2020 Exam 1: Forecasting of Quarterly Gross Domestic Product for the United States of America

BAN430: Forecasting

NHH NORGES HANDELSHØYSKOLE

submitted in March 2020

Contents

List of Figures

List of Tables

1 Data description

For this assignment, we chose to use the quarterly unadjusted gross domestic product (GDP) of the United States of America (USA) from the first quarter of 1990 to the forth quarter of 2019 as our time series. We obtained the data from Fred (2020a). The quarterly unadjusted data were given in nominal values. For this reason, we downloaded the GDP deflator from Fred (2020b) to adjust the series to real values. The base year is 2012, we did the adjustment because the inference from the economic analysis, beyond the statistical one, is clearer with real figures. We define the training horizon as the window from the first quarter of 1990 to the fourth quarter of 2015. Consequently, the test set encompasses the first quarter of 2015 to the fourth quarter of 2019. The time series in Figure 1 is the real GDP. It shows a constant upward trend and thus mostly positive growth and no stationarity. During the Great Recession from 2008 to 2009, the quarterly GDP fell sharply. The series exhibits a seasonal pattern.

After the fourth quarter of each year, the value for the following first quarter of the next year is usually lower

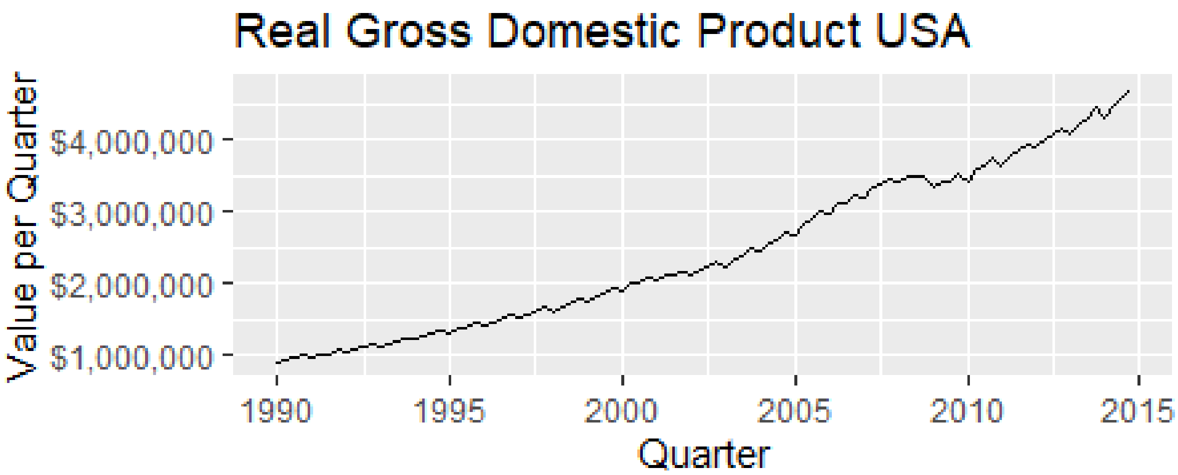

Figure 1: Quarterly Real GDP USA

Due to this trend, the first entry in 1990 has the lowest value in this time series as is shown in Table 1. Consequently, the highest number in the training time series is the last entry in the fourth quarter of 2014. The median of the quarterly GDP values is $2,218,684,000, which is aligned with the first quarter of 2004. Since minimum and maximum are at the extremes and the 1st and 3rd quantiles and the median are fairly evenly distributed over the time horizon, it can be assumed that no extreme outline disturbs the growth trend. In general, this descriptive analysis only allows the conclusion that the GDP is growing. Any further economic inference is not possible.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Date	1990-Q1	1996-Q1	2004-Q1	calculated	2010-Q1	2014-Q4
GDP in $ Thousand	900128	1457459	2218684	2451147	3431409	4715080

Table 1: Summarize statistics of quarterly real GDP USA

In Figure 2, we plot the quarterly GDP separately per quarter to see the differences in the mean. The (expenditure) GDP is defined as
$GDP_t = Consumption_t + Government\ Expenditures_t + Investments_t + Current\ Account_t$ (Mankiw, 2014) with consumption usually as the most important component (for the USA in particular). This is why the entire GDP is affected when Christmas shopping strongly drives up consumption figures in the fourth quarter. It also explains why the GDP is always the highest in the fourth quarter. In the first quarter of each year, consumption is usually weak. This makes sense because most purchases were made before Christmas and construction activity is impaired by poor weather conditions. Both of these factors result in a low mean GDP compared to the other quarters.

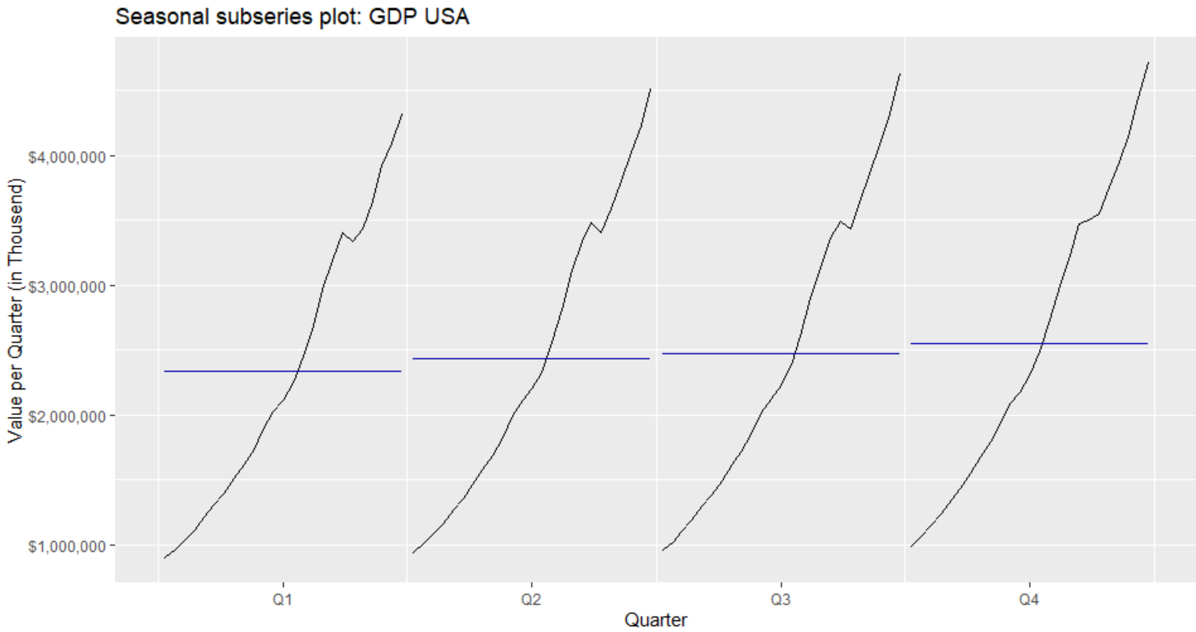

Figure 2: Real GDP USA development per Quarter

In order to test the trend and seasonality, we use a simple regression:

$$GDP_t = \beta_0 + \beta_1 * trend_t + \beta_2 * Quarter_2 + \beta_3 * Quarter_3 + \beta_3 * Quarter_4 + \varepsilon_t$$

We regress the dependent variable GDP_t on the independent trend variable and quarterly dummies. Quarter 1 acts as a reference for the chosen dummies. In R, we use the function tslm dedicated to linear regressions of time series.
Table 2 shows us a strongly significant trend coefficient of 37664.3, which accounts for most

of the explanatory power of this regression. Furthermore, the table shows a significant seasonality at least between quarter 1 and 4.

Depended Variable:	GDP_t			
Coefficients	Estimate (in $ thd.)	Std. Error	t value	Pr(>|t|)
constant	496741.4	42356.9	11.728	<2e-16 ***
trend	37664.5	559.6	67.305	<2e-16 ***
2nd Quarter Dummy	56039.5	45658.7	1.227	0.2227
3th Quarter Dummy	56775.6	45669.0	1.243	0.2169
4th Quarter Dummy	96569.9	45686.1	2.114	0.0372 *
Signif. codes:	0 ’***’ 0.001 ’**’	0.01 ’*’ 0.05 ’.’	0.1 ’ ’ 1	
Residual standard error:	161400 on 95 degrees of freedom			
Multiple R-Squared	0.9796	Adjusted R-squared:	0.9787	
F-statistic:	1138 on 4 and	95 DF, p-value:	<2.2e-16	

Table 2: Time series regression on quarterly dummies

We use the autocorrelation functions to further investigate the pattern of the time series. The partial autocorrelation function shows that the quarterly GDP has a strong direct influence on the following quarter. Nonetheless, there is no statistically significant direct link between the quarterly GDP and a lagged quarter of an order higher than one. Because of the strong relationship with the first delayed quarter, the autocorrelation function shows that indirect effects are statistically significant until the 29th lag. This shows that the persistence is strong.

