

Spis treści

1 Szereg niesezonowy	3
1.1 Opis danych	3
1.2 Dekompozycja szeregu	3
1.3 Stacjonarność	3
1.4 ARIMA	10
1.5 Prognoza	12
1.6 Podsumowanie	15
2 Szereg sezonowy	15
2.1 Opis danych	15
2.2 Dekompozycja szeregu	15
2.3 Stacjonarność	16
2.4 SARIMA	20
2.5 Prognozy	23
2.6 Podsumowanie	26
3 Bibliografia	27
4 Spis rysunków	28
5 Spis tabel	29

1 Szereg niesezonowy

1.1 Opis danych

Szereg niesezonowy, który został wybrany do tej pracy, to szereg zawierający dane o wskaźniku cen towarów i usług konsumpcyjnych (CPI) pobrane z bazy danych Narodowego Banku Polskiego, do których uzyskałam dostęp 13 marca 2018 roku ze strony <http://www.nbp.pl/home.aspx?f=/statystyka/bazowa/bazowa.htm> . Dane te pokazują zmiany CPI do analogicznego miesiąca poprzedniego roku. Dane są miesięczne z pierwszą datą styczeń 2001 rok.

1.2 Dekompozycja szeregu

Pierwszą wykonaną czynnością, była dekompozycja szeregu za pomocą programu DEMETRA. Poniższy rysunek przedstawia szereg.

Rysunek 1 Dekompozycja szereg niesezonowy



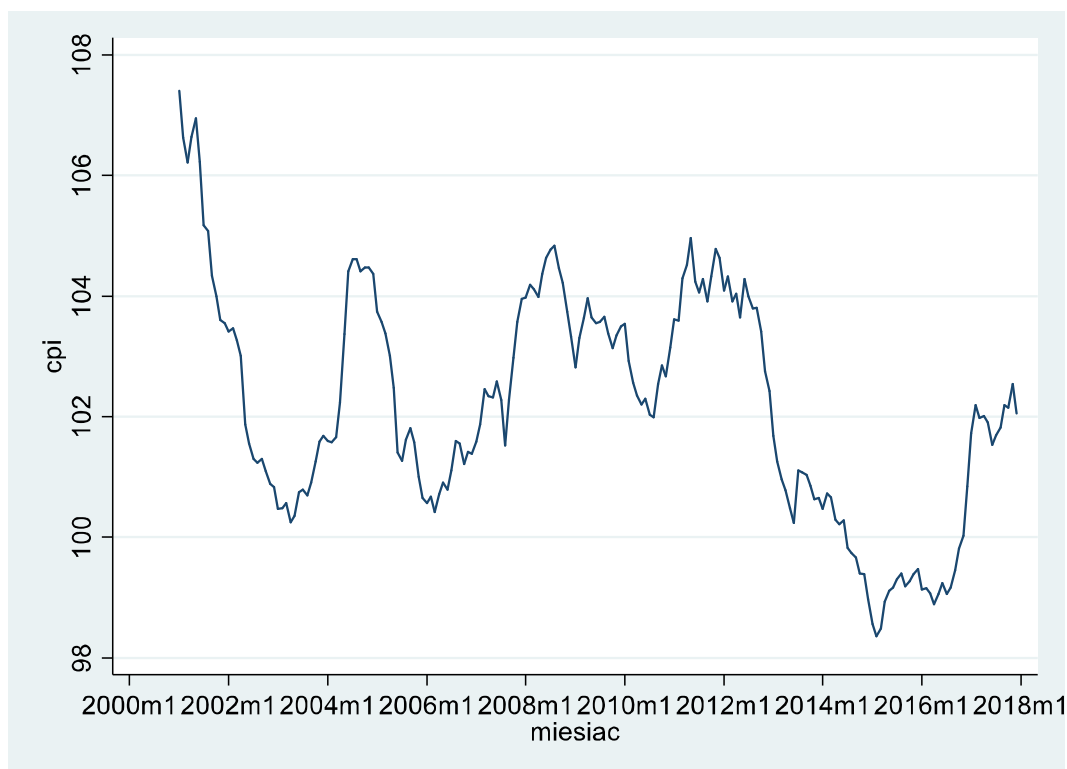
Źródło: Opracowanie własne

Jak widać, szereg ma trend stochastyczny i nie ma sezonowości. Demetra nie przekształciła szeregu ani nie pokazała obserwacji odstających.

1.3 Stacjonarność

Aby była możliwość skorzystania z ARIMY trzeba mieć pewność, że badany szereg jest stacjonarny. Poniższy rysunek przedstawia wykres szeregu .

Rysunek 2 Szereg niesezonalny



Źródło: Opracowanie własne

Interpretując powyższy wykres można stwierdzić, że szereg jest niestacjonarny. Niestacjonarność występuje według wartości oczekiwanej, gdyż szereg nie oscyluje wokół stałego poziomu. Można również stwierdzić niestacjonarność według wariancji. Odchylenia między słupkami znacząco się od siebie różnią. Jednak zanim rozpoczęto testy nad stacjonarnością, przeprowadzono też Boxa-Coxa, aleby dowiedzieć się, czy pierwotny szereg powinien zostać przekształcony. Poniższy rysunek pokazuje wyniki otrzymane po przeprowadzeniu wspomnianego wyżej testu.

Rysunek 3 Test Boxa-Coxa szereg niesezonalny

Test H0:	Restricted log likelihood	LR statistic chi2	P-value Prob > chi2
theta = -1	-418.87142	0.05	0.816
theta = 0	-418.97683	0.26	0.607
theta = 1	-419.16251	0.64	0.425

Źródło: Opracowanie własne

Wyniki otrzymane przy tym teście nie są jednoznaczne, dlatego też zdecydowano o niezmiennianiu formy szeregu.

Pomimo podejrzeń o stacjonarność szeregu na podstawie wykresu, przeprowadzono odpowiednie testy, aby potwierdzić lub odrzucić hipotezę o niestacjonarności szeregu. Na początku przeprowadzono test Dickeya- Fullera zarówno ze stałą jak i bez stałej.

Rysunek 4 Test Dickeya-Fullera bez stałej

```
. dfuller cpi, nocons reg /*test DF dla zmiennej x*/
```

Dickey-Fuller test for unit root Number of obs = 203

Test Statistic	Interpolated Dickey-Fuller		
	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-1.056	-2.586	-1.617

D.cpi	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
cpi L1.	-.0002705	.0002562	-1.06	0.292	-.0007756 .0002346

Źródło: Opracowanie własne

Wartość statystyki jest większa od wartości krytycznej na poziomie 5%, więc nie ma podstaw do odrzucenia hipotezy H_0 o niestacjonarności. Test ten przeprowadzony dla regresji ze stałą pokazał taką samą interpretację- o braku podstaw do odrzucenia hipotezy H_0 o niestacjonarności szeregu. Jednakże, przy przeprowadzeniu testu Dickeya- Fullera należy również sprawdzić, czy nie wstępnie występuje autokorelacja. W celu takiego sprawdzenia przeprowadzono test Breuscha- Godfrey.

Rysunek 5 Test Breuscha- Godfreya szereg niesezonowy

```
. reg d.cpi l.cpi, nocons
```

Source	SS	df	MS	Number of obs	=	203
Model	.155143707	1	.155143707	F(1, 202)	=	1.12
Residual	28.1014631	202	.139116154	Prob > F	=	0.2922
Total	28.2566068	203	.139195107	R-squared	=	0.0055
				Adj R-squared	=	0.0006
				Root MSE	=	.37298

D.cpi	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
cpi						
L1.	-.0002705	.0002562	-1.06	0.292	-.0007756	.0002346

```
.
. bgodfrey, lag(1/4)
```

Breusch-Godfrey LM test for autocorrelation

lags (p)	chi2	df	Prob > chi2
1	25.233	1	0.0000
2	25.236	2	0.0000
3	25.567	3	0.0000
4	26.327	4	0.0000

H0: no serial correlation

Źródło: Opracowanie własne

Jak widać na powyższym rysunku, hipoteza zerowa o braku autokorelacji została odrzucona. Oznacza to, że nie powinno się interpretować wyników otrzymanych z testu Dickeya- Fullera. Dlatego też kolejnym krokiem było przeprowadzenie rozszerzonego testu, który eliminuje autokorelację reszt.

