Time Series Analysis of Italian and EU GDP

- **Datasets:** the data consists of the Quarterly GDP of Italy and Quarterly GDP of Europe.
- **Time range:** from the first quarter of 1996 to the fourth quarter of 2016.
- **Source of the data:** OECD Data site https://data.oecd.org/gdp/quarterly-gdp.htm

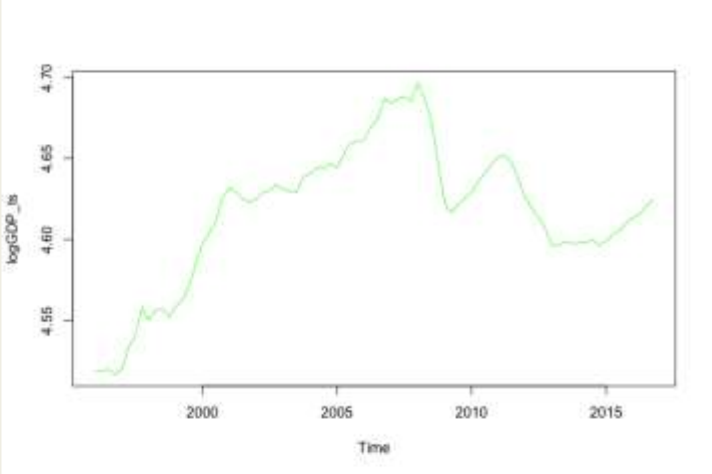
- Univariate Time Series Analysis of Italian GDP
- Creation of the best model for forecasting
- Multivariate Time Series Analysis of Italian and EU GDP
- Analysis of the relationship between the two series
- ADLM(1), VAR(1), VECM(1) models, cointegration tests and forecast

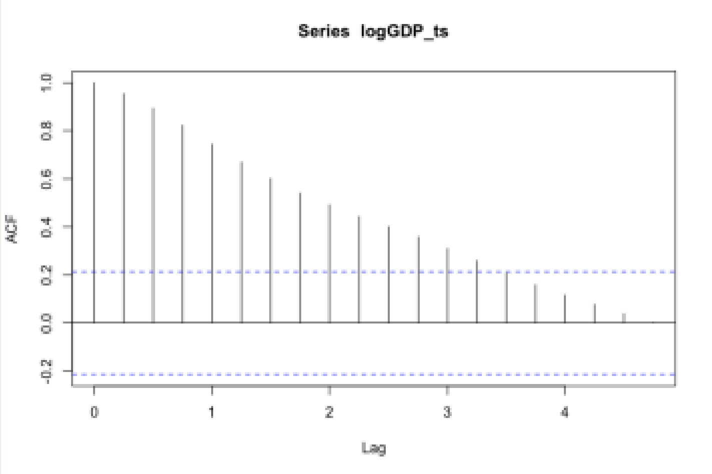
KU LEUVEN

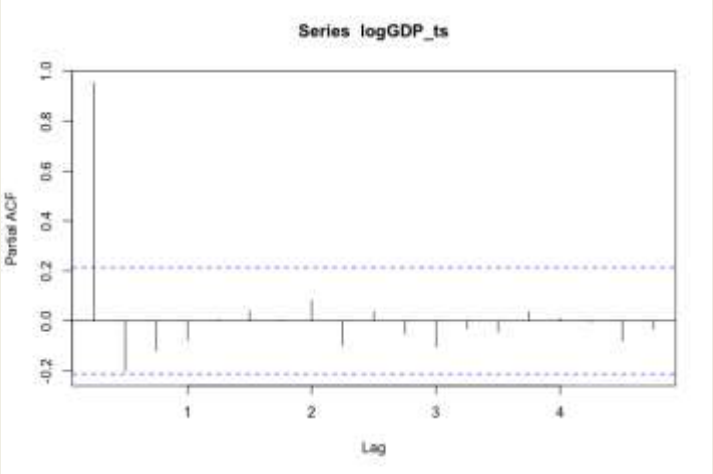
Advanced Time Series Analysis - D0M63B
KU Leuven
December 2023

Student: Alessandro Salvatori (r0926525)
Professor: Christophe Croux

UNIVARIATE TIME SERIES ANALYSIS: Italian GDP (logGDP_ts)