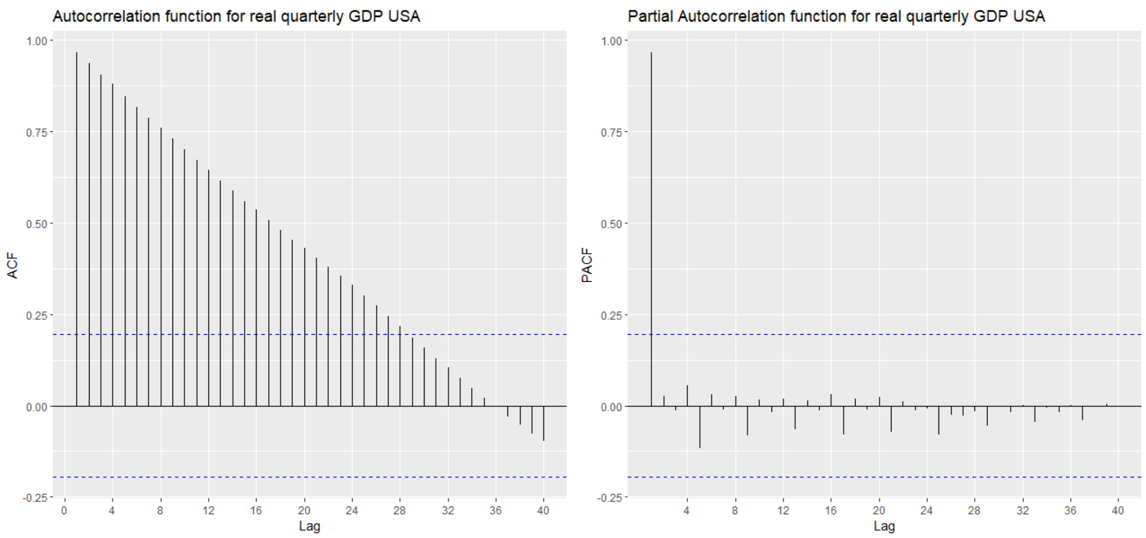

Figure 3: Autocorrelations for Quarterly Real GDP

2 Time series decomposition

We use the STL method for the decomposition of the time series. This choice is based on the robust features of this method, allowing us to focus our analysis more on economic fundamentals. It allows the seasonality to change over time (Hyndman & Athanasopoulo, 2018) and smooths over outliers well. Therefore, we are able to construct a robust decomposition. This means that the outline situation of the Great Recession of 2008 to 2009 does not affect the estimates of the trend cycle and seasonal components. In order to offset that (non-)effect, the remainder component nevertheless contains bigger values than the other decomposition methods. Furthermore, the disadvantages of the STL method are not as problematic for out purpose of analyzing real US GDP data. It does not consider trading days or variation in the calendar. But because we use quarterly data, these issues do not affect our analysis. Neither does the weak handling of additive composition since our level data has to be multiplicative decomposed (Hyndman & Athanasopoulo, 2018).

The decomposition in Figure 4 shows that the highest fraction of the data is explained by the trend component. This corresponds to the regression before. The gradient of the trend remains very steady. It only gets slightly steeper after 2005 but was corrected immediately in the period from 2008 to 2009. The variance of seasonal fluctuation is stable, but this is one reason why we have chosen the STL method. Consequently, the larger deviation from the long-term trend, especially in the period from 2005 to 2010, goes into the remainder.

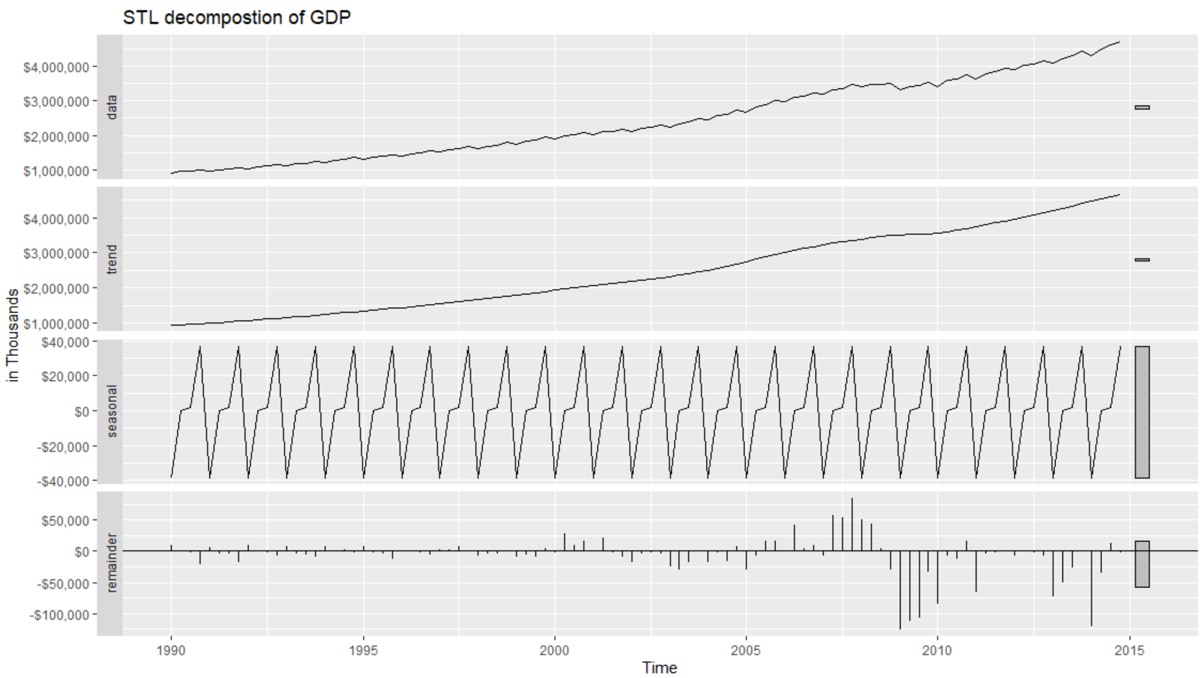

Figure 4: STL decomposition of quarterly real GDP

Figure 5 plots the trend and the real GDP series. The actual data fluctuates with a constant variance around the STL trend. The deviation only gets larger after 2008, but goes into the remainder, as we discussed earlier.

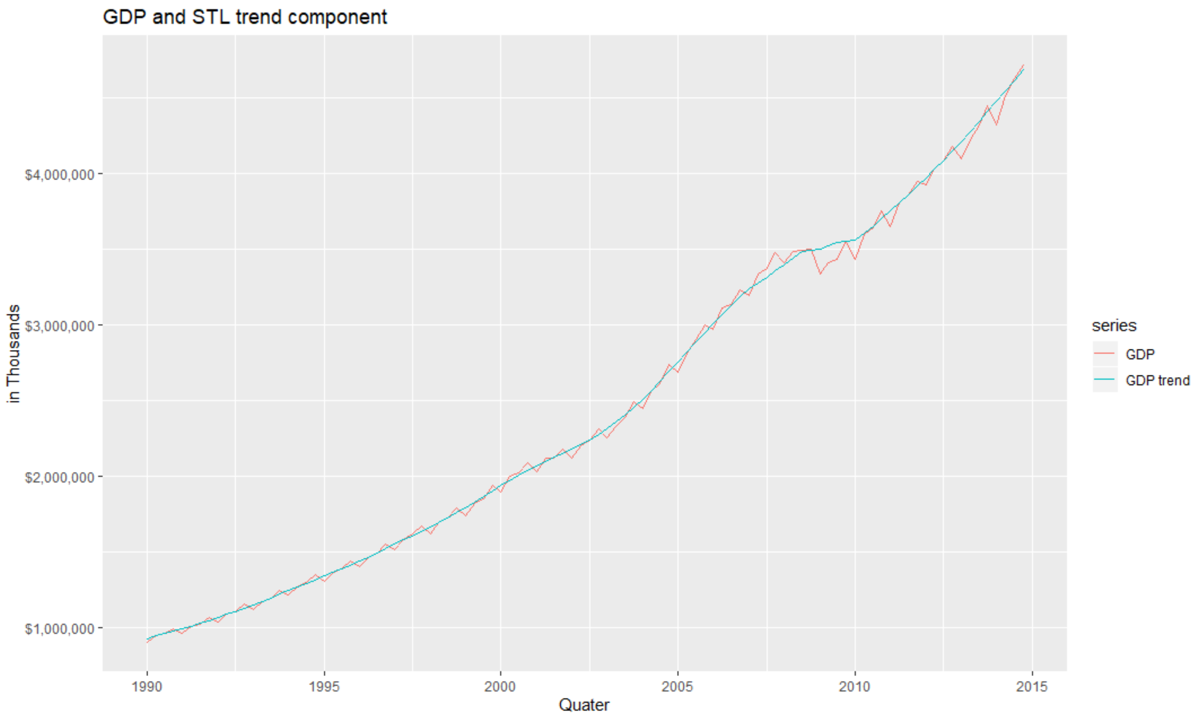

Figure 5: STL decomposed trend and quarterly real GDP

We use the Hodrick-Prescott filter to determine business cycles in a more differentiated manner. The business cycle is the difference between the long-term trend and actual data, hence the short-term fluctuation beyond the seasonal pattern during the year. Due to the recommendation for quarterly data (Phillips & Shi, 2019), we choose 1600 for λ.

