

# Prognozowanie stopy bezrobocia rejestrowanego

Patryk Szymkowiak

Krzysztof Ignasiak

## 1. Tematyka

Stopa bezrobocia określa jaki procent osób w wieku produkcyjnym, gotowych oraz chętnych do podjęcia pracy nie jest zatrudnionych ani nie prowadzi własnych działalności gospodarczych. Możliwość zasadnego prognozowania wyżej określonego współczynnika może być pożyteczna zarówno dla instytucji państwowych jak i przedsiębiorstw prywatnych. W pierwszym przypadku należy mieć na uwadze, że jest to jedna z najważniejszych zmiennych makroekonomicznych, a otrzymane prognozy mogą wpływać na decyzje podejmowane przez m.in. rząd i Narodowy Bank Polski w szczególności dotyczące poziomu stóp procentowych. Ponad to dane o prawdopodobnym przyszłym poziomie bezrobocia mogą wpływać na charakter i skalę wykorzystywanych instrumentów aktywizacji zawodowej. Natomiast przedsiębiorstwa prywatne wykorzystując prognozy dotyczące bezrobocia mogą uaktualniać swoje oczekiwania w stosunku do przewidywanego popytu, a także wysokości pensji jakie należałoby zaproponować nowo pozyskanym pracownikom.

## 2. Dane

Wszystkie wykorzystane dane pochodzą z zasobów Głównego Urzędu Statystycznego. W kontekście zmiennej objaśnianej został wykorzystany szereg czasowy zawierający informacje o stopie bezrobocia rejestrowanego w każdym z miesięcy w latach 1990 – 2020.

## 3. Wykorzystane narzędzia.

Większość obliczeń wykonano z użyciem Excela, którego funkcjonalności zostały rozszerzone z wykorzystaniem BRETA. Jest to oprogramowanie rozpowszechniane na otwartej licencji GLP v.3, które umożliwia uruchamianie w Excelu funkcji z języka R. Opis instalacji i sposobu weryfikacji obliczeń znajduje w aneksie nr. 1.

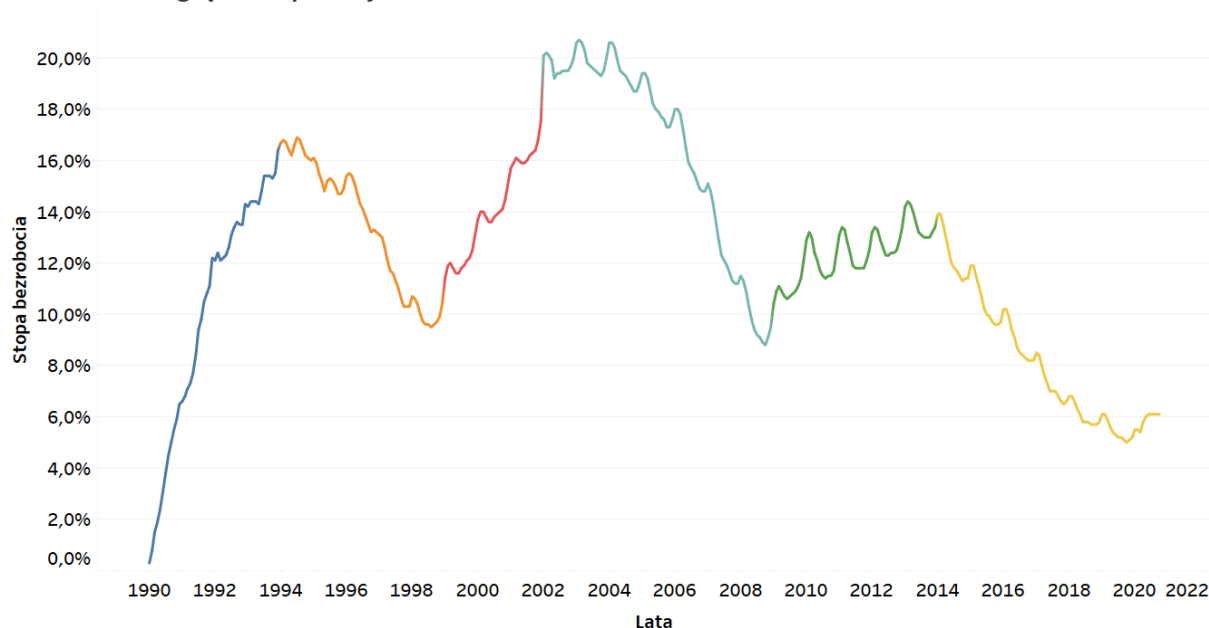
## 4. Test Chow'a

Test Chow'a służy do badania stabilności parametrów liniowego modelu prognostycznego. Jego idea polega na obliczeniu trzech modeli regresyjnych: 1. dla wszystkich obserwacji, 2. dla obserwacji poprzedzający prawdopodobny punkt zwrotnych, 3. dla obserwacji następujących po punkcie zwrotnym. Jeżeli nie występuje zmiana trendu to parametry wszystkich modeli nie powinny różnić się od siebie w sposób istotnie statystyczny. Na podstawie przygotowanej wizualizacji zostało wybranych 5 możliwych punktów zwrotnych, znajdujących się na przełomie lat:

1. 1994 – 1995
2. 1998 – 1999
3. 2001 – 2002
4. 2008 – 2009
5. 2013 - 2014

Po każdorazowym przeprowadzeniu testu Chow'a okazało się, że w każdym przypadku miała miejsce istotna zmiana trendu w związku z czym do dalszych analiz zostały wykorzystane tylko najnowsze obserwacje tj. z lat 2014 – 2020.

## Podział ze względu na punkty zwrotne



## Testy Chow'a

	Test1	Test2	Test3	Test4	Test5
Numer obserwacji 1	1-48	49-108	109-144	145-228	229-288
Numer obserwacji 2	49-108	109-144	145-228	229-288	288-360
Lata 1	1990-1994	1995-1998	1999 - 2001	2002 - 2008	2009 - 2013
Lata 2	1995-1998	1999-2001	2002 - 2008	2009 - 2013	2014 - 2019
Wartość F	875,47	776,57	386,05	153,39	451,22
Stopnie swobody 1	2	2	2	2	2
Stopnie swobody 2	104	92	116	141	129
Wartość p	8,5615E-66	2,44858E-58	5,34305E-52	4,17139E-36	5,8363E-59
czy punkt zwrotny	Tak	Tak	Tak	Tak	Tak

## 5. Metody analizy szeregu czasowego

Wykorzystaliśmy poniższe metody prognozowania szeregów czasowych:

1. Metoda naiwna
2. Średnie ważone
3. Średnie ruchome
4. Średnie ruchome z tendencją rozwojową bezwzględną/względną
5. Proste wygładzanie wykładnicze
6. Podwójne wygładzanie wykładnicze
7. Model Wintersa w wersji multiplikatywnej.

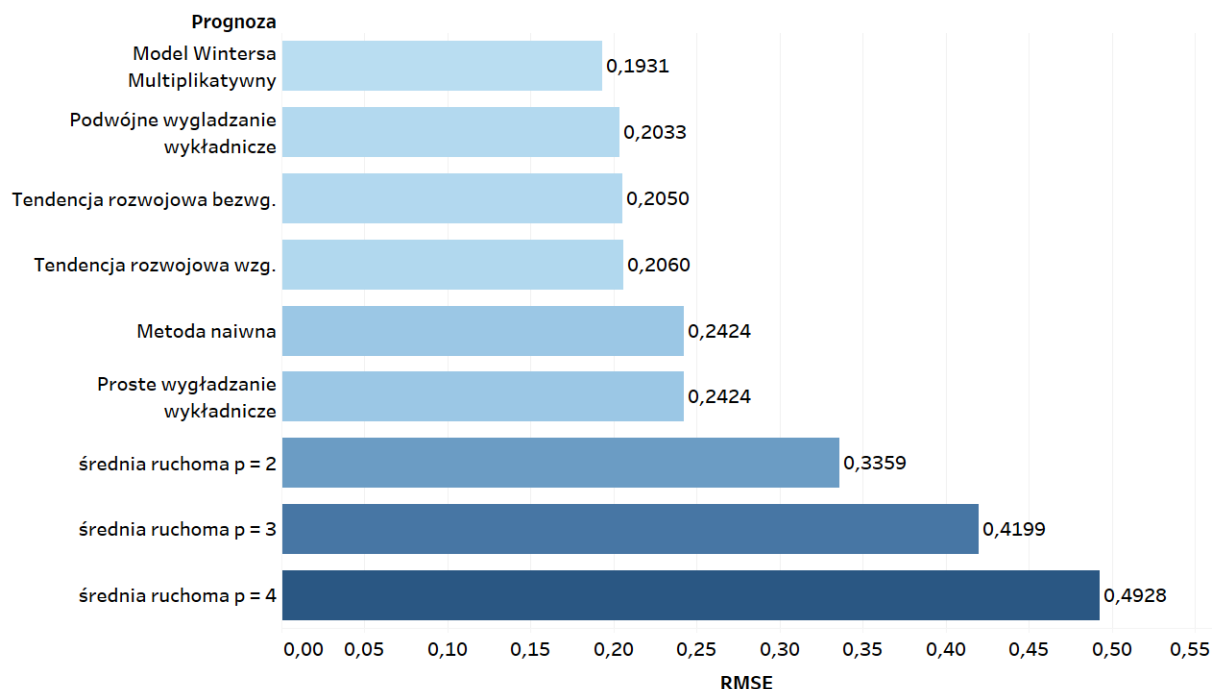
Punktem wyjścia stanowi oczywiście metoda naiwna, która polega na przyjęciu założenia, że następna obserwacja będzie równa poprzedniej. Natomiast kolejne modele są coraz bardziej rozbudowane w celu weryfikacji, czy bardziej rozbudowana postać analityczna umożliwi otrzymanie dokładniejszych oszacowań.