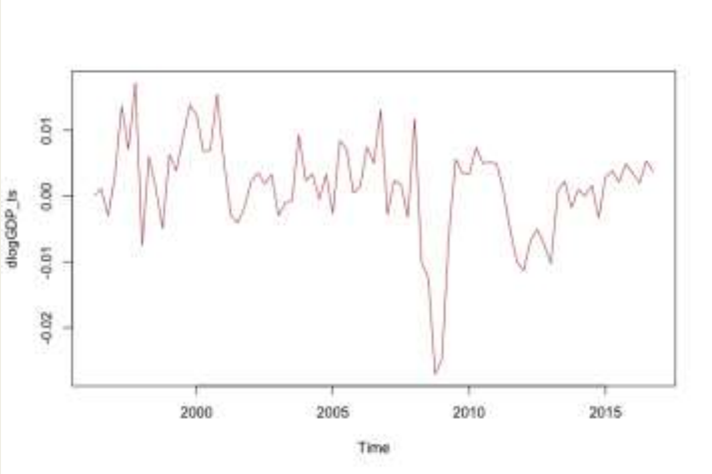


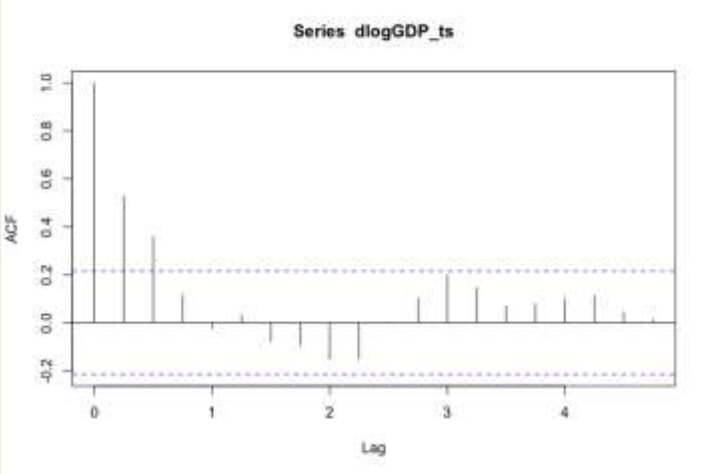
- The autocorrelations show strong persistency.
- Two test where applied to check for stationarity (**ADF test**) and if the series is white-noise (**Ljung-Box test**).

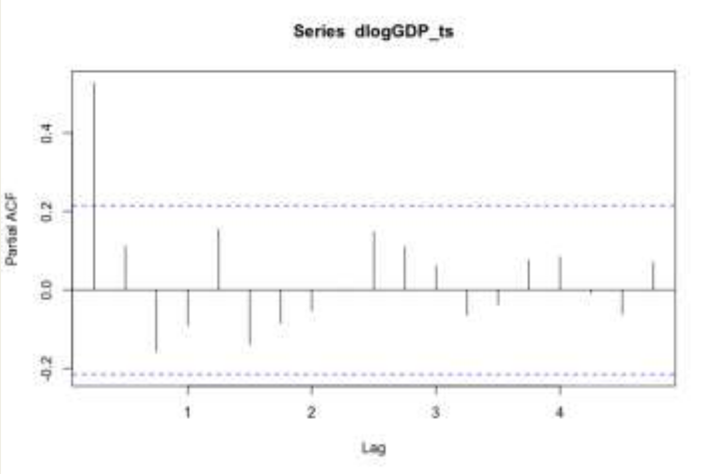
Augmented Dickey Fuller (ADF)	Ljung-Box test
p-value = 0.136	p-value < 2.2e-16

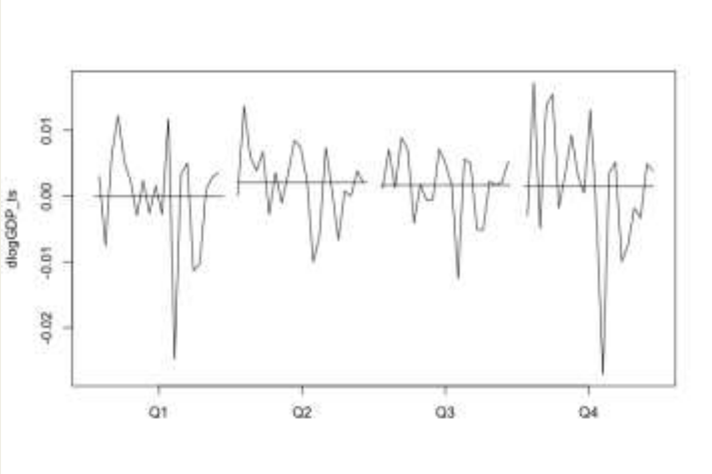
- The **ADF test** produces a p-value = **0.136**, thus it can be concluded that the time-series contains unit roots and hence it is necessary to go into differences.
- The **Ljung-Box test** produces a p-value < 2.2e-16, indicating that the time series is not a white noise.