Rysunek 6 Rozszerzony test szereg niesezonowy

```
. dfuller cpi, nocons reg lags(1) /*test ADF dla pierwszych roznic random walk*/
```

Augmented Dickey-Fuller test for unit root Number of obs = 202

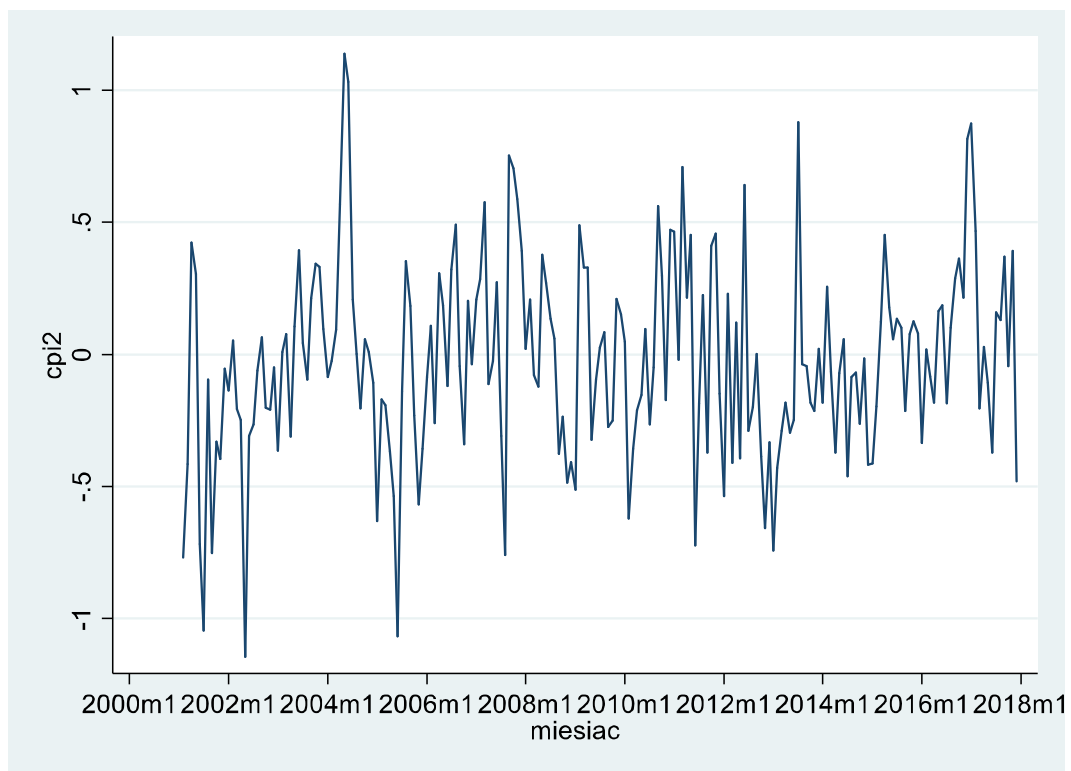
Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-0.625	-2.586	-1.950	-1.617

D.cpi	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
cpi					
L1.	-.0001495	.0002394	-0.62	0.533	-.0006215 .0003225
LD.	.3536748	.0656512	5.39	0.000	.2242174 .4831322

Źródło: Opracowanie własne

Wyniki testu rozszerzonego ADF pokazały, że szereg jest niestacjonarny, gdyż wartość statystyki testu jest większa od wartości krytycznej, więc brak podstaw do odrzucenia hipotezy zerowej o niestacjonarności szeregu, a test Breuscha- Godfrey'a potwierdził brak autokorelacji. Aby więc pozbyć się niestacjonarności, zastosowano pierwsze różnice szeregu.

Rysunek 7 Pierwsze różnice szereg niesezonowy



Źródło: Opracowanie własne

Z powyższego rysunku, który przedstawia wykres szeregu po jednokrotnym zróżnicowaniu jest trudny do interpretacji, jednak można zauważyć, że szereg oscyluje wokół jednego poziomu i odległości między odchyleniami są zbliżonej wielkości. Jednak aby być pewnym, co do stacjonarności tego szeregu, również przeprowadzono odpowiednie testy.

Pierwszym z nich był test Dickeya-Fullera.

Rysunek 8 Pierwsze różnice Dickeya-Fuller szereg niesezonowy

Dickey-Fuller test for unit root		Number of obs = 202		
Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-9.840	-2.586	-1.950	-1.617

D.cpi2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
cpi2 L1.	-.6436774	.0654149	-9.84	0.000	-.7726648	-.51469

Źródło: Opracowanie własne

Test ten pokazuje, że wartość statystyki testowej jest mniejsza od wartości krytycznej, więc odrzucamy hipotezę zerową o niestacjonarności szeregu, na rzecz jego stacjonarności. Takie same wyniki otrzymano, kiedy przeprowadzono ten test uwzględniając stałą. Przeprowadzono również test na autokorelację, ale sprawdzić, czy interpretowanie wyników testu Dickeya- Fullera jest sensowne.

Rysunek 9 Test Breuscha- Godfrey dla pierwszych różnic szereg niesezonowy

. bgodfrey, lag(1/4)

Breusch-Godfrey LM test for autocorrelation

lags (p)	chi2	df	Prob > chi2
1	0.059	1	0.8080
2	0.092	2	0.9551
3	0.133	3	0.9876
4	1.043	4	0.9032

H0: no serial correlation

.

Źródło: Opracowanie własne

Jak widać na powyższym rysunku nie ma podstaw do odrzucenia hipotezy zerowej o autokorelacji. Można więc interpretować test przeprowadzony wcześniej.

Dla upewnienia się, że po jednokrotnym zróżnicowaniu otrzymano szereg stacjonarny, przeprowadzono test KPSS, którego wyniki przedstawia poniższy rysunek.

Rysunek 10 Test KPSS szereg niesezonowy

```
KPSS test for D.cpi

Maxlag = 14 chosen by Schwert criterion
Autocovariances weighted by Bartlett kernel

Critical values for H0: D.cpi is level stationary

10%: 0.347  5% : 0.463  2.5%: 0.574  1% : 0.739

Lag order      Test statistic
  0              .355
  1              .263
  2              .229
  3              .209
  4              .195
  5              .185
  6              .177
  7              .168
  8              .16
  9              .154
 10              .148
 11              .143
 12              .142
 13              .141
 14              .14
```

Źródło: Opracowanie własne

Wartość statystyki testowej jest mniejsza od wartości krytycznej, także brak podstaw do odrzucenia hipotezy zerowej o stacjonarności szeregu.

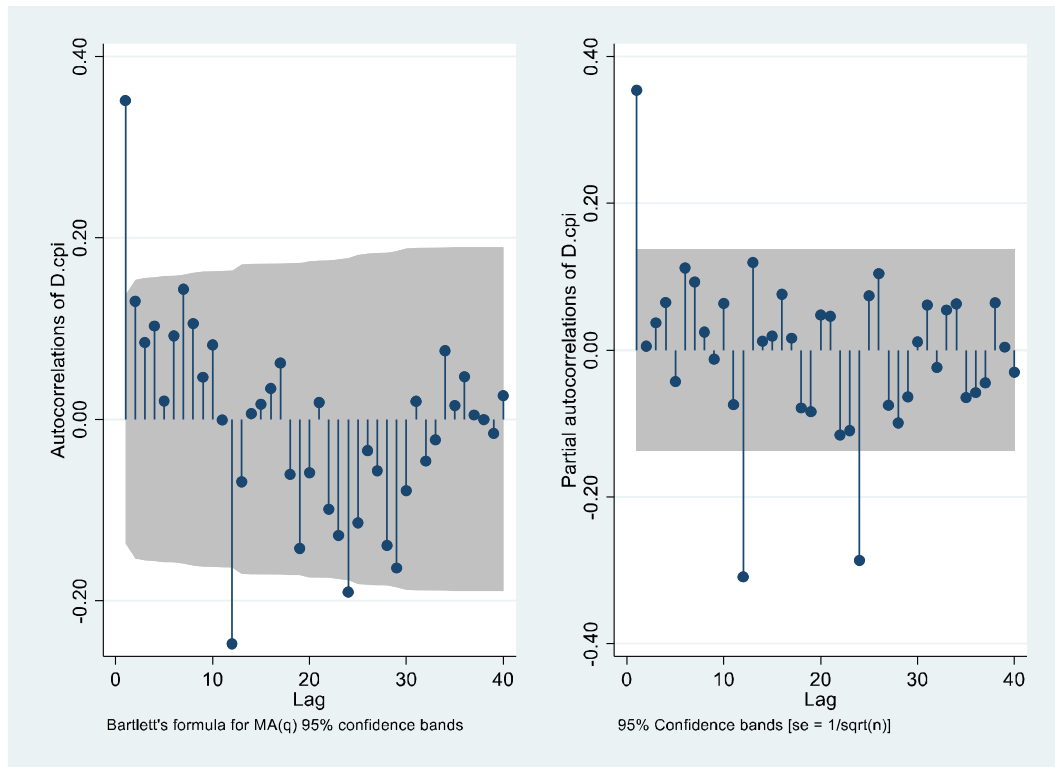
Wystarczyło więc jednokrotne zróżnicowanie szeregu niestacjonarnego, aby otrzymać szereg stacjonarny. Badany szereg jest więc szeregiem zintegrowanym w stopniu pierwszym. Kolejną czynnością, było sprawdzenie, czy otrzymany szereg nie jest białym szumem. Skorzystano z testu Portmanteau, którego wyniki pozwoliły na odrzucenie hipotezy zerowej, że badany szereg jest białym szumem.

