zdecydowanie najlepsze wyniki uzyskano przy wygładzeniu długiego szeregu metodą TRAMO/SEATS. Zauważyć należy, że średnia arytmetyczna nie jest w tym przypadku pełnym odzwierciedleniem jakości metody, ponieważ na wartość tej miary wpływają znaki poszczególnych jej składników (tzn. ujemne wartości korekt kompensują się z dodatnimi). Z tego względu decydującym kryterium jakościowym jest tu wielkość odchylenia standardowego.

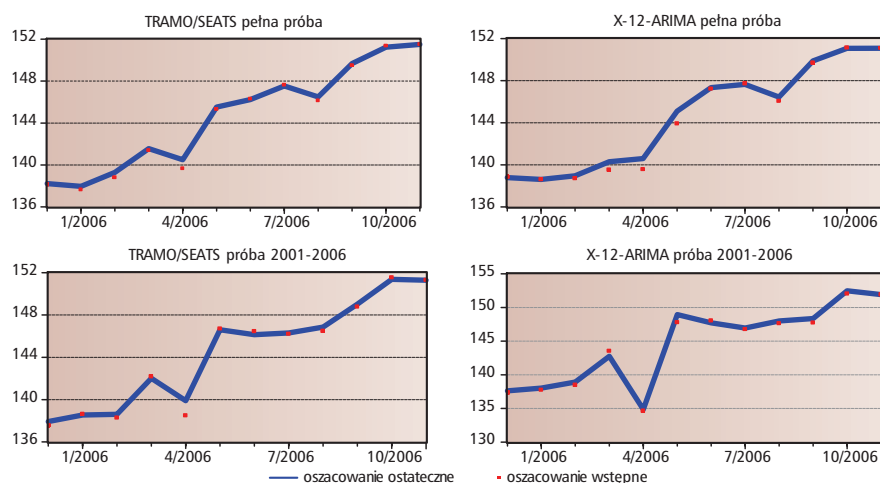
Tabela 5. Rewizje historyczne – dane odsezonowane

Metoda	Wynik	Średnia w %	Odchylenie standardowe w %
TRAMO/SEATS pełna próba		0,15	0,20
X-12-ARIMA pełna próba		0,24	0,32
TRAMO/SEATS 2001-2006		0,14	0,34
X-12-ARIMA 2001-2006		0,19	0,34

Źródło: opracowanie własne.

Reasumując, wyniki otrzymane dla kryterium rewizji wskazują, że rekomendowaną metodą odsezonowania długiego szeregu czasowego w analizowanym przypadku powinna być metoda TRAMO/SEATS.

Wykres 4. Rewizje szeregu odsezonowanego – oszacowanie wstępne vs ostateczne



Źródło: opracowanie własne.

Na powyższych wykresach przedstawiono oszacowanie wstępne i ostateczne odsezonowanych wartości indeksu produkcji sprzedanej przemysłu dla metod TRAMO/SEATS i X-12-ARIMA zarówno dla pełnej, jak i krótkiej próby obejmującej okres od grudnia 2001 r. do listopada 2006 r. Oszacowania wstępne oznaczono punktami, natomiast liniami ciągłymi przedstawiono oszacowanie ostateczne. Dla pełnego szeregu czasowego mniejsze wartości rewizji otrzymano, stosując metodę TRAMO/SEATS. Okresami, w których wartości wstępnie oszacowane tą metodą odbiegają nieznacznie od wartości ostatecznych, są kwiecień i sierpień 2006 r. Dla X-12-ARIMA w marcu, kwietniu, maju i sierpniu 2006 r. rewizje są istotnie większe od 0. Metoda ta systematycznie zaniża wartość ostatecznego oszacowania. TRAMO/SEATS w porównaniu z X-12-ARIMA powinna być rekomendowana do wyrównania długiego szeregu czasowego. Warto podkreślić, że wygładzanie sezonowe krótkich szeregów czasowych ma mniejsze wartości rewizji w przypadku obu metod. Wizualna ocena nie pozwala jednak na wyłonienie lepszej metody. Wartości odchylenia standardowego w przypadku rewizji danych odsezonowanych również nie pozwalają na sformułowanie jednoznacznej oceny. Metoda X-12-ARIMA okazała się minimalnie lepszą w przypadku krótkich szeregów czasowych (wartości odchylenia standardowego są w tym przypadku bardzo zbliżone – różnica wynosi zaledwie 0,001). Wnioski te potwierdzają wyniki przedstawione w aneksie w tabelach 5.1A oraz 5.2A. Podobnie zatem, jak w przypadku kryterium *sliding spans*, arbitralny dobór okresu krótkiej próby nie wpłynął na porównanie jakości odsezonowania.

Analogiczne wnioski można wysnuć z analizy statystyk dla komponentu trend-cykl przedstawionych w tabeli 6.

Tabela 6. Rewizje historyczne – trend

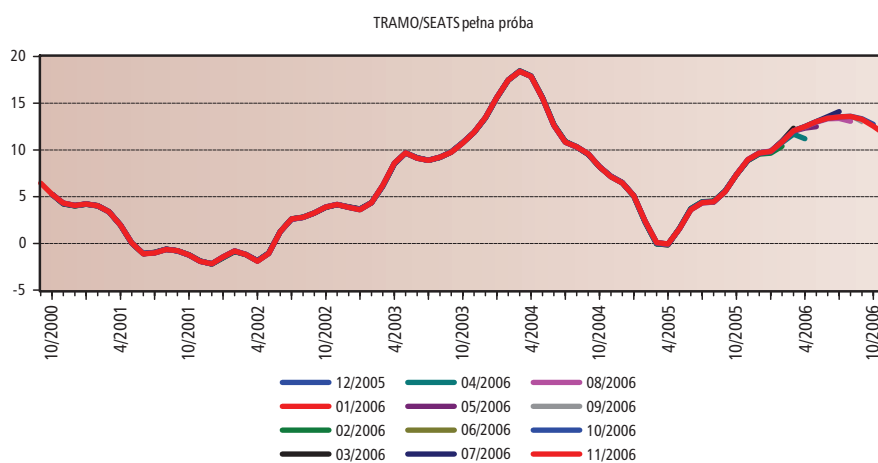
Metoda	Wynik	Średnia w %	Odchylenie standardowe w %
TRAMO/SEATS pełna próba		0,24	0,46
X-12-ARIMA pełna próba		0,29	0,54
TRAMO/SEATS 2001-2006		0,20	0,64
X-12-ARIMA 2001-2006		0,36	0,98

Źródło: opracowanie własne.

Na wykresach 5-8 przedstawiono rewizje dla okresu od grudnia 2005 r. do listopada 2006 r. uzyskane dla stóp wzrostu komponentu trend-cykl. Poszczególne kolorowe, ciągłe linie przedstawiają estymacje dokonane na podstawie próby obejmującej odpowiednio: wrzesień 1994 – grudzień 2005; wrzesień 1994 – styczeń 2005; wrzesień 1994 – luty 2005; wrzesień 1994 – marzec 2005; wrzesień 1994 – kwiecień 2006; wrzesień 1994 – maj 2006; wrzesień 1994 – czerwiec 2006; wrzesień 1994 – lipiec 2006; wrzesień 1994 – sierpień 2006; wrzesień 1994 – wrzesień 2006; wrzesień 1994 – październik 2006; wrzesień 1994 – listopad 2006.

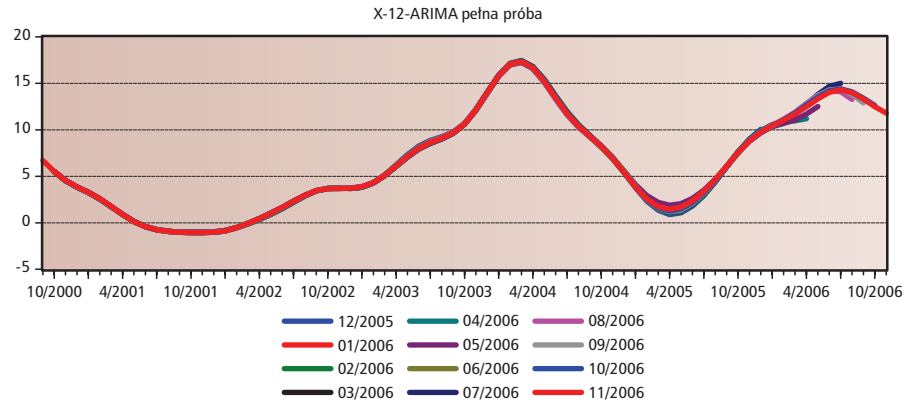
Comiesięcznej estymacji dokonywano na podstawie postaci modelu ustalonej w procesie odsezonowania szeregów danych obejmujących okres od września 1994 r. do grudnia 2005 r. W obu metodach można zaobserwować wzrost wartości korekt w końcowej części próby. Ze względu na występujące rozbieżności w okresach od maja do lipca 2003 r. oraz w okresie od marca do czerwca 2005 r. (wykres 6) należy uznać, że dla składnika trend-cykl w przypadku pełnej próby stosunkowo lepszą metodą wygładzania jest TRAMO/SEATS. Szczególnie duże wartości rewizji zaobserwowano w otoczeniu punktów przegięcia wykresu rocznych stóp wzrostu (kwiecień–lipiec 2003 oraz marzec–sierpień 2005).

Wykres 5. Rewizje TRAMO/SEATS pełna próba – roczne stopy wzrostu trendu-cyklu



Źródło: opracowanie własne.

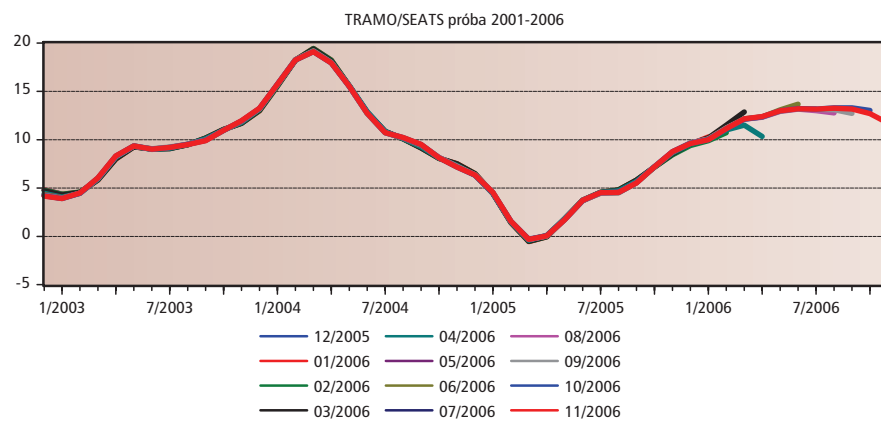
Wykres 6. Rewizje X-12-ARIMA pełna próba – roczne stopy wzrostu trendu-cyklu



Źródło: opracowanie własne.

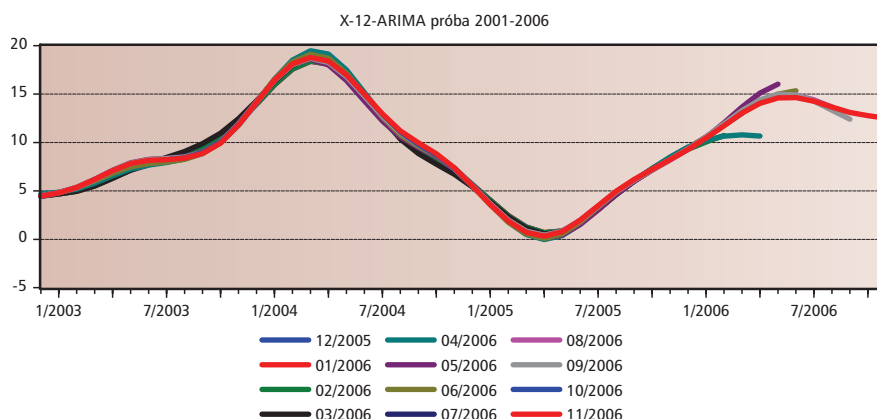
Analiza rewizji stóp wzrostu komponentu trend-cykl dla krótkich szeregów czasowych przedstawionych na wykresach 7-9 potwierdza poczynione wcześniej spostrzeżenia dotyczące rewizji danych dla pełnej próby. Ponownie zaobserwowano, że metoda TRAMO/SEATS ma mniejsze wartości korekt. Rewizje uzyskane dla krótkiego szeregu są jednak większe niż w przypadku analizy pełnej próby indeksu produkcji sprzedanej przemysłu. Zauważmy, że podobnie jak w przypadku pełnej próby, większe rewizje dla niepełnej próby były widoczne w otoczeniach punktów przegięcia wykresu (kwiecień – sierpień 2003, luty – czerwiec 2004 oraz luty – lipiec 2005), co może świadczyć o gorszych właściwościach dopasowania metody X-12-ARIMA w okresach zmiennej tendencji wzrostowej szeregu. Wnioski te potwierdza także wizualna ocena rewizji stóp wzrostu komponentu trend-cykl dla pozostałych siedmiu krótkich szeregów czasowych zamieszczonych w aneksie.