UNIVARIATE TIME SERIES ANALYSIS: Italian GDP (dlogGDP_ts)









Augmented Dickey Fuller (ADF)	Ljung-Box test
p-value = 0.001333	p-value = 2.988e-06

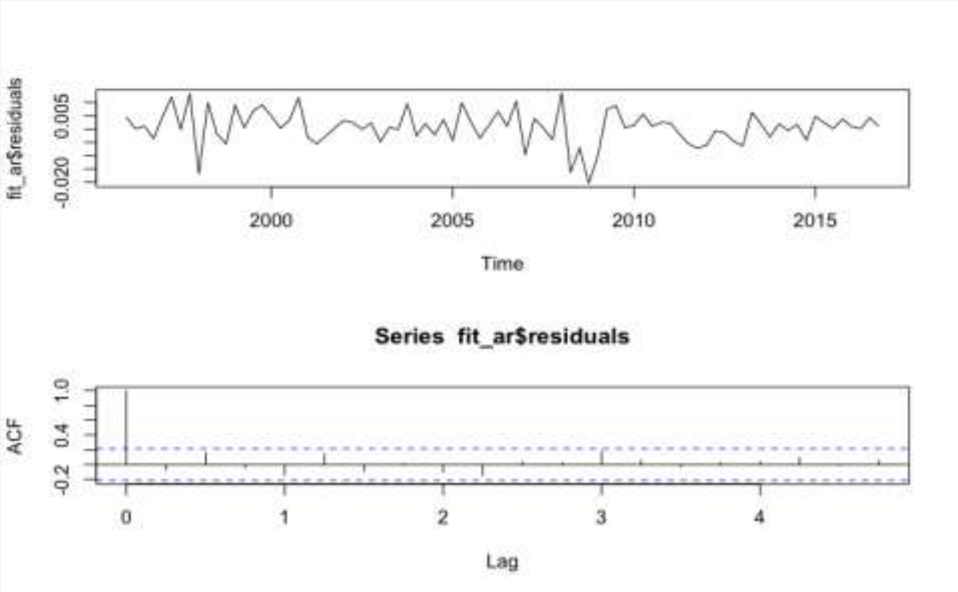
- Two significant correlations can be seen from the ACF, which suggests an MA(2).
- One significant correlations can be seen in the PACF, so it might suggest an AR(1) model.

- The ADF test produces a p-value < 5%, thus it can be concluded that the time-series is now stationary.
- The Ljung-Box test produces a p-value < 0.05, indicating that the time series is not a white noise.

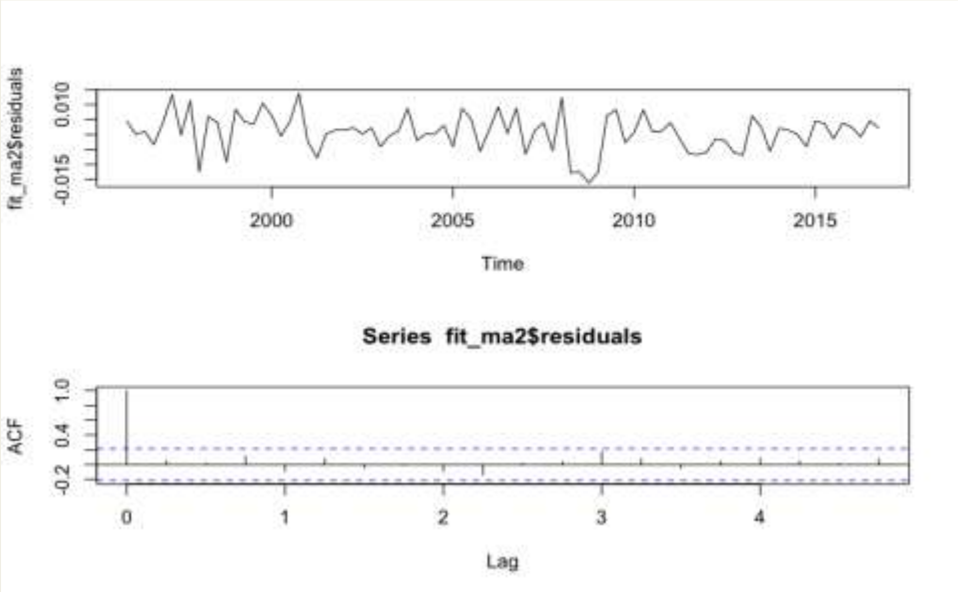
- The plot by quarter of the log differences (bottom right plot) of the time series shows that there is **no seasonality**.

