2020 Exam 2: Forecasting quarterly Consumption Expenditures and Disposable Income of France

BAN430: Forecasting

NHH NORGES HANDELSHØYSKOLE

submitted in April 2020

Contents

List of Figures

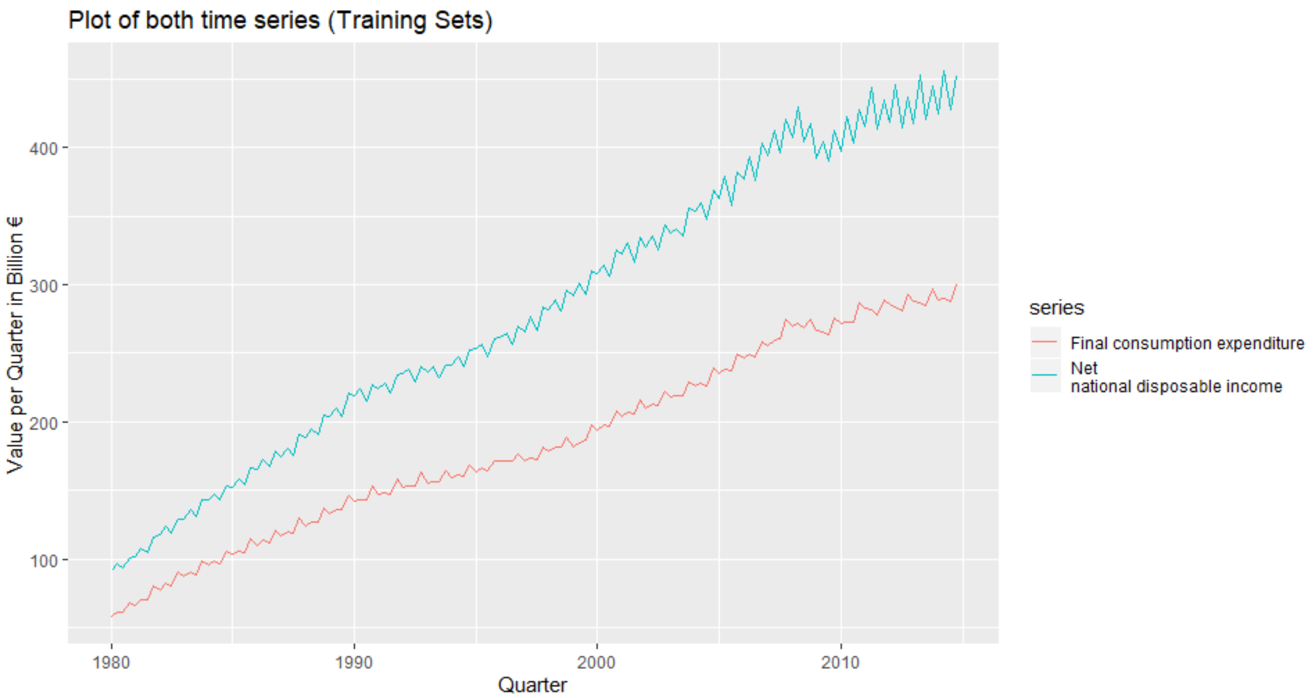
List of Tables

1 Data description

For this Exam, we use the time series of consumption expenditure and net disposable income of France. We have obtained our data set from OCED (2020), which covers the first quarter of 1980 to the third quarter of 2019. We have defined the training sets up to the fourth quarter of 2014 so that we have 19 quarters available for the test sets.

The training sets for both series are shown in Figure 1. Both charts are tilted upwards, but the disposable income graph is steeper. Quarterly volatility is clearly visible, indicating seasonality. In later periods, volatility appears to increase in both graphs. Disposable income only went through a downturn after the 2008 crisis. Consumption, on the contrary, did not experience such a downturn, but its growth rate seems to flatten out afterwards.

Figure 1: Quarterly development over time