Figure 6 shows a relatively small deviation in the 1990s. In the early 2000s, growth slowed down. Then, it picked up even more strongly, leading to an overshooting of the trend until the financial crisis occurred. During the crisis, growth fell well below its long-term trend level. It started to recover in 2010 but flattened out again in 2015. Since then, growth has been developing strongly.

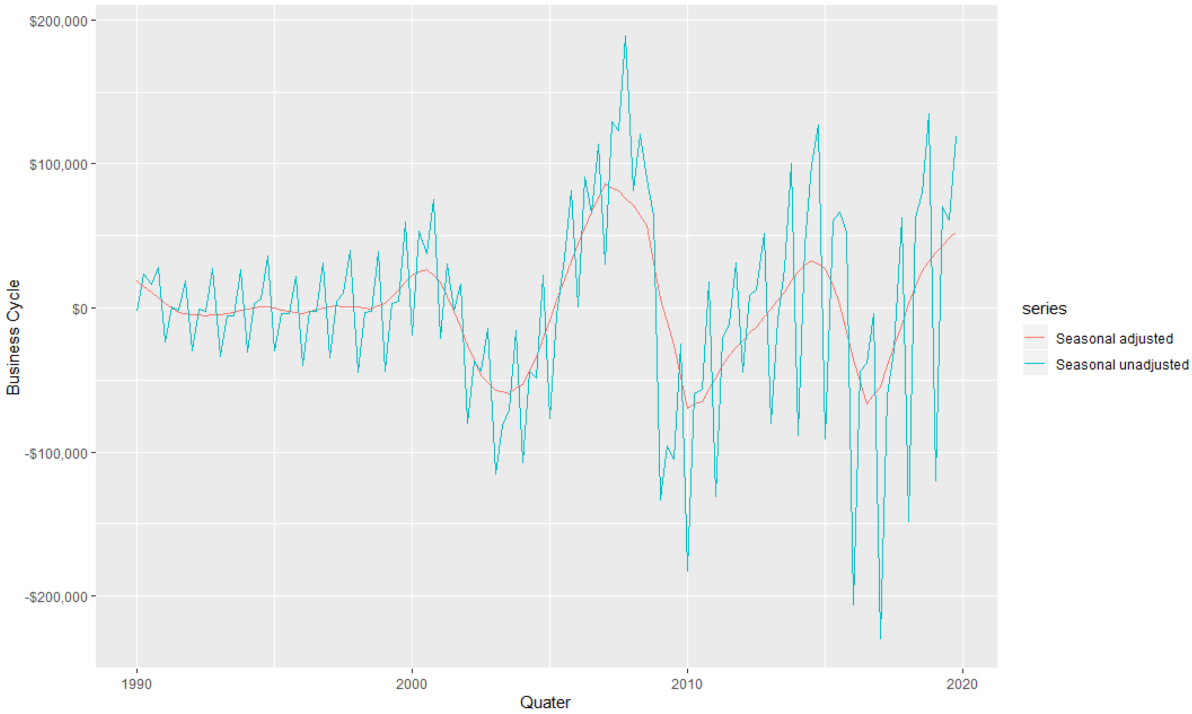

Figure 6: Business cycle for real GDP USA

3 Forecast individual components

We decompose the GDP over the entire time horizon and then extract the test period from these values. This allows us to compare the test figures to the forecast figures . We apply the fpp2 function forecast(). It uses h = h20 because the test horizon encompasses 20 quarters. In addition, we set allow.multiplicative.trend = TRUE in order to take the multiplicative trend characteristics of our GDP data into account.

Figure 7 shows the forecast for the trend component. The forecast function uses an ETS(M,A,N) model. In the forecast, the actual trend is overestimated initially and underestimated after. However, the actual trend lies well within the confidence intervals of the forecast. .

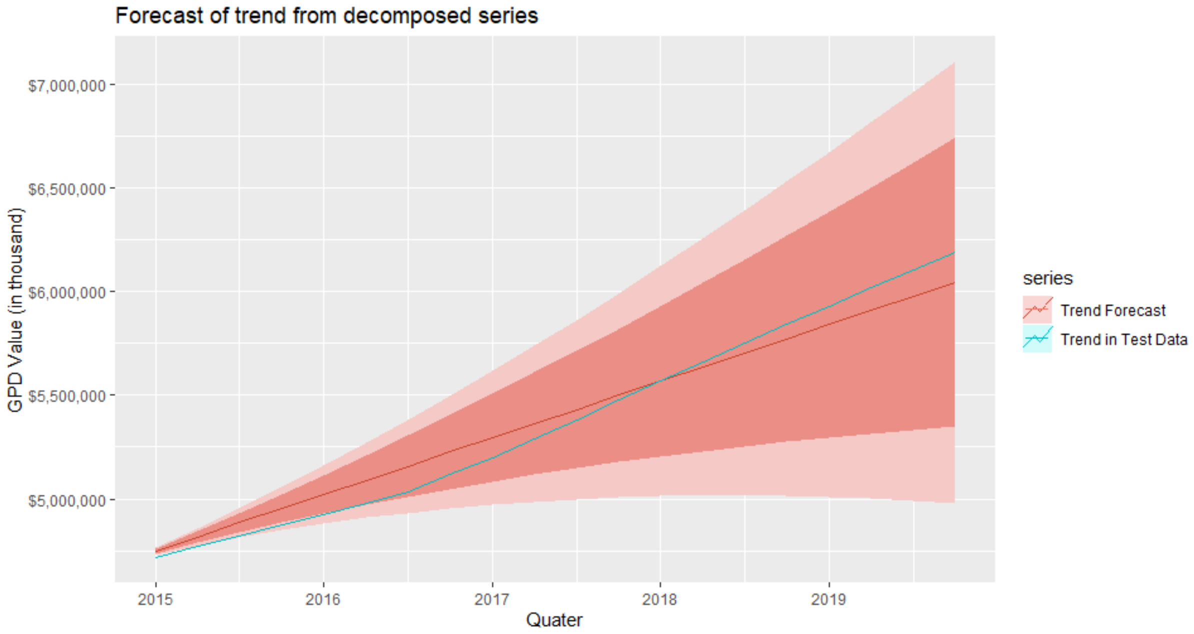

Figure 7: Forecast of trend component

To predict seasonality components, the forecast function uses an ETS(A,N,A) model. The forecast is very accurate, as shown in Figure 8. Only the seasonality of the fourth quarter is always underestimated slightly, but only very marginally.