UNIVARIATE TIME SERIES ANALYSIS: Fitting models

AR(1)



MA(2)



Model	Highest Order Term Significance	Box Test	Residual Correlations	AIC	SIC
ARIMA(1,1,0)	Significant	0.306	White-Noise	-601.1765	-596.3149
ARIMA(0,1,2)	Significant	0.788	White-Noise	-602.1725	-594.8801

- Both models have highest order terms significant, are validated by the box test and ACF of residuals.
- The AIC measure indicates that the **MA(2)** model is slightly better, but the SIC measure indicates that AR(1) model is slightly lower, hence better, so we prefer that. Also the AR(1) model is more parsimonious.
- Since results are still really close we still try to forecast using both models and then compare results.

UNIVARIATE TIME SERIES ANALYSIS: Forecasting

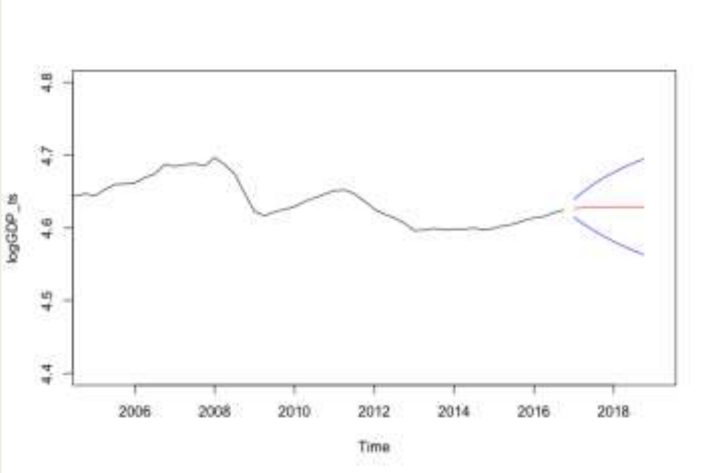
- The forecast of both models are very similar.
- We compare the performance of the models using the **MAE** (Mean Absolute Error) and **MSE** (Mean Squared Error) and perform a Diebold-Mariano test to see if the forecast performance of the two models is not significantly different.

Model	MAE	MSE
AR(1)	0.003013793	1.410299e-05
MA(2)	0.003608187	1.729048e-05

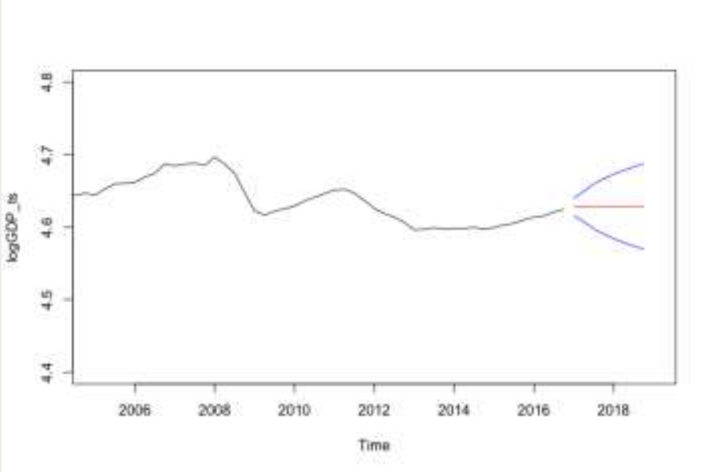
- The AR(1) model is preferred because has lower **SIC**, **MAE** and **MSE** and is more parsimonious.
- According to Diebold-Mariano Test at horizon=1, we do not reject H0 and conclude that the forecast performance of the two models, using the squared value loss, is not significantly different.

- Prediction intervals are used to provide a range where the forecast is likely to be with a specific degree of confidence (in this case it is 95%).

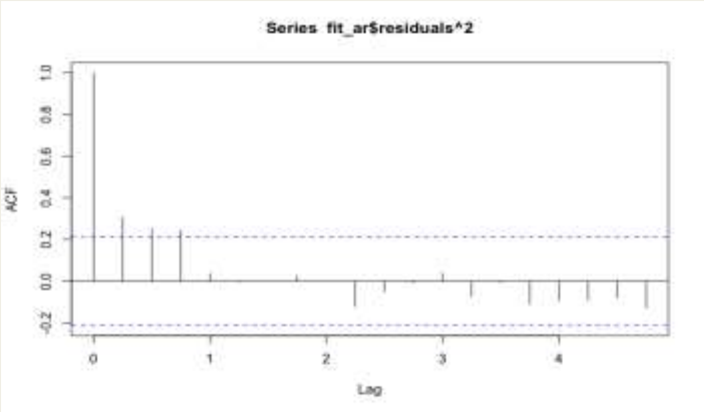
Forecast AR(1)

