UNEMPLOYMENT RATE & ECONOMIC PERSPECTIVE

Ana Beatriz Pires (M20210396) | Mariana Novo (R20181106) | Paula Eveling (M20210331)

Time Series Analysis 2021/2022

In order to make accurate predictions it is imperative that we have a good perception of the country's economic growth. The unemployment rate is the most used indicator for understanding conditions in the labour market. It is measured by the number of people who are unemployed as a percentage of the labour force (the total number of people employed added to those unemployed). It is important to note that this is a lagging indicator, meaning that it generally rises or falls in the wake of changing economic conditions, rather than anticipating them. When the economy is in poor shape and jobs are scarce, the unemployment rate can be expected to rise. When the economy is growing at a healthy rate and jobs are relatively plentiful, it can be expected to fall (Anderson, Sommer. Clarine, Skylar. 2021). With that said, this poster focuses on the interpretation and forecast of the time series of unemployment rate in Portugal between 1999 and 2021 and its possible trends, seasonality, long-run cycles, variance, outliers, and abrupt changes that could impact the analysis, in order to accurately forecast the next four years of data using the ARIMA (Autoregressive Integrated Moving Average) model. Since the unemployment rate is a lagging indicator of the economic perspective, this project aims to understand how the economic perspective affects the time series under study using the periodogram and cross-correlation functions. The economic perspective consists of a set of analytical and econometric forecasts for carrying out estimates and evaluations of current and future economic indicators. The analysis was done in R, and both time series studied in this project are available on the INE (Instituto Nacional de Estatística) website.

Literature Review

The unemployment rate is a useful measure of the underutilization of the labour supply. It reflects the inability of an economy to generate employment for those who want to work but are not doing so, even though they are available for employment and actively seeking opportunities (ILO,2021).

Concerning Portugal, it is now known that it was one of the most affected countries by the 2007 Europe economic crisis, and therefore, the unemployment rate rapidly suffered its consequences. The unemployment rate increased from 7.6% in 2008 to 16.2% in 2013, being one of the highest in Europe (INE, PORDATA, 2021). Until the beginning of the COVID-19 outbreak, Portugal was slowly recovering from the crisis. With a progressive decrease in the unemployment rate. In the first trimester of 2021 the unemployment rate was 7.1%. This high rate can be explained by the pandemic outbreak restrictions. However, it has been decreasing evolving to 6.7% in the second trimester and to 6.1% in the third trimester. The latter is already lower than the one verified before the pandemic (6.3% in third trimester of 2019) (Instituto Nacional de Estatística, 2021).

Portugal has been addressing labour market policies that can be divided into passive and active interventions. Passive LMPs' purpose is to provide income support to unemployed individuals, including unemployment benefits, redundancy compensation, bankruptcy compensation and also early retirement. While active LMPs are interventions used by the government to 'actively' increase the jobseekers' employability (individual's potential propensity to find/be placed in a job) and decrease unemployment (ILO, 2018).

Data from the INE (Instituto Nacional de Estatística) shows that, compared to the previous quarter, Portugal grew by 4.9%, approaching the pre-pandemic level at the end of the year. It was the moment when Portugal reached the highest level of GDP since the beginning of the pandemic. This confirmed a strong economic recovery in the country and that the recovery of economic activity reflects the control of the pandemic, mainly through the vaccination process, which showed positive effects on the confidence of economic agents. Through this recovery process, it can be considered that the initial shock of the pandemic was relieved temporary, despite the longer impact on some sectors and companies.

In light of the inflation, in 2021, this economic indicator stands at 0.9%, up from the 0.7% projected in June. According to the October Economic Bulletin, Banco de Portugal pointed out that the unemployment rate should be 6.8% in 2021, when in June it was expected to be close to 7.2%. The Public Finance Council's projections on invariant policies for the Portuguese economy point to a recovery in real economic growth to 3.3% in 2021 and to 4.9% in 2022, compared to a contraction of 7.6% in 2020 This dynamic is mainly due to the evolution of private consumption (2.7% in 2021 and 6.0% in 2022) and exports (8.9% in 2021 and 11.3% in 2022).