Po doprowadzeniu szeregu do stacjonarności oraz sprawdzeniu czy szereg nie jest białym szumem, podjęto próbę oszacowania szeregu za pomocą modelu klasy ARIMA.

1.4 ARIMA

Po wyznaczeniu stopnia zintegrowania szeregu, należy jeszcze wyznaczyć parametry p i q odpowiednio dla procesów AR i MA. Pierwszym przeprowadzonym krokiem była analiza wykresów autokorelacji (ACF) oraz cząstkowej autokorelacji (PACF) dla szeregu zróżnicowanego.

Rysunek 11 Wykresy ACF i PACF szereg niesezonowy



Źródło: Opracowanie własne

Wykresy te pokazują, że istotnie różne od zera są tylko pierwsze słupki. Oznacza to, że $p=q=1$ w przypadku tego szeregu. Wyjściowa postać to ARIMA(1,1,1). Przeprowadzono więc test sprawdzający, czy reszty są białym szumem. Test Portmanteau potwierdził hipotezę zerową, że reszty są białym szumem, więc następnym krokiem była próba zmniejszenia rzędu modelu.

Ponieważ zaczynamy zmniejszanie od ARIMA(1,1,1), możliwe zmniejszania to ARIMA(0,1,1) bądź ARIMA(1,1,0). W celu porównania zmniejszonych modeli zastosowano test ilorazu największej wiarygodności a oraz Kryteria Informacyjne. Najpierw porównano ARIMA(0,1,1). W pierwszym z wymienionych testów otrzymano:

- Prob > chi2 = 0.1051

Test ten pokazał, że brak podstaw do odrzucenia hipotezy zerowej o poprawności zmniejszenia rzędu p. Otrzymane kryteria informacyjne są następujące:

Rysunek 12 Kryteria informacyjne ARIMA(0,1,1)

```
. estimates stats ar111 ar011/*porownanie modeli*/
```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
<u>ar111</u>	203	.	-73.68597	4	155.3719	168.6248
<u>ar011</u>	203	.	-75.22363	3	156.4473	166.3869

Źródło: Opracowanie własne

Jak widać, są one niejednoznaczne. AIC dla ARIMA(1,1,1) są niższe, jednak BIC jest niższy dla ARIMA (0,1,1) co potwierdza wynik otrzymany w teście ilorazu największej wiarygodności. Różnica między wartościami jest bardzo niewielka, dlatego stwierdzono, że ponieważ BIC jest korzystniejszy dla ARIMA(0,1,1) i test ilorazu największej wiarygodności również potwierdza słuszność zastosowanego ograniczenia oraz reszty po przeprowadzeniu testu Portmanteau okazały się być białym szumem, to ARIMA(0,1,1) będzie wykorzystana do oszacowania prognoz.

Te same testy zostały przeprowadzone dla ARIMA(1,1,0). Otrzymując wyniki dla testu ilorazu największej wiarygodności, stwierdzono, że nie ma podstaw do odrzucenia hipotezy zerowej stanowiącej o poprawności narzuconego ograniczenia. Kryteria informacyjne wyglądają następująco:

Rysunek 13 Kryteria Informacyjne ARIMA(1,1,0)

```
. estimates stats ar111 ar110/*porownanie modeli*/
```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
<u>ar111</u>	203	.	-73.68597	4	155.3719	168.6248
<u>ar110</u>	203	.	-73.6894	3	153.3788	163.3184

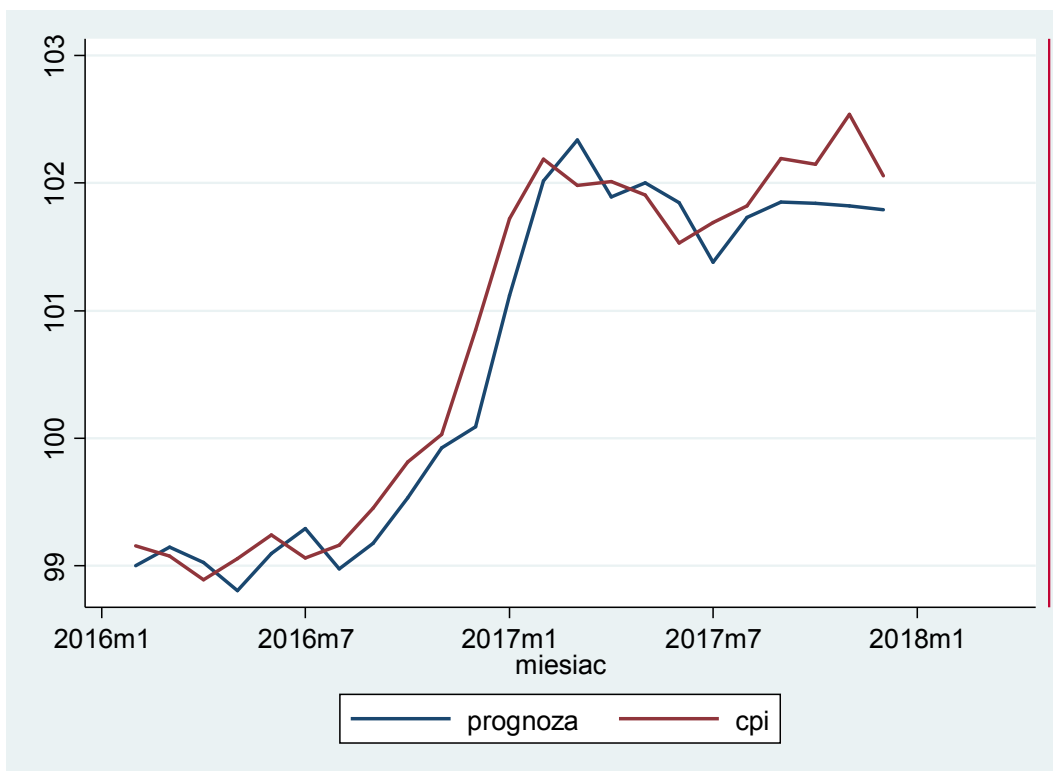
Źródło: Opracowanie własne

W tym przypadku nie ma wątpliwości, że $ARIMA(1,1,0)$ jest modelem efektywniejszym od pierwotnego. Potwierdzają to wartości AIC i BIC, gdyż w obydwóch przypadkach są niższe niż dla $ARIMA(1,1,1)$. Również w tym przypadku reszty okazały się być białym szumem, więc również $ARIMA(1,1,0)$ będzie wykorzystana do oszacowania prognoz.

1.5 Prognoza

Po zmniejszeniu rzędu p i q ostatecznie przyjęto, że do prognozowania zostanie użyta $ARIMA(0,1,1)$ oraz $ARIMA(1,1,0)$, na których zastosowano prognozę dynamiczną. Prognozy te dodatkowo porównano do prognozy otrzymanej poprzez wykładnicze wygładzenie. Na początku przeprowadzono prognozę dynamiczną dla $ARIMA(1,1,0)$.