Metody oparte na średnich ważonych zostały uznane za niedopuszczalne, gdyż oszacowane błędy RMSE stanowiły ponad 20% wartości prognozy.

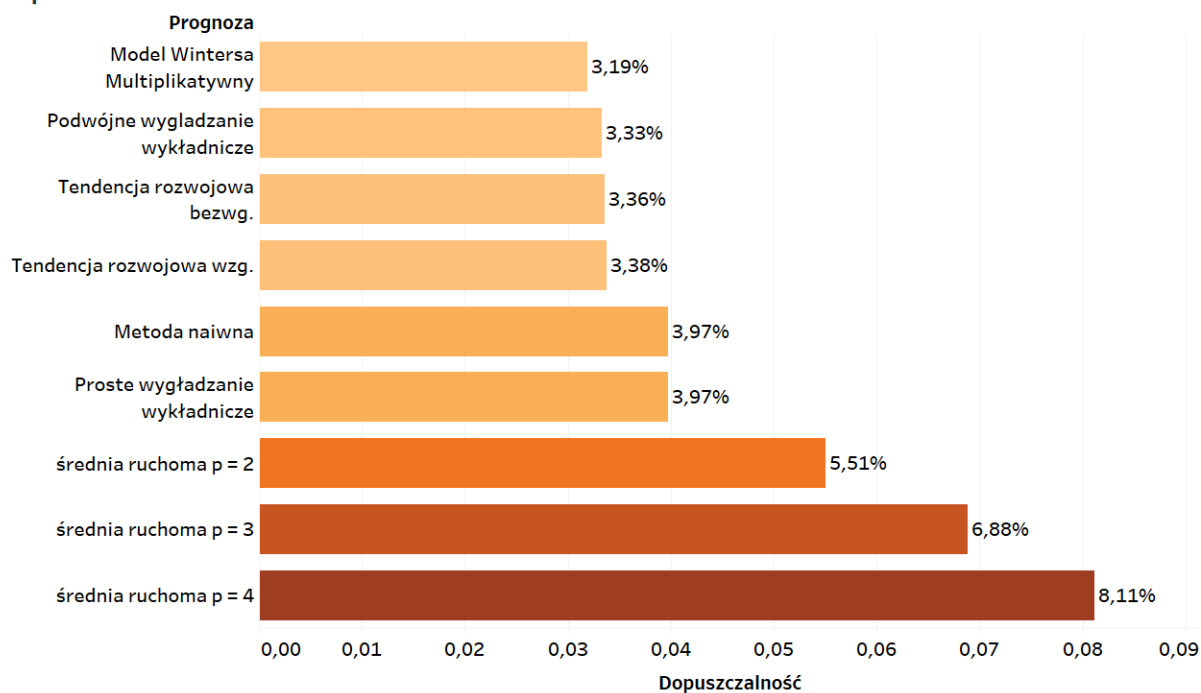
Rok	Miesiąc	wartość rzeczywista	Prognoza	MSE	RMSE	Dopuszczalność
2020	7	6,1	5,3	1,592	1,262	24,0%
2020	8	6,1	5,8	2,156	1,468	25,3%
2020	9	6,1	6	2,306	1,519	25,3%

Natomiast modele oparte na średnich ruchomych i wykorzystujących wygładzanie wykładnicze okazały się o wiele dokładniejsze, w szczególności w przypadku włączenia do modelu tendencji rozwojowej oraz wahań sezonowych.