Wykres 7. Rewizje TRAMO/SEATS krótki szereg – roczne stopy wzrostu trendu-cyklu



Źródło: opracowanie własne.

Wykres 8. Rewizje X-12-ARIMA krótki szereg – roczne stopy wzrostu trendu-cyklu



Źródło: opracowanie własne.

4.3. Kryterium zgodności

Dobrym przybliżeniem krótkookresowych zmian w tendencjach wielkości ekonomicznych mogą być stopy wzrostu. Ważnym jest, aby poszczególne metody odsezonowania dawały spójny obraz tendencji rozwojowej danego zjawiska, a zatem aby analiza znaków stóp wzrostu wskazywała na jak największy odsetek zgodności. Należy jednak pamiętać, że przy analizowaniu szeregów, które mają silny trend, problem różnicy znaków dynamik praktycznie nie istnieje. Miary niezgodności danych odsezonowanych metodami TRAMO/SEATS i X-12-ARIMA przedstawione zostały w tabeli 7.

Największy odsetek zgodnych znaków zanotowano dla szeregów skróconych. W przypadku metody TRAMO/SEATS kierunek zmian stóp wzrostu w 97,92% zgadzał się z oryginalnym szeregiem. Podobne wnioski wypływają z analizy wartości odchylenia standardowego i maksimum różnic rocznych stóp wzrostu pomiędzy danymi surowymi a odsezonowanymi. Z porównania odsezonowanych różnymi metodami długiego i krótkiego szeregu można wywnioskować, że lepsze wyniki daje metoda TRAMO/SEATS na krótkim szeregu. Należy jednak zaznaczyć, iż rezultatu tego nie można uogólnić na pozostałe analizowane próby. Zgodność znaków wynosiła średnio 91,41% w przypadku TRAMO/SEATS i była niemal identyczna ze średnią otrzymaną dla X-12-ARIMA (patrz aneks). Zauważalny spadek stopnia zgodności znaków jest prawdopodobnie spowodowany szokami, jakim poddawane było badane zjawisko przed 2001 r.

Tabela 7. Różnica rocznych stóp wzrostu danych surowych i odsezonowanych

Metoda	Wynik	Średnia różnica stóp wzrostu	Odchylenie stand. różnic stóp wzrostu	Maksimum różnic stóp wzrostu	Procent zgodnych znaków
TRAMO/SEATS pełna próba		0,000	0,021	0,051	93,01%
X-12-ARIMA pełna próba		0,000	0,022	0,046	92,31%
TRAMO/SEATS 2001-2006		0,001	0,017	0,042	97,92%
X-12-ARIMA 2001-2006		0,001	0,015	0,037	95,83%

Źródło: opracowanie własne.

4.4. Gładkość dopasowania komponentów

Miary gładkości dopasowania – R_1 , R_2 , $MAR (TC)$ przedstawione zostały tabeli 8. Odnoszą się one odpowiednio do odsezonowanych danych, komponentu nieregularnego oraz komponentu trend-cykl. Jak już wspomniano, miara określająca gładkość dopasowania jest istotnym wyznacznikiem poprawności zastosowanej metody odsezonowania. Jej mniejsza wartość świadczy o lepszym wygładzeniu szeregu. Na podstawie uzyskanych wyników nie można jednoznacznie wskazać, która z zastosowanych metod jest lepsza. Najbardziej zadowalające dopasowanie odsezonowanego szeregu uzyskano dla pełnych prób (miara R_1). Nieznacznie lepszymi wynikami charakteryzowała się X-12-ARIMA przy szacowaniu miary R_1 dla ostatnich 3 lat próby, jednak TRAMO/SEATS uzyskało przewagę w przypadku obliczenia miary na całej próbie. Analogiczne wnioski nasuwają się w przypadku analizy miary gładkości dla komponentu trend-cykl ($MAR (TC)$). Miara gładkości komponentu nieregularnego (R_2) osiągnęła niższe wartości dla długich szeregów, ze znaczną przewagą metody TRAMO/SEATS. Zauważyć należy także, że miary zastosowane dla ostatnich 3 lat, przyjmują niejednokrotnie większe wartości niż miary stosowane dla całego okresu. Świadczyć to może o słabszym dopasowaniu szeregu na końcu próby przy użyciu każdej z metod. Generalnie, w przypadku miar R_1 i R_2 wyższą jakość dopasowania osiągnięto w przypadku długich szeregów, co wskazuje na występowanie większych zaburzeń i mniejszej przewidywalności zachowania szeregu na końcu próby. Natomiast średnie wygładzenie składnika trendu ($MAR (TC)$) jest lepsze dla krótkiego szeregu.

Tabela 8. Miary gładkości dopasowania komponentów

Metoda	Wynik	R_1	R_1 (ostatnie 3 lata)	R_2	R_2 (ostatnie 3 lata)	MAR (TC)
TRAMO/SEATS pełna próba		5,3	8,3	1,0	1,0	111,6
X-12-ARIMA pełna próba		5,6	8,1	1,8	2,5	70,1
TRAMO/SEATS 2001-2006		6,8	8,0	2,3	2,9	83,7
X-12-ARIMA 2001-2006		8,1	10,3	2,7	3,6	72,8

Źródło: opracowanie własne.

4.5. Względny błąd prognozy

Prognozy *ex post* na okres grudzień 2005 – listopad 2006 dokonano na podstawie próby obejmującą okres od września 1994 r. do listopada 2005 r. w przypadku pełnej próby. Dla krótkiej próby prognozę *ex post* na okres grudzień 2005 – listopad 2006 wykonano na podstawie obserwacji od grudnia 2000 r. do listopada 2005 r. Aby wyznaczyć względny błąd prognozy, posłużono się statystyką *MAPD* (*mean absolute percentage deviation*) obliczoną dla szeregu odsezonowanego oraz dla komponentu trendu-cykl na podstawie obserwacji z ostatnich dwunastu miesięcy badanej próby. *MAPD* mierzy względny absolutny średni błąd prognozy, tzn. dla szeregu odsezonowanego i trendu-cyklu obliczana jest odpowiednio jako:

$$MAPD_{(SA)} = \frac{1}{12} \sum_{t=N_F-11}^{N_F} \left| \frac{X_{-SA_t} - X_{-SA_t}^F}{X_{-SA_t}} \right|, \quad (25)$$

$$MAPD_{(TC)} = \frac{1}{12} \sum_{t=N_F-11}^{N_F} \left| \frac{TC_t - TC_t^F}{TC_t} \right|. \quad (26)$$

gdzie:

$X_{-SA_t}^F$ – prognozowana wartość odsezonowanego szeregu na okres t ,

TC_t^F – prognozowana wartość składnika trend-cykl na okres t ,

N_F – horyzont prognozy.

Tabela 9. Względny błąd prognozy dla szeregu odsezonowanego i komponentu trend-cykl

Metoda \ Wynik	$MAPD_{(SA)}$	$MAPD_{(TC)}$
TRAMO/SEATS pełna próba	2,6	2,6
X-12-ARIMA pełna próba	2,9	2,7
TRAMO/SEATS 2001-2006	2,2	2,1
X-12-ARIMA 2001-2006	0,0	0,0

Źródło: opracowanie własne.

Średnie względne błędy prognozy dla krótkiej próby przyjmują relatywnie małe wielkości dla X-12-ARIMA, co świadczy o dobrych własnościach predykcyjnych tej metody. W odniesieniu do szeregu odsezonowanego i komponentu trend-cykl zdecydowanie mniejsze błędy względne uzyskano stosując metodę X-12-ARIMA dla krótkiego szeregu. Wartości miary $MAPD$ obliczonej dla skróconego szeregu wskazują na ponad dwie jednostki większy błąd względny prognozy przy zastosowaniu metody TRAMO/SEATS w porównaniu z X-12-ARIMA. Analiza zamieszczonych w aneksie średnich wartości $MAPD$ również wskazuje na wyraźną przewagę X-12-ARIMA nad TRAMO/SEATS. Należy jednak zauważyć, że średnie zamieszczone w aneksie pochodzą z wyników uzyskanych dla prób obejmujących różne okresy, tym samym porównywalność tej miary jest w tym przypadku ograniczona.

Dla długiego szeregu uzyskano mniejszy błąd prognozy danych odsezonowanych przy zastosowaniu TRAMO/SEATS niż X-12-ARIMA. Wartości $MAPD_{(TC)}$ są bardzo zbliżone dla obu metod, różnica między nimi wynosi 0,1.

4.6. Kryterium idempotencji

Z wyjątkiem modelu TRAMO/SEATS wyestymowanego na pełnej próbie, we wszystkich analizowanych przypadkach oczyszczenie szeregu z wahań sezonowych nie było w pełni satysfakcjonujące. Automatyczna procedura ponownej estymacji szeregów jednokrotnie odsezonowanych w każdym z przypadków wskazywała bowiem na sezonowy model ARIMA. Dla modeli oszacowanych na pełnym zakresie danych wykryto również obecność outlierów oraz efektów dni roboczych. Rezultaty te mogą być wynikiem zarówno nieprawidłowego odsezonowania szeregu jak i niedoskonałości procedury estymacyjnej. W przypadku odsezonowania metodą X-12-ARIMA dobrany został model sezonowy. Niezadowolające wyniki otrzymano również po zastosowaniu metody TRAMO/SEATS dla skróconego szeregu. W szeregu tym wykryta została sezonowość. Jak wskazuje C. Hood, algorytm programu SEATS często wykrywa sezonowość w danych pozbawionych komponentu sezonowego⁴⁹. Zarówno X-12-ARIMA, jak i TRAMO/SEATS nie są w pełni kompletne pod względem wykrywania i usuwania sezonowości. Eksperymenty polegające na automatycznym wyrównaniu procesów błędzenia przypadkowego wskazują na przewagę TRAMO/SEATS, przy czym skuteczność metody wynosi około 70%⁵⁰. Dlatego też oprócz analizy dokonanej w pakiecie *Demetra* posłużono się dodatkowo opisaną poniżej, bardziej wiarygodną miarą idempotencji⁵¹ (R) oraz analizą wariancji składnika sezonowego dla szeregu jednokrotnie i dwukrotnie odsezonowanego.

⁴⁹ Hood (2002), s. 4.

⁵⁰ European Central Bank (2000), s. 57.

⁵¹ Kryterium to jest stosowane m.in. przez Fischer (1995), s. 42.

Tabela 10. Ponowne odsezonowanie danych jednokrotnie wyrównanych sezonowo

Wynik Metoda	Model	Sezonowość	Outliery			Efekt dni roboczych	Efekt Wielkono- cności	Test Ljunga – Boxa	R	Warian- cja S/S(S)
			TC	LS	AO					
TRAMO/ SEATS pełna próba	(1,1,0) × (0,0,0)	nie- istotna	-	-	maj-02	nie	tak	16,53** [0; 35,2]	0,021	0,002 /0,000
X-12- ARIMA pełna próba	(0,1,2) × (0,1,1)	nie- istotna	-	-	-	nie	nie	25,65* [0; 51,20]	0,022	0,002 /0,000
TRAMO/ SEATS 2001- 2006	(1,1,0) × (0,1,1)	istotna	-	-	-	nie	nie	11,78** [0; 33,90]	0,225	24,287 /0,229
X-12- ARIMA 2001- 2006	(2,1,0) × (0,1,1)	nie- istotna	-	-	-	nie	nie	20,4* [0; 51,2]	0,068	0,002 /0,069

Uwagi: *oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej na poziomie istotności 5%.

** oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej na poziomie istotności 10%.

W nawiasach kwadratowych podano przedział ufności danego estymatora

Źródło: opracowanie własne.

Statystyka R mierzy stopień rozbieżności danych jedno- i dwukrotnie wyrównanych. Jej konstrukcja opiera się na założeniu, że powtórne zastosowanie procedury odsezonowania nie powinno zasadniczo zmienić obserwacji wynikowego szeregu. Statystyka obliczana jest ze wzoru:

$$R = \frac{1}{T} \sum_{t=1}^T (X_SA_t - X_SA_{tSA})^2, \quad (27)$$

gdzie:

X_SA_t – dane wyrównane sezonowo,

X_SA_{tSA} – dane poddane ponownemu odsezonowaniu.