Rysunek 14 Prognoza dynamiczna $ARIMA(1,1,0)$

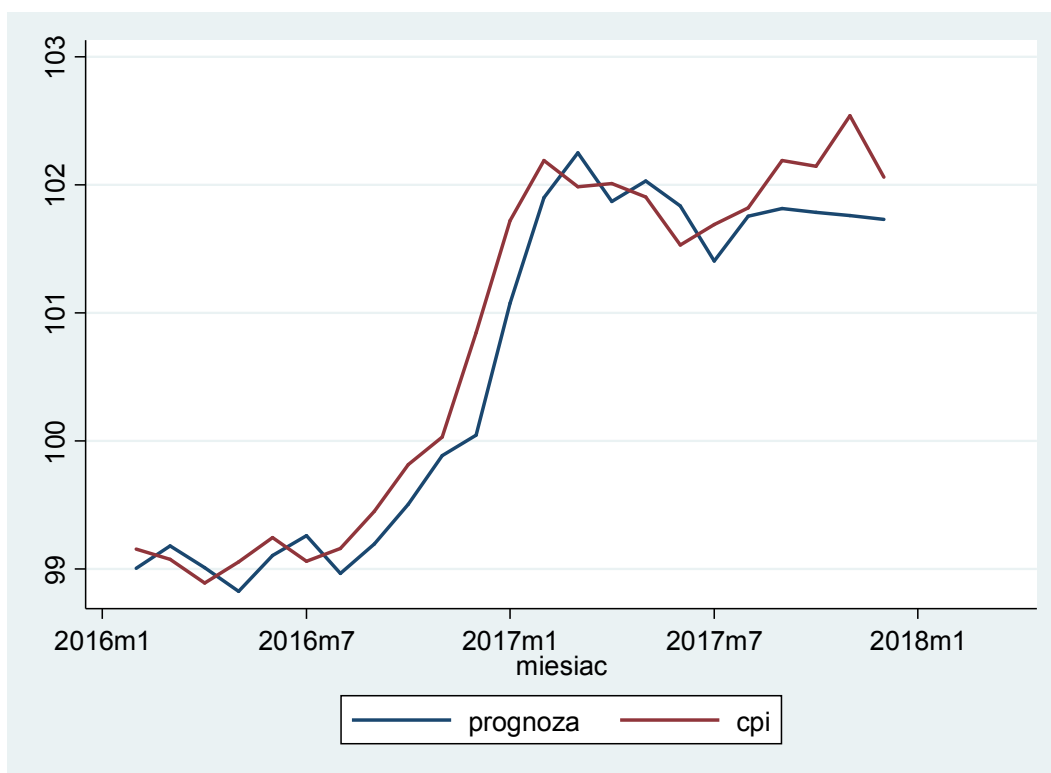


Źródło: Opracowanie własne

Niebieską linią jest zaznaczona prognoza dynamiczna (out-of-sample to okres 4 ostatnich miesięcy), jak widać prognozy oraz wartości znacząco różnią się od siebie.

Prognoza dynamiczna oszacowana na podstawie $ARIMA(0,1,1)$ jest pokazana na poniższym rysunku.

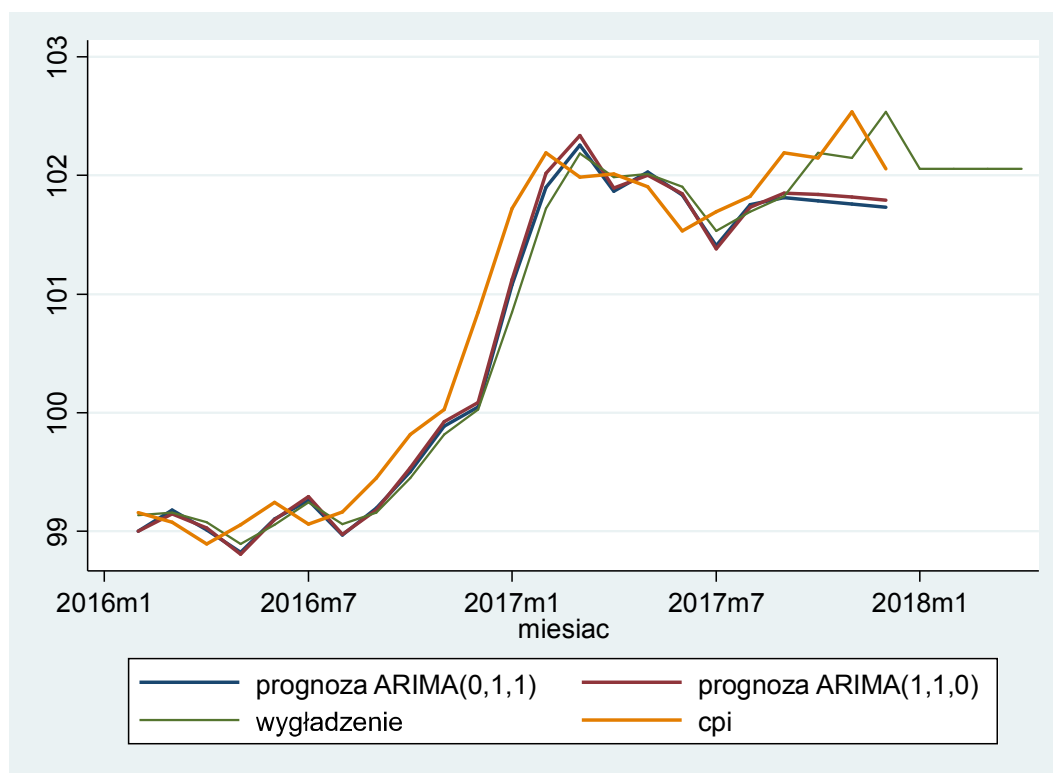
Rysunek 15 Prognoza dynamiczna ARIMA(0,1,1)



Źródło: Opracowanie własne

Prognoza ta jest zbliżona to prognozy pierwszej, co doskonale widać na poniższym rysunku, na którym są oby dwie przeprowadzone prognozy oraz prognoza ekstrapolacyjna oszacowaną za pomocą wykładniczego.

Rysunek 16 Porównanie prognoz ARIMA



Źródło: Opracowanie własne

Niebieska linia to prognoza oszacowana na podstawie ARIMA(0,1,1), czerwona na podstawie ARIMA(1,1,0), zielona to za pomocą wygładzenia wykładniczego. Widać, że najbardziej zbliżone wartości tylko przesunięte są dla prognoz z wygładzenia wykładniczego, ale oszacowano błędy MAE (średni błąd bezwzględny) oraz MSE (błąd średniokwadratowy)

Tabela 1 Porównanie prognoz

	ARIMA(1,1,0)	ARIMA(0,1,1)	Wygładzenie
MAE	.4610424	.4085751	.3214741
MSE	.2466584	.1998655	.1305393

Źródło: Opracowanie własne

Jak pokazują wyniki zawarte w Tabeli 1 błędy prognoz są wysokie w każdym przypadku. Potwierdziła się natomiast hipoteza, że to prognozy otrzymane za pomocą wygładzenia wykładniczego pokazują najmniejsze błędy. Tak wysokie błędy mogą być spowodowane tym, że dla szeregów, dla których wartości są mniejsze od 1 lepiej jest zastosować względne błędy, a w tym szeregu dużo obserwacji jest mniejsza od 1.

Na końcu została przeprowadzona jeszcze prognoza za pomocą metody Holta, jednak błędy tej prognozy były jeszcze wyższe od prognozy oszacowanej metodą wygładzenia wykładniczego.

1.6 Podsumowanie

Szereg niesezonowy okazał się być zintegrowany w stopniu pierwszym i już pierwsze różnice wystarczyły, aby otrzymać szereg stacjonarny. Oszacowano dwa modele ARIMA-ARIMA(1,1,0) oraz ARIMA (0,1,1). Następnie oszacowano prognozy za pomocą tych dwóch modeli oraz za pomocą metody ekstrapolacyjnej. Wszystkie prognozy są obarczone dużym błędem. Najwyższy błąd został obliczony dla prognozy modelu ARIMA(1,1,0) a najniższy dla metody ekstrapolacyjnej.