In terms of economic perspective, the pandemic triggered a deep recession in Portugal and the crisis is likely to leave scars, with increased poverty and inequality (OECD, 2021). According to the Focus Economic (2021), Portugal's economy should reach its pre-pandemic level in 2022. EU funds will likely boost capital spending. However, political instability and the lack of a budget—necessary for the arrival of EU funds—threaten investment in the first half of the year. The instability caused by the pandemic certainly adds a certain level of unpredictability to any future economic scenario of the country.

Time Series Analysis

Unemployment Rate

The time series decomposition plot (figure 1) shows a clear uptrend between 1999 and 2013, where the unemployment rate reaches its peak and then starts decreasing. In August 2020, the rate increases slightly, likely as a consequence of the pandemic.

The dataset is not seasonally adjusted as usual official government statistics on unemployment, therefore, there is seasonality. On average, the unemployment rate is lower between May and August.

In terms of variance, it can be observed on the decomposition plot that it is not constant over time. Considering that there is trend, seasonality and non-constant variance, it can be said that this is a non-stationary time series.

The seasonality and stationarity were treated using the differencing method, while the variance was treated using the Box-Cox transformation where lambda is equal to 0.449.

un_rate_treated<-ts(diff(diff(ds^lambda.est, lag=12, differences = 1)), frequency=12)

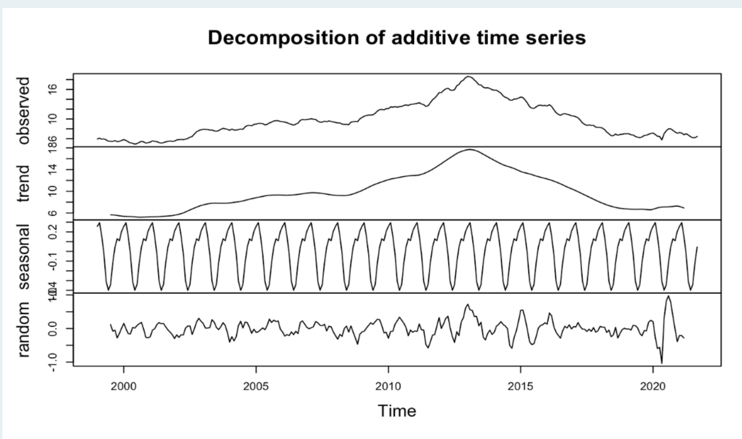

Figure 1: Decomposition of the Unemployment Rate Time Series

Economic Perspective

The Economic Perspective time series decomposition plot (figure 2), shows a non-constant variation with a steep decrease in 2020, likely due to COVID-19.

The autocorrelation and partial autocorrelation functions show that the lags approach to zero very slowly, indicating that the series is not stationary.

In terms of seasonality, a test was applied which evidenced that the time series is seasonally adjusted, meaning that there is no seasonality.

This time series was treated with one order differencing.

treatedperspectiva<-ts(diff(perspective),frequency=12)

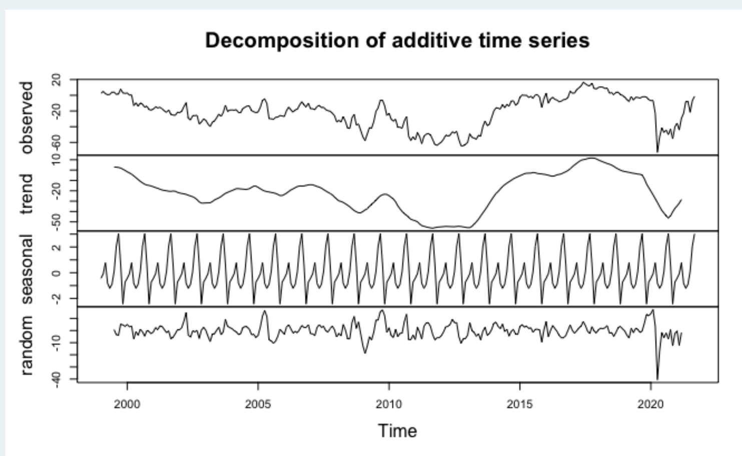

Figure 2: Decomposition of the Economic Perspective Time Series

Forecast

This project focuses on the Unemployment Rate forecast. Once the time series was properly treated, the best possible ARIMA models were identified using the observed ACF and PACF charts and the auto.arima function was run. The results of the models were compared and the one that presented the best results, considering both fit and residuals, was the ARIMA(1,1,0)(2,1,0)[12]. Its coefficients can be found on figure 3.

```
Series: taxa
ARIMA(1,1,0)(2,1,0)[12]
Box Cox transformation: lambda= 0.449

Coefficients:
         ar1     sar1     sar2
      0.2766  -0.6764  -0.3967
s.e.  0.0599   0.0618   0.0684

sigma^2 estimated as 0.005045:  log likelihood=316.51
AIC=-625.03   AICc=-624.87   BIC=-610.79
```