Wyniki przedstawione w tabeli 10 pozwalają stwierdzić, że powtórnie odsezonowanie w znacznie mniejszym stopniu wpływa na ostateczny szereg w przypadku długiego niż krótkiego szeregu czasowego. Niskie wartości miary R wskazują, że wartości poszczególnych obserwacji otrzymane przy jedno- i dwukrotnym zastosowaniu procedury odsezonowania są do siebie zbliżone, co jest zjawiskiem pożądanym. Minimalny poziom R ($R=0,021$) uzyskano dla metody TRAMO/SEATS zastosowanej na pełnej próbie. Porównanie w obszarze krótkich szeregów czasowych prowadzi do niejednoznacznych wniosków. Z jednej strony miara rozbieżności R wskazuje na przewagę metody X-12-ARIMA, lecz po uwzględnieniu wariancji składników sezonowych należy stwierdzić, że lepszą metodą jest TRAMO/SEATS. Wariancja składnika sezonowego dwukrotnie wyrównanego szeregu zmniejszyła się bowiem ponad 10-krotnie, podczas gdy w przypadku metody X-12-ARIMA, wariancja ta uległa zwiększeniu po dokonaniu dwukrotnego odsezonowania (z 0,002 do 0,069).

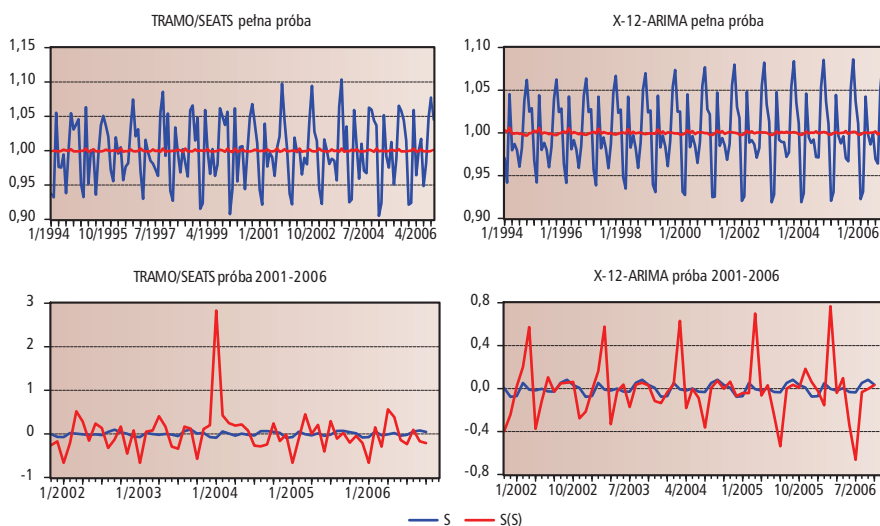
Ponownie, wnioski dotyczące krótkiego szeregu czasowego z okresu 2001–2006 są zbieżne z tymi uzyskanymi dla średnich statystyk obliczonych dla siedmiu podprób (patrz Aneks). Zgodność ta dotyczy zarówno miary rozbieżności R (X-12-ARIMA: $R = 0,266$; TRAMO/SEATS: $R = 0,174$), jak i zachowania się wariancji, która w przypadku TRAMO/SEATS zmniejszyła się ponad 11-krotnie, a przy zastosowaniu X-12-ARIMA uległa zwiększeniu.

Reasumując, kryterium idempotencji wskazuje na przewagę metody TRAMO/SEATS nad X-12-ARIMA zarówno dla krótkich, jak i długich szeregów czasowych. W kontekście długości rozpatrywanej próby wyższą jakość wyrównania sezonowego uzyskano dla długiej próby. Badanie przeprowadzone przez B. Fishera również wykazało, że ze względu na

kryterium idempotencji znacznie lepsze wyniki przynosi zastosowanie TRAMO/SEATS niż każdej innej metody odsezonowania⁵².

Analiza wyników graficznych (wykres 9) wskazuje, że w przypadku oszacowania metodą X-12-ARIMA i TRAMO/SEATS krótkiego szeregu czasowego, otrzymano komponent sezonowy dla danych poddanych ponownemu wyrównaniu sezonowemu (czerwona linia) charakteryzujący się wielokrotnie większą amplitudą wahań niż składnik sezonowy szeregu oryginalnego (niebieska linia). Pomimo to testy diagnostyczne nie wskazywały na istnienie sezonowości w szeregu wyrównanym metodą X-12-ARIMA. Prawdopodobną przyczyną tego zjawiska jest duży udział składnika ruchomego w komponencie sezonowym, powodujący jego nadmierną nieregularność. W pozostałych rozważanych przypadkach oszacowany komponent sezonowy skorygowanych szeregów cechował się znacznie niższą wariancją niż analogiczny komponent oryginalnych szeregów, co potwierdza zawodność testu na obecność sezonowości w pakiecie *Demetra*.

Wykres 9. Komponent sezonowy jednokrotnie (S) i dwukrotnie wyrównanego szeregu (S(S))



Źródło: opracowanie własne.

4.7. Analiza komponentu nieregularnego

Odpowiednia dekompozycja szeregu czasowego powinna prowadzić do otrzymania komponentu nieregularnego, który nie odzwierciedla żadnego identyfikowalnego procesu, takiego jak autokorelacja lub wahania sezonowe. Czynniki nieregularne powinny być realizacją procesu białego szumu (*white noise*). Na podstawie wartości testów Fishera-Kappa przedstawionych w tabeli 11 należy stwierdzić, że w przeprowadzonym badaniu komponent nieregularny tylko w przypadku krótkiej próby odsezonowanej metodą TRAMO/SEATS nie był białym szumem.

⁵² Fischer (1995), s. 43. Badanie przeprowadzono na szeregach pochodzących z rachunków narodowych, handlu zagranicznego, statystyki bezrobocia i produkcji przemysłowej dla wybranych krajów europejskich.

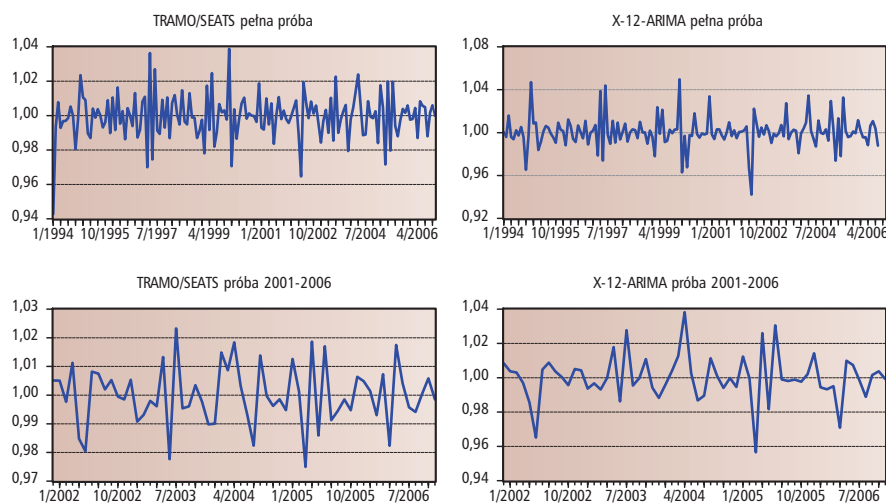
Tabela 11. Analiza komponentu nieregularnego – test na biały szum

Metoda	Opis	M-1	Max(J_k)	Suma(J_k)	Test Fishera– Kappa ⁵³
TRAMO/SEATS pełna próba		77	0,00	0,03	4,85* [7,122]
ARIMA X– 12 pełna próba		77	0,00	0,03	4,5* [7,122]
TRAMO/SEATS 2001-2006		29	0,00	0,00	7,80 [5,935]
X-12-ARIMA 2001-2006		29	0,00	0,01	3,95* [5,935]

Uwagi: *oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej na poziomie istotności 5%. W nawiasach kwadratowych podano wartość krytyczną testu.
Źródło: opracowanie własne.

Poniżej przedstawiono wykres komponentu nieregularnego dla poszczególnych metod odsezonowania (wykres 10). Należy zauważyć, że postać graficzna komponentu nieregularnego potwierdza wnioski otrzymane dzięki analizie powyższych statystyk, testujących hipotezę o występowaniu białego szumu. W przypadku metody X-12-ARIMA komponent ten zarówno dla długiego, jak i krótkiego szeregu jest nieprzewidywalny. Wahania w przypadku metody TRAMO/SEATS dla krótkiego szeregu, począwszy od drugiego kwartału 2003 do drugiego kwartału 2005 r., wykazują pewną regularność, a co za tym idzie także określoną przewidywalność. Składnik sezonowy w tym okresie jest powtarzalny, np. ten sam wzorzec sezonowości zaobserwować można w okresach: czerwiec 2003 – luty 2004 i luty – październik 2005, co najprawdopodobniej zaważyło na odrzuceniu hipotezy zerowej testu Fishera-Kappa zakładającej, że badany proces jest białym szumem.

Wykres 10. Komponent nieregularny TRAMO/SEATS vs X-12-ARIMA (pełna próba vs próba okrojona)



Źródło: opracowanie własne.

Statystyki diagnostyczne dla komponentu nieregularnego przedstawiono w tabeli 12. Analiza opierała się na badaniu średniej, odchylenia standardowego, skośności i kurtozy składnika nieregularnego. Ponadto przeprowadzono test Jarque-Berra na normalność rozkładu komponentu, test na występowanie procesu autokorelacji w komponencie, oraz testy na obecność efektu dni roboczych i sezonowości. Należy zauważyć, że sezonowość została zdiagnozowana w komponencie nieregularnym w przypadku modelu TRAMO/SEATS dla krótkiego i długiego szeregu, a ponadto dla długiego szeregu czasowego przy zasto-

⁵³ Test Fishera-Kappa ma postać $\frac{(M-1) \cdot (\text{Max}(J_k))}{\text{Suma}(J_k)}$, gdzie $M-1$ oznacza liczbę obserwacji, na podstawie której przeprowadzany jest test. W przypadku parzystej liczby wszystkich dostępnych obserwacji pierwotnego szeregu o długości T , $M = \frac{T}{2}$, w przypadku nieparzystej liczby $M = \frac{T-1}{2}$. J_k oznacza amplitudę periodogramu liczoną zgodnie ze wzorem zamieszczonym na stronie 15.

sowaniu metody X-12-ARIMA. W przypadku X-12-ARIMA dla krótkiej próby nie wykryto sezonowości w komponencie nieregularnym, aczkolwiek został dobrany sezonowy model ARIMA. Przyczyną tego zjawiska jest niezależność procesu identyfikacji sezonowości i doboru modelu ARIMA. W żadnym z badanych komponentów nie stwierdzono występowania procesu autokorelacji. W przypadku metody TRAMO/SEATS nie udało się wyeliminować efektu dni roboczych z komponentu nieregularnego. Najniższą wartością odchylenia standardowego odznaczały się komponenty nieregularne, które uzyskano w wyniku dekompozycji przeprowadzonej na krótkich szeregach czasowych.

Tabela 12. Analiza komponentu nieregularnego

Metoda \ Wynik	Średnia ¹	Odchylenie stand.	Skośność/ Kurtoza	Test Jarque–Berra	Autokorelacja (Test Ljunga – Boxa)	ADF ²	Efekt dni roboczych	Wykrycie sezon./model ARIMA ³
TRAMO/SEATS pełna próba	1	0,013	-0,4/5,5	43,50 [0,00]	25,04* [0; 32,70]	-19,21 [0,00]	tak (7 dni)	tak (1,0,1) (0,1,1)
X-12-ARIMA pełna próba	1	0,014	0,2/6,5	80,61 [0,00]	48,11* [0; 51,20]	-15,37 [0,00]	nie	tak (0,1,1) (0,1,1)
TRAMO/SEATS 2001-2006	1	0,010	-0,1/2,9	0,15 [0,93]	24,96** [0; 33,90]	-10,91 [0,00]	tak (7 dni)	tak (0,1,1) (0,1,1)
X-12-ARIMA 2001-2006	1	0,01	1,21/7,69	69,6 [0,00]	37,59** [0,51,20]	-13,31 [0,00]	nie	nie (2,1,0) (0,1,1)

1) Dla modelu multiplikatywnego średnia komponentu nieregularnego wynosi 1, co wynika z konstrukcji tego modelu.

2) Test ADF (*Augmented Dickey-Fuller test*) bada obecność pierwiastka jednostkowego w danym szeregu. Statystyka tego testu przyjmuje wartości ujemne (wraz ze zmniejszaniem wartości statystyki rośnie prawdopodobieństwo odrzucenia hipotezy zerowej o występowaniu pierwiastka jednostkowego).