2 Szereg sezonowy

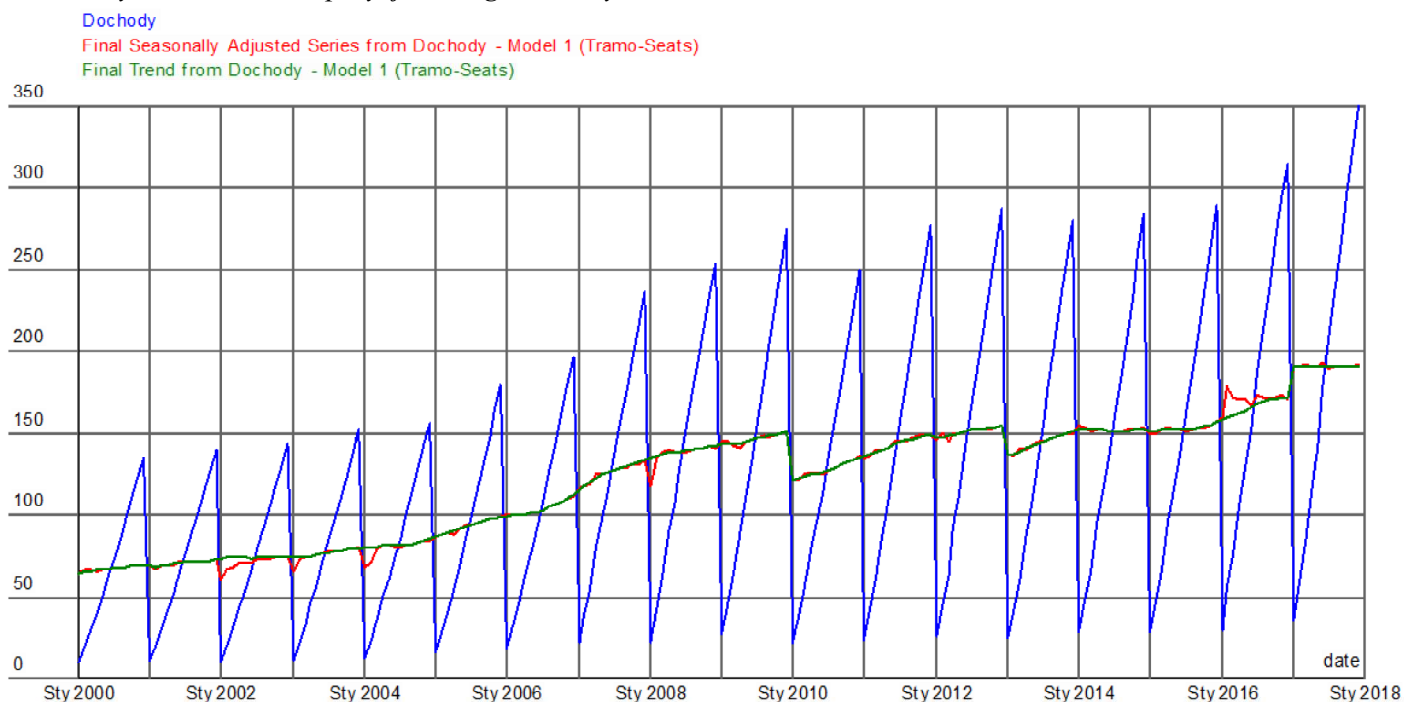
2.1 Opis danych

Szereg sezonowy został uzyskany przez pobranie danych ze strony <https://bdm.stat.gov.pl/> ministerstwa finansów w milionach złotych. Dostęp do strony uzyskano 13 marca 2018 roku. Dane dotyczą budżetu państwa od początku roku do grudnia. Są to dane miesięczne, gdzie pierwszą datą jest styczeń 2000 roku a ostatnią grudzień 2017.

2.2 Dekompozycja szeregu

Dekompozycja szeregu została dokonana przy użyciu programu DEMETRA. Jak się okazało, program ten zlogarytmował szereg. Jest to wskazówka do zlogarytmowania szeregu przed przeprowadzeniem na nim testów, dopasowania postaci ARIMA SARIMA. Dekompozycja nie wykazała obserwacji odstających.

Rysunek 17 Dekompozycja szereg sezonowy



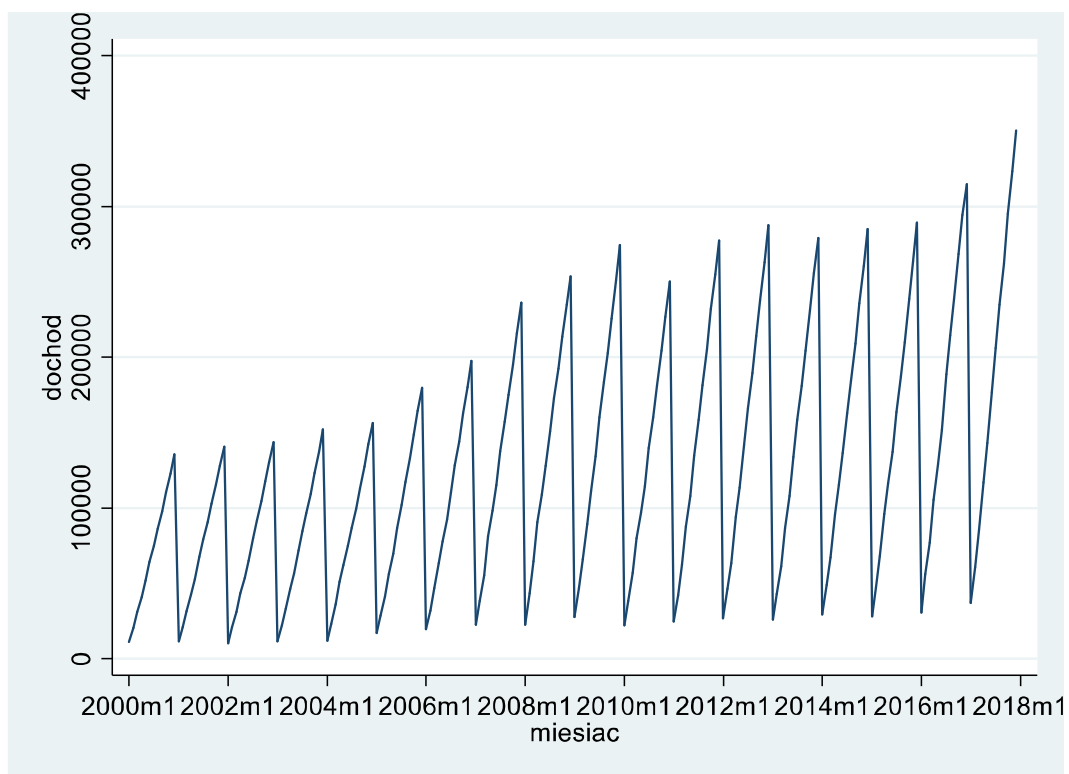
Źródło: Opracowanie własne

Kolor niebieski to szereg sezonowy, czerwonym kolorem zaznaczono odsezonowany szereg a kolor zielony jest przypisany trendowi. Można zauważyć, że sezonowość multiplikatywną. Trend w dużej części pokrywa się z szeregiem odsezonowanym.

2.3 Stacjonarność

Zanim będzie można przejść do dopasowania postaci ARIMA SARIMA należy sprawdzić, czy szereg jest stacjonarny. Na poniższym rysunku jest surowy szereg. Z rysunku można odczytać, że szereg nie jest stacjonarny co najmniej ze względu na średnią i wariancję.

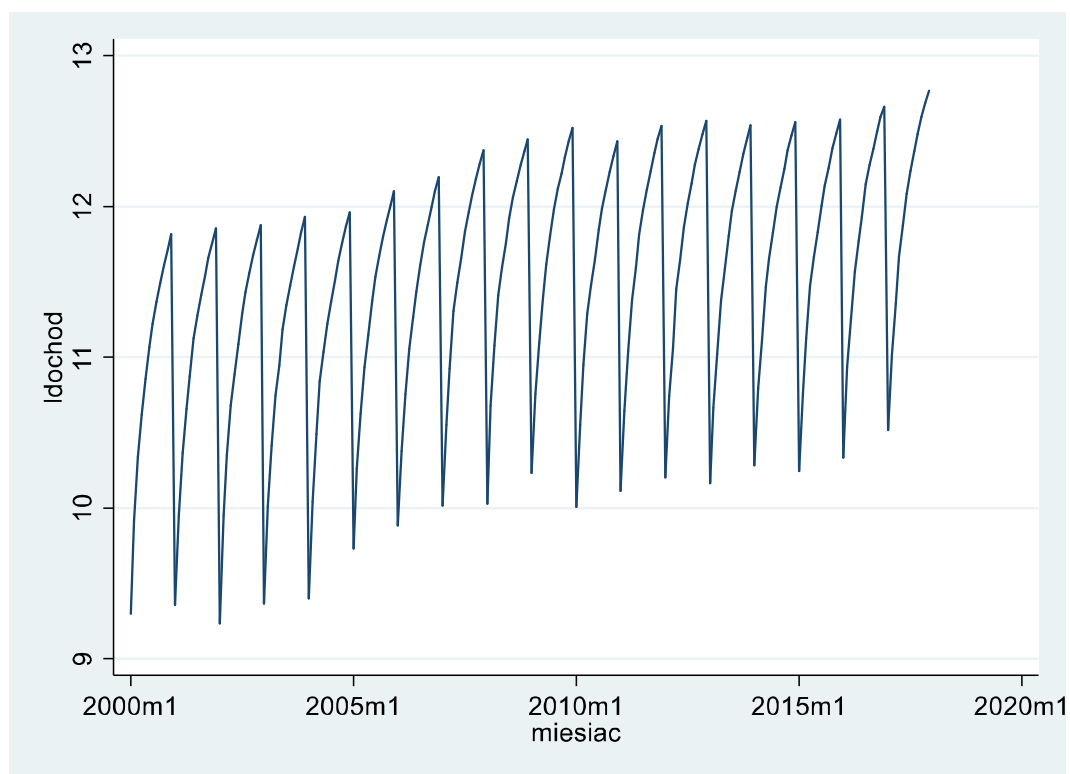
Rysunek 18 Szereg sezonowy surowy



Źródło: Opracowanie własne

Ponieważ szereg wykazuje sezonowość multiplikatywną oraz program DEMETRA zlogarytmował szereg, pierwszą dokonaną czynnością było również zlogarytmowanie szeregu. Na następnym rysunku jest pokazany szereg po zlogarytmowaniu.