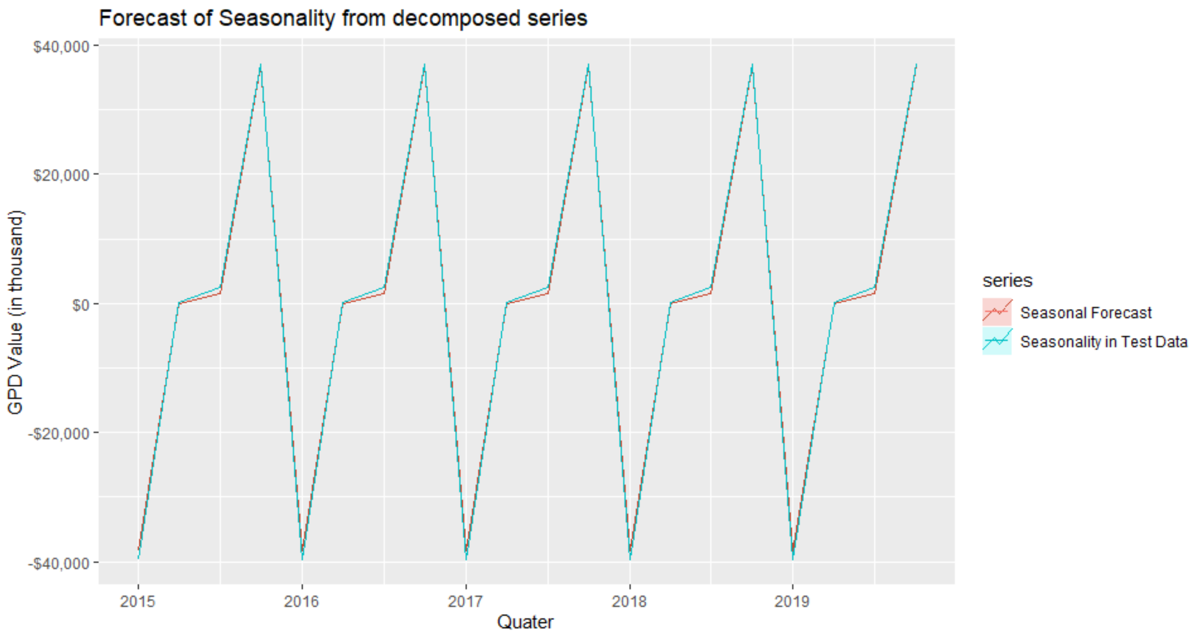

Figure 8: Forecast of seasonality component

The remaining component is predicted by an ETS(A,N,A) model as shown in figure 9. This time, the ETS(A,N,A) model is less accurate. In 2016 and 2017, the actual data in the first quarter are outside the confidence interval, and in general the actual data are mostly overestimated.

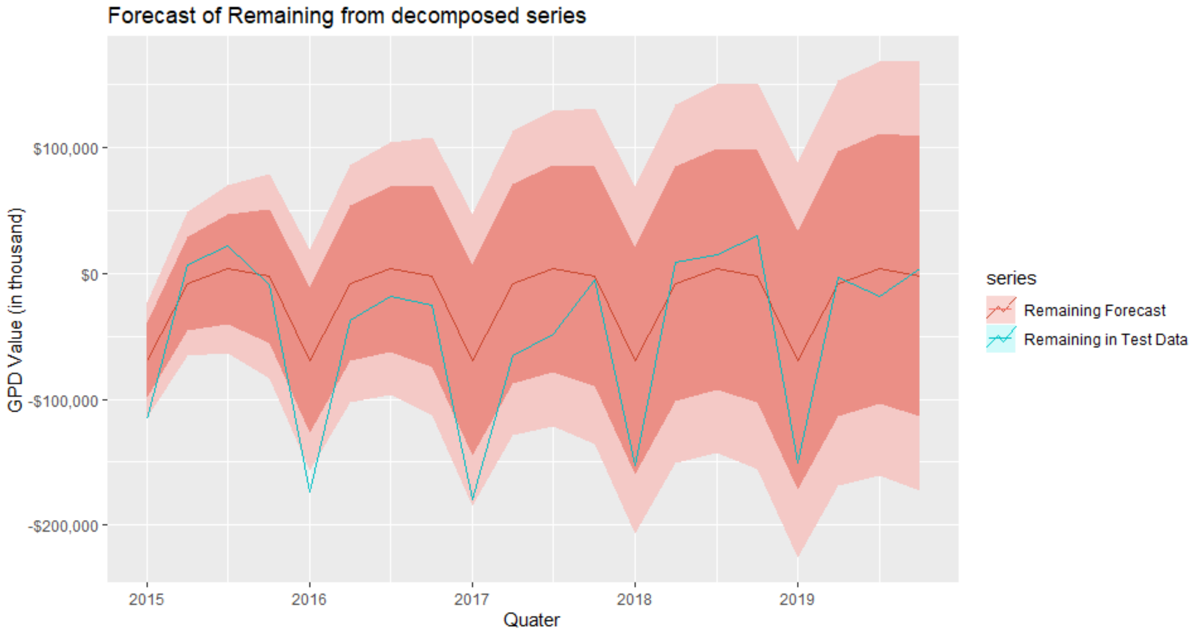

Figure 9: Forecast of remainder component

In Figure 10, we combine all the components into one forecast. The mean forecast initially overestimates the actual data but follows the seasonality seamlessly. In the last quarters it underestimates the actual figures. In 2016, the actual data lie just outside our confidence interval, but otherwise develop close to our mean trend.

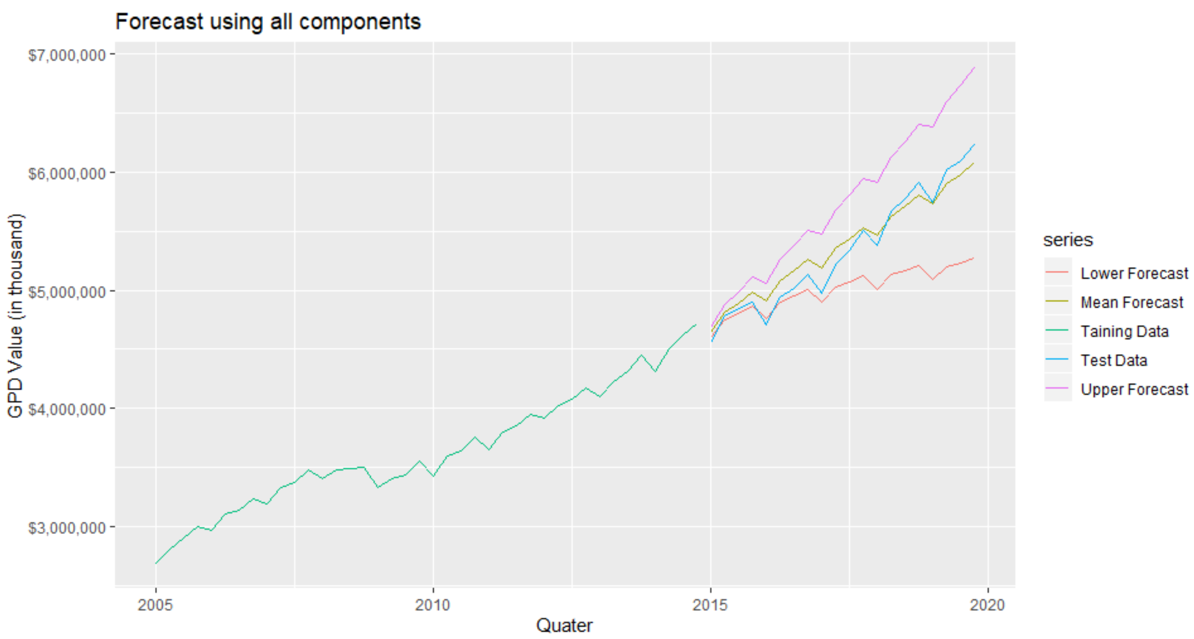

Figure 10: Forecast of combined components

4 Selection of appropriate ETS Model

For the determination of the ETS model, we use the ets() function in R. If the model specifications are on "ZZZ", the function chooses the model with the best Akine Information Criteria . For our training set of data, the model selects an ETS(M,A,M) model, which is a multiplicative Holt-Winters method with multiplicative errors.
In more detail, we receive the following estimators for the smoothing parameters:

$\hat{\alpha} = 0.8478, \quad \hat{\beta} = 0.3471, \quad \hat{\gamma} = 0.1522$ which yields following model:

$$y_t = (l_{t-1} + b_{t-1}) s_{t-m} (1 + \varepsilon_t)$$

$$l_t = (l_{t-1} + b_{t-1})(1 + 0.8478\varepsilon_t)$$

$$b_t = b_{t-1} + 0.3471(l_{t-1} + b_{t_1})\varepsilon_t$$

$$s_t = s_{t-m}(1 + 0.1522\varepsilon_t)$$

With these parameters we obtain a solid forecast, which is shown in the Figure 11. Initially, the forecast overestimates the actual GDP data a little. In recent years, however, the estimate has been very accurate.