Figure 3: Unemployment Rate Forecast Results

Once the model was chosen, the ARIMA was implemented and the results were analyzed. On figure 4, the observed and the forecasted values are represented in black and red lines respectively. It can be observed that the model was able to capture the variability and seasonality of the data.

fit2<- Arima(ds, order=c(1,1,0), seas=list(order=c(2,1,0), period=12), include.drift=TRUE, lambda=lambda.est)

Considering the volatility of the observed results caused by the pandemic and to confirm if the model is indeed well-adjusted, the year of 2019 was used to test the model, where the data between 1999 and 2018 was used to predict the upcoming year (2019). In this project, the results presented were satisfactory: the model was able to predict the unemployment rate in 2019 within the 95% confidence intervals.

With this result, the next four years (48 months) were estimated. The forecast of the upcoming years (2021-2025) are represented in blue on figure 4, where the grey shadow indicates the forecast ranges (upper and lower tails). As we forecast further out into the future, it is natural to become less confident in the values. This is reflected by the confidence intervals generated by the model, which grow larger as we move further out into the future.

To conclude, the model is well-fitted and presented good results within the 95% confidence interval.

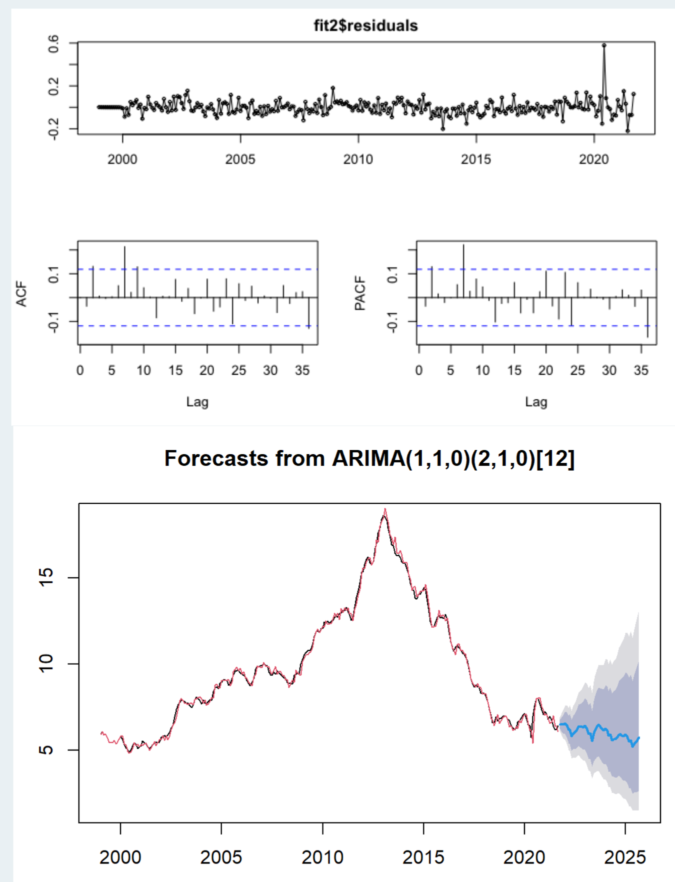

Figure 4: Unemployment Rate Forecast

Periodogram

Unemployment Rate in Portugal between January 1999 and September 2021

A periodogram is an estimate of the spectral density of a signal, and it can be used to identify dominant frequencies in a time series.

