University of Warsaw
Faculty of Economic Sciences

Forecasting the unemployment rate in Poland using the Box-Jenkins method

Natalia Graczyk Joanna Gawlik

Warsaw, June 2022

Introduction

Unemployment is a socio-economic problem that occurs in every country in the world and affects both the economic state of nations and people's living standards. Forecasting the future unemployment rate is crucial in the context of planning a country's economic policy, as it allows us to identify any changes in this indicator and design measures to stop its possible increase. The aim of this paper is to find the most appropriate forecasting model to analyse the unemployment rate in Poland using the Box-Jenkins methodology. Based on the selected model, the expected level of the unemployment rate in Poland in May-July 2022 will be projected.

Literature review

For forecasting macroeconomic variables, including the unemployment rate, a frequently used methodology is the Box-Jenkins method, using ARIMA models, or autoregressive integrated moving average models.

Mladenovic et al. (2017) analysed the unemployment problem on a value basis over the period from 2000 to 2015 in 28 European Union countries. The aim of the article was to create a statistical model to forecast the monthly unemployment rate in the studied EU countries. The authors used the Box-Jenkins methodology and conducted the study in three phases: 1) model identification, 2) model estimation and 3) model diagnostics and forecasting. During the study, they tested the accuracy of 100 models with different combinations of AR and MA variables and, according to the AIC and SC criterion, the model that best predicted the phenomenon under study was the SARIMA (2,1,1)(12,1,12) model.

Davidescu et al. (2021) performed a comparative analysis of the forecasting performance of the unemployment rate using different time series methods, i.e. the seasonal autoregressive moving average model (SARIMA), the self-excited threshold autoregressive model (SETAR), the Holt-Winters model, the ETS, and the NNAR (neural network autoregressive model). The predictive performance of the aforementioned models was assessed on unemployment rate data from Romania from January 2000 to December 2017 (for the in-sample) and data from 2018-2020 (for the out-of-sample).

As the results showed, for in-sample forecasts, forecast measures such as mean absolute error (MAE), mean absolute percentage error (MAPE) and error mean squared error (RMSE) indicated the superiority of the Holt-Winters model over the other models. However, for the out-of-sample, the RMSE and MAE values indicated an advantage of the NNAR model, while according to MAPE, it was the SARIMA model that showed the the highest forecast accuracy.

Dritsaki (2016) attempted to find the most appropriate model fitted to the unemployment rate in Greece using the Box-Jenkins methodology and to investigate the precision of the estimated model. As the author points out, the use of ARIMA models is a very flexible tool for forecasting the mentioned indicator and the results of the of the study carried out proved that the SARIMA (0,2,1) (1,2,1) model is a suitable model for forecasting the unemployment rate in Greece. The projected value for September 2015 was 24.62 per cent, which was very close to the actual value (24.60 per cent).

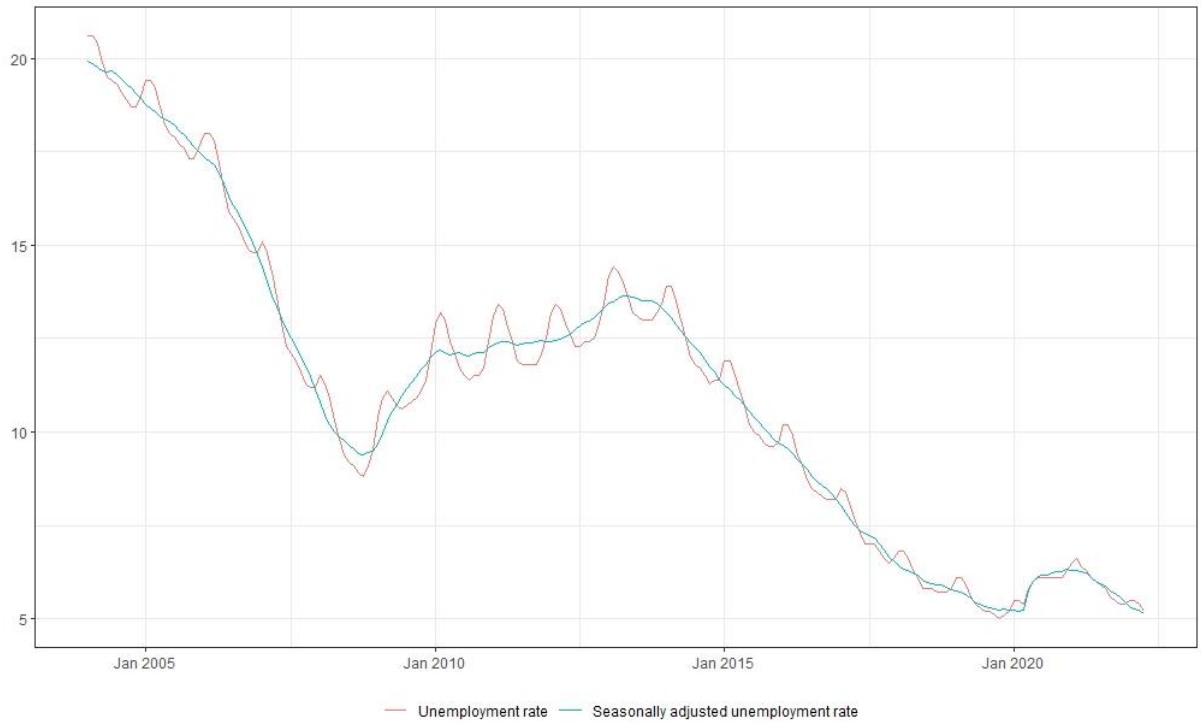
A similar analysis to the forecasting model estimate of the unemployment rate unemployment was also made by, among others, Stoklasová (2012), forecasting the rate for the Czech Republic using the using the SARIMA model (1,1,0) (1,1,0) or Dobre and Alexandru (2008), conducting a study for Romania using the ARIMA model (2,1,2); Syukri and Didiharyono (2020).