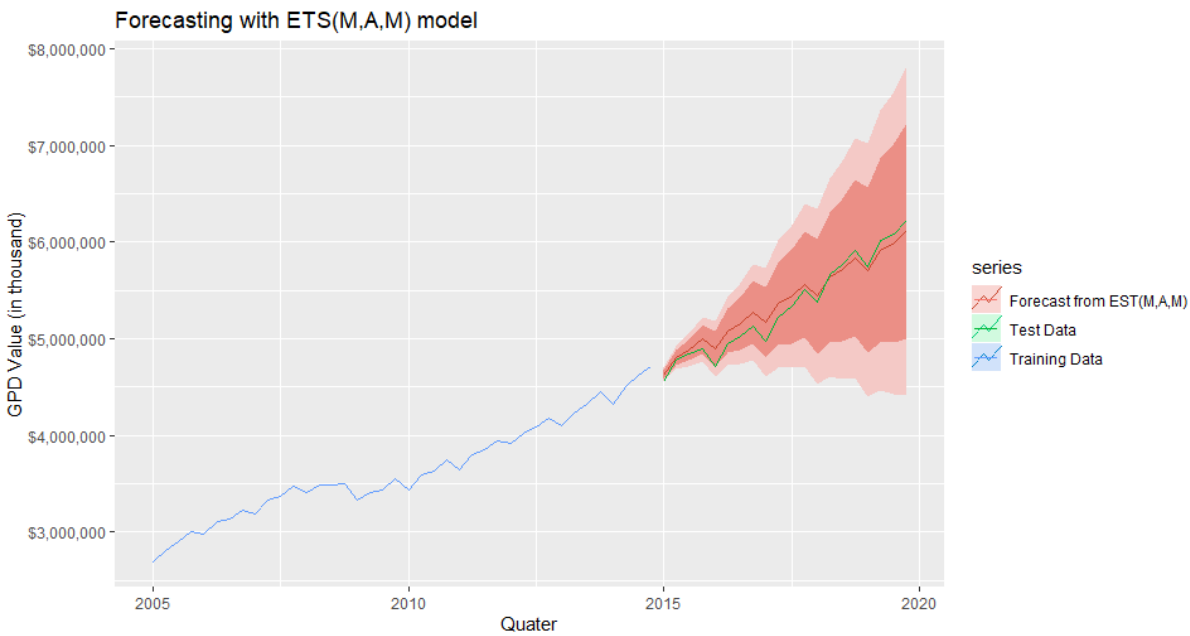

Figure 11: Forecast with ETS(M,A,M)

In order to test for autocorrelation, we use the Box-Pierce test. Because we only slightly exceed the 100 times series entries required for the Box-Pierce test, we additionally apply the Box-Ljung test.

H_0: $\rho_i \ = 0$, for all i, hence no autocorrelation
H_1: $\rho_i \ \neq 0$

Since the p-values shown in Table 3 for both tests are above all critical significance levels, we cannot reject the null hypothesis of no autocorrelation. Thus, we cannot obtain any further statistical inference.

	χ^2	Degrees of freedom	p-value
Box-Pierce test	10.517	10	0.3964
Box-Ljung test	11.586	10	0.3137

Table 3: Autocorrelation tests

The residuals, which we analyze in Figure 12 have no clear graphical pattern. Additionally, the autocorrelation function does not show any significant lag. The density function approximately follows a normal distribution. The Shapiro-Wilk test for a normal distribution of the residuals gives us a p-value of 0.1029. Thus, we can assume normality for the residuals.

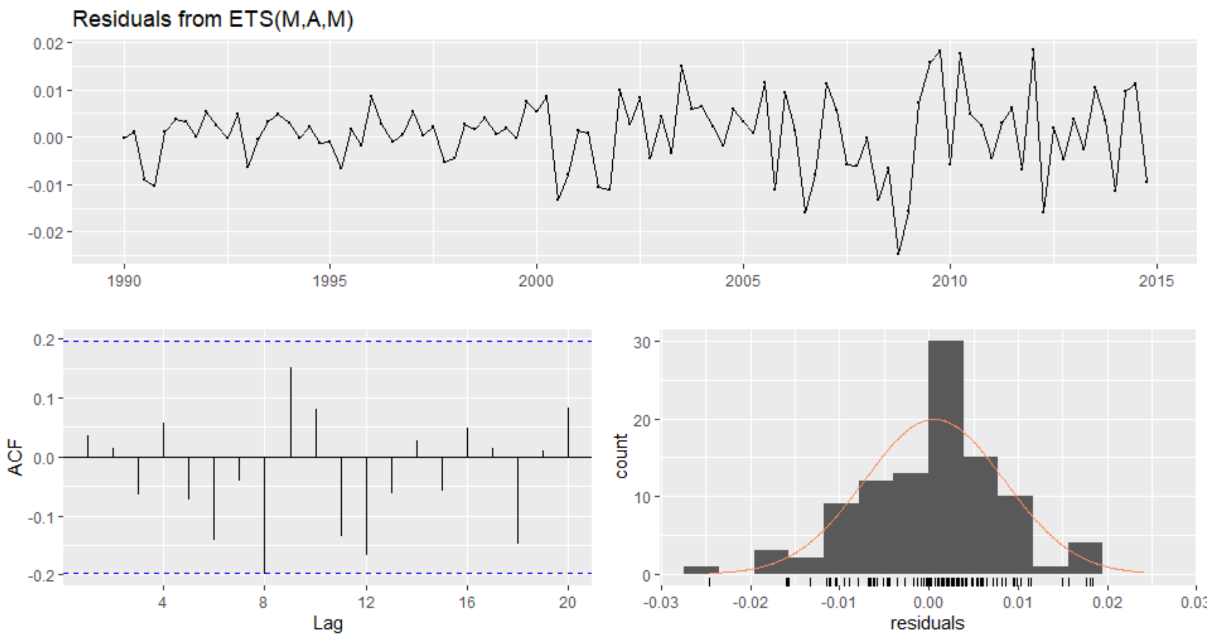

Figure 12: Residuals of ETS(M,A,M) Forecast

Given these results, the residuals do not indicate any problems or suggest further improvements of the ETS(M,A,M) model.

5 Forecast Comparison

Among the basic prediction techniques, random walk with drift prediction seems to be the best description of our test set. Similarly, the multiplicative Holt-Winter method from the class of Holt methods is best suited to predict our data. The selection of both methods is shown in the appendix. In this section we compare the quality of the predictions for the test period of these two methods and the ETS(M,A,M) model.

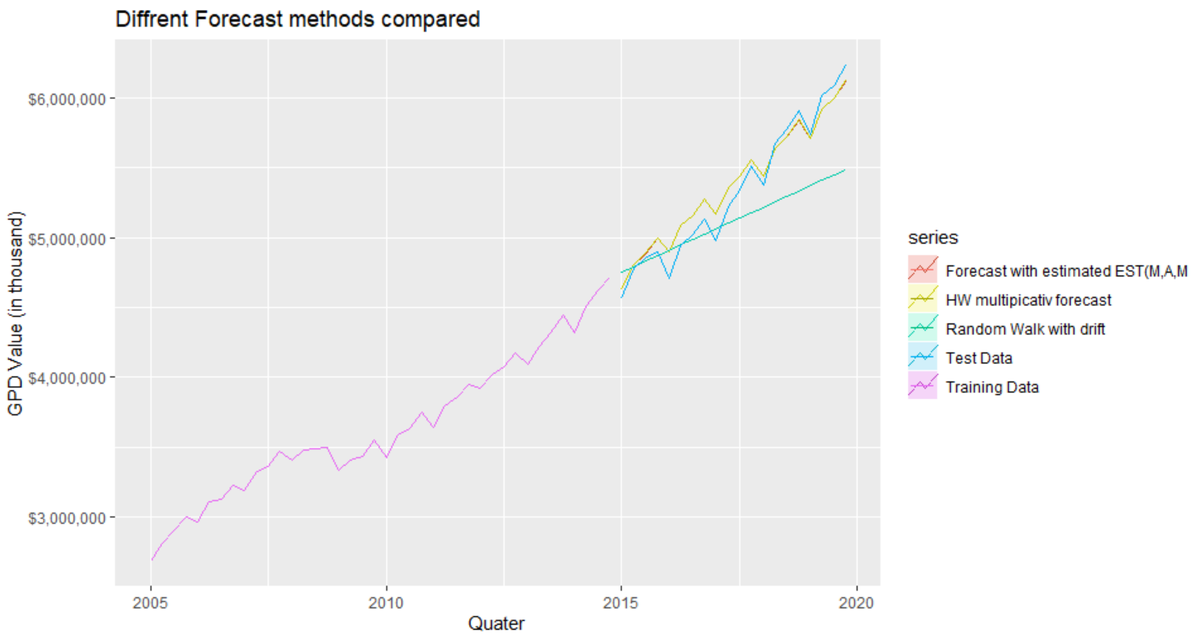

Figure 13: Different forecasts in comparison

Visually, there is almost no difference between the ETS(M,A,M) and the multiplicative Holt-Winter estimates. This is not surprising, considering that the only differences are in the use of multiplication errors for the former method. Both methods are noticeably better than the drifted random walk, as shown in Figure 13. Since the first two are not graphically distinguishable, we review the residuals and evaluate their magnitude in order to determine the best prediction model:
Based on the Root mean squared error (RMSE) in Table 4 we interfere that the ETS(M,A,M) is slightly more precise than the Holt-Winter multiplicative model.