The unemployment rate is a monthly dataset where n is equal to 260. This time series was properly treated prior to this analysis, as detailed on section Time Series Analysis. Its periodogram plot on the right (figure 5), shows that the highest peak occurs on the lag thirty six, which corresponds to a frequency of 0.1333. By dividing one by this frequency, the most dominant period can be identified. In this case, the most dominant period occurs every 7.5 months.

order(new$spec) # approximate the Fourier frequencies
(1/new$freq[36])

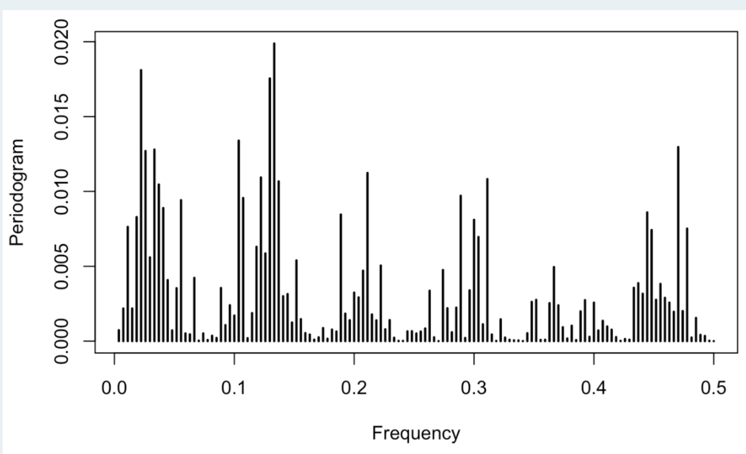

Figure 5: Unemployment Rate Periodogram

Economic Perspective in Portugal between January 1999 and September 2021

Since both datasets, the economic perspective and unemployment rate, have different lengths due to the transformation applied, a periodogram was plot for n=272 (transformed data) and n=260 (12 observations removed to match the unemployment dataset). The expected result was that both analysis presented approximately the same results.

Where n=272, the corresponding frequency is equal to 0.2326, and the dominant period occurs every 4.30 months. Where n= 260, the frequency is equal to 0.2481 and the dominant period 4.0298. As expected, the results are similar and the dominant period is around 4 months.

new = periodogram(x)
(1/new$freq[67])

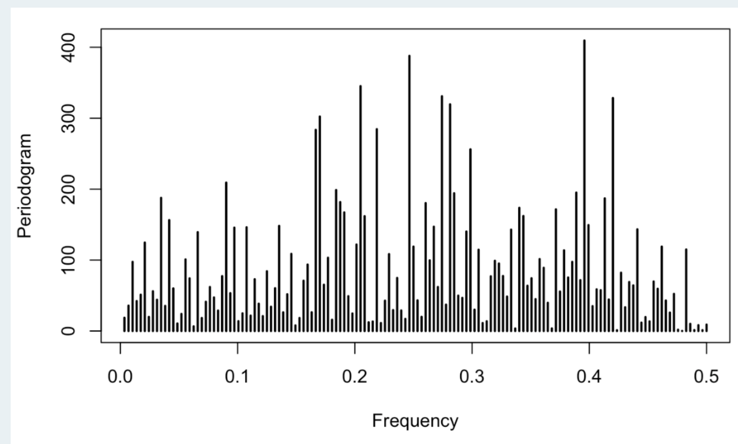

Figure 6: Economic Perspective Pediodogram

Cross-Correlation Function

As seen on the previous section, Periodogram, the unemployment rate time series is indeed a lagging indicator of the economic perspective, meaning that it rises or falls after 4.30 months of changing economic conditions, rather than anticipating them. The cross-correlation function is helpful for identifying lags of the x-variable (economic perspective) that might be useful predictors of y (unemployment rate).

In order to apply the cross-correlation function, both time series must be stationary, with no seasonality and stable variance. The treatment of each time series was described on the Time Series section. With both series treated, the CCF (Cross Correlation Function) can be analyzed, as per figure 7.