3) Wybrany model ARIMA, to dobrany automatycznie przez pakiet *Demetra* model ARIMA.

Źródło: opracowanie własne.

5 Podsumowanie

Literatura przedmiotu wskazuje na istnienie wielu kryteriów porównawczych, które służą ocenie jakości wyrównywania danych metodami X-12-ARIMA i TRAMO/SEATS. Problemem staje się uznanie jednego z tych kryteriów za priorytetowe. Głównym celem odsezonowywania danych jest eliminacja czynnika sezonowego. Dlatego też ważnym kryterium porównawczym jest kryterium idempotencji. Z kolei kryterium (nie)zgodności pozwala na przeprowadzenie prostej i ważnej diagnostyki, która decyduje o poprawności kierunku rozwojowego danego szeregu po odsezonowaniu. Ponadto autorzy tacy jak R. Feldpausch, C. Hood i K. Wills⁵⁴, skłaniają się ku twierdzeniu, że kryteria stabilności modelu są ważniejsze od kryteriów dopasowania. Dlatego też, zgodnie z tym podejściem, w niniejszym opracowaniu skupiono się przede wszystkim na ocenie jakości dopasowania na podstawie analizy *sliding spans* i korekt (rewizji).

W zamieszczonej poniżej tabeli 13 zawarto zestawienie najbardziej istotnych wyników z przeprowadzonej analizy porównawczej pod kątem długości próby. Kolorem zielonym wyróżniono przypadki, w których lepsze wyniki uzyskano dla długiego szeregu czasowego, jasnozielonym – sytuacje, w których odnotowano przewagę krótkiego szeregu obejmującego okres 2001–2006 lub średniej dla krótkich szeregów czasowych. Sytuacje, w których nie otrzymano jednoznacznego wyniku oznaczono znakiem zapytania.

Wyniki analizy porównawczej zaprezentowane w tabeli 13, pozwalają stwierdzić, że wyższą jakość odsezonowania uzyskano dla długich szeregów czasowych. Wskazują na to jednoznacznie kryteria uznane przez autorki pracy za kluczowe przy porównywaniu metod wyrównania, tj. *sliding spans*, kryterium rewizji i idempotencji.

W kolejnej tabeli zestawiono wyniki porównania pomiędzy metodami odsezonowania. Kolor zielony wskazuje na lepszą jakość wyrównania uzyskaną metodą TRAMO/SEATS, natomiast jasnozielony – X-12-ARIMA. Tylko w jednym przypadku nie uzyskano rezultatów pozwalających na jednoznaczny wybór metody.

Przypomnijmy, że na podstawie kryterium idempotencji, zgodności, gładkości i rewizji, otrzymano wnioski o przewadze metody TRAMO/SEATS nad X-12-ARIMA. Ponowna procedura odsezonowania i analiza wykresu składnika sezonowego nie wskazała w tym przypadku na występowanie istotnej sezonowości. Zatem wielokrotne zastosowanie procedury odsezonowania nie powodowało błędów w postaci wykrywania składników sezonowych o dużej wariancji. Ponadto analiza graficzna składnika sezonowego jednokrotnie i dwukrotnie wyrównanego szeregu wskazała, że dla krótkich szeregów mniejsze odchylenia od średniej występowały przy zastosowaniu metody TRAMO/SEATS niż X-12-ARIMA. Kryterium rewizji i zgodności wskazuje na analogiczne wnioski. Z tego względu za rekomendowaną do wyrównania sezonowego indeksu produkcji sprzedanej w niniejszym opracowaniu uznano metodę TRAMO/SEATS, pomimo że nie wszystkie przedstawione kryteria w jednoznaczny sposób na nią wskazywały. Przypomnijmy, że podczas analizy czynnika nieregularnego dla metody TRAMO/SEATS zastosowanej na krótkim szeregu pojawiły się problemy diagnostyczne. Odrzucono, między innymi, hipotezę o tym, że jest on białym szumem. Podobne spostrzeżenia sformułował M. Ongan, analizując turecki indeks cen. Jego badanie, ograniczone jedynie do długich szeregów czasowych, wskazało, że rekomendowaną metodą ich odsezonowania jest TRAMO/SEATS⁵⁵.

⁵⁴ Feldpausch, Hood i Wills (2004), s. 5.

⁵⁵ Ongan (2002), s. 25.

Tabela 13. Podsumowanie – pełna próba vs próba skrócona

Kryterium – długość próby		X-12-ARIMA		X-12-ARIMA		TRAMO/SEATS		TRAMO/SEATS	
		Długi szereg	Krótki szereg 2001-2006	Długi szereg	Krótki szereg średnia	Długi szereg	Krótki szereg 2001-2006	Długi szereg	Krótkie szeregi średnia
Kryterium									
1. Sliding spans		Długi szereg		Długi szereg		Długi szereg		Długi szereg	
2. Rewizje	Szereg odsezonowany	Długi szereg		Długi szereg		Długi szereg		Długi szereg	
	Trend-cykl	Długi szereg		Długi szereg		Długi szereg		Długi szereg	
3. Zgodność		Krótki szereg		Długi szereg		Krótki szereg		Długi szereg	
4. Gładkość		Długi szereg		Długi szereg		Długi szereg		Długi szereg	
5. Względny błąd prognozy		Krótki szereg		Długi szereg		Krótki szereg		Długi szereg	
6. Idempotencja (na podstawie statystyki R)		Długi szereg		Długi szereg		Długi szereg		Długi szereg	
7. Komponent nieregularny		Krótki szereg		?		Długi szereg		?	

Źródło: opracowanie własne.

Tabela 14. Podsumowanie – porównanie metod

Długość próby – metoda		Długi szereg		Krótki szereg okres 2001-2006		Krótkie szeregi Średnia	
		TRAMO/SEATS	X-12-ARIMA	TRAMO/SEATS	X-12-ARIMA	TRAMO/SEATS	X-12-ARIMA
Kryterium							
1. Sliding spans		X-12-ARIMA		TRAMO/SEATS		TRAMO/SEATS	
2. Rewizje	Szereg odsezonowany	TRAMO/SEATS		X-12-ARIMA		X-12-ARIMA	
	Trend-cykl	TRAMO/SEATS		TRAMO/SEATS		TRAMO/SEATS	
3. Zgodność		TRAMO/SEATS		TRAMO/SEATS		X-12-ARIMA (minimalna przewaga rzędu 0,15 punktów procentowych)	
4. Gładkość		TRAMO/SEATS		TRAMO/SEATS		TRAMO/SEATS	
5. Względny błąd prognozy		TRAMO/SEATS		X-12-ARIMA		X-12-ARIMA	
6. Idempotencja (na podstawie statystyki R)		TRAMO/SEATS		TRAMO/SEATS		TRAMO/SEATS	
7. Komponent nieregularny		TRAMO/SEATS (minimalna przewaga pod względem odchylenia standardowego)		X-12-ARIMA		?	

Źródło: opracowanie własne.

Otrzymane wyniki wykazują, że w przypadku indeksu produkcji sprzedanej przemysłu w Polsce sezonowe wyrównanie szeregu metodą TRAMO/SEATS daje wyniki o wyższej jakości niż przy użyciu X-12-ARIMA, nawet w przypadku znacznej redukcji długości próby. Co więcej, losowy dobór okresu dla krótkiej próby nie wpływa zasadniczo na wyniki analizy jakości odsezonowania. Wskazuje na to zgodność statystyk pomiędzy krótkim

szeregiem czasowym z okresu 2001–2006 a średnią statystyk liczonych z 8 krótkich szeregów czasowych, odpowiednio z okresów: 2001–2006, 2000–2005, 1999–2000, 1998–2003, 1997–2002, 1996–2001, 1995–2000. Tylko w przypadku jednego z kryteriów jakościowych wyniki te różniły się od siebie, jednak jak zaznaczono w tabeli 14, różnica ta była bardzo mała, na poziomie błędu statystycznego.

Badanie porównania jakości odsezonowania wykonano również dla nominalnego szeregu produkcji sprzedanej⁵⁶. Porównując wyrównanie pełnej i pięcioletniej (2001–2006) próby, zastosowano analogiczny zestaw kryteriów. Dla tego szeregu uzyskano lepsze rezultaty, stosując metodę TRAMO/SEATS niż X-12-ARIMA zarówno dla krótkiej, jak i długiej próby. Wyniki poszczególnych kryteriów częściowo się pokrywały z rezultatami otrzymanymi dla szeregu realnego. Dotyczyło to m.in. badania idempotencji i komponentu nieregularnego. Analizując jakość wyrównania zarówno szeregu nominalnego, jak i realnego widać, że lepsze wyniki otrzymano przy zastosowaniu TRAMO/SEATS bez względu na długość próby.

Powyższe wnioski wskazują, iż ocena przewagi jednej z metod może być determinowana przez sam dobór szeregu czasowego. Przykładowo, G. Mazzi i G. Savio, badając szeregi eksportu dla różnych krajów Unii Europejskiej stwierdzili, że metoda X-12-ARIMA prowadzi do mniejszego pogorszenia jakości odsezonowania w wyniku redukcji próby niż TRAMO/SEATS⁵⁷.

Na podstawie zebranych w tabeli 14 zauważyć należy, że metoda TRAMO/SEATS powinna być rekomendowana w większości przypadków, ponieważ w 15 z 23 przypadków uzyskano dla niej lepszą jakość dopasowania.

⁵⁶ Wyników tych nie zamieszczono w pracy ze względu na ich znaczną objętość.

⁵⁷ Mazzi i Savio (2005), s. 13.

6 Aneks

Zestawienie wyników otrzymanych dla krótkich szeregów czasowych:

Tabela 1.1A. Informacje podstawowe – TRAMO/SEATS

Wynik Zakres próby	Model	Trans.	Efekt dni roboczych	Efekt Wielkano- nocy	Outliery	Skośność/ Kurtoza	Test Jarque- Berra	Test Ljunga -Boxa
2001-2006	(0,1,1) × (0,1,1)	log	7 regr.	nie	-	-0,22/ 2,56	0,67 [0,72]	15,79** [0,33,9]
2000-2005	(0,1,1) × (0,1,1)	log	1	nie	-	0,01/ 2,46	0,56 [0,76]	12,35** [0,33,9]
1999-2004	(0,1,1) × (0,1,1)	log	nie	nie	-	-0,15/ 2,86	0,21 [0,90]	23,47** [0,33,9]
1998-2003	(2,2,1) × (0,1,1)	brak	nie	nie	gru-02 (AO)	0,05/ 2,68	0,21 [0,90]	19,34** [0,31,4]
1997-2002	(0,1,1) × (0,1,1)	brak	2	nie	gru-02 (AO)	-0,28/ 3,88	2,00 [0,37]	11,81** [0,33,9]
1996-2001	(0,1,1) × (0,1,1)	brak	1	nie	-	-0,03/ 3,41	0,33 [0,85]	29,32** [0,33,9]
1995-2000	(0,1,1) × (0,1,1)	log	7 regr.	nie	-	-0,39/ 3,39	1,88 [0,39]	34,21 [0,33,9]
1994-1999	(0,1,1) × (0,1,1)	log	7	tak (6 dni)	-	0,25/ 3,40	0,69 [0,71]	29,94** [0,33,9]
2001-2006	(0,1,1) × (0,1,1)	log	7 regr.	nie	-	-0,22/ 2,56	0,67 [0,72]	15,79** [0,33,9]

** oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej o braku autokorelacji na poziomie istotności 5%.
Źródło: opracowanie własne.

Tabela 1.2A. Informacje podstawowe – X-12-ARIMA

Wynik Zakres próby	Model	Trans.	Efekt dni roboczych	Efekt Wielkano- nocy	Outliery	Skośność/ Kurtoza	Test Jarque- Berra	Test Ljunga -Boxa
2001-2006	(2,1,2) × (0,1,1)	log	1 regr.	tak (6dni)	-	-0,14/2,72	0,38 [0,83]	14,13** [0,51,20]
2000-2005	(2,1,0) × (0,1,1)	log	0 regr.	tak (6 dni)	-	-0,35/2,99	1,22 [0,54]	19,6** [0,51,20]
1999-2004	(2,1,0) × (0,1,1)	log	1 regr.	tak (6dni)	-	0,12/2,86	0,20 [0,91]	23,14** [0,51,20]
1998-2003	(2,1,0) × (0,1,1)	log	1 regr.	nie	-	-0,11/2,74	0,280 [0,87]	21,72** [0,51,20]
1997-2002	(2,1,0) × (0,1,1)	log	6 regr.	nie	maj-02 (AO)	-0,04/2,85	0,07 [0,97]	22,45** [0,51,20]
1996-2001	(2,1,0) × (0,1,1)	log	1 regr.	tak (6dni)	-	0,14/3,33	0,47 [0,80]	23,14** [0,51,20]
1995-2000	(2,1,0) × (0,1,1)	log	0 regr.	tak (6 dni)	-	0,53/4,04	5,55 [0,06]	19,60** [0,51,20]
1994-1999	(2,1,2) × (0,1,1)	log	1 regr.	tak (6dni)	-	0,48/3,73	3,61 [0,17]	14,13** [0,51,20]

** oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej o braku autokorelacji na poziomie istotności 5%.
Źródło: opracowanie własne.