	ME	RMSE	MAE	MPE	MASE	ACF1
Random Walk	217021.16	352972.9	266359.18	3.6339	1.7317	0.7102
HW multiplicativ forecast	-44684.35	105339.6	93083.56	-0.9723	0.6052	0.8227
EST(M,A,M)	-43173.58	105509.1	93462.83	-0.9457	0.6076	0.8216

Table 4: Comparison of Residuals

6 Box-Cox transformation and Forecasts

The GDP data can be well explained by an exponential process. Since it is easier to work with linear relationships, exponential data are transformed to linearize them. Among the techniques that make this possible is the Box-Cox transformation. They smooth the data towards a linear process. We obtain a well-suited λ by using the BoxCox.lambda function, which yields a λ equal to 0.19. A λ is well suited if the seasonality of the transformed time series has the same variance over the entire period (Hyndman & Athanasopoulo, 2018). To be able to compare the predicted point value of the transformed series with the original values, we have to reverse the transformation. In this way, we obtain the median of the predicted distribution. Since this is not the median, we sometimes need to adjust the forecast to compensate for a bias. In our case, the RMSE is 1.1439 for the adjusted and 1.1301 for the unadjusted forecast, so we continue with the unadjusted forecast.
Similarly, one can use splines to regress nonlinear data, as exponential data like GDP, in pieces. We use cubic splines, so that the pieces are contentious to each other .

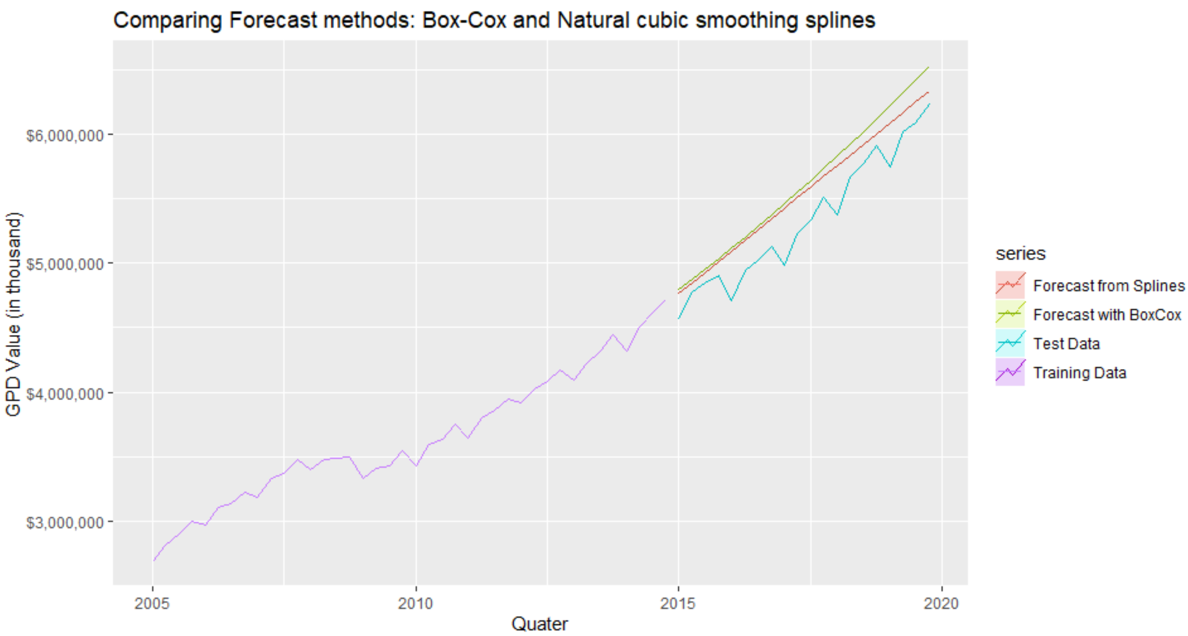

Figure 14: Box Cox transformation and spline forecast

Out-of-sample predictions of splines can be unreliable (Hyndman & Athanasopoulo, 2018) and in the the first quarters of 2015, the test data lie outside the spline prediction interval (See Appendix). However, the graph in Figure 14 indicates that the spline forecast still performs better than the Box-Cox transformation, thus we will examine the residuals of this forecast . As one can see in Figure 15, the lags of the residuals show a clear and significant pattern. In addition, the density is screwed to the right and the Shapiro-Wilk normality test yields a p-value of 0.001205, so we cannot assume normality for the residuals.

Both forecasts in this section give weaker results than the ETS(M,A,M) or the Holt-Winter multiplicative model.

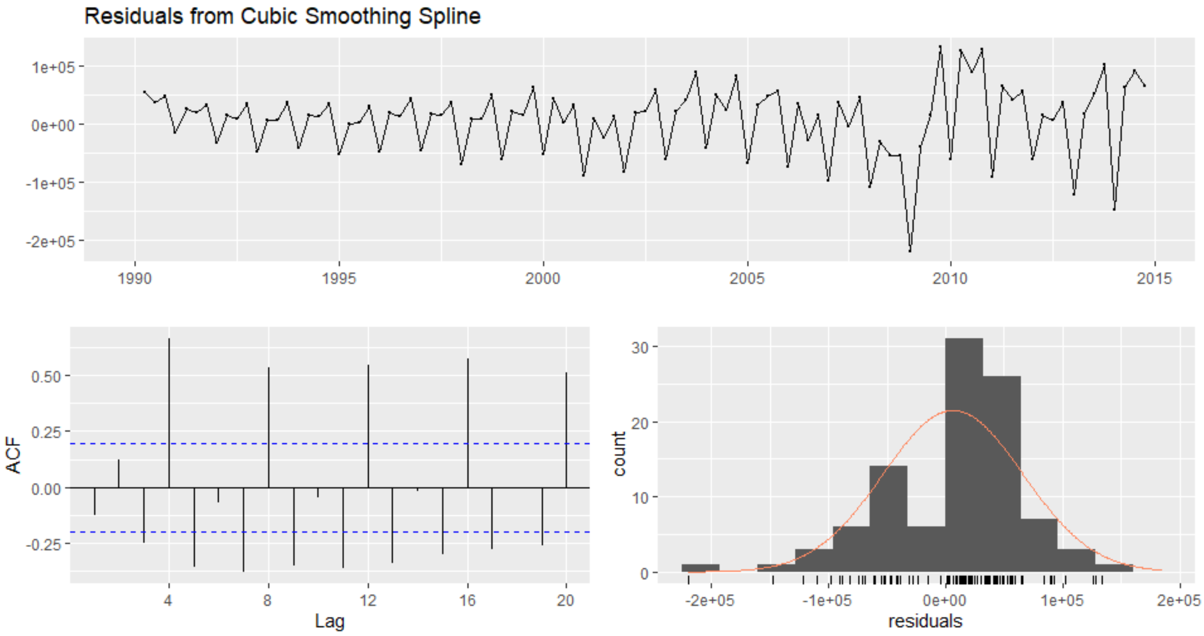

Figure 15: Residuals of spline Forecast

7 Modeling the tsCV function

We decided the best way to show our work in this section is to show the R-Code with the in-code comments:

```
#time series cross-validation
GDP.tsCV = tsCV(GDP.train, rwf, h = 2)
summary(GDP.tsCV)

#creating the tsCV_self function which takes an timeseries,
#a forecasting function and h as attributes
tsCV_self = function(ts, fcfunction, h=1){
#creating a time series object
ts = as.ts(ts)
#defining the length of the output matrix
n = length(ts)
#constrction of the output matrix, in the first place filled
#with "NA" which will be overwritten by the function
m = ts(matrix(NA, nrow = n, ncol = h))
#Ensures that the final matirx has the same time series
#attributes like the given timeseries, (start, end, frequency)
tsp(m) = tsp(ts)
indx = seq(1, n - 1L)
for (i in indx) {
#defining the subset
ts.subset = subset(
ts, start = 1L, end = i)
#applying the forecast function successively on the subset
fc = fcfunction(ts.subset, h = h)
#making sure that no error has occured
if (!is.element("try-error", class(fc))) {
#Calculating the errors and writing them into the output matrix
m[i, ] = ts[i + (1:h)] - fc$mean
}
}
#if h=1 the output is given as a vector
if (h == 1) {
return(m[, 1L])
} else {
colnames(m) = paste("h=", 1:h, sep = "")
return(m)
}
}
t = tsCV_self(GDP.train, rwf, h = 2)
```

8 Regression analysis of first difference Log(GDP)

We define as depended variable the first log difference of the GDP data. The first log difference is approximately (for values <|0.1|) the same as percentage change between the two periods. In our case, this is Quarter-to-Quarter growth.