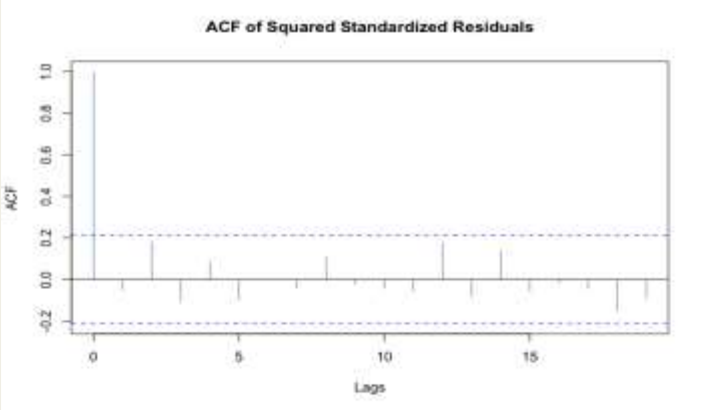


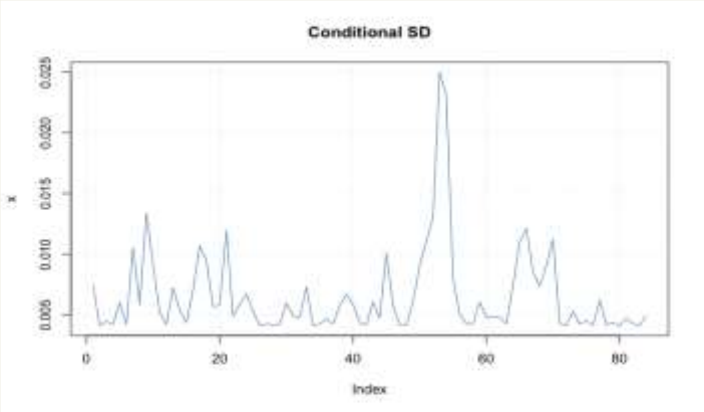
Forecast MA(2)



UNIVARIATE TIME SERIES ANALYSIS: GARCH Model







- After choosing the AR(1) model we checked for heteroscedasticity by looking at the squared standardized residuals and we can see a few significant autocorrelations in the squared residuals (top plot) in the selected model which suggests presence of heteroscedasticity so we decide to fit a Garch model.

ARMA(0,1) GARCH(1,1)

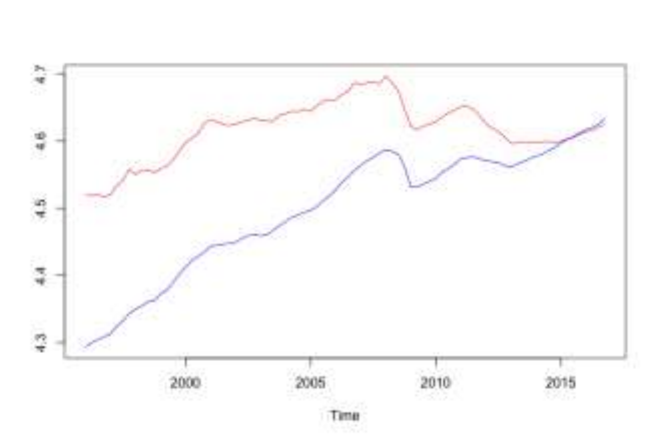
- We fit **ARMA(0,1) GARCH(1,1)** model to the series in differences and found significant values of omega and alpha1.

- The **Q-tests** on the standardized residuals and on the squared standardized residuals have p-values > 5% for Q(10), Q(15) and Q(20) (only for standardized residuals Q(10) we have a 0.046), thus, we conclude that there is no structure left in the (squared) standardized residuals, hence, the model is valid.

- We observe the **conditional standard deviation** (bottom plot) is not constant over time and that there are clusters of high volatility.

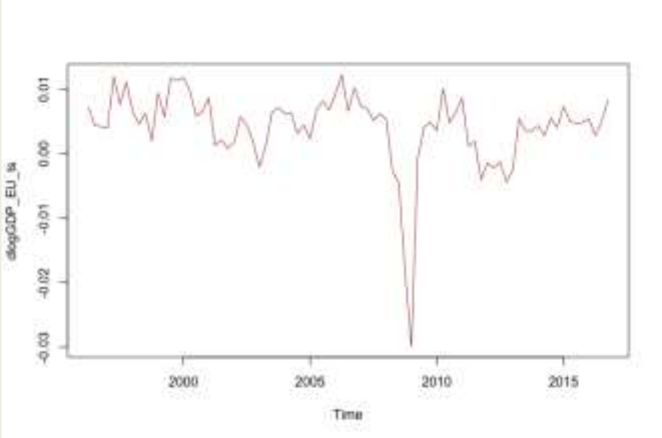
MULTIVARIATE TIME SERIES ANALYSIS: EU GDP, ADLM(1) and testing for Granger Causality