Rysunek 19 Szereg zlogarytmowany



Źródło: Opracowanie własne

Interpretując powyższy rysunek nadal nie można się dopatrzeć stacjonarności. Przeprowadzono więc test Dickeya-Haszy-Fullera, aby sprawdzić czy konieczne jest odsezonowanie szeregu. Przeprowadzony test pokazał, że wystarczy jednokrotne odsezonowanie szeregu. Jednak aby móc interpretować otrzymane wyniki sprawdzono za pomocą testu Breuscha-Godfrey autokorelację. Jak widać na rysunku 19 brak podstaw do odrzucenia hipotezy zerowej o autokorelacji.

Rysunek 20 Test na autokorelację

Breusch-Godfrey LM test for autocorrelation

lags (p)	chi2	df	Prob > chi2
1	2.736	1	0.0981
2	2.943	2	0.2296
3	3.704	3	0.2952
4	3.705	4	0.4474

H0: no serial correlation

Źródło: Opracowanie własne

W ten sposób można już zdefiniować wielkości D i S, gdzie D=1 i S=12, gdyż są to dane miesięczne.

Po przeprowadzeniu testu Dickeya- Haszy- Fullera, kolejnym krokiem było zbadanie stopnia zintegrowania szeregu. W tym celu przeprowadzono test Dickeya-Fullera. Jak widać na rysunku 20, należy odrzucić hipotezę zerową o niestacjonarności na rzecz stacjonarności szeregu.

Rysunek 21 Dickey- Fuller szereg sezonowy

Test Statistic	Interpolated Dickey-Fuller		
	1% Critical Value	5% Critical Value	10% Critical Value
Z (t)	-4.625	-3.476	-2.883

Źródło: Opracowanie własne

Po przeprowadzeniu testu na autokorelację, okazało się, że nie ma powodu do odrzucenia hipotezy zerowej o braku autokorelacji. Jednak przeprowadzono również test KPSS, który dla pierwszych dwóch opóźnień pokazuje niestacjonarność, dlatego też zróżnicowano raz zmienną i ponownie przeprowadzono test Dickeya- Fullera, Breuscha- Godfrey'a i test KPSS. Test Dickeya- Fullera dla zróżnicowanych zmiennych dał wyniki pozwalające na odrzucenie hipotezy zerowej o niestacjonarności, ponieważ wartość statystyki testowej wynosiła -17, 591 a wartość krytyczna -2,833. Test Breuscha- Godfrey'a po raz kolejny pokazał brak korelacji.

Rysunek 22 Test Breuscha- Godfrey'a po zróżnicowaniu

Breusch-Godfrey LM test for autocorrelation

lags (p)	chi2	df	Prob > chi2
1	2.463	1	0.1166
2	3.645	2	0.1616
3	4.205	3	0.2402
4	5.145	4	0.2727

H0: no serial correlation

Źródło: Opracowanie własne

W tym przypadku interpretacja testu KPSS jest taka sama, jak dla Dickeya- Fullera.

Rysunek 23 Test KPSS szereg sezonowy

Critical values for H0: DS12.ldochod is trend stationary

10%: 0.119 5% : 0.146 2.5%: 0.176 1% : 0.216

Lag order	Test statistic
0	.0135
1	.0172
2	.0213
3	.0251
4	.0282
5	.0281
6	.0286
7	.0304
8	.0319
9	.0338
10	.0336
11	.0314
12	.0353
13	.0387
14	.0414

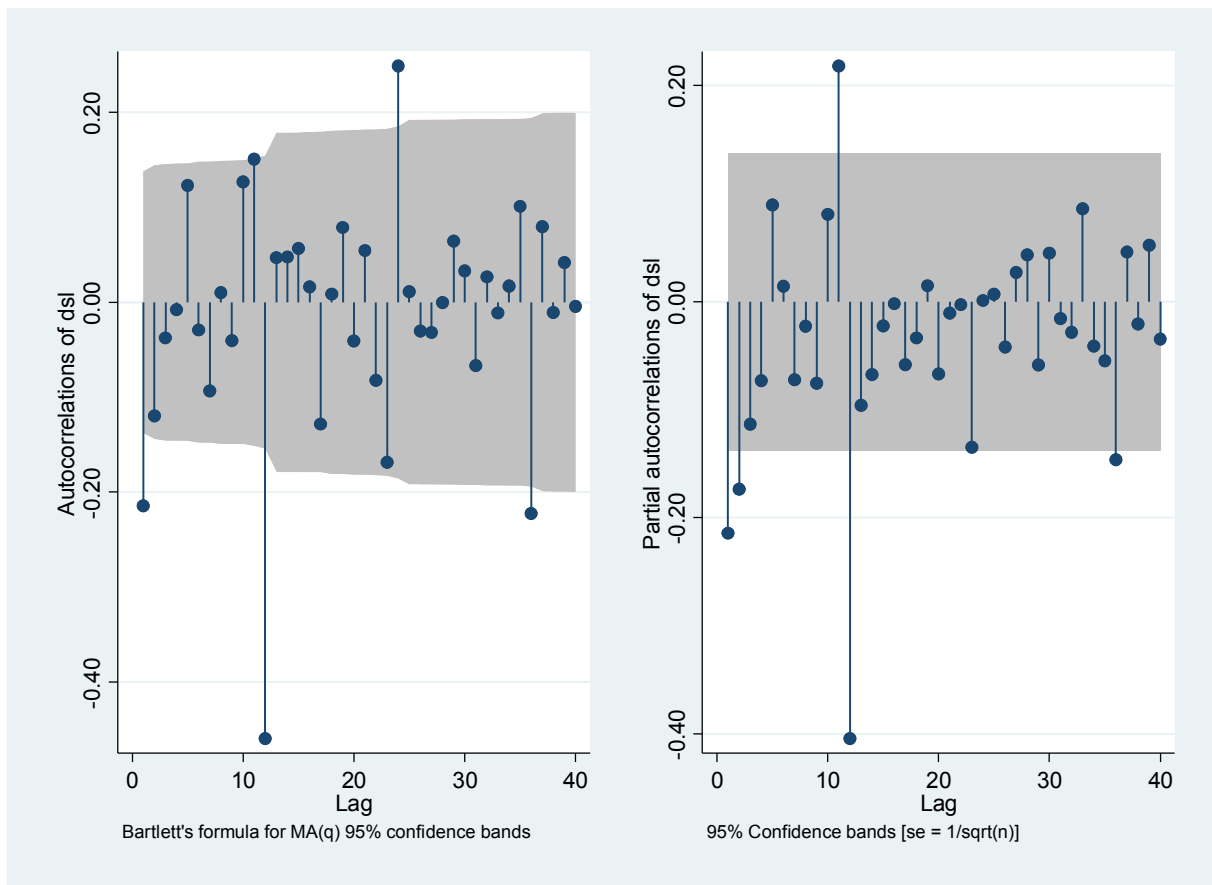
Źródło: Opracowanie własne

Wszystkie wartości statystyczne są mniejsze od wartości krytycznej, dlatego nie ma podstaw do odrzucenia hipotezy zerowej o stacjonarności. Dzięki przeprowadzonym testom, okazało się, że szereg jest zintegrowany w stopniu pierwszym.

2.4 SARIMA

Wiadomo już, że szereg jest zintegrowany w stopniu pierwszym. Można więc teraz podjąć próbę określenia kolejnych parametrów w modelu ARIMA SARIMA. Na początku dokonano graficznej analizy wykresów autokorelacji (ACF) oraz cząstkowej autokorelacji (PACF) dla szeregu zróżnicowanego.