For this regression, we are using the logged difference of M3, the value of the money stock in the economy. Since the log only is likely to be stationary, we use the first difference. Second, we regress on the unemployment rate. Both data series are also obtained from FRED (2020c&d). As we have seasonal data, we also control for that by using seasonal dummies.
We deploy following regression model:

$$1.Diff\ log(GDP)_t = \beta_0 + \beta_1 * 1.Diff\ log(M3)_t + \beta_2 * Unemployment\ ratio_t + \beta_3 * Quarter_2 + \beta_4 * Quarter_3 + \beta_5 * Quarter_4 + \varepsilon_t$$

By running our model, we get following estimations:

Depended Variable:	1.Diff log(GDP)			
Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
constant	-0.010655	0.004133	-2.578	0.01150 *
1.diff Log(M3)	-0.335026	0.108003	-3.102	0.00255 **
Unemployment ratio	-0.160366	0.051160	-3.135	0.00230 **
2nd Quarter Dummy	0.065624	0.002294	28.605	<2e-16 ***
3th Quarter Dummy	0.041085	0.002290	17.941	<2e-16 ***
4th Quarter Dummy	0.058004	0.002310	25.109	<2e-16 ***

Signif. codes:	0 '***' 0.001 '**'	0.01 '*' 0.05 '.'	0.1 ' 1
Residual standard error:	0.007965 on 93 degrees of freedom		
Multiple R-Squared	0.9178	Adjusted R-squared:	0.9133
F-statistic:	207.6 on 5 and	93 DF, p-value:	<2.2e-16

Table 5: Regression on control variables

All our estimated coefficients in Table 5 are significant. The coefficient for 1.diff Log(M3) means, that if the **growth** of the money stock rises by one percent, we would expect GDP **growth** to change by -0.34 percent. So the relationship is significant, but tiny. We assume the negative sign is due to the extreme developments in 2008-2009, as the FED increased the money stock significantly (so maybe causality is reversed) in order to cope with the crisis.

The coefficient for the unemployment ratio is more clear. If the unemployment rate increases by 1 percent, we would expect GDP growth to fall by 16 percent. The major inference from the seasonal dummies is, that in the 1st quarter growth is significantly smaller than in the other ones.

Figure 16 indicates a small autocorrelation between the residuals but Shapiro-Wilk normality test yields a p-value of 0.2102, so we can assume normality for the residuals.