logGDP_ts and logGDP_EU_ts



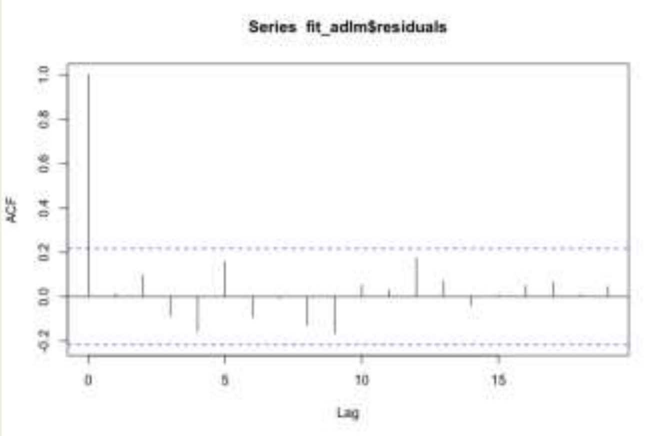
- ADF test on logGDP_EU_ts: (p-value = 0.495) the TS is not stationary.
- Thus, dlogGDP_EU_ts is created using the diff() function.

dlogGDP_EU_ts



- ADF test on dlogGDP_EU_ts : (p-value = 0.003) the TS is now stationary.
- Thus, logGDP_ts and logGDP_EU_ts are both integrated of order 1.

ADLM(1)



An **Autoregressive Dynamic model of order 1** is estimated and then we test for **Granger causality** comparing the ADLM(1) with the model without lagged explanatory variables.

- The model is validated using the Box test (p-value: 0.2936) and the plot of the autocorrelations suggest that there are no significant autocorrelation, and hence, residuals are white-noise.

- The p-value = 0.0326 < 5%, thus we reject H0 of no Granger Causality. We conclude that dlogGDP_EU has incremental explanatory power in predicting dlogGDP.

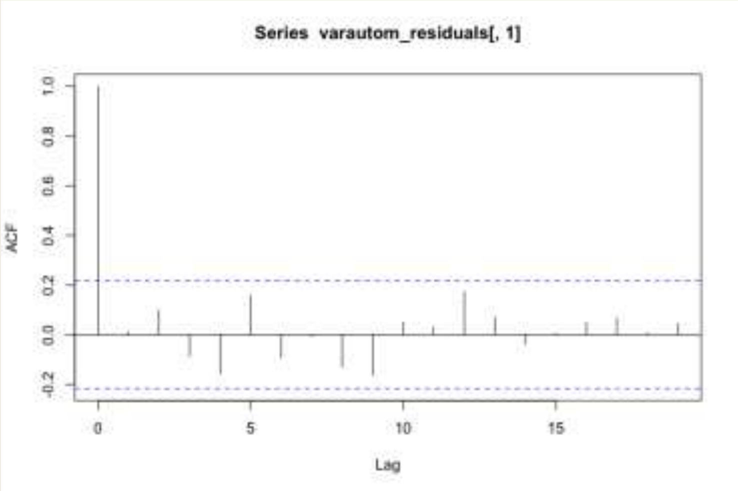
MULTIVARIATE TIME SERIES ANALYSIS: VAR Model

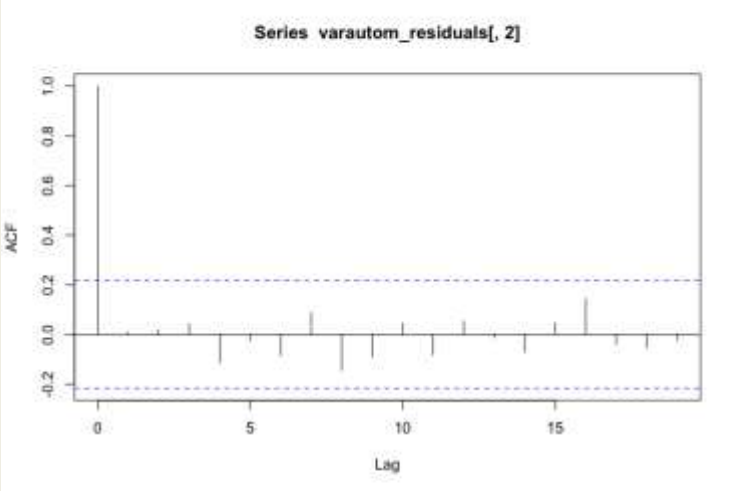
- The **selectvar()** function was used with dlogGDP_ts and dlogGDP_UE_ts.
- The order of the VAR model selected by **Schwarz's information criterion** is 1, so a VAR(1) was estimated.