Tabela 2.1A. *Sliding spans* – analiza czynnika sezonowego SS –TRAMO/SEATS

Zakres próby \ Wynik	Średnia	Maksimum	Odchylenie stand.	SS (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach czynników sezonowych	
					w miesiącach	w %
2001-2006	0,006	0,016	0,004	0,0%	16/36	44,4%
2000-2005	0,006	0,021	0,005	0,0%	19/36	52,8%
1999-2004	0,009	0,033	0,007	2,8%	4/36	11,1%
1998-2003	-0,348	1,363	1,038	36,1%	3/36	8,3%
1997-2002	0,119	2,147	0,685	44,4%	0/36	0,0%
1996-2001	-0,215	1,497	0,996	41,7%	3/36	8,3%
1995-2000	0,008	0,026	0,006	0,0%	0/36	0,0%
1994-1999	0,010	0,035	0,008	2,8%	3/36	8,3%
Średnia	-0,051	0,642	0,344	16,0%	6/36	16,7%

Źródło: opracowanie własne.

Tabela 2.2A. *Sliding spans* – analiza czynnika sezonowego SS –X-12-ARIMA

Zakres próby \ Wynik	Średnia	Maksimum	Odchylenie stand.	SS (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach czynników sezonowych	
					w miesiącach	w %
2001-2006	0,010	0,024	0,006	0,0%	18/36	50,0%
2000-2005	0,009	0,016	0,005	0,0%	19/36	52,8%
1999-2004	0,011	0,028	0,008	0,0%	21/36	58,3%
1998-2003	0,010	0,026	0,009	0,0%	21/36	58,3%
1997-2002	0,010	0,028	0,006	0,0%	21/36	58,3%
1996-2001	0,010	0,024	0,007	0,0%	20/36	55,6%
1995-2000	0,013	0,040	0,011	8,3%	21/36	58,3%
1994-1999	0,006	0,011	0,004	0,0%	21/36	58,3%
Średnia	0,010	0,025	0,007	1,0%	20,25/36	56,3%

Źródło: opracowanie własne.

**Tabela 3.1A. Analiza miesięcznych stóp wzrostu danych odsezonowanych
MM – TRAMO/SEATS**

Zakres próby \ Wynik	Średnia	Maksimum	Odchylenie stand.	MM (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach miesięcznych stóp wzrostu	
					w miesiącach	w %
2001-2006	0,788	2,312	0,620	0,0%	5/35	14,3%
2000-2005	1,015	2,661	0,724	0,0%	6/35	17,1%
1999-2004	0,905	3,373	0,833	2,9%	8/35	22,9%
1998-2003	1,328	5,422	1,082	5,7%	10/35	28,6%
1997-2002	1,265	4,224	1,119	11,4%	8/35	22,9%
1996-2001	1,206	5,981	1,136	2,9%	12/35	34,3%
1995-2000	1,037	3,895	0,976	5,7%	5/35	14,3%
1994-1999	1,468	4,677	1,123	8,6%	13/35	37,1%
Średnia	1,1265	4,0680	0,9515	4,6%	8,4/35	23,9%

Źródło: opracowanie własne.

**Tabela 3.2A. Analiza miesięcznych stóp wzrostu danych odsezonowanych
MM – X-12-ARIMA**

Zakres próby \ Wynik	Średnia	Maksimum	Odchylenie stand.	MM (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach miesięcznych stóp wzrostu	
					w miesiącach	w %
2001-2006	2,103	6,491	1,763	28,6%	13/35	37,1%
2000-2005	1,491	2,742	0,855	0,0%	7/35	20,0%
1999-2004	2,534	6,711	1,750	34,3%	14/35	40,0%
1998-2003	2,548	8,877	2,173	34,3%	14/35	40,0%
1997-2002	2,385	6,795	1,680	28,6%	16/35	45,7%
1996-2001	1,816	6,172	1,435	20,0%	13/35	37,1%
1995-2000	2,415	10,588	2,193	28,6%	14/35	40,0%
1994-1999	1,115	3,024	0,800	2,9%	10/35	28,6%
Średnia	2,051	6,425	1,581	22,1%	12,6/35	36,1%

Źródło: opracowanie własne.

Tabela 4.1A. Analiza rocznych stóp wzrostu danych odsezonowanych
YY – TRAMO/SEATS

Zakres próby \ Wynik	Średnia	Maksimum	Odchylenie stand.	YY (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach rocznych stóp wzrostu	
					w miesiącach	w %
2001-2006	0,277	9,545	2,254	4,2%	2/24	8,3%
2000-2005	1,031	15,045	3,674	16,7%	2/24	8,3%
1999-2004	-0,556	23,160	5,861	4,2%	9/24	37,5%
1998-2003	-8,560	2,092	40,359	0,0%	4/24	16,7%
1997-2002	0,322	16,260	3,781	8,3%	3/24	12,5%
1996-2001	-0,737	6,911	3,077	4,2%	8/24	33,3%
1995-2000	0,703	20,619	4,423	4,2%	3/24	12,5%
1994-1999	-0,036	1,981	1,156	0,0%	2/24	8,3%
Średnia	-0,944	11,952	8,073	5,7%	4,1/24	17,2%

Źródło: opracowanie własne.

Tabela 4.2A. Analiza rocznych stóp wzrostu danych odsezonowanych
YY – X-12-ARIMA

Zakres próby \ Wynik	Średnia	Maksimum	Odchylenie stand.	YY (odsetek obserwacji niestabilnych)	Rozbieżności w oszacowaniach rocznych stóp wzrostu	
					w miesiącach	w %
2001-2006	-1,111	41,677	12,358	12,5%	10/24	41,7%
2000-2005	-2,252	7,042	11,337	12,5%	6/24	25,0%
1999-2004	-7,902	17,130	39,051	8,3%	9/24	37,5%
1998-2003	-1,807	1,565	3,210	0,0%	13/24	54,2%
1997-2002	-1,020	0,800	1,357	0,0%	9/24	37,5%
1996-2001	42,029	1326,635	276,943	8,3%	13/24	54,2%
1995-2000	-1,841	17,621	10,354	12,5%	10/24	41,7%
1994-1999	-1,451	4,460	4,456	4,2%	10/24	41,7%
Średnia	3,081	177,116	44,883	7,3%	10/24	41,7%

Źródło: opracowanie własne.

Tabela 5.1A. Rewizje historyczne – dane odsezonowane – TRAMO/SEATS

Zakres próby \ Wynik	Średnia w%	Odchylenie standardowe w %
2001-2006	0,136	0,341
2000-2005	0,115	0,476
1999-2004	-0,265	1,579
1998-2003	1,039	1,444
1997-2002	0,089	0,291
1996-2001	-0,182	0,349
1995-2000	0,067	0,671
1994-1999	0,239	0,757
Średnia	0,155	0,739

Źródło: opracowanie własne.

Tabela 5.2A. Rewizje historyczne – dane odsezonowane – X-12-ARIMA

Zakres próby \ Wynik	Średnia w%	Odchylenie standardowe w %
2001-2006	0,187	0,340
2000-2005	-0,175	0,741
1999-2004	0,079	1,276
1998-2003	0,500	0,656
1997-2002	0,267	0,920
1996-2001	-0,198	0,431
1995-2000	0,035	0,338
1994-1999	0,318	0,610
Średnia	0,127	0,664

Źródło: opracowanie własne.

Tabela 6.1A. Rewizje historyczne – trend – TRAMO/SEATS

Zakres próby \ Wynik	Średnia w%	Odchylenie standardowe w %
2001-2006	0,205	0,639
2000-2005	0,204	1,100
1999-2004	-0,265	1,579
1998-2003	2,924	3,958
1997-2002	0,245	0,810
1996-2001	-0,289	0,631
1995-2000	-0,433	1,352
1994-1999	2,983	4,224
Średnia	0,697	1,787

Źródło: opracowanie własne.

Tabela 6.2A. Rewizje historyczne – trend – X-12-ARIMA

Zakres próby	Wynik	Średnia w%	Odchylenie standardowe w %
2001-2006		0,359	0,976
2000-2005		-0,598	1,273
1999-2004		-0,632	2,153
1998-2003		1,982	0,816
1997-2002		-1,504	2,693
1996-2001		-0,107	0,993
1995-2000		-1,443	2,990
1994-1999		2,657	4,202
Średnia		0,089	2,012

Źródło: opracowanie własne.

Tabela 7.1A. Różnica rocznych stóp wzrostu danych surowych i odsezonowanych – TRAMO/SEATS

Zakres próby	Wynik	Średnia różnica stóp wzrostu	Odchylenie stand. różnic stóp wzrostu	Maksimum różnic stóp wzrostu	Procent zgodnych znaków
2001-2006		0,001	0,017	0,042	97,92%
2000-2005		0,000	0,016	0,038	89,58%
1999-2004		0,000	0,014	0,034	91,67%
1998-2003		0,000	0,022	0,060	83,33%
1997-2002		0,000	0,019	0,060	87,50%
1996-2001		0,001	0,020	0,048	89,58%
1995-2000		0,000	0,023	0,058	95,83%
1994-1999		0,000	0,025	0,057	95,83%
Średnia		0,000	0,020	0,050	91,41%

Źródło: opracowanie własne.

Tabela 7.2A. Różnica rocznych stóp wzrostu danych surowych i odsezonowanych – X-12-ARIMA

Zakres próby \ Wynik	Średnia różnica stóp wzrostu	Odchylenie stand. różnic stóp wzrostu	Maksimum różnic stóp wzrostu	Procent zgodnych znaków
2001-2006	0,001	0,015	0,037	95,83%
2000-2005	0,000	0,012	0,028	95,83%
1999-2004	0,000	0,023	0,052	87,50%
1998-2003	0,000	0,019	0,043	87,50%
1997-2002	0,000	0,022	0,045	85,42%
1996-2001	0,001	0,023	0,055	87,50%
1995-2000	0,000	0,015	0,048	97,92%
1994-1999	0,000	0,025	0,055	95,83%
Średnia	0,000	0,019	0,045	91,67%

Źródło: opracowanie własne.

Tabela 8.1A. Miary gładkości dopasowania komponentów – TRAMO/SEATS

Zakres próby \ Wynik	R ₁	R ₁ (ostatnie 3 lata)	R ₂	R ₂ (ostatnie 3 lata)	MAR (TC)
2001-2006	6,8	8,0	2,3	2,9	83,7
2000-2005	4,4	5,6	1,3	1,6	54,6
1999-2004	5,5	5,7	1,5	1,5	50,0
1998-2003	2,8	0,8	1,1	0,2	30,7
1997-2002	4,4	5,8	1,7	2,4	22,5
1996-2001	4,2	4,3	1,1	1,4	19,9
1995-2000	5,4	5,5	1,7	1,8	27,8
1994-1999	2,7	3,7	0,8	1,1	35,2
Średnia	4,5	4,9	1,4	1,6	40,6

Źródło: opracowanie własne.

Tabela 8.2A. Miary gładkości dopasowania komponentów – X-12-ARIMA

Zakres próby \ Wynik	R ₁	R ₁ (ostatnie 3 lata)	R ₂	R ₂ (ostatnie 3 lata)	MAR (TC)
2001-2006	8,1	10,3	2,7	3,6	72,8
2000-2005	6,6	7,8	1,9	1,9	54,7
1999-2004	5,7	5,9	1,8	2,1	36,3
1998-2003	5,8	4,1	1,9	1,4	28,1
1997-2002	5,7	7,1	1,9	2,5	19,5
1996-2001	5,5	6,2	1,7	2,2	23,2
1995-2000	8,9	10,0	3,1	3,7	27,3
1994-1999	3,4	4,9	1,0	1,5	34,6
Średnia	6,2	7,0	2,0	2,4	37,1

Źródło: opracowanie własne.

Tabela 9.1A. Względny błąd prognozy dla szeregu odsezonowanego i komponentu trend-cykl – TRAMO/SEATS

Zakres próby \ Wynik	MAPD (SA)	MAPD (TC)
2001-2006	2,2	2,1
2000-2005	1,2	0,9
1999-2004	3,3	3,2
1998-2003	12,7	12,8
1997-2002	1,7	1,2
1996-2001	.	5,0
1995-2000	2,2	1,7
1994-1999	3,5	3,1
Średnia	3,8	3,8

Źródło: opracowanie własne.