Data and research methodology

To estimate the forecasting model, monthly data on the of the unemployment rate in Poland between January 2004 and April 2022 in Poland. The data were downloaded from the CSO website (2022).

In the first stage of the analysis, the series were seasonally adjusted using the procedure of Tramo-Seats. The values of the original variable, as well as the seasonally adjusted variable are shown in Figure 1.

Fig. 1. Unemployment rate and seasonally adjusted unemployment rate for the period January 2004 - April 2022 in Poland.



Source: Own elaboration based on CSO data (2022).

Then, in order to determine the stationarity of the series, a Dickey-Fuller test was performed. The results showed that the unemployment rate is stationary in the first difference, thus the degree of integration of the model is 1 (d=1).

The next step was to determine the p and q parameters in the ARIMA model (p, d, q). This was done based on an analysis of correlograms, and also based on a comparison of the Akaike information criterion (AIC) of 20 different ARIMA models. The value of the AIC criterion led to the conclusion that the model with the best fit to the data was the ARIMA model (1, 1, 2).

Results

Using the ARIMA model (1, 1, 2), a forecast of the unemployment rate in Poland was made for the period May-July 2022 (3-month forecast horizon). The forecast unemployment rate was 5.04% in May, 4.97% in June and 4.90% in July (Figures 2 and 3).

Fig. 2. Forecast of the unemployment rate in Poland for the period May-July 2022.