It can be observed that the lag -3 is the most relevant for this analysis because it is further out of the confidence intervals compared to the other lags. This means that the most dominant correlation occurs in time t-3. So, it can be concluded that xt + h, with negative h, are predictors of yt. Hence, it can be said that the variable economic perspective leads the variable unemployment rate. To obtain the regression with the lag -3, the auto.arima function was applied. This is a function that automatically deals with the series' residuals. It is possible to observe that the final model obtained will be:

$$\widehat{Unemployment}\ rate_t = 0.7017 Unemployment\ rate_{t-1} - 0.0298 Unemployment\ rate_{t-12} + 0.0209 Unemployment\ rate_{t-13} - 0.4497\omega_{t-1} - 0.8743\omega_{t-12} + 0.3931\omega_{t-13} - 0.0004 * Economic\ Perspetctive_{t-3} + \omega_t$$

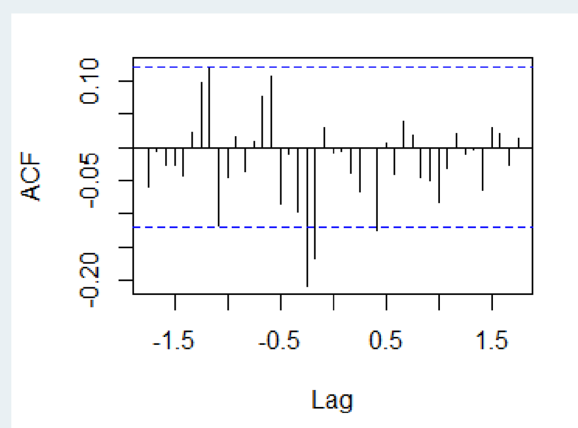

Figure 7: Cross Correlation Function

Conclusion

The unemployment rate is the proportion of unemployed people in the labor force. When workers are unemployed, their families lose wages, and the nation as a whole loses their contribution to the economy in terms of the goods or services that could have been produced. In addition, unemployed people also lose their purchasing power, which can lead to unemployment for other workers, creating a cascading effect that ripples through the economy. In other words, the unemployment even impacts those who are still employed. These are some of the many reasons why the unemployment rate is constantly in the entre of many political discussions. Considering this scenario, accurate forecasting is primordial to the leaders and investors to determine in which sectors to invest.

In this study, the Auto Regressive Moving Average (ARIMA), a class of statistical models for analyzing and forecasting time series data was applied. The exceptionally adverse effects of the recession, in conjunction with relatively sizable revisions to the real-time data, have made this analysis more challenging. Nevertheless, the ARIMA model applied presented good results: the model is statistically well-adjusted, with a relatively uniform distribution where errors have a zero mean and constant variance. The model suggest that the unemployment rate will decrease to 5-6% between 2021 and 2025. However, it is important to highlight that the current pandemic scenario adds in a level of uncertainty to the predictions.

One important factor when analysing the unemployment rate is to consider that this is a lagging indicator: it measures the effect of economic events, such as a recession. The unemployment rate does not rise until after a recession has already started. It also means the unemployment rate will continue to rise even after the economy has started to recover. That said, this study has investigated the link between unemployment rate and the economic perspective of the country using the periodogram and the cross-correlation function (CCF). Based on the CCF output, it can be concluded that, as expected, the variable economic perspective does lead the unemployment rate occurring on the third lag, which means that, approximately three months after an increase or drop in the economic perspective, the unemployment rate is affected.

In conclusion, the methods applied in this study, ARIMA, Periodogram, and Cross-Correlation, can help to understand the unemployment rate in Portugal and estimate how this scenario will look in the upcoming years, so the competent organs can define an economic plan for the country.

LIMITATIONS

- The recession in 2007 and the impact of the pandemic in 2020 caused a change in the behaviour of the data, which impacts the adjustment of the model, and therefore, future predictions.
- Parameters p, d, q need to be manually defined, therefore, finding the most accurate fit can be a long trial-and-error process.

REFERENCES

"Indicator Description: Unemployment Rate - ILOSTAT". 2021. ILOSTAT. https://ilostat.ilo.org/resources/concepts-and-definitions/description-unemployment-rate/.; "Banco De Portugal Mantém Perspetiva De Crescimento Da Economia Em 4,8% - Renascença". 2021. Rádio Renascença. https://rr.sapo.pt/noticia/economia/2021/10/06/banco-de-portugal-mantem-perspetiva-de-crescimento-da-economia-em-48/255810/.; MathWorks 2021. https://www.mathworks.com/help/signal/ref/periodogram.html.; "Portugal economic".2021. Portugal Economy - GDP, Inflation, CPI and Interest Rate (focus-economics.com) ; "Portugal Economic Snapshot". 2021. Portugal Economic Snapshot - OECD