Tabela 9.2A. Względny błąd prognozy dla szeregu odsezonowanego i komponentu trend-cykl – X-12-ARIMA

Zakres próby \ Wynik	MAPD (SA)	MAPD (TC)
2001-2006	0,0	0,0
2000-2005	0,1	0,1
1999-2004	0,0	0,0
1998-2003	0,1	0,1
1997-2002	0,0	0,0
1996-2001	0,0	0,0
1995-2000	0,0	0,0
1994-1999	0,1	0,1
Średnia	0,0	0,0

Źródło: opracowanie własne.

Tabela 10.1A. Ponowne odsezonowanie danych jednokrotnie wyrównanych sezonowo – TRAMO/SEATS

Wynik Zakres próby	Model	Sezono- wość	Outliery	Efekt dni roboczych	Efekt Wielka- nocy	Test Ljunga – Boxa	R	Wariancja S/S(S)
2001-2006	(1,1,0) × (0,1,1)	istotna	-	6 reg.	nie	11,78** [0; 33,90]	0,225	24,287 /0,229
2000-2005	(0,0,1) × (0,1,1)	istotna	-	1 reg.	nie	18,23** [0; 33,90]	0,006	24,400 /0,006
1999-2004	(0,1,1) × (0,1,1)	istotna	-	6 reg.	nie	24,61** [0; 33,90]	0,830	24,732 /0,843
1998-2003	(0,1,1) × (0,1,1)	istotna	gru- 99	1 reg.	nie	30,93** [0; 33,90]	0,007	28,231 /0,007
1997-2002	(0,1,1) × (0,1,1)	istotna	gru- 99	6 reg.	nie	15,80** [0; 33,90]	0,015	22,318 /0,016
1996-2001	(0,1,1) × (0,1,1)	istotna	gru- 99	2 reg.	nie	22,53** [0; 33,90]	0,240	20,463 /0,244
1995-2000	(3,1,0) × (0,1,1)	istotna	-	2 reg.	nie	18,11** [0; 33,90]	0,057	20,711 /0,058
1994-1999	(1,1,0) × (0,1,1)	istotna	-	2 reg.	nie	30,12** [0; 33,90]	0,008	18,280 /0,014
Średnia							0,174	22,928 /0,177

Źródło: opracowanie własne.

Tabela 10.2A. Ponowne odsezonowanie danych jednokrotnie wyrównanych sezonowo – X-12-ARIMA

Wynik Zakres próby	Model	Sezono- wość	Outliery	Efekt dni roboczych	Efekt Wielka- nocy	Test Ljunga – Boxa	R	Wariancja S/S(S)
2001-2006	(2,1,0) × (0,1,1)	nieistotna	-	nie	nie	20,5** [0; 51,2]	0,068	0,002 /0,069
2000-2005	(2,1,2) × (0,1,1)	nieistotna	-	nie	nie	30,3** [0; 51,2]	0,068	0,002 /0,069
1999-2004	(2,1,0) × (0,1,1)	nieistotna	-	nie	nie	16,7** [0; 51,2]	0,046	0,002 /0,000
1998-2003	(2,1,0) × (0,1,1)	nieistotna	-	nie	nie	24,15** [0; 51,2]	0,047	0,002 /0,047
1997-2002	(2,1,0) × (0,1,1)	nieistotna	-	nie	nie	22,33** [0; 51,2]	0,017	0,002 /0,002
1996-2001	(2,1,0) × (0,1,1)	nieistotna	-	nie	nie	27,58** [0; 51,2]	0,039	0,002 /0,004
1995-2000	(0,1,1) × (0,1,1)	nieistotna	maj-02	7 reg.	nie	29,27** [0; 51,2]	1,834	0,002 /0,002
1994-1999	(2,1,2) × (0,1,1)	nieistotna	-	nie	nie	18,78** [0; 51,2]	0,010	0,001 /0,034
Średnia							0,266	0,002 /0,028

Źródło: opracowanie własne.

**Tabela 11.1A. Analiza komponentu nieregularnego – test na biały szum
– TRAMO/SEATS**

Zakres próby \ Wynik	M-1	Max (J_K)	Suma (J_K)	Test Fishera– Kappa
2001-2006	29	0,00	0,00	7,80 [5,935]
2000-2005	29	0,00	0,01	4,66* [5,935]
1999-2004	29	0,00	0,01	5,75* [5,935]
1998-2003	29	0,00	0,01	2,38* [5,935]
1997-2002	29	0,00	0,01	2,99* [5,935]
1996-2001	29	0,00	0,01	5,29* [5,935]
1995-2000	29	0,00	0,01	8,13 [5,935]
1994-1999	29	0,00	0,00	7,80 [5,935]

Uwaga: *oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej na poziomie istotności 5%. W nawiasach kwadratowych podano wartość krytyczną testu.
Źródło: opracowanie własne.

**Tabela 11.2A. Analiza komponentu nieregularnego – test na biały szum
– X-12-ARIMA**

Zakres próby \ Wynik	M-1	Max (J_K)	Suma (J_K)	Test Fishera– Kappa
2001-2006	29	0,00	0,01	3,95 *[5,935]
2000-2005	29	0,00	0,01	4,94* [5,935]
1999-2004	29	0,00	0,01	3,49* [5,935]
1998-2003	29	0,00	0,01	3,51* [5,935]
1997-2002	29	0,00	0,01	3,43* [5,935]
1996-2001	29	0,00	0,01	6,78 [5,935]
1995-2000	29	0,01	0,02	10,21 [5,935]
1994-1999	29	0,00	0,01	4,43* [5,935]

Uwaga: *oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej na poziomie istotności 5%. W nawiasach kwadratowych podano wartość krytyczną testu.
Źródło: opracowanie własne.

Tabela 12.1A. Analiza komponentu nieregularnego – TRAMO/SEATS

Wynik Zakres próby	Śred- nia	Odchyle- nie stand.	Skośność/ Kurtoza	Test Jarque- Berra	Autokorelacja (Test Ljunga – Boxa)	ADF	Efekt dni roboczych	Wykrycie sezon./ model ARIMA
2001-2006	1	0,010	-0,10/ 2,85	0,15 [0,93]	24,96** [0; 33,90]	-10,91 [0,00]	tak (7 dni)	tak (0,1,1)×(0,1,1)
2000-2005	1	0,009	0,08/ 2,64	0,37 [0,82]	29,93** [0; 33,90]	-11,89 [0,00]	tak (1 dzień)	tak (0,1,1)×(0,1,1)
1999-2004	1	0,012	-0,17/ 2,35	1,35 [0,51]	31,79** [0; 33,90]	-11,38 [0,00]	nie	tak (0,1,1)×(0,1,1)
1998-2003	1	0,011	6,52/ 48,07	5504,77 [0,00]	25,61** [0; 33,90]	-7,78 [0,00]	tak (1 dzień)	tak (3,1,1)×(0,1,1)
1997-2002	1	0,013	3,43/ 22,27	1045,97 [0,00]	17,30** [0; 33,90]	-8,68 [0,00]	nie	tak (0,1,1)×(0,1,1)
1996-2001	1	0,012	0,03/ 3,06	0,017 [0,99]	15,73** [0; 33,90]	-14,12 [0,00]	tak (2 dni)	tak (0,1,1)×(0,1,1)
1995-2000	1	0,014	0,20/ 2,98	0,41 [0,81]	31,42** [0; 33,90]	-13,67 [0,00]	tak (2 dni)	tak (2,1,1)×(0,1,1)
1994-1999	1	0,009	0,05/ 3,67	1,18 [0,56]	17,77** [0; 33,90]	-8,27 [0,00]	tak (7 dni)	tak (3,1,1)×(0,1,1)

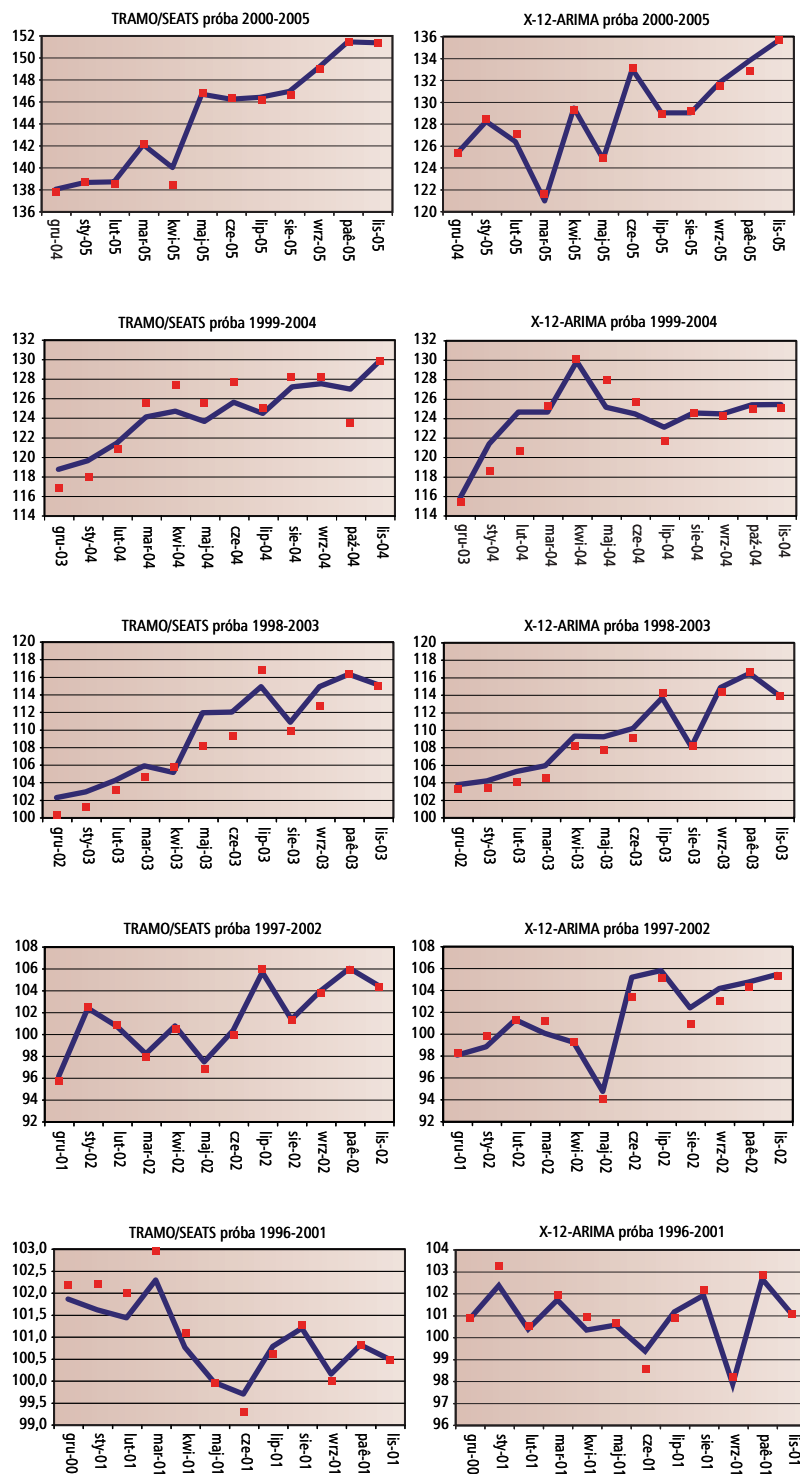
Źródło: opracowanie własne.