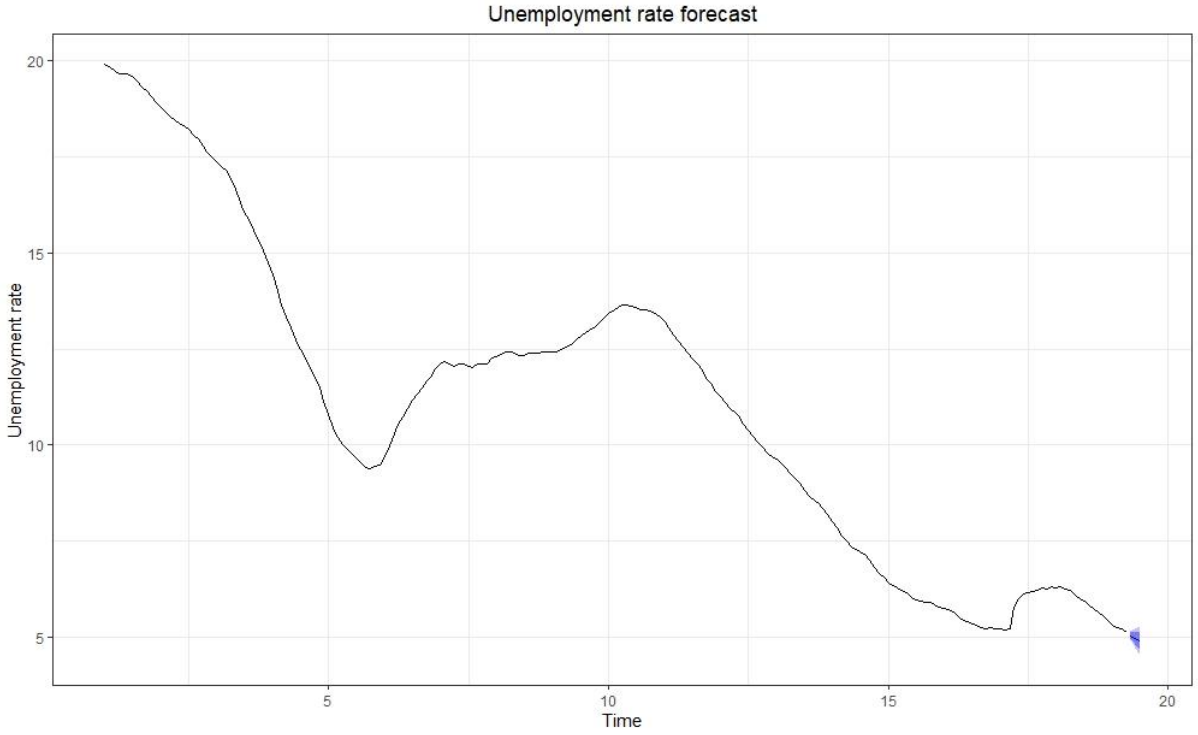
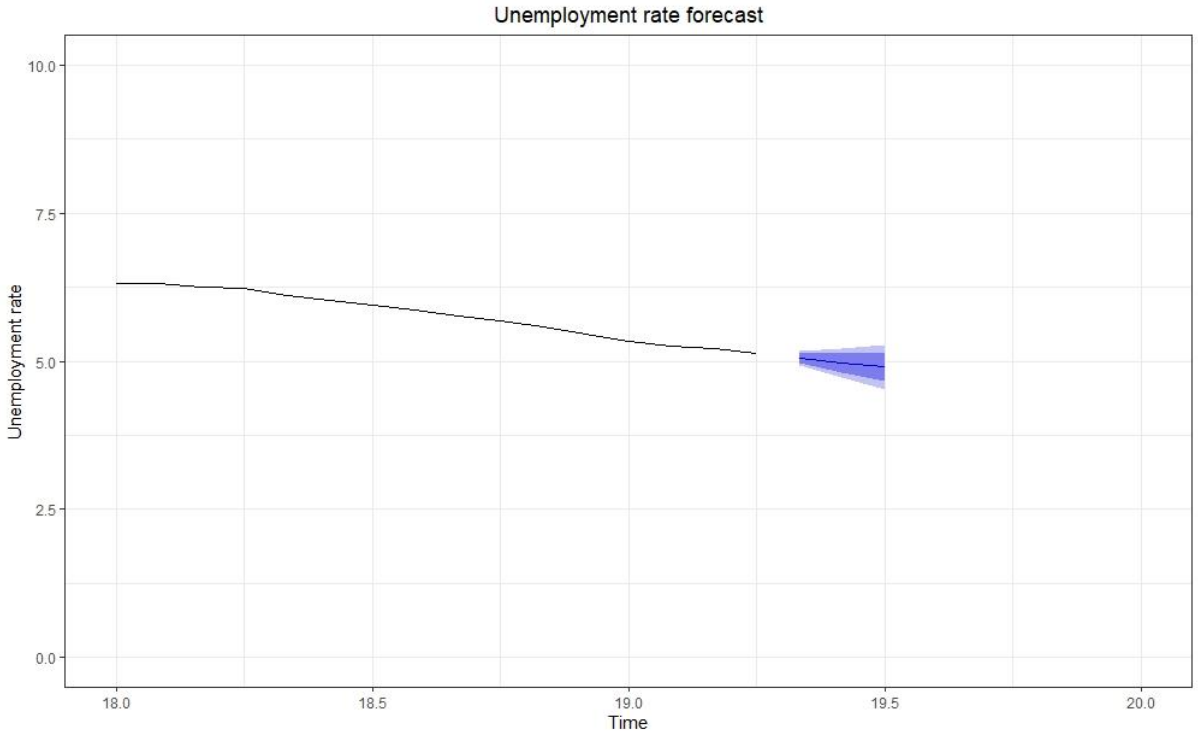


Fig. 3. Forecast unemployment rate in Poland for May-July 2022 zoomed.



Summary

Unemployment is a phenomenon found in every country today and the ability to forecast its value is important from the perspective of socio-economic policy planning. In this paper, using the Box-Jenkins method, a forecasting model was selected to study the phenomenon being described. The analysis carried out allowed us to conclude that the model best describing the unemployment rate in Poland is the ARIMA model (1, 1, 2). Using this model, it was possible to forecast future unemployment rates for May-July 2022, amounting to 5.04 %, 4.97 % and 4.90 % respectively.

Bibliography

Davidescu, AM. (2021), Comparative Analysis of Different Univariate Forecasting Methods in Modelling and Predicting the Romanian Unemployment Rate for the Period 2021–2022, Entropy 2021, 23, 325

Dobre, I., Alexandru, A. A. (2008), Modelling unemployment rate using Box–Jenkins procedure, Journal of Applied Quantitative Methods, 3(2), 156–166

Dritsaki, Ch. (2016), Forecast of Sarima Models: An Application to Unemployment Rates of Greece, American Journal of Applied Mathematics and Statistics, Vol. 4, No. 5, 136-148

GUS (2022), Stopa bezrobocia rejestrowanego w latach 1990-2022, https://stat.gov.pl/obszary tematyczne/rynek-pracy/bezrobocie-rejestrowane/stopa-bezrobocia-rejestrowanego-w latach-1990-2022,4,1.html

Mladenovic, J., Ilic I., Kostic, Z. (2017). Modeling the unemployment rate at the EU level by using Box–Jenkins methodolkneogy. KnE Social Sciences, 1(2), 1–13

Stoklasová, R. (2012), Model of the unemployment rate in the Czech Republic, Proceedings of 30th International Conference Mathematical Methods in Economics, 836-841