Norwegian School of Economics
Bergen, spring semester 2020

ECN430: Assignment 1

Univariate time-series analysis

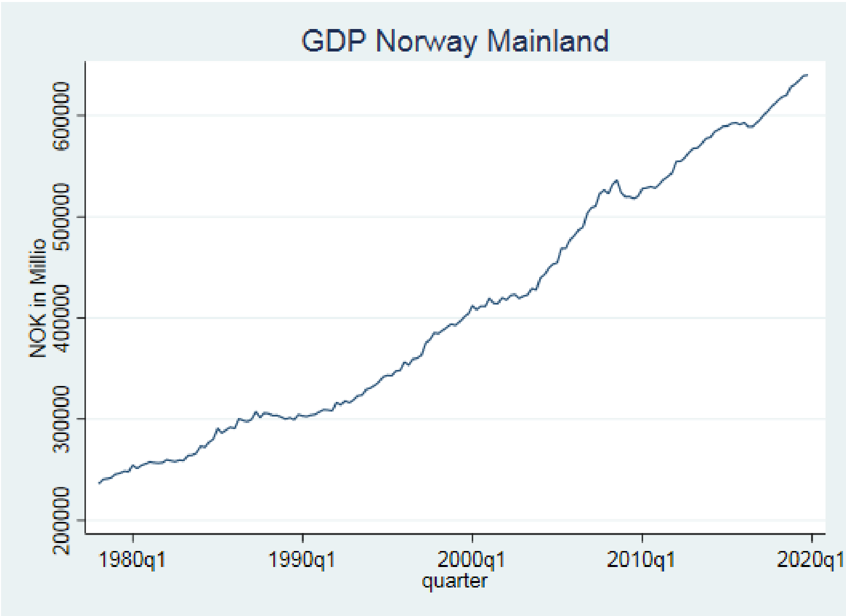
Veronika Gottwald (s195169),
Benjamin C. Herbert (s202881)

NORWEGIAN SCHOOL OF ECONOMICS

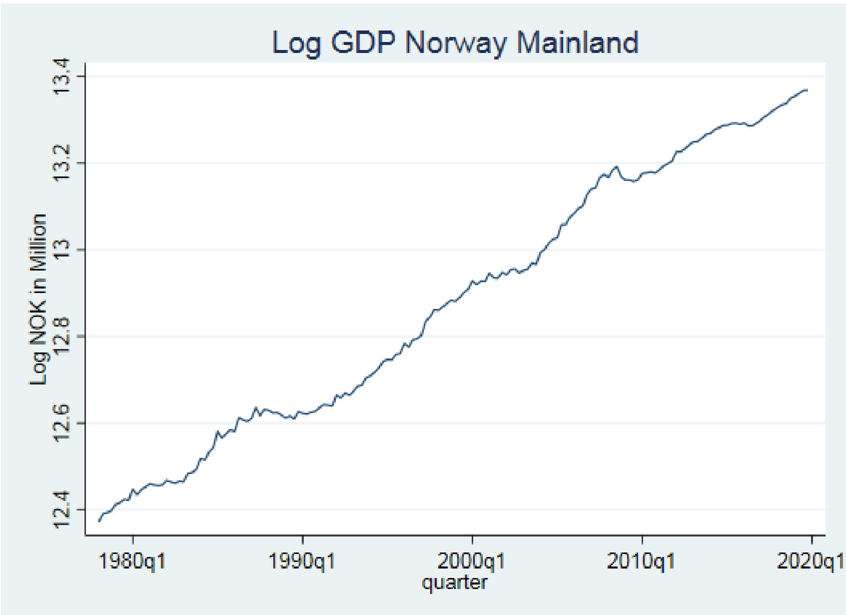
Content

1. Box-Jenkins

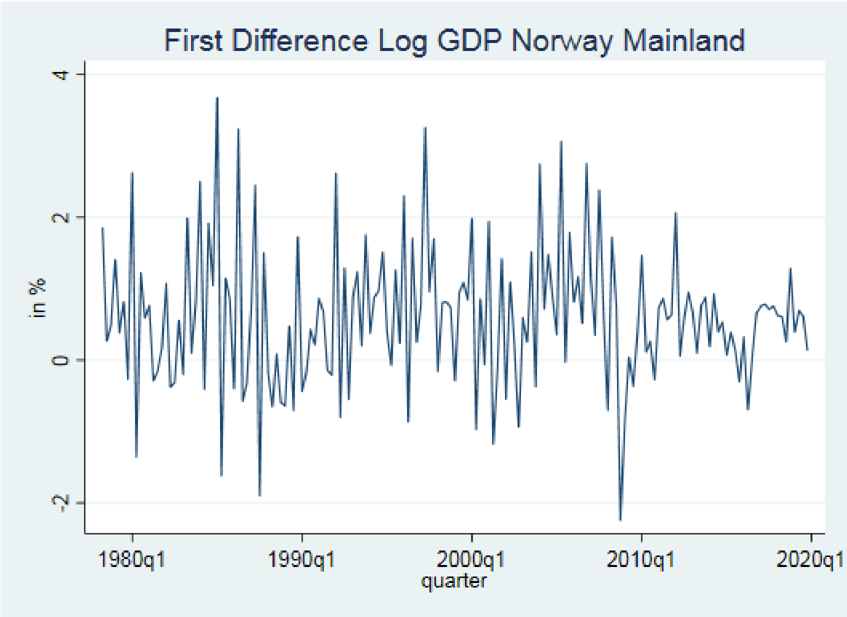
1.1 Task A



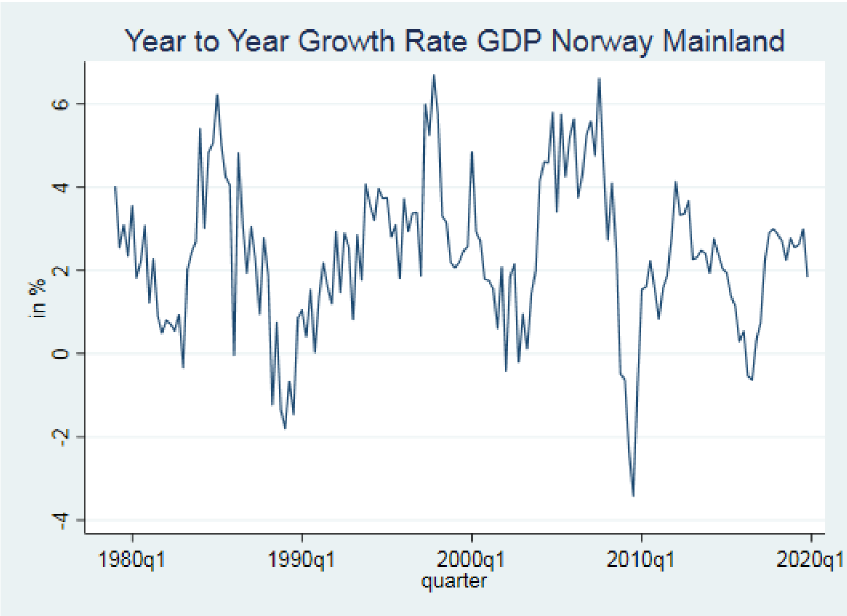
The GDP series has a visible upward trend. The growth is slightly convex. We assume that individual quarters do not have any significant influence here as we are already using seasonally adjusted data to generate the series. Therefore, there are no quarterly effects. The mean of the series is not constant over time from which it can be inferred that the series is not stationary.



The log-transformed GDP series also has a visible upward trend, but the growth is nearly linear. As for the GDP series, we assume no quarterly effects. The mean is not constant either and the series therefore must be non-stationary.



The first difference of the log-transformed GDP series does not show any visible trend nor growth. Instead, it seems to fluctuate around a constant mean with a similar variance and could be stationary. As above, we assume no quarterly effects.



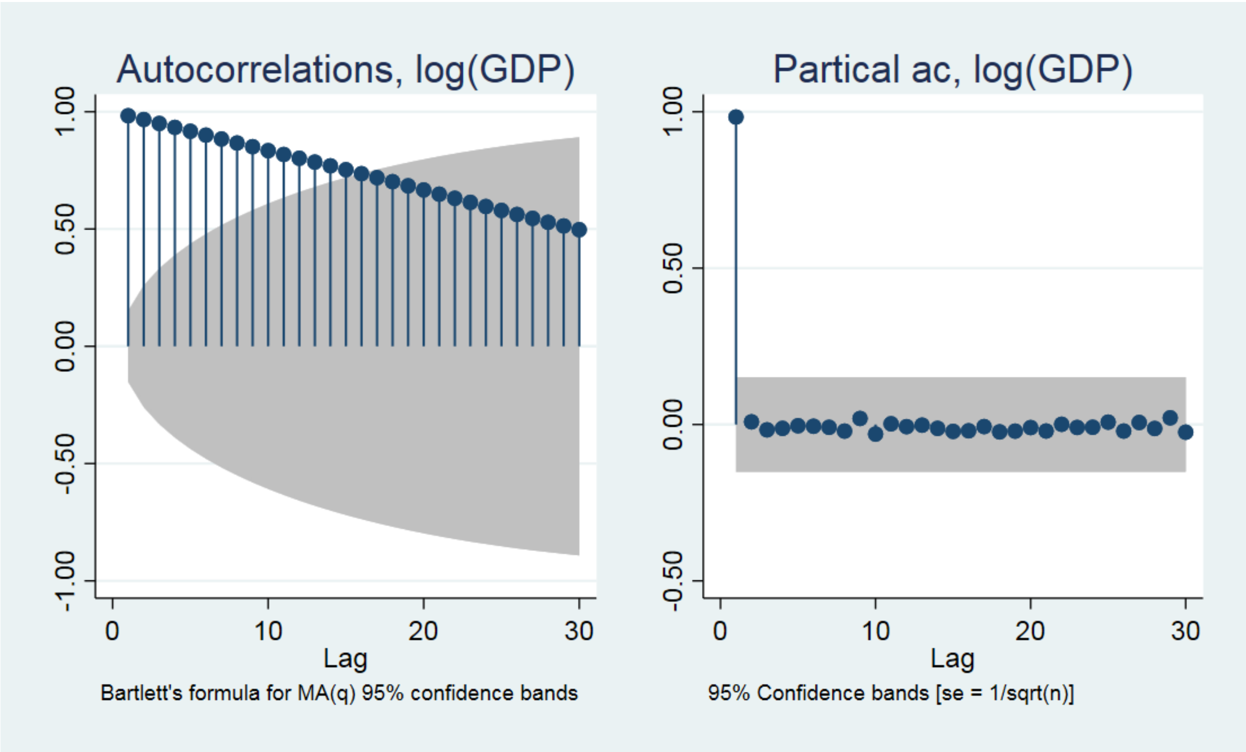
The year-on-year growth also does not seem to have a clear trend or growth. It is possibly fluctuating around a constant mean, but it needs to be investigated further. The series tends to revert to the mean after shocks which leads to the conclusion that it might be trend stationary. As above, we assume no quarterly effects.

1.2 Task B

In order to determine whether the series follow an MA(.) or an AR(.) process, we used the following criteria:

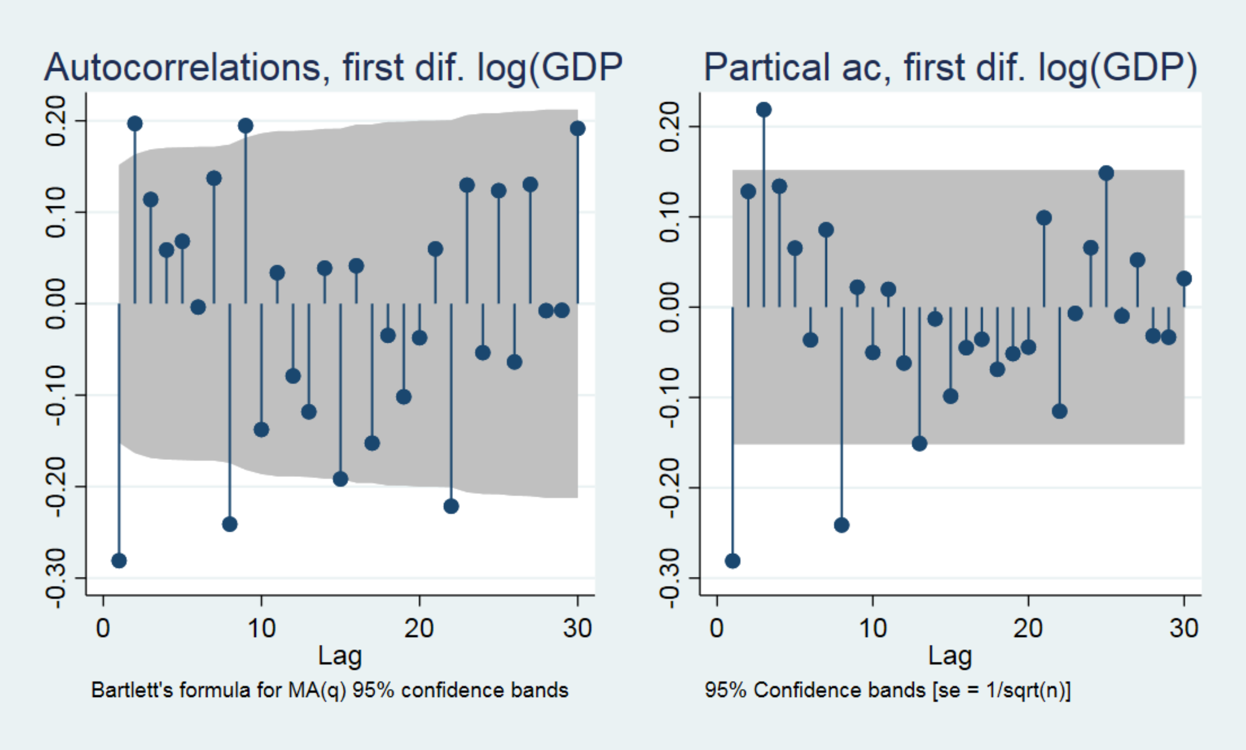
Table 3.1. Behavior of the ACF and PACF for ARMA Models

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off



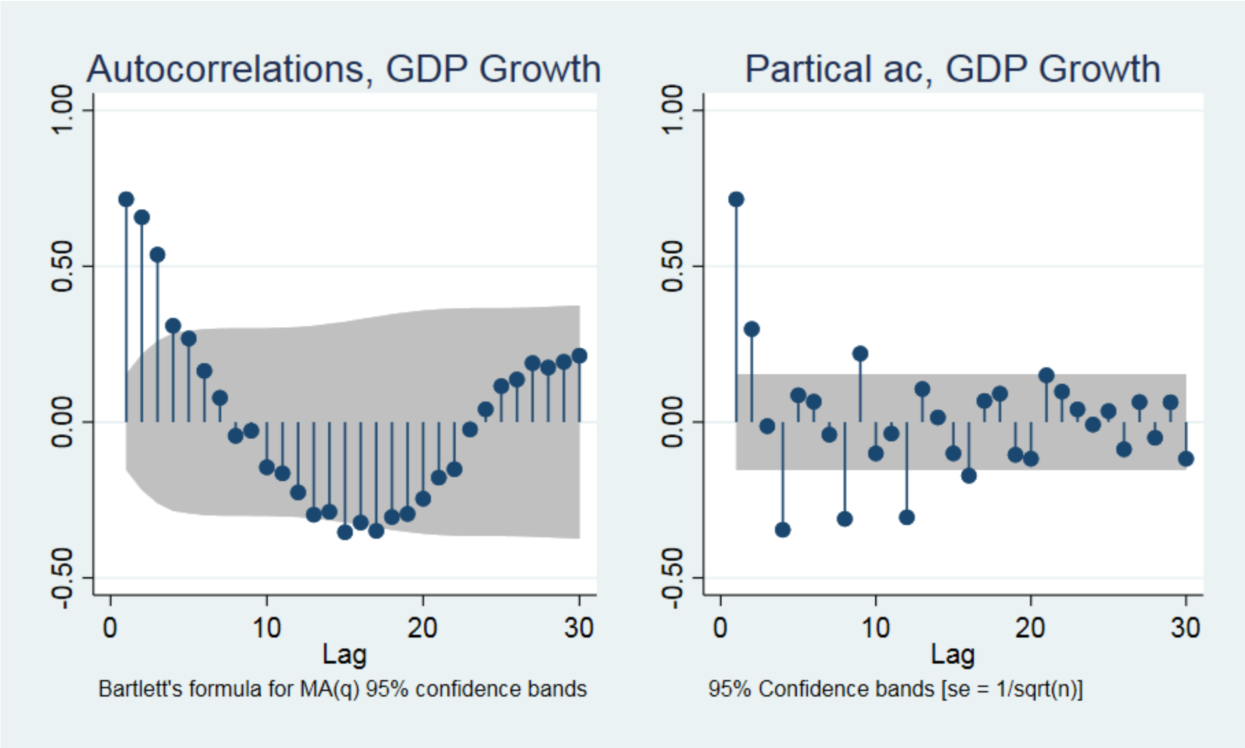
Log-transformed GDP:

Here it is clearly visible that the autocorrelation function (ACF) slowly tails off, while the partial autocorrelation function (PACF) has a clear cut-off after the first (significant) lag. Therefore, the log-transformed GDP series follows an AR(.) process. Because only the first lag in the PACF is significant, it likely is an AR(1) process.



First difference of the log-transformed GDP series:

Both the ACF and the PACF tail off and there is no clear cut-off. The first difference of the log-transformed GDP series therefore likely follows an ARMA process.



<u>Year-on-year growth:</u>

For the year-on-year growth, the ACF tails off relatively quickly. The PACF decays more slowly, but also tails off. Therefore, the only viable option for this series is an ARMA process.

1.3 Task C

In order to determine the necessity of seasonally adjusting the data, we regressed the three time series from above on quarterly dummies. We started out by regressing on a trend variable. For the series did not show a significance of the trend, we also regressed without the trend component.

In an effort to validate our results from the OLS regression, we decided to apply the Newey-West estimator to our series which is specifically constructed for panel data. The results match in every case and thus confirm the results of the regression.

The constant is significant for every single test.

Regression with seasonally unadjusted data on quarter dummies:

Log-transformed GDP:

OLS and Newey-West Regression on logGDP without seasonal adjustment

	(1) reg logGDP	(2) newey logGDP
trend	0.00624***	0.00624***
	(0.000)	(0.000)
dum1	-0.0240**	-0.0240***
	(0.001)	(0.000)
dum2	-0.0466***	-0.0466***
	(0.000)	(0.000)
dum3	-0.0590***	-0.0590***
	(0.000)	(0.000)
_cons	12.38***	12.38***
	(0.000)	(0.000)
N	168	168

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- With trend
  - Reg: trend and all three dummies are significant
  - NW: trend and all three dummies are significant

First difference of the log-transformed GDP:

OLS and Newey-West Regression on First Difference LogGDP without seasonal adjustment

	(1) reg fdlogGDP	(2) newey fdlo~P	(3) reg fdlogGDP	(4) newey fdlo~P
trend	-0.000580		-0.000580	
	(0.881)		(0.759)	
dum1	-8.370***	-8.369***	-8.370***	-8.369***
	(0.000)	(0.000)	(0.000)	(0.000)
dum2	-8.169***	-8.167***	-8.169***	-8.167***
	(0.000)	(0.000)	(0.000)	(0.000)
dum3	-7.148***	-7.147***	-7.148***	-7.147***
	(0.000)	(0.000)	(0.000)	(0.000)
_cons	6.578***	6.528***	6.578***	6.528***
	(0.000)	(0.000)	(0.000)	(0.000)
N	167	167	167	167

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- With trend
  - Reg: only dummies significant (trend not significant)
  - NW: only dummies significant (trend not significant)
- Without trend
  - Reg: all three dummies are significant
  - NW: all three dummies are significant

Year-on-year growth:

OLS and Newey-West Regression on Year-to-Year Growth without seasonal adjustment

	(1) reg growth	(2) newey growth	(3) reg growth	(4) newey growth
trend	-0.000439		-0.000439	
	(0.917)		(0.927)	
dum1	0.0579	0.0592	0.0579	0.0592
	(0.918)	(0.916)	(0.899)	(0.897)
dum2	-0.0762	-0.0753	-0.0762	-0.0753
	(0.892)	(0.893)	(0.846)	(0.847)
dum3	0.0411	0.0415	0.0411	0.0415
	(0.942)	(0.941)	(0.861)	(0.859)
_cons	2.415***	2.376***	2.415***	2.376***
	(0.000)	(0.000)	(0.000)	(0.000)
N	164	164	164	164

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- With trend
  - Reg: nothing significant (trend, dum 1-3)
  - NW: nothing significant (trend, dum 1-3)
- Without trend
  - Reg: nothing significant (trend, dum 1-3)
  - NW: nothing significant (trend, dum 1-3)

Regression with seasonally adjusted data on quarter dummies:

Log-transformed GDP:

OLS and Newey-West Regression on logGDP with seasonal adjustment

	(1) reg logGDP	(2) newey logGDP
trend	0.00621***	0.00621***
	(0.000)	(0.000)
dum1	0.00284	0.00284
	(0.670)	(0.368)
dum2	0.00179	0.00179
	(0.789)	(0.565)
dum3	0.000835	0.000835
	(0.900)	(0.784)
_cons	12.35***	12.35***
	(0.000)	(0.000)
N	168	168

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- with trend
  - Reg: only the trend is significant
  - NW: only the trend is significant

<u>First difference of the log-transformed GDP:</u>

OLS and Newey-West Regression on First Difference LogGDP with seasonal adjustment

	(1) reg fdlogGDP	(2) newey fdlo~P	(3) reg fdlogGDP	(4) newey fdlo~P
trend	-0.000329		-0.000329	
	(0.837)		(0.807)	
dum1	0.278	0.278	0.278	0.278
	(0.206)	(0.204)	(0.209)	(0.208)
dum2	-0.0226	-0.0219	-0.0226	-0.0219
	(0.917)	(0.920)	(0.914)	(0.917)
dum3	-0.0121	-0.0118	-0.0121	-0.0118
	(0.956)	(0.957)	(0.947)	(0.949)
_cons	0.566**	0.538***	0.566***	0.538***
	(0.007)	(0.001)	(0.000)	(0.000)
N	167	167	167	167

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- With trend
  - Reg: nothing significant (trend, dum 1-3)
  - NW: nothing significant (trend, dum 1-3)
- Without trend
  - Reg: nothing significant (trend, dum 1-3)
  - NW: nothing significant (trend, dum 1-3)

Year-on-year growth:

OLS and Newey-West Regression on Year-to-Year Growth with seasonal adjustment

	(1) reg growth	(2) newey growth	(3) reg growth	(4) newey growth
trend	-0.000399		-0.000399	
	(0.896)		(0.936)	
dum1	0.0273	0.0285	0.0273	0.0285
	(0.947)	(0.944)	(0.911)	(0.908)
dum2	-0.000686	0.000112	-0.000686	0.000112
	(0.999)	(1.000)	(0.997)	(1.000)
dum3	0.00823	0.00863	0.00823	0.00863
	(0.984)	(0.983)	(0.969)	(0.968)
_cons	2.404***	2.369***	2.404***	2.369***
	(0.000)	(0.000)	(0.000)	(0.000)
N	164	164	164	164

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- With trend
  - Reg: nothing significant (trend, dum 1-3)
  - NW: nothing significant (trend, dum 1-3)
- Without trend
  - Reg: nothing significant (trend, dum 1-3)
  - NW: nothing significant (trend, dum 1-3)

Summary:

Log-transformed GDP:

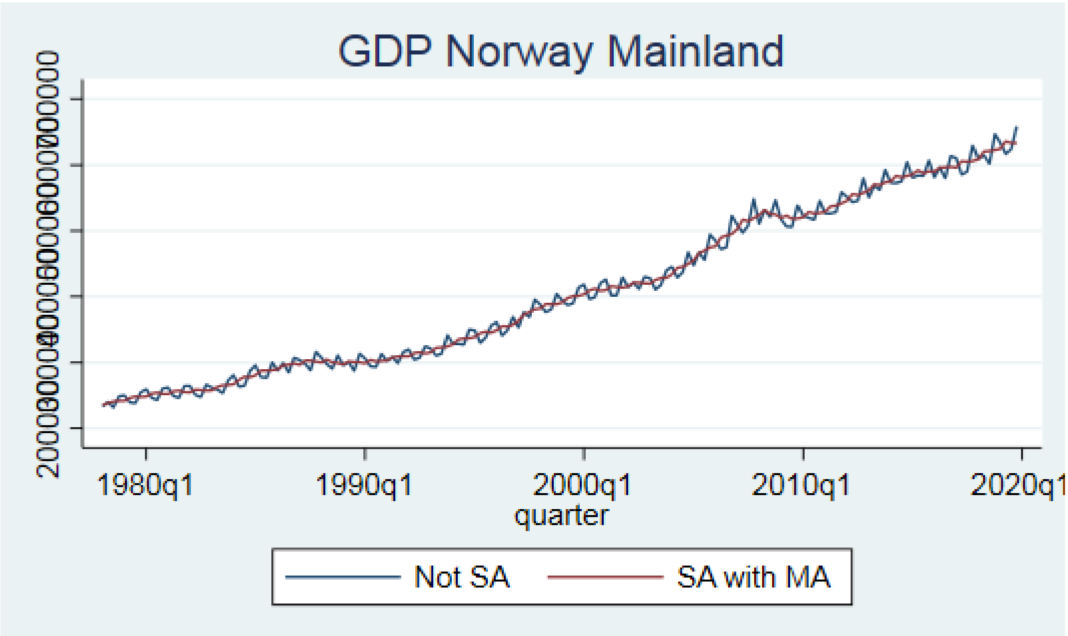
When using the seasonally unadjusted data, the trend as well as all dummies are statistically significant. When using the seasonally adjusted data, on the other hand, only the trend is signicant. Therefore, the log-transformed GDP needs to be seasonally adjusted to be able to properly analyze the series.

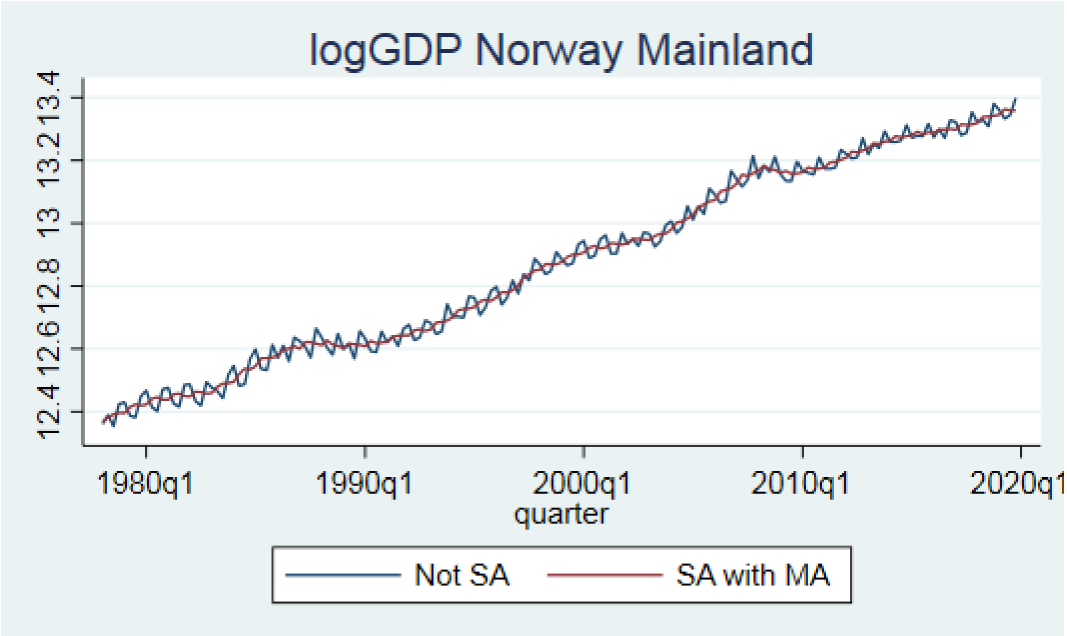
First difference of the log-transformed GDP:

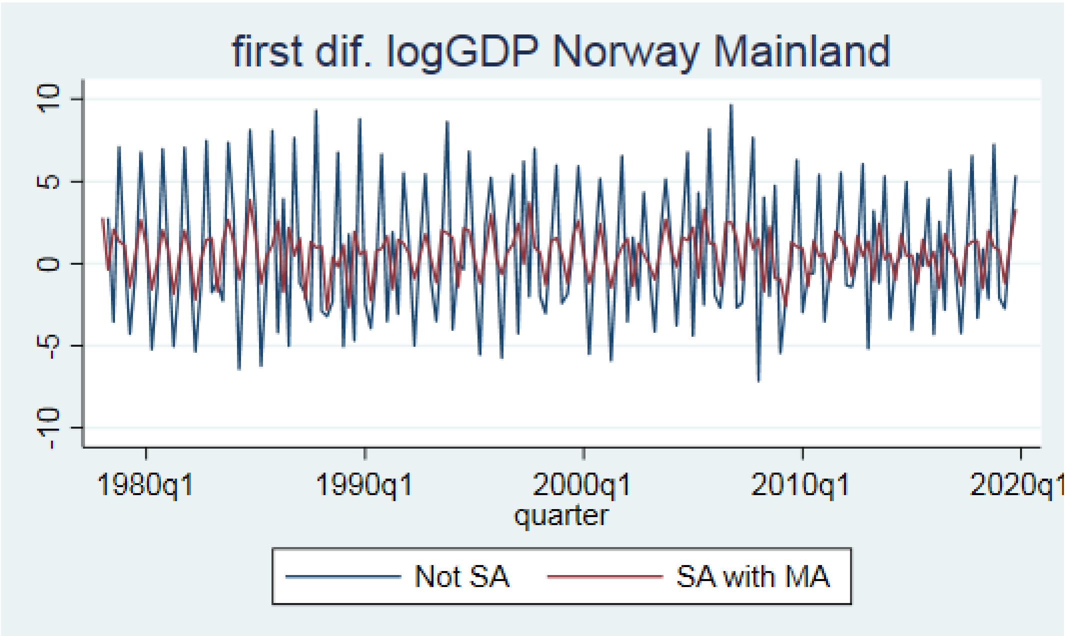
The trend is statistically significant both using the seasonally adjusted and the seasonally unadjusted data, but the dummy variables only test significant when using the unadjusted data. Thus, the first difference of the log GDP needs to be seasonally adjusted.

Year-on-year growth:

For both sets of data, the year-on-year growth does not show neither any significant trend nor significant dummy variables. This makes perfect sense since the year-on-year growth compares from year to year and does not regard anything that happens over the course of a year, i. e. seasonal effects. The year-on-year growth therefore does not need to be adjusted for further analysis.







The algorithm that we use to adjust the data works well for the GDP as well as the log-transformed GDP series, but does not work perfectly for the first difference of the logGDP as can be seen in the table below. We therefore continue to use the seasonally adjusted data from the SSB in the following tasks.

OLS Regression on seasonal adjusted data

	(1) GDPSAma	(2) logGDPSAma	(3) fdlogGDPSAma
trend	2503.0***	0.00623***	
	(0.000)	(0.000)	
dum1	-525.7	-0.00147	-0.365*
	(0.905)	(0.821)	(0.040)
dum2	4088.9	0.0101	-2.752***
	(0.353)	(0.122)	(0.000)
dum3	1717.7	0.00490	0.0264
	(0.696)	(0.451)	(0.881)
_cons	195203.2***	12.34***	1.396***
	(0.000)	(0.000)	(0.000)
N	168	168	168

p-values in parentheses
* p<0.05, ** p<0.01, *** p<0.001

1.4 Task D

According to the results from task c), we control both for trend and constant in this task. We used the ‘varsoc’ command to determine the ideal number of lags using the Aikaike Information Criterion (AIC) and the Bayesian information criterion (BIC). We chose to implement a maxlag number of 8 for ‘varsoc’. We generally use the latter as it has become the norm in economics to use less lags. After that, we use the Augmented Dickey-Fuller (ADF) test and the KPSS test to test for stationarity.

<u>Log-transformed GDP:</u>

- The AIC calculates 5 lags while the BIC only determines 4. We chose the latter as mentioned above.
- The results look as follows:
    - ADF: H0 cannot be rejected → non-stationary
    - KPSS: H0 cannot be rejected → trend stationary

First difference of the log-transformed GDP:

- The AIC determines 8 lags, the BIC 3. We therefore use 3 lags.
- The results:
  - ADF: H0 can be rejected → stationary
  - KPSS: H0 cannot be rejected → stationary

Year-on-year growth:

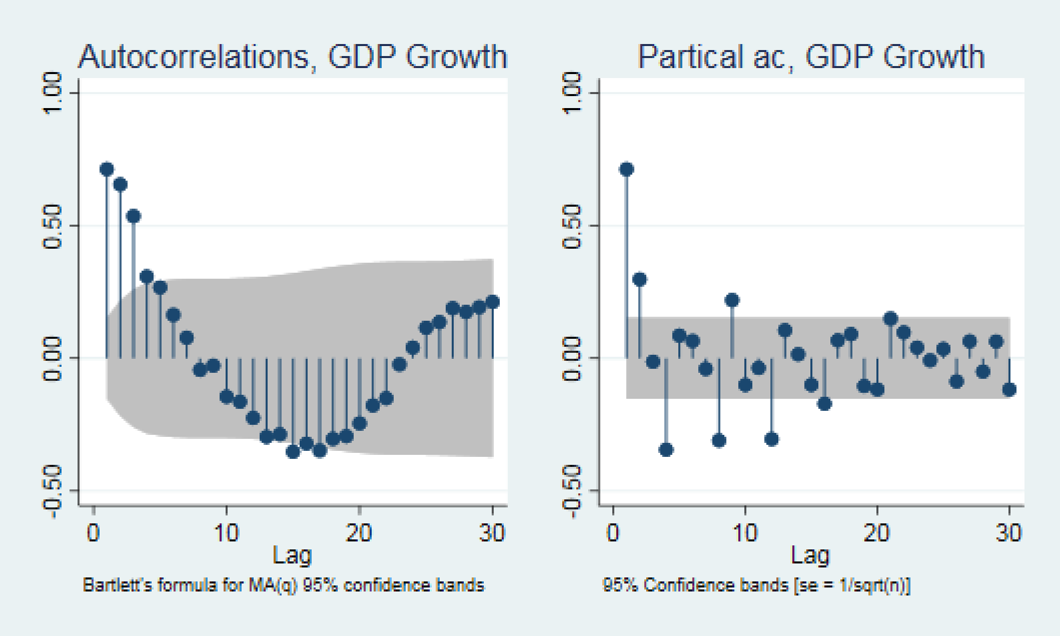
- The AIC determines 8 lags, the BIC 4. We therefore use 4 lags.
- The results:
  - ADF: H0 can be rejected → stationary
  - KPSS: H0 cannot be rejected → level stationary

Interestingly, KPSS gives the same result as ADF for the first difference of the log series and the year-on-year growth, but not for the log GDP. Here, the log GDP is non-stationary according to ADF, but stationary according to KPSS. This is likely due to the makeup of both tests: the null hypothesis of the KPSS test assumes stationarity, while the null hypothesis of the ADF test assumes a unit root. The result was not statistically significant enough to reject either of the null hypotheses.

As our preferred series, we chose the year-on-year growth. It is uniform and easy to predict. The series is stationary, which is a main requirement for the analysis later on. Additionally, both the ADF and the KPSS come to the same conclusion of stationarity.

1.5 Task E

A partial correlation is the correlation between two variables with the effect of a set of random variables removed. Autocorrelation means that something correlates with itself at another point in time. A partial autocorrelation function (PACF) therefore gives us information about the correlation of a time series with itself at prior time steps with the influence of intervening observations of the series itself removed.



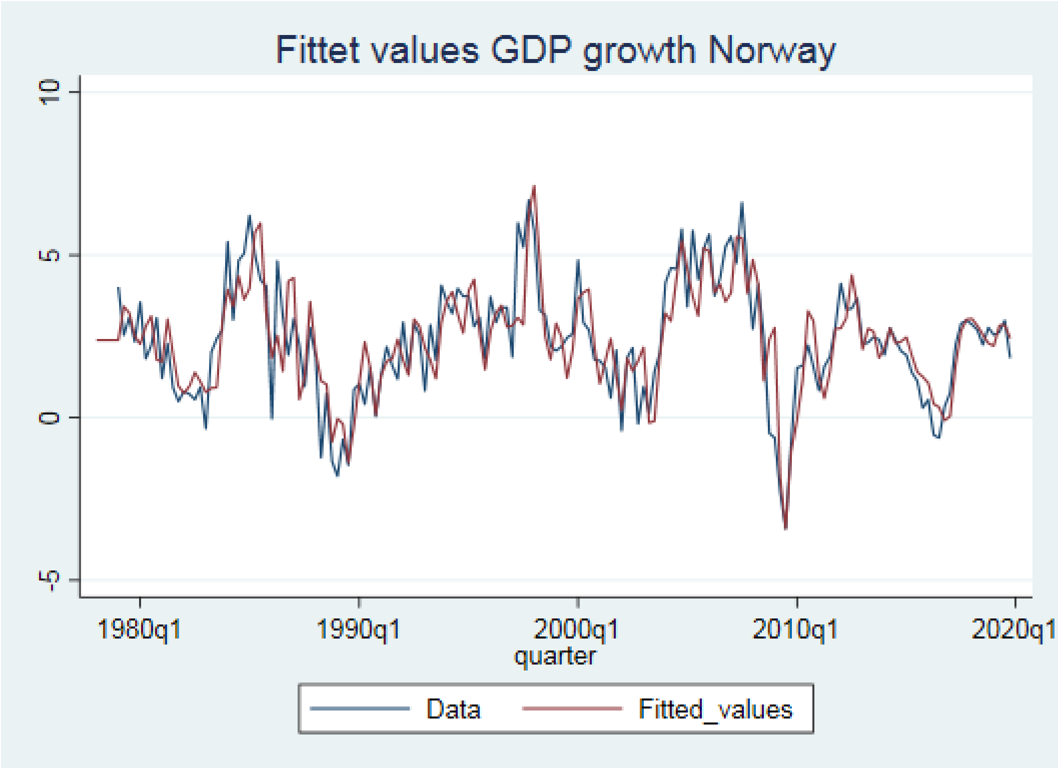
As is visible in our correlogram, both the autocorrelation and the partial autocorrelation oscillate and tail off, the former more quickly and the latter more slowly. This is a clear signal that the series follows an ARMA process.

It is interesting to see that the partial autocorrelation is statistically significant at every fourth lag until lag 20. This shows a quarterly pattern in the series as that means that every fourth quarter has a strong partial correlation with the corresponding quarter in the respective years before.

1.6 Task F & G

To determine the best ARMA/ARIMA model for the year-on-year growth series, we looped through combinations of AIC and BIC from ARMA(0 0) to ARMA(8 8) to find the model where they are minimized. As a result, we obtained ARIMA(1 0 5), i.e. an ARMA(1 5) process. When testing this model, it turned out than none of the MA lags were significant. We therefore changed the model to AR(1) MA(1 2 5) where all lags are statistically significant.

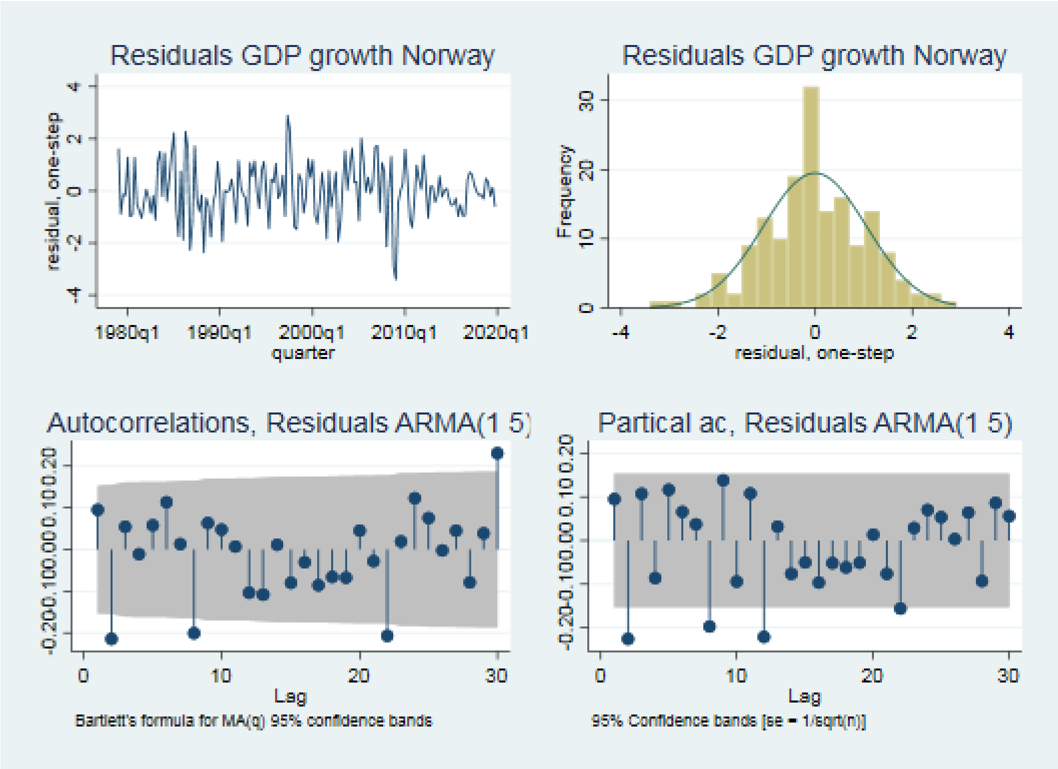
To start our evaluation of the model, we plotted the fitted values against the in-model values.



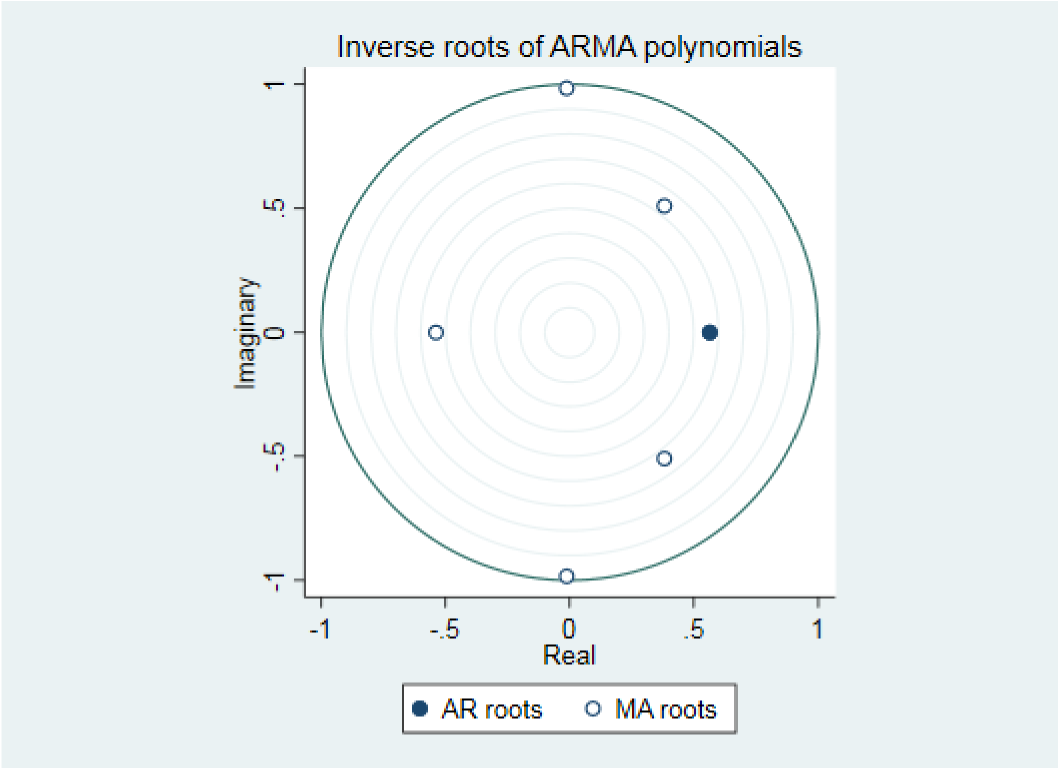
In order to further determine the quality of our model, we then tested the residuals by creating a correlogram, a histogram and a line plot.

The line plot of the residuals shows that the variance is relatively large in the beginning, but slightly decreases in the early 2010s. This means that it is heteroscedastic and might make the prediction more inaccurate.

As shown in the histogram, the residuals are approximately normally distributed which is beneficial for the calculation of the prediction intervals. Finally, the correlogram of the residuals shows that there is a small number of significant lags. While this is not ideal, it does not mean that the model is wrong. Instead, it leads to larger prediction intervals than necessary.



The eigenvalues, i.e. roots of the characteristic equation, all lie inside the unit circle. This means that the AR parameters satisfy the stability condition and the MA parameters satisfy the invertibility condition.



The model

$$y_t = 2{,}3938 + 0{,}5647 y_{t-1} + \varepsilon_t - 0{,}2018\varepsilon_{t-1} + 0{,}9578\varepsilon_{t-2} + 0{,}2118\varepsilon_{t-5}$$

leads to a characteristic equation of:

$$(1 - 0{,}5647L)Y_t = 2{,}3938 + (1 - 0{,}2018L + 0{,}9578L^2 + 0{,}2118L^5)\varepsilon_t$$

2. Forecasting

2.1 Task H

In the ARMA(2 2) model, only the AR L1 and the MA L2 are statistically significant, whereas in the AR(1) MA(1 2 5) model, all lags are significant. This only means that the corresponding coefficients are not significantly different from 0. The significant coefficients of the ARMA(2 2) model are very similar to ours.

```
Sample:  1979q1 - 2019q4                        Number of obs     =        164
                                                Wald chi2(4)      =     615.75
Log likelihood = -246.0259                      Prob > chi2       =     0.0000
```

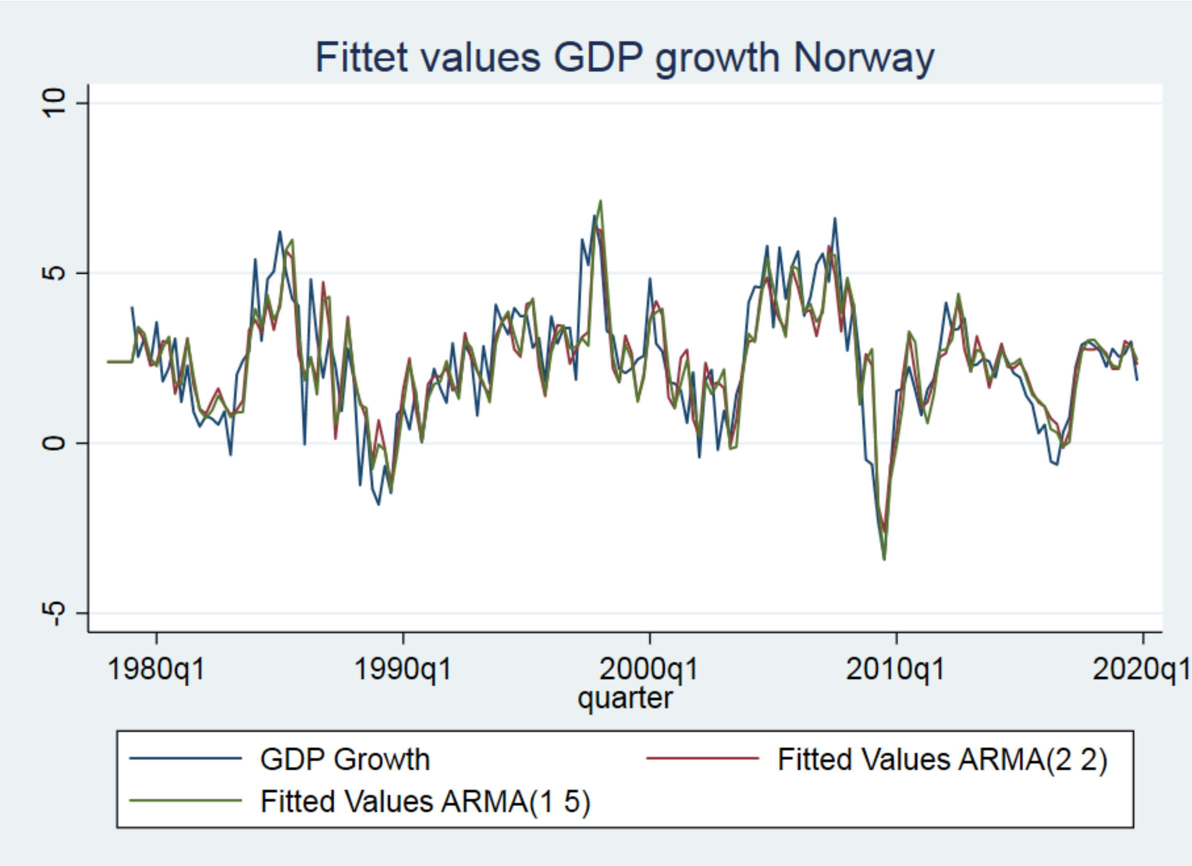
growth	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
growth						
_cons	2.390638	.2930894	8.16	0.000	1.816193	2.965082
ARMA						
ar						
L1.	.5020441	.0700836	7.16	0.000	.3646827	.6394055
L2.	-.0580287	.0837841	-0.69	0.489	-.2222426	.1061851
ma						
L1.	-.039721	.0257171	-1.54	0.122	-.0901256	.0106837
L2.	.972664	.043267	22.48	0.000	.8878623	1.057466
/sigma	1.063985	.0638661	16.66	0.000	.9388099	1.18916

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

```
Sample:  1979q1 - 2019q4                    Number of obs     =        164
                                            Wald chi2(4)      =    1102.96
Log likelihood = -241.7715                  Prob > chi2       =     0.0000
```

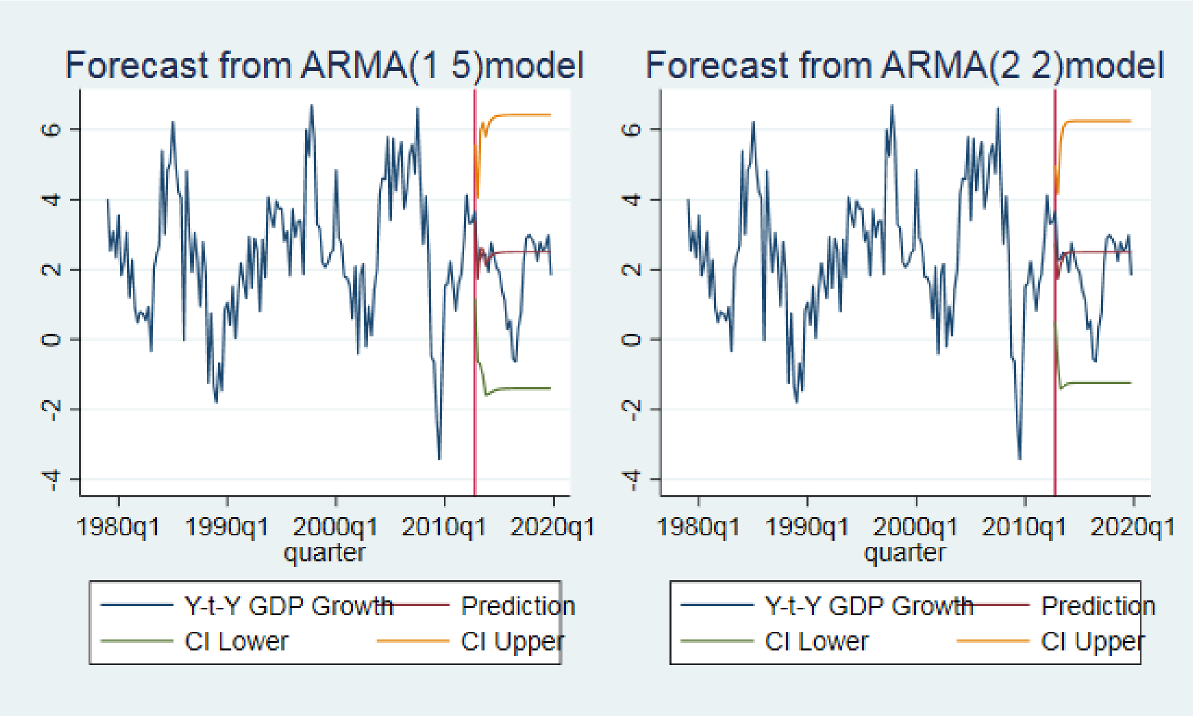
growth	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
growth						
_cons	2.393819	.3728865	6.42	0.000	1.662975	3.124663
ARMA						
ar						
L1.	.5647054	.0742558	7.60	0.000	.4191668	.7102441
ma						
L1.	-.2018061	.0546456	-3.69	0.000	-.3089096	-.0947027
L2.	.9578555	.0298946	32.04	0.000	.8992632	1.016448
L5.	.2117557	.0548163	3.86	0.000	.1043177	.3191936
/sigma	1.037321	.0535558	19.37	0.000	.9323535	1.142289

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.



2.2 Task I

Both forecasts, like the models they are based on, are quite similar. The confidence interval for our model of AR(1) MA(1 2 5) is a bit wider, but the realized values of GDP growth lie within the interval for both forecasts. The predicted GDP growth is larger than the observed growth. Around 2015, the GDP growth plummeted, but this development is still contained in the prediction interval.

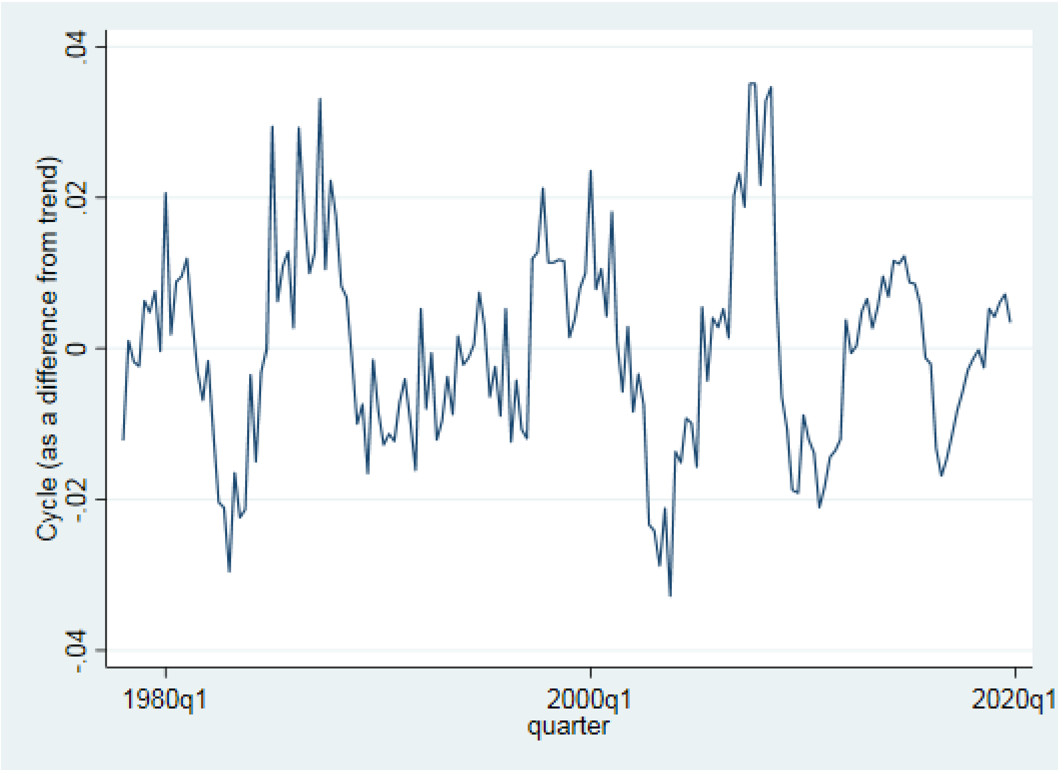


2.3 Task J

Model	RMSE	BIAS	SE	AIC	BIC
ARMA(2 2)	1,9076	-0,2071	0,1024	504,0518	522,651
ARMA(1 5)	1,9972	-0,1258	0,0890	495,5431	514,1423

The evaluation of the forecast of both models suggests that the ARMA(1 5) is preferable. This can be concluded from the comparison of the values in the table above. An example of that is that our model minimizes the information criteria.

The Norwegian business cycle is strongly dependent on the price development of oil, more specifically Brent Crude oil as that is exported by Norway. This becomes very apparent in the slump at the end of 2015 and the beginning of 2016. Around that time, both the price of Brent Crude oil as well as the Norwegian GDP growth fell steeply.



To evaluate the current state of the Norwegian economy, literature recommends to use the Hodrick-Prescott filter. It separates the data in a trend and a cycle component. The current cycle suggests that the Norwegian growth has been above the trend in the last quarters. This can also be observed visually in the graph above as the growth currently is above 0.

3. Appendix

3.1 Data:

- https://www.ssb.no/en/statbank/sq/10018113 (Table 2. Final expenditure and gross domestic product. Chained volume figures/fixed prices from base year. NOK million [xlsx])

- https://www.ssb.no/en/statbank/sq/10018120 (Table 6. Final expenditure and gross domestic product. Seasonally adjusted figures. Chained volume figures/fixed prices from base year. NOK million [xlsx])

3.2 STATA-Packages:

- Kpss
- Estout
- Hprescott