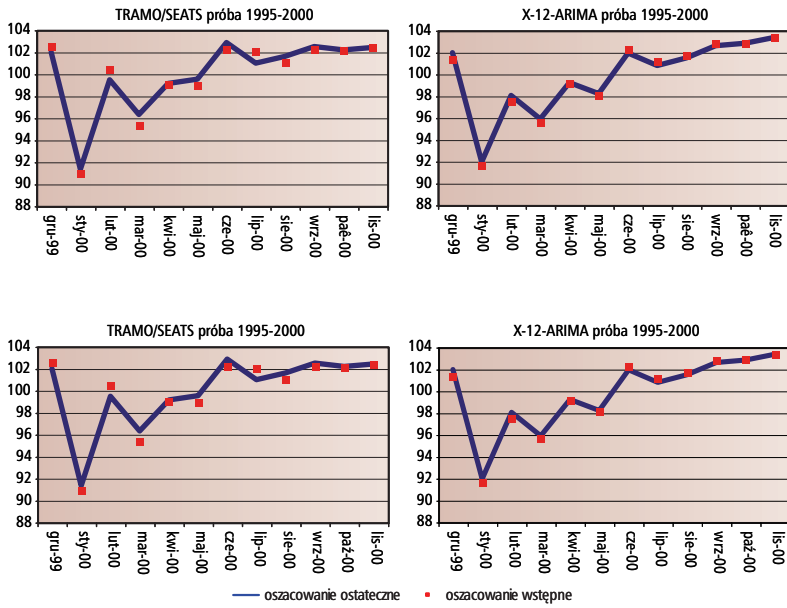
Tabela 12.2A. Analiza komponentu nieregularnego X – 12-ARIMA

Wynik Zakres próby	Śred- nia	Odchyle- nie stand.	Skośność/ Kurtoza	Test Jarque- Berra	Autokorelacja (Test Ljunga – Boxa)	ADF	Efekt dni roboczych	Wykrycie sezon./ model ARIMA
2001-2006	1	0,01	-0,19/ 5,15	11,97 [0,00]	15,02** [0,51,20]	-9,03 [0,00]	nie	nie (2,1,0)×(0,1,1)
2000-2005	1	0,02	0,55/ 5,74	21,77 [0,00]	33,23** [0,51,20]	-8,64 [0,00]	nie	nie (0,1,1)×(0,1,1)
1999-2004	1	0,01	0,89/ 8,03	71,17 [0,00]	34,53** [0,51,20]	-9,60 [0,00]	nie	nie (2,1,0)×(0,1,1)
1998-2003	1	0,02	-0,18/ 6,10	24,38 [0,00]	32,01** [0,51,20]	-10,33 [0,00]	nie	nie (2,1,0)×(0,1,1)
1997-2002	1	0,01	-2,09/ 11,59	228,20 [0,00]	20,62** [0,51,20]	-9,12 [0,00]	nie	nie (2,1,0)×(0,1,1)
1996-2001	1	0,02	0,47/ 5,44	17,12 [0,00]	14,5** [0,51,20]	-12,61 [0,00]	nie	nie (0,1,1)×(0,1,1)
1995-2000	1	0,01	0,06/ 6,38	28,66 [0,00]	29,62** [0,51,20]	-7,91 [0,00]	tak (7 dni)	nie (2,1,0)×(0,1,1)
1994-1999	1	0,01	1,21/ 7,69	69,68 [0,00]	37,59** [0,51,20]	-13,31 [0,00]	nie	nie (2,1,0)×(0,1,1)

Źródło: opracowanie własne.

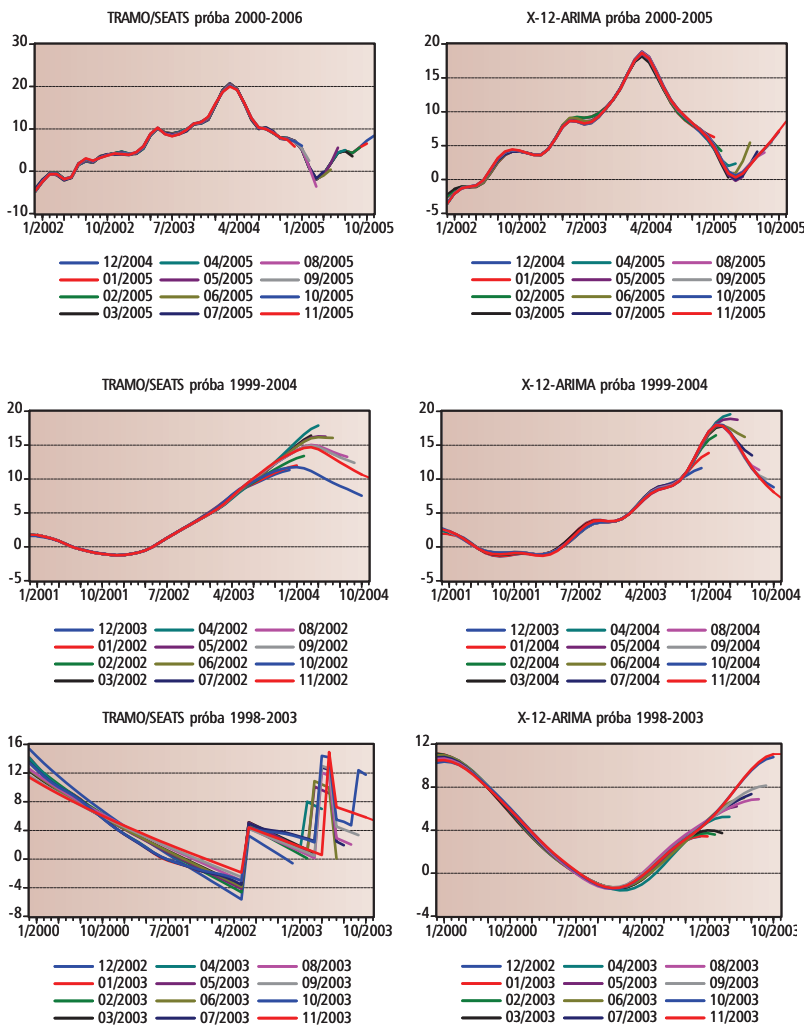
Wykres 1A. Rewizje danych odsezonowanych – próby 5-letnie

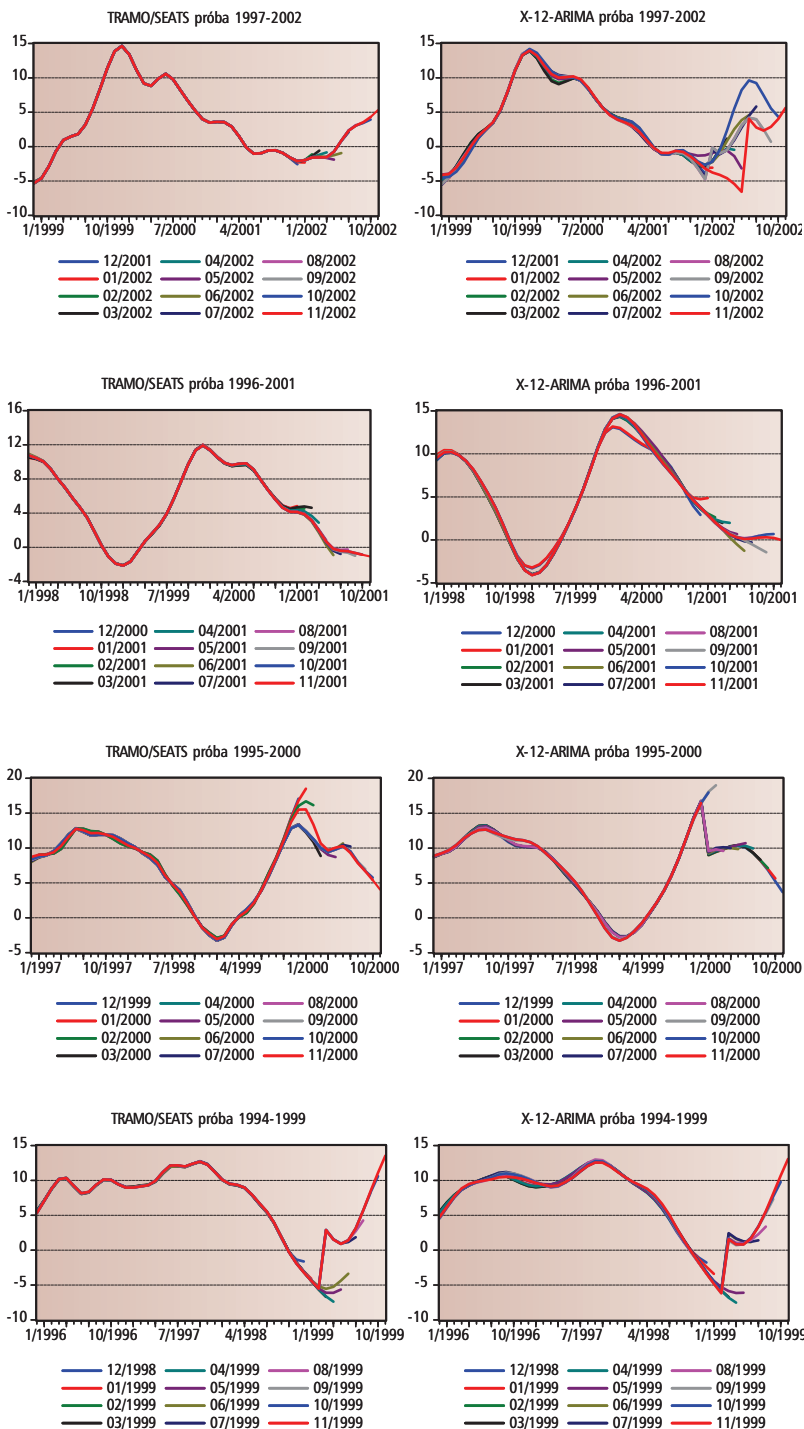




Źródło: opracowanie własne.

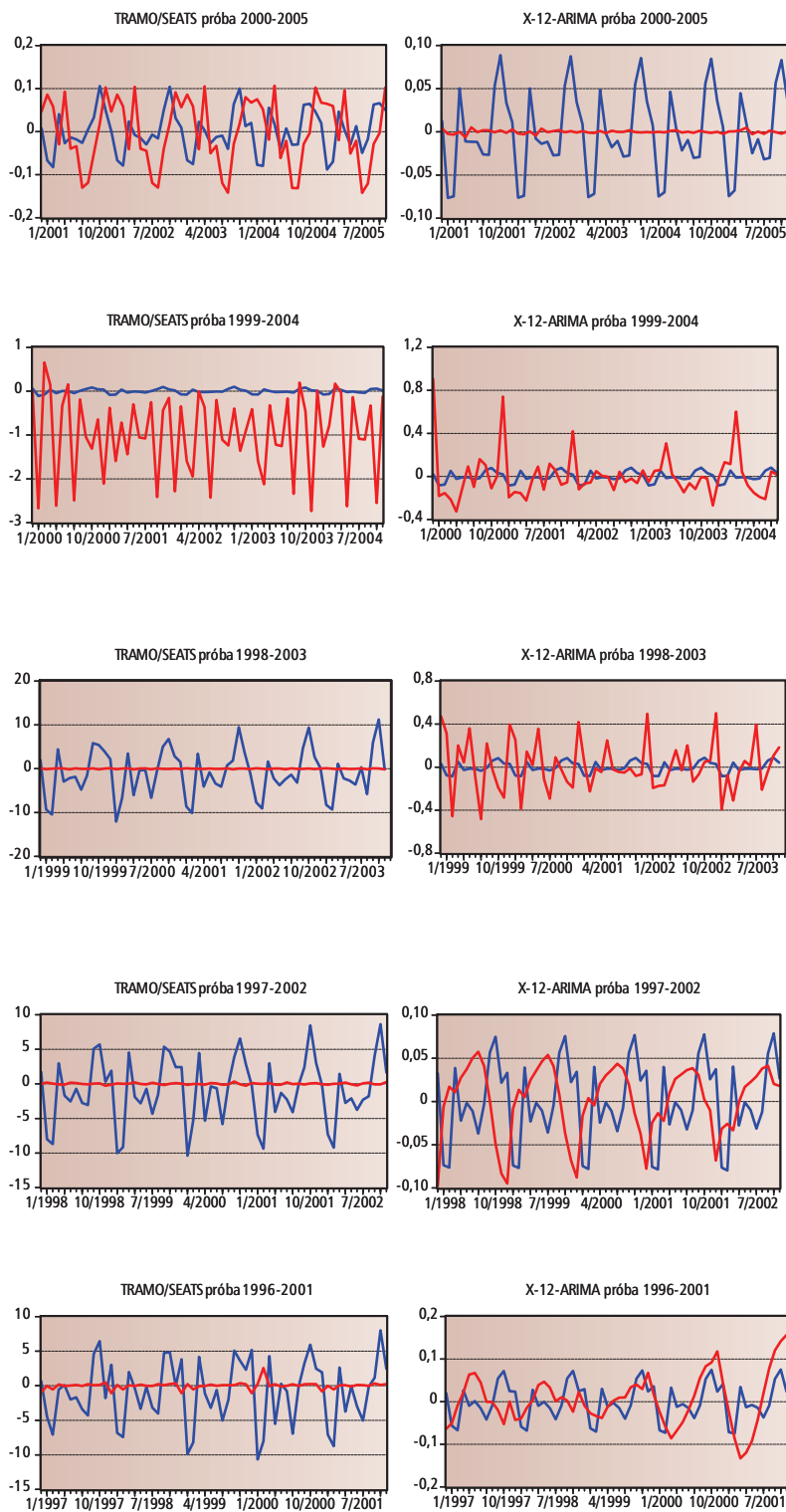
Wykres 2A. Rewizje komponentu trend-cykl – próby 5-letnie

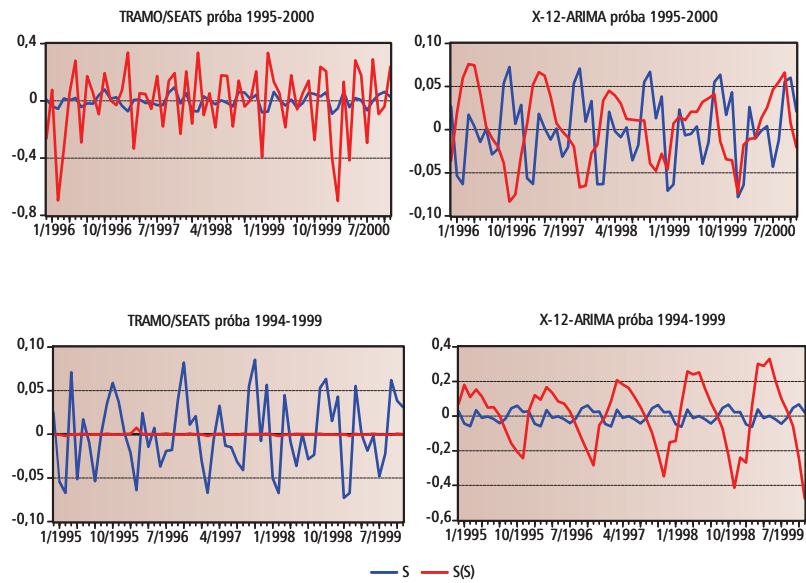




Źródło: opracowanie własne.