## RMSE



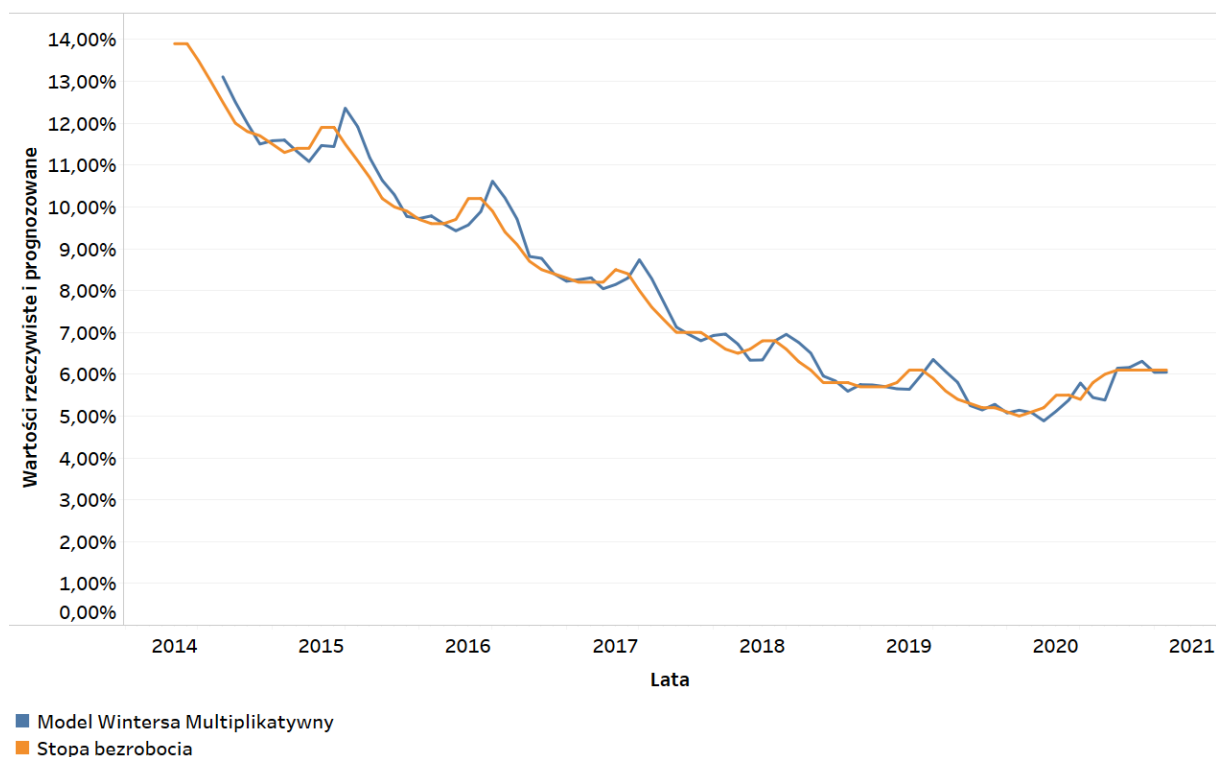
## Dopuszczalność



Ostatecznie za najlepszy do prognozowania został uznany model Wintersa w wersji multiplikatywnej.

Wynika to przede wszystkim z najmniejszych wartości błędów prognozy oraz z tego, że uwzględnia on główne charakterystyki szeregu czasowego tj. trend spadkowy oraz sezonowość.

## Model Wintersa



## Wnioski końcowe:

Bezrobocie jest zjawiskiem złożonym, a otrzymane prognozy nawet jeżeli w pełni dopuszczalny pod względem miar błędów i dopuszczalności należy traktować z rezerwą, w szczególności mając na uwadze, że w danych historycznych można było wielokrotnie zaobserwować zdecydowane zmiany trendów. Prognozy otrzymane z wykorzystaniem wygładzania wykładniczego, a w szczególności multiplikatywnego modelu Wintersa można uznać za odpowiednie narzędzie do przeprowadzania krótkoterminowych predykcji, jednakże w długim okresie konieczna bardziej rozbudowana i niekoniecznie ściśle statystyczna analiza sytuacji makroekonomicznej, zwłaszcza w kontekście sytuacji, które mogły doprowadzić do kolejnej zdecydowanej zmiany trendu.

## Aneks nr.1

Oprogramowanie BERT można pobrać z tej strony: <https://bert-toolkit.com/>. Instalator automatycznie doda BERTa do dodatków programu Excel. Po uruchomieniu dodatku po prawej stronie będzie widoczna konsola, a po lewej plik functions.r. Następnie konieczne jest zainstalowanie biblioteki gap, w której znajduje się implementacja testu Chow'a. Aby tego dokonać należy w konsoli wprowadzić następujące polecenie :

```
install.packages("gap")
```

W ostatnim kroku do pliku functions.r należy dodać następujące linijki:

```
library(gap)
```

```
ChowTest <- function(y1,x1,y2,x2,x=NULL)
```

```
{
```

```
  chow.test(y1,x1,y2,x2,x);
```

```
}
```

Oraz zapisać plik (File -> save). Po zapisaniu funkcja będzie dostępna w Excelu pod nazwą R.chow.test.

W przesłanym pliku Excel w arkuszu dotyczącym punktów zwrotnych znajduje się także „kopia zapasowa”, która umożliwia zapoznanie się z wynikami bez instalowania powyższego oprogramowania.