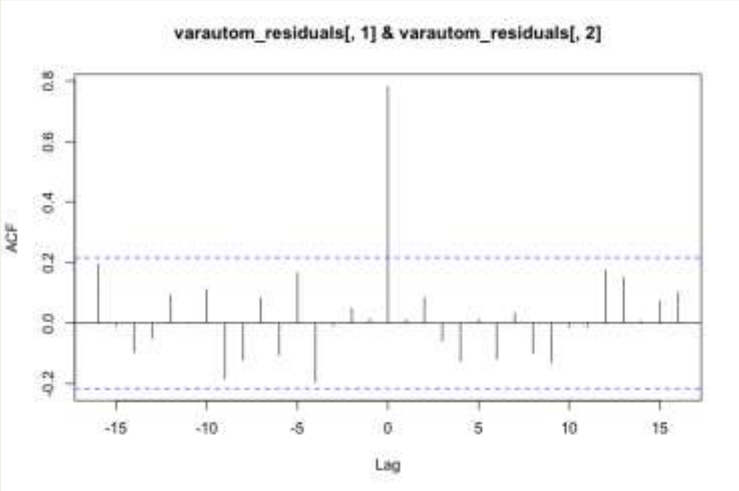
```
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     1      1      1      1
```

VAR(1)

- The **R2** = 0.4453, thus 44.53% of the variance of dlogGDP is explained by the lagged observations of dlogGDP and of dlogGDP_EU at lag 1.
- The **F-statistics** has p-value (7.789e-11) < 5%, thus we reject H0 and conclude that the regressors are jointly significant.
- In the **correlograms** we observe no significant correlations, while a strong significant cross-correlation without lag is observed in the **cross-correlogram**. There is no problem with contemporaneous correlations in a VAR model, as it does not allow for explicit modeling of contemporaneous interdependence. Thus, the residuals look multivariate white noise and the model is valid.



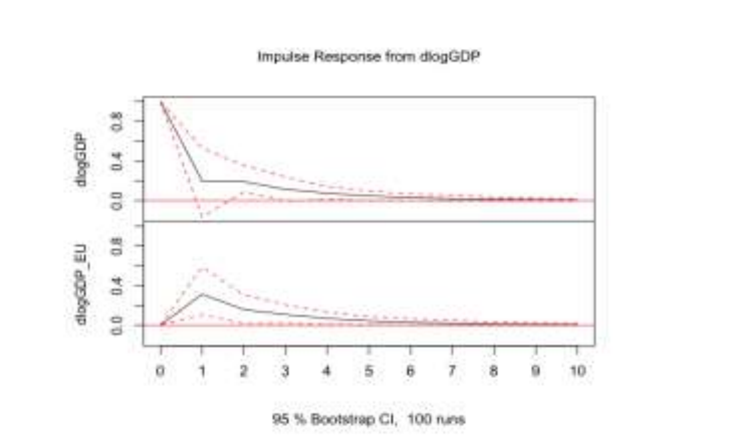


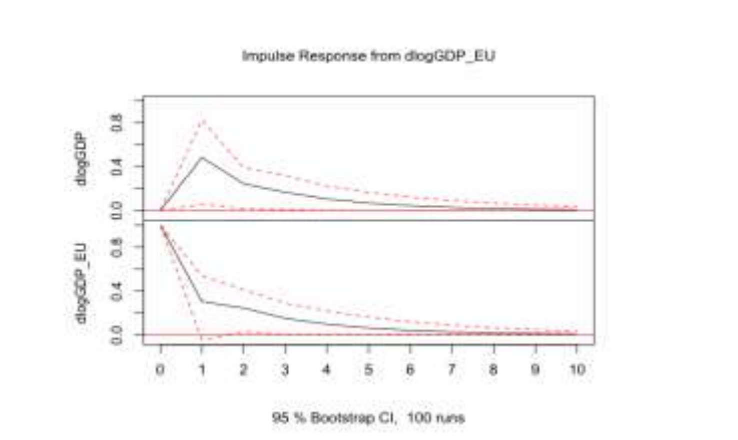


MULTIVARIATE TIME SERIES ANALYSIS: Impulse Response Functions and Cointegration Tests

IMPULSE RESPONSE FUNCTIONS

- The IRFs provide an easy way to interpret the estimated coefficients of the VAR model.
- Given a **unitary impulse** in dlogGDP at time t, we observe a significant positive response of dlogGDP_EU at t + 1.
- Given a **unitary impulse** in dlogGDP_EU at time t, we observe a significant positive response of dlogGDP at time t + 1.