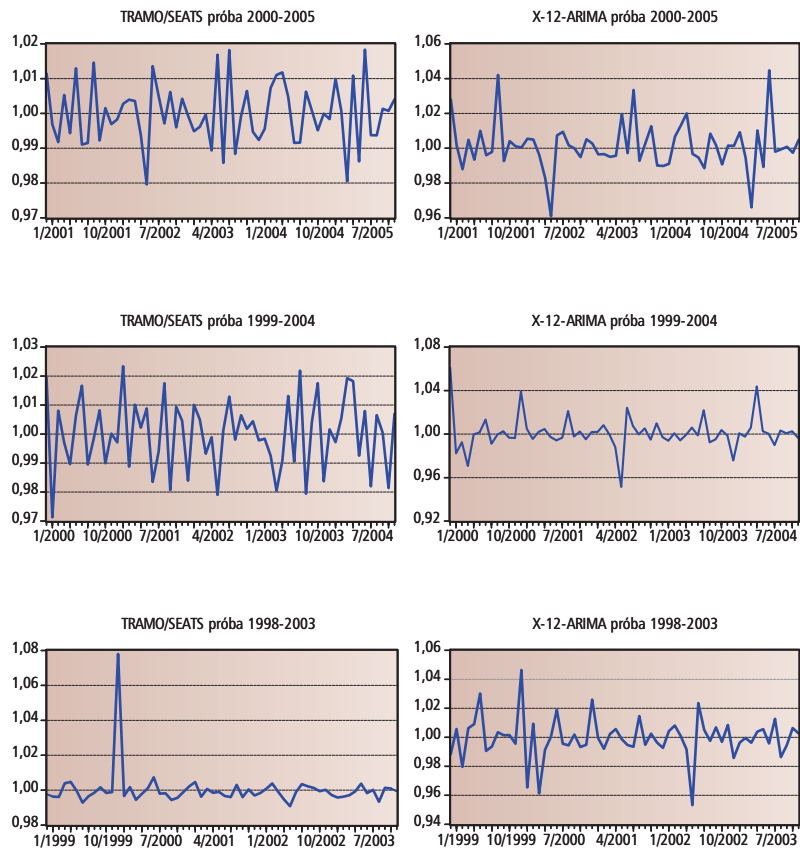
Wykres 3A. Komponent sezonowy – próby 5-letnie

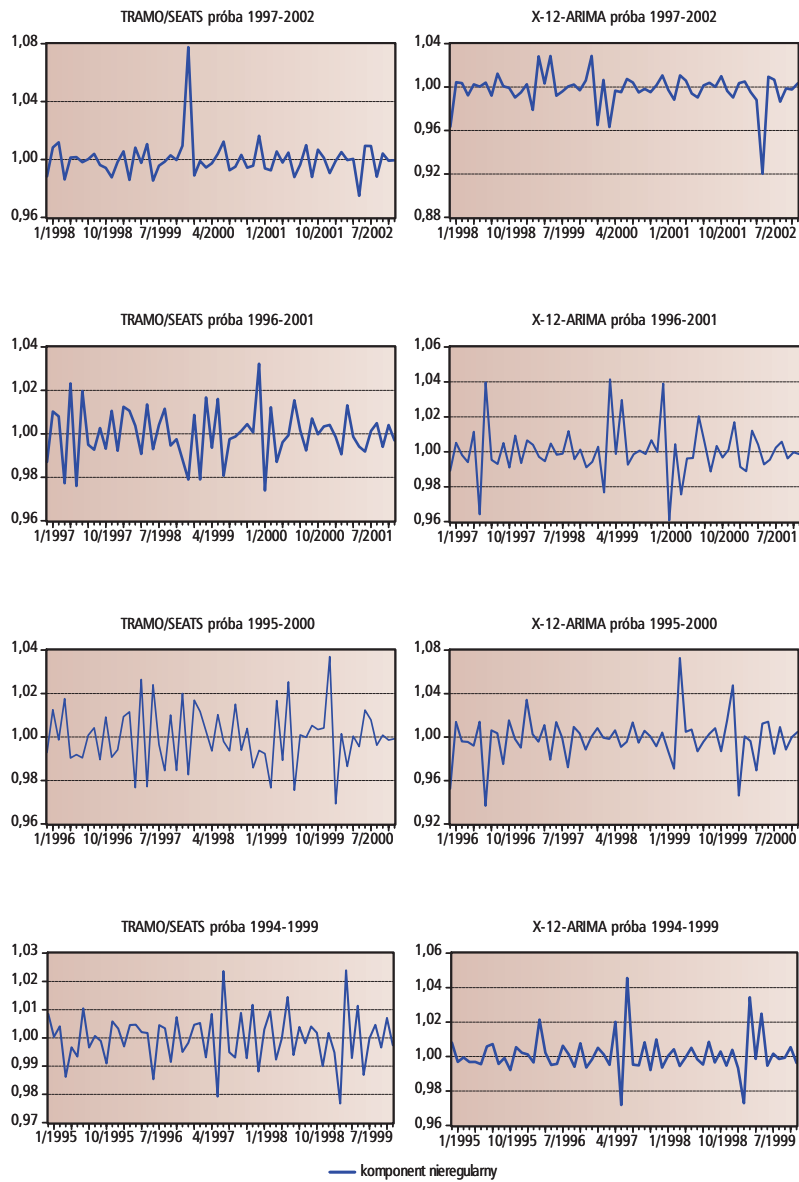




Źródło: opracowanie własne.

Wykres 4A. Komponent nieregularny – próby 5-letnie





Źródło: opracowanie własne.

7

Bibliografia

Astolfi R., Ladiray D., Mazzi G. (2003): *Seasonal Adjustment of European aggregates: direct versus indirect approach*. „*Seasonal Adjustment*”, European Central Bank, s. 37-66.

Atuk O., Ural B. (2002): *Seasonal Adjustment in Economic Time Series*. Central Bank of Turkey, Discussion Paper No. 1.

Bratckoviene N. (2006): *The influence on changes in model on seasonal adjusted data*. Conference on Seasonality, Seasonal Adjustment and Their Implications for Short-Term Analysis and Forecasting, 10-12 May.

Bruce A., Jurke S. (1992): *Empirical Comparison of Two Methods for Non-Gaussian Seasonal Adjustment*. Institute of Statistics and Operations Research, Victoria University, New Zealand, May 27.

Cabra C., Esteban C., Afonso A. (2003): *Monthly re-estimation of the parameters of once-a-year fixed model: an assessment*. „*Seasonal Adjustment*”, European Central Bank, s. 109-127.

Dose J., Planas C. (1996): *Pre-adjustment in Seasonal Adjustment Methods: A Comparison of REGARIMA & TRAMO*. Luksemburg.

Dose J., Planas C. (1996): *Revisions in Seasonal Adjustment Methods: an Empirical Comparison of X-12-ARIMA & SEATS*. Luksemburg.

European Central Bank (2000): *Seasonal Adjustmant of Monetary Aggregates and HICP for Euro Area*.

Feldpausch R., Hood C., Wills K. (2004): *Diagnostics for Model-Based Seasonal Adjustment*. US Census Bureau, Washington, DC 20233-6200.

Fischer B. (1995): *Decomposition of Time Series Comparing Different Methods in Theory and Practice*.

Findley D., Monsell B., Bell W., Otto M., Bor-Chung Chen (1998): *New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program*. „*Journal of Business and Economic Statistics*”.

Findley D., Monsell B., Shulman H., Pugh M. (1990): *Sliding spans for seasonal and related adjustments*. „*Journal of the American Statistical Association*”.

Fuller W. (1976): *Introduction to Statistical Time Series*. New York: John Wiley & Sons, Inc.

Gatto R. (2006): *Series Revision and Seasonal Adjustment of Short Time Series in Presence of a Major Methodological Back*. Seasonal Adjustment and Their Implications for Short-Term Analysis and Forecasting, 10-12 May.

Hood C. (2002): *Comparison of Time Series Characteristics for Seasonal Adjustments from SEATS and X-12-Arima*. US Census Bureau, Washington, DC 20233.

Hood C., Findley D. (2003): *Comparing direct and indirect seasonal adjustments of aggregate series*. „*Seasonal Adjustment*”, European Central Bank, s. 9-22.

Kaiser R., Maravell M. (2001): *An application of TRAMO/SEATS: Changes in seasonality and current trend-cycle assessment*. Banco de España – Servicio de Estudios, Documento de Trabajo Nr° 0011.

Maurin L. (2003): *Criteria to determine the optimal revision policy: a case based on euro zone monetary aggregates data*. „Seasonal Adjustment”, European Central Bank, s. 67-84.

Mazzi G., Savio G. (2005): *The Seasonal Adjustment of Short Time Series*. European Commission, Luksemburg.

Miller L. (1956): *Tables of Percentage Points of Kolmogorov Statistics*. „Journal of American Statistic Association”, No. 111.

Muirhead C. (1986): *Distinguishing Outlier Types in Time Series*. „Journal of Royal Statistical Society”, Series B (Methodological), Vol. 48, No. 1.

Ongan M. (2002): *The Seasonal Adjustment of the Consumer and Wholesale Prices: a Comparison of Census X-11, X-12 ARIMA and TRAMO/SEATS*. The Central Bank of the Republic of Turkey.

Peronaci R. (2003): *The seasonal adjustment of euro area monetary aggregates: direct versus indirect approach*. „Seasonal Adjustment”, European Central Bank, s. 91-107.

Planas C. (1997): *The Analysis of Seasonality in Economic Statistics: A Survey of Recent Developments*. Eurostat Working Group Document.

Quenneville B., Fortier S., Chenzad Z., Latendresse E. (2006): *Recent Developments in Benchmarking to Annual Totals in X-12-ARIMA and Statistics Canada*. Seasonal Adjustment and Their Implications for Short-Term Analysis and Forecasting, 10-12 May.

SAS/ETS Software: *Applications Guide 1*, 1991.

Seasonal adjustment (2006) w: *Seasonal and Prior Adjustment Factors for 2006, Trends for 2002-2006*, Jerusalem, Version 3.0.

Skrzypczyński P. (2006): *Analiza synchronizacji cykli koniunkturalnych w strefie euro*. „Materiały i Studia” nr 210, NBP.

Täht M. (2006): *Seasonal Adjustment and Forecasting of Quarterly Gross Domestic Product: Estonian Experience*. Conference on Seasonality, Seasonal Adjustment and Their Implications for Short-Term Analysis and Forecasting, 10-12 May.

Willoughby D. (2003): *Direct versus indirect seasonal adjustment of UK monetary statistics: preliminary discussion*. „Seasonal Adjustment”, European Central Bank, s. 85-90.

Woźniak P. (2003): *Korygowanie danych inflacyjnych z tytułu sezonowości na podstawie miesięcznych wskaźników cen towarów i usług konsumpcyjnych ogółem za lata 1991-2004*. Centrum Analiz Społeczno-Ekonomicznych.

X-12-Arima Reference Manual v 0.3. U.S. Census Bureau Washington, DC, December 2006.