Rysunek 24 Wykres ACF PACF szereg sezonowy



Źródło: Opracowanie własne

Z wykresu ACF można odczytać, że $Q=3$, bo istotnie różne od zera słupki są na 12, 24 i 36 miejscu. Natomiast wykres PACF wygasa, więc $P=0$. Na podstawie tych samych rysunków można odczytać, że $p=2$, a $q=1$. Wynika z tego, że potencjalny efektywny model wygląda następująco: ARIMA(2,1,1) SARIMA(0,1,3,12).

Stata IC nie poradziła sobie z wielkością $Q=3$, dlatego też porównywano modele zaczynając od $Q=2$. Porównania zostały przeprowadzone wykorzystując test ilorazu największej wiarygodności oraz Kryteriów Informacyjnych. Zaczęto od zmniejszanie p do 1. Wywołując model pierwotny zauważono, że oprócz stałej inne zmienne też są nie istotne, więc na pewno któryś z modeli ze zmniejszonymi parametrami okaże się bardziej efektywny. Z wyników testu ilorazu największej wiarygodności okazało się, że nie ma podstaw do odrzucenia hipotezy zerowej o poprawnie narzuconym ograniczeniu. Test sam test porównujący model pierwotny z modelem z wartością p równą 0 również pokazał, że nie odrzucamy test hipotezy zerowej o kolejnym narzuconym ograniczeniu. Następni próbowano zmniejszyć wielkość q , jednak test ilorazu największej wiarygodności pokazał, że należy odrzucić hipotezę zerową o

poprawności narzuconego ograniczenia. Test ten przeprowadzono również dla zmniejszonej wielkości Q do 1. Każdy z modeli okazał się być efektywniejszy od pierwotnego. Aby wyodrębnić najefektywniejsze modele, zastosowano Kryteria Informacyjne. Na poniższym rysunku przedstawiono wyniki tego testu.

Rysunek 25 Kryteria Informacyjne szereg sezonowy

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
<u>ar211</u>	203	.	344.1574	7	-674.3148	-651.1224
<u>ar111</u>	203	.	344.1541	6	-676.3082	-656.429
<u>ar011</u>	203	.	342.4941	5	-674.9882	-658.4221
<u>ar110</u>	203	.	340.617	5	-671.234	-654.6679
<u>ar211 s1</u>	203	.	344.1067	6	-676.2133	-656.3341
<u>ar111 s1</u>	203	.	344.1024	5	-678.2049	-661.6388
<u>ar0111 s1</u>	203	.	342.3925	4	-676.7849	-663.5321
<u>ar210 s1</u>	203	.	342.8431	5	-675.6861	-659.1201
<u>ar110 s1</u>	203	.	340.4517	4	-672.9035	-659.6506
<u>ar110 s1</u>	203	.	340.4517	4	-672.9035	-659.6506

Źródło: Opracowanie własne

Jak można zauważyć najniższą wartość AIC i drugą najniższą wartość BIC ma model SARIMA(1,1,1)(0,1,1,12)- dalej nazywany modelem 1, natomiast model SARIMA(0,1,1)(0,1,1,12)- dalej nazywany modelem2- ma najniższą wartość BIC i drugą najniższą wartość AIC. Dlatego też prognozy zostały przeprowadzone na oby dwóch modelach. Sprawdzone jeszcze czy reszty w tych dwóch modelach są białym szumem za pomocą testu Portmanteau

Rysunek 26 Biały szum SARIMA(1,1,1)(0,1,1,12)

Portmanteau test for white noise

Portmanteau (Q) statistic = 29.3228
 Prob > chi2(40) = 0.8932

Źródło: Opracowanie własne

Rysunek 27 Biały szum SARIMA(0,1,1)(0,1,1,12)

Portmanteau test for white noise

Portmanteau (Q) statistic =	34.0314
Prob > chi2(40) =	0.7350

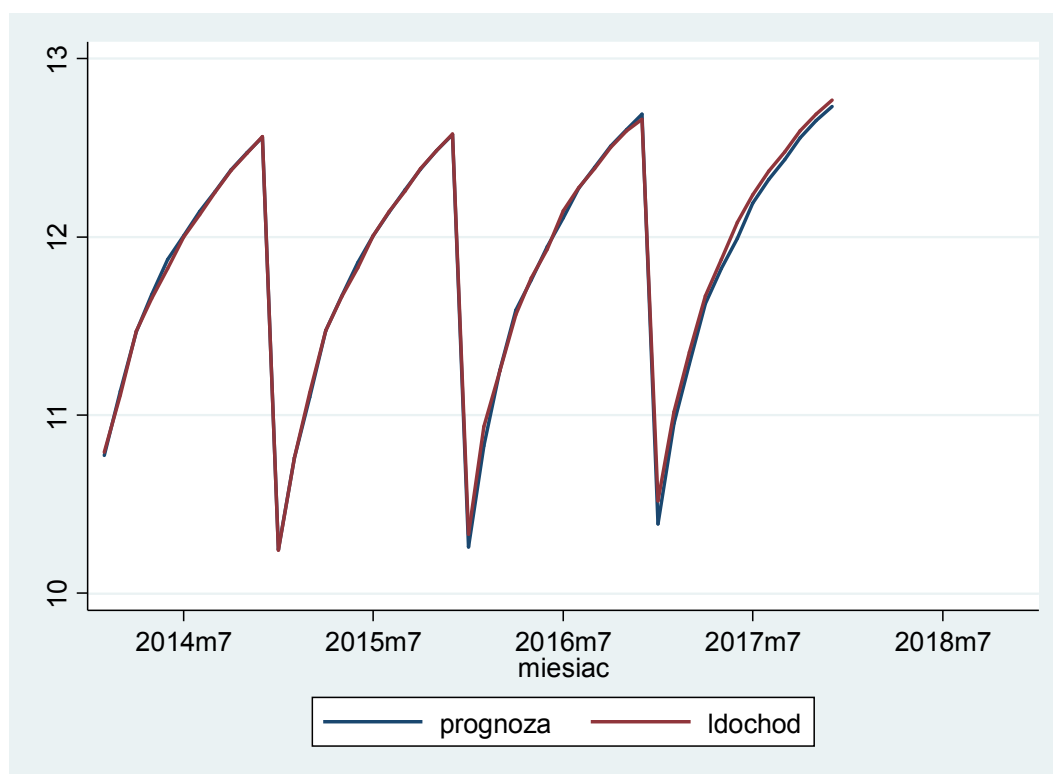
Źródło: Opracowanie własne

Jak widać brak podstaw do odrzucenia hipotezy zerowej, że reszty są białym szumem.

2.5 Prognozy

Na kolejnych dwóch rysunkach zostały pokazane prognozy dynamiczne dla modelu 1 i modelu 2.

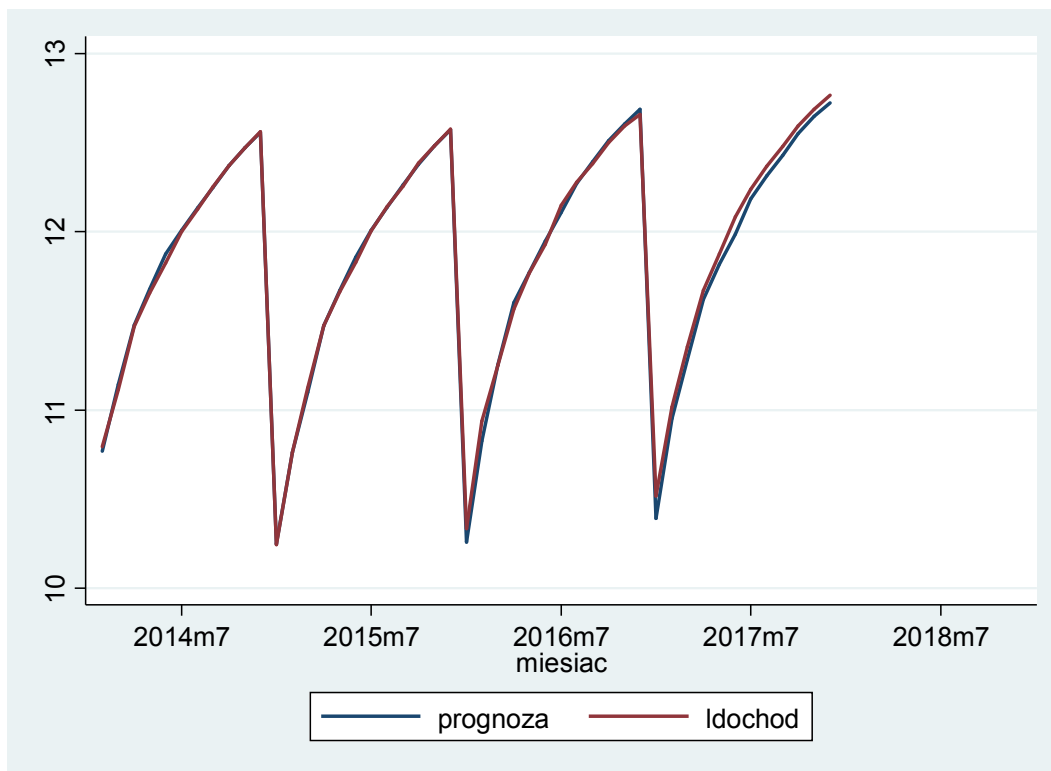
Rysunek 28 Prognoza dynamiczna model 1 szereg sezonowy



Źródło: Opracowanie własne

Kolorem niebieskim jest zaznaczona prognoza. Out-of-sample to okres od stycznia 2017 roku. Jak widać prognoza a zbliżone wielkości do szeregu.