- Checking for cointegration first with the Engle-Granger Test then with Johansen Test.

ENGLE-GRANGER TEST

- The Engle-Granger test indicates no cointegration since the test statistic of the ADF test (= - 1.12) is larger than the critical value (= - 3.41) for one explanatory variable. Thus, H0 of no cointegration is not rejected and we conclude that there is **no cointegration**.

JOHANSEN TEST

- According to Johansen's **trace test**, the test statistics is larger than then the critical value (28.05 > 19.96), thus there is at least one cointegrating relation. Hence, logGDP and logGDP_EU are cointegrated.
- The **cointegrating equation** is:

$$0.2189629 + \log(GDPt) - 1.2631723 \log(GDP\_EUt) = \delta t$$

- After repeating the cointegration test using the Johansen's **maximum eigenvalue** statistics we conclude that logGDP and logGDP_EU are cointegrated.

MULTIVARIATE TIME SERIES ANALYSIS: VECM and Forecast

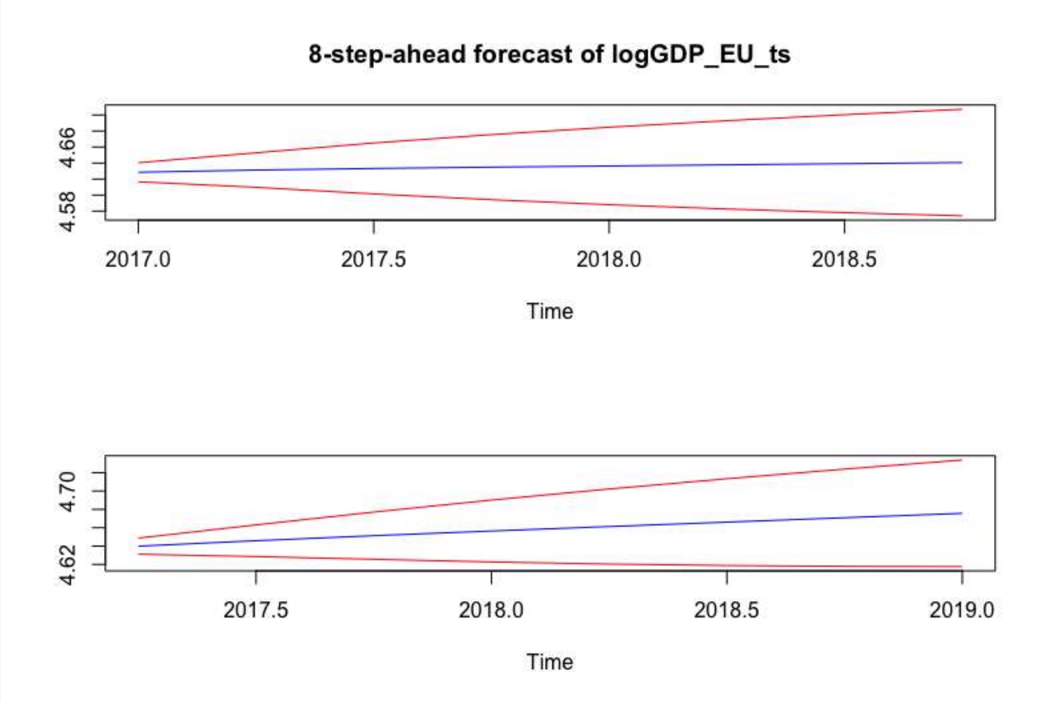
VECM(1)

- The cointegrating equation, which corresponds to the one while testing with the Johansen procedure, is a stationary linear combination of logGDP and logGDP_EU.

$$0.2189629 + \log(GDPt) - 1.2631723 \log(GDP\_EUt) = \delta t$$

FORECAST

- The figures in the right plots the 8 step-ahead forecast of logGDP and logGDP_EU based on the VECM(1).



Conclusion: In this assignment two time-series were used: **Italian and EU GDPs**. From the ADLM(1) it is seen that **Δlog(GDP_EU)** has an incremental explanatory power over **Δlog(GDP)**, which means that the EU GDP is helpful in forecasting the Italian GDP. Most notably, our cointegration analysis using the Johansen Test and then building a VECM(1) revealed a long-term equilibrium relationship between the Italian and EU GDPs. The VECM also enabled us to forecast future trends with a high degree of confidence.