Figure 16: Residuals of log(GDP) regression

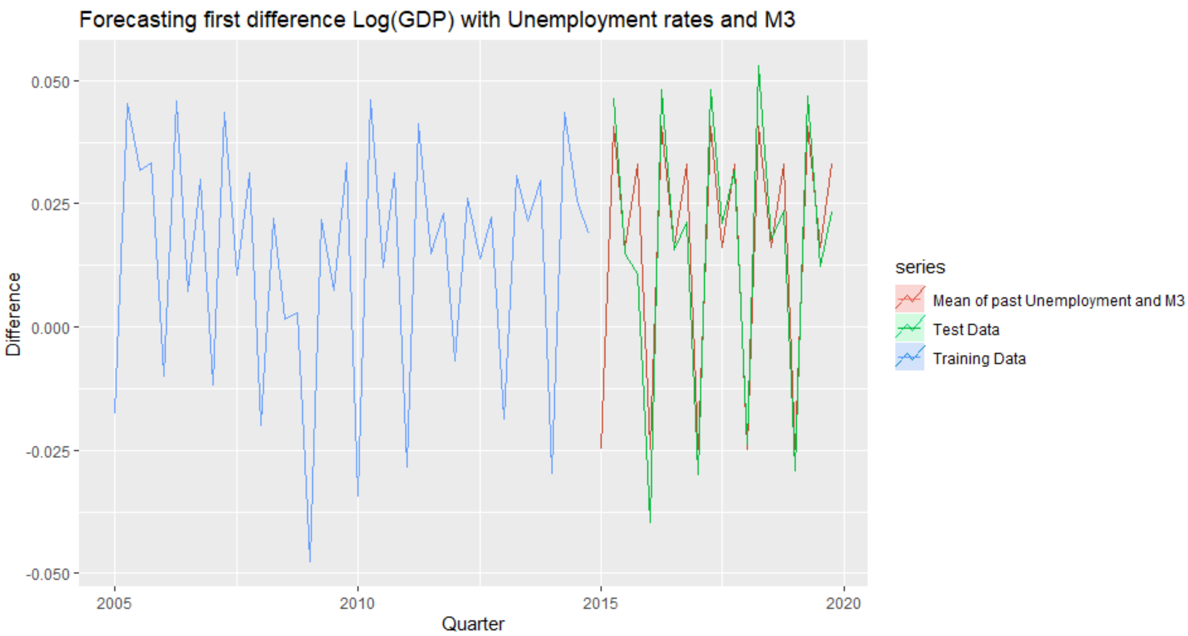

Figure 17: Forecasting based on regression results

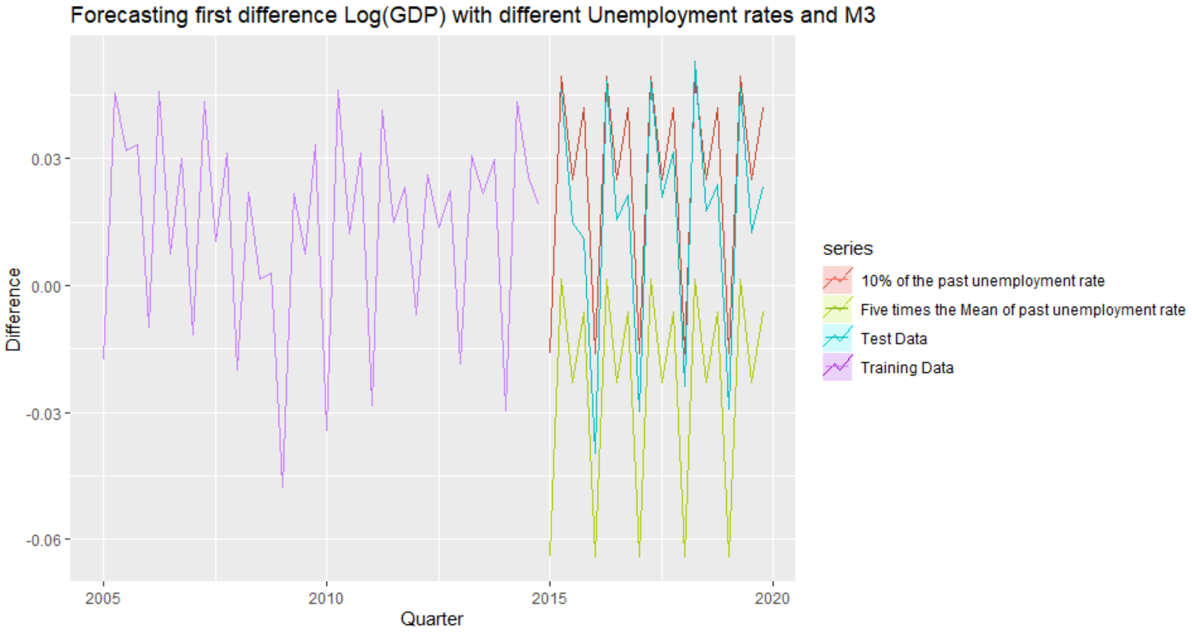

Figure 18: Forecasting based on regression results with diff. scenarios

Appendix

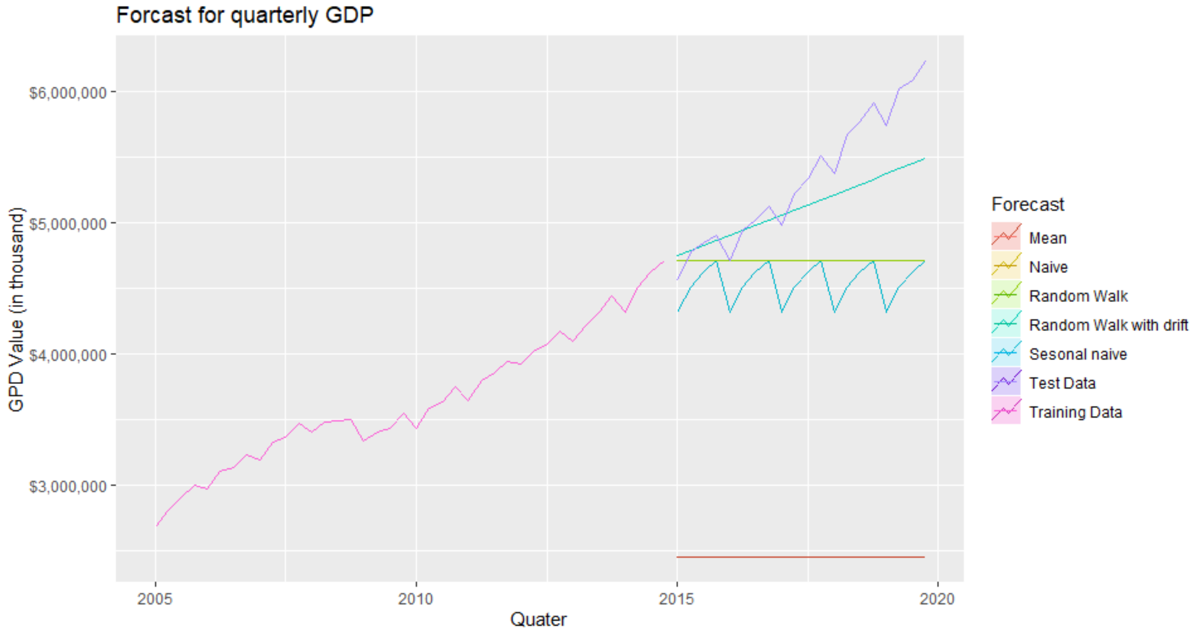

Figure 19: Forecasting with basic techniques

Figure 19 shows, that a random walk with a drift is best suited among the basic forecast methods to predict the data, hence we use this in the comparison in section 5.

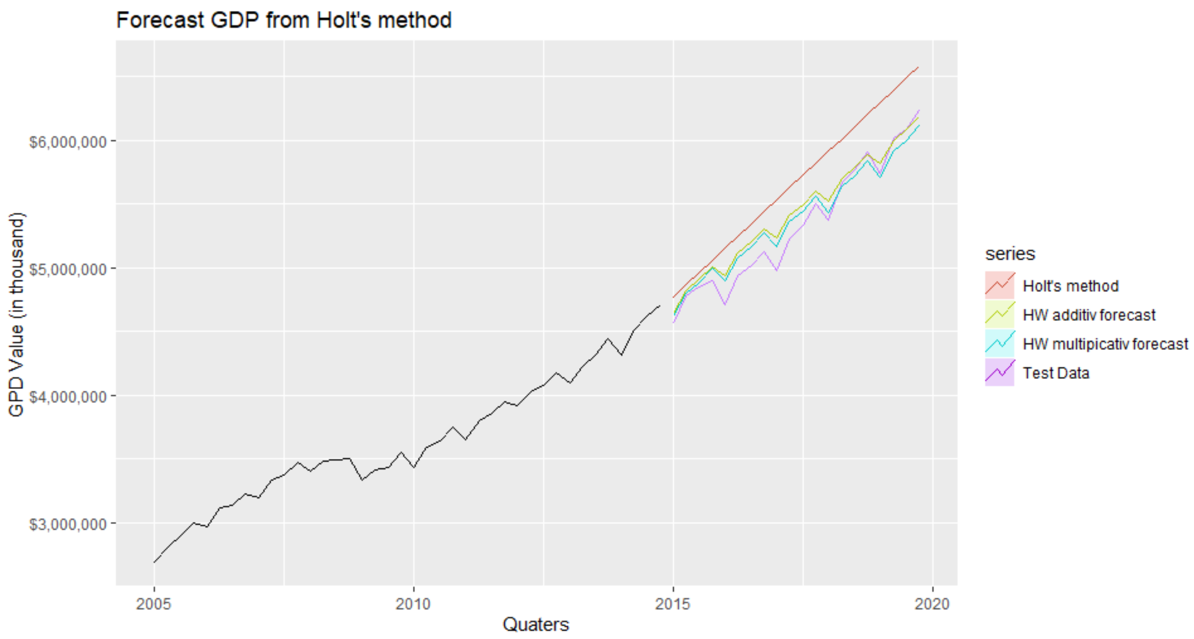

Figure 20: Forecasting by using Holts method

Since it is not graphically clear, whether Holt-Winter additive or Holt-Winter multiplicative is the more precise method, we use the Forecast errors to determine the superior method. Table 6 shows, that Holt-Winter multiplicative is the method with the lowest error terms, hence we use this method in the main analysis in section 5.

	ME	RMSE	MAE	MPE	MASE	ACF1
Holt's method	217021.16	352972.9	266359.18	3.6339	1.7317	0.7102
Holt addiditve	-97452	129810	105889	2.0745	0.6884	0.6618
HW multiplicativ forecast	-44684.35	105339.6	93083.56	-0.9723	0.6052	0.8227

Table 6: Comparison of Residuals of Holt methods

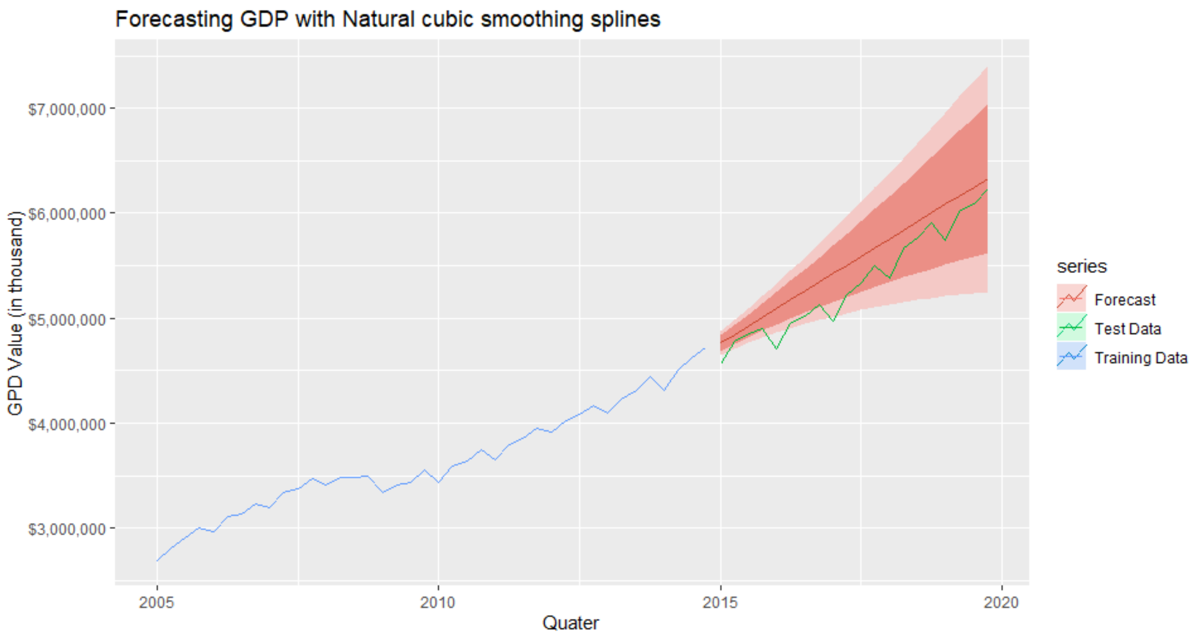

Figure 21: Spline Forecast

We used following packages in R:
"forecast", "ggplot2","ggpubr", "fma", "expsmooth", "fpp2", "seasonal", "mFilter", "tseries", "scales"

References

Board of Governors of the Federal Reserve System (2020a). Gross Domestic Product, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/. Accessed on 11.03.2020.

Board of Governors of the Federal Reserve System (2020b). Gross Domestic Product: Implicit Price Deflator, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/. Accessed on 11.03.2020.

Board of Governors of the Federal Reserve System (2020c). M3 for the United States, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/. Accessed on 11.03.2020..

Board of Governors of the Federal Reserve System (2020d). US Unemployment Rate, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/. Accessed on 11.03.2020.

Hyndman, R.J., & Athanasopoulos, G. (2018). Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on 15.03.2020.

Mankiw, N. G. (2018). Principles of Economics 7th edition, South-Western College Publishing: Cincinnati, USA.

Phillips, P.C., & Shi, Z. (2019). Boosting: Why you Can Use the HP Filter, Working paper, Cornell University.