Rysunek 29 Prognoza dynamiczna model 2 szereg sezonowy

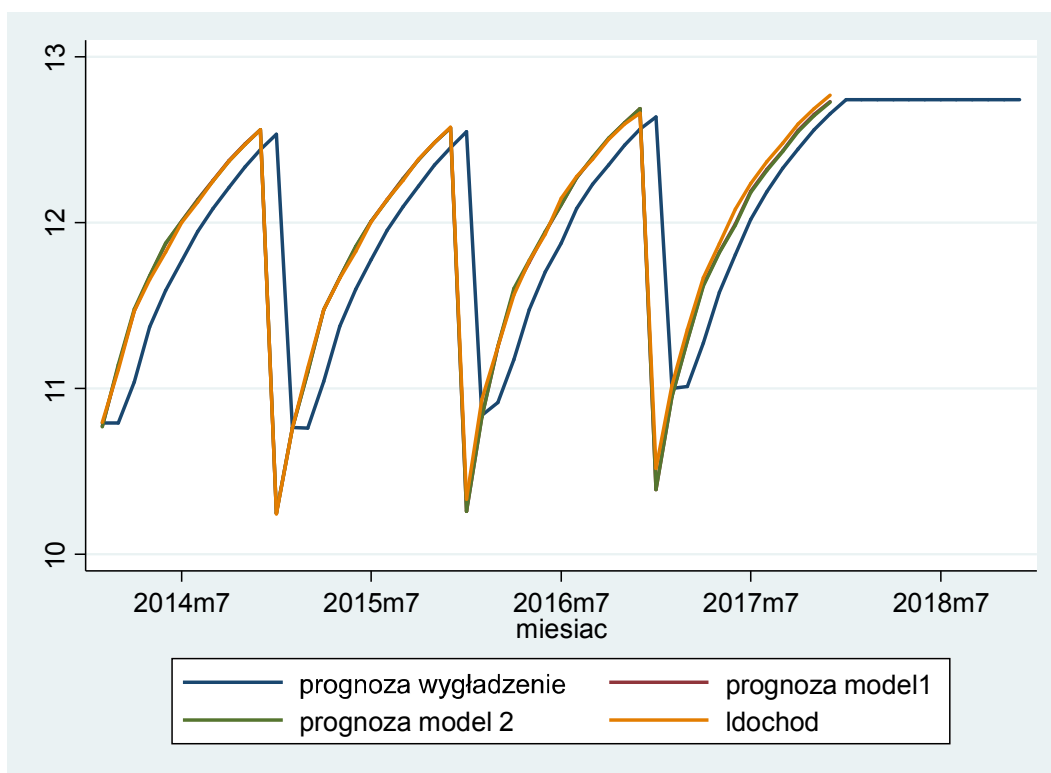


Źródło: Opracowanie własne

Prognoza przeprowadzona za pomocą modelu 2 również pokazuje zbliżone wartości do szeregu.

Następną przeprowadzoną prognozą, była prognoza zastosowana za pomocą wykładzenia wykładniczego. Na kolejnym rysunku pokazano wszystkie 3 prognozy oraz badany szereg.

Rysunek 30 Prognozy szereg sezonowy



Źródło: Opracowanie własne

Jak można zauważyć, najmniej z logarytmem dochodu pokrywa się prognoza oszacowana za pomocą wygładzenia wykładniczego. Jednak aby pokazać dokładne różnice obliczono błędy MAE (średni błąd bezwzględny) oraz MSE (błąd średniokwadratowy).

Tabela 2 Błędy prognoz szereg sezonowy

	Model 1	Model 2	Wygładzenie
MAE	.05756092	.06264361	.36463365
MSE	.00396371	.0045073	.42339426

Źródło: Opracowanie własne

Jak można zauważyć z tabeli 2, najniższe wartości błędy prognoz są dla modelu 1, czyli SARIMA(1,1,1)(1,1,1,12). Najwyższe wielkości błędów są dla prognoz oszacowanych za pomocą wygładzenia wykładniczego. Tutaj również, podobnie jak w szeregu niesezonowym, oszacowano dodatkowe prognozy za pomocą metody Holta, jednak błędy były wyższe niż te otrzymane z metody wykładniczej.

2.6 Podsumowanie

Szereg sezonowy okazał się być zintegrowany w pierwszym stopniu. Wyodrębniono dwa najefektywniejsze modele SARIMA- SARIMA(0,1,1)(0,1,1,12) oraz SARIMA(1,1,1)(1,1,12). Następnie oszacowano prognozy na podstawie tych modeli, metodą wygładzenia wykładniczego oraz metodą Holta. Najniższe błędy prognoz uzyskała prognoza oszacowana za pomocą modelu 1(SARIMA(1,1,1)(0,1,1,12). Modele ekstrapolacyjne okazały się być mniej efektywne od SARIMY.

3 Bibliografia

<http://www.nbp.pl/home.aspx?f=/statystyka/bazowa/bazowa.htm>, 13.03.2018

<https://bdm.stat.gov.pl/>, 12.03,2018

4 Spis rysunków

Rysunek 1 Dekompozycja szereg niesezonowy	3
Rysunek 2 Szereg niesezonowy	4
Rysunek 3 Test Boxa-Coxa szereg niesezonowy	4
Rysunek 4 Test Dickeya-Fullera bez stałej	5
Rysunek 5 Test Breuscha- Godfrey'a szereg niesezonowy	6
Rysunek 6 Rozszerzony test szereg niesezonowy	7
Rysunek 7 Pierwsze różnice szereg niesezonowy	7
Rysunek 8 Pierwsze różnice Dickeya-Fuller szereg niesezonowy	8
Rysunek 9 Test Breuscha- Godfrey'a dla pierwszych różnic szereg niesezonowy	8
Rysunek 10 Test KPSS szereg niesezonowy	9
Rysunek 11 Wykresy ACF i PACF szereg niesezonowy	10
Rysunek 12 Kryteria informacyjne ARIMA(0,1,1)	11
Rysunek 13 Kryteria Informacyjne ARIMA(1,1,0)	11
Rysunek 14 Prognoza dynamiczna ARIMA(1,1,0)	12
Rysunek 15 Prognoza dynamiczna ARIMA(0,1,1)	13
Rysunek 16 Porównanie prognoz ARIMA	14
Rysunek 17 Dekompozycja szereg sezonowy	16
Rysunek 18 Szereg sezonowy surowy	17
Rysunek 19 Szereg zlogarytmowany	18
Rysunek 20 Test na autokorelację	18
Rysunek 21 Dickey- Fuller szereg sezonowy	19
Rysunek 22 Test Breuscha- Godfrey'a po zróżnicowaniu	19
Rysunek 23 Test KPSS szereg sezonowy	20
Rysunek 24 Wykres ACF PACF szereg sezonowy	21
Rysunek 25 Kryteria Informacyjne szereg sezonowy	22
Rysunek 26 Biały szum SARIMA(1,1,1)(0,1,1,12)	22
Rysunek 27 Biały szum SARIMA(0,1,1)(0,1,1,12)	23
Rysunek 28 Prognoza dynamiczna model 1 szereg sezonowy	23
Rysunek 29 Prognoza dynamiczna model 2 szereg sezonowy	24
Rysunek 30 Prognozy szereg sezonowy	25

5 Spis tabel

Tabela 1 Porównanie prognoz	14
Tabela 2 Błędy prognoz szereg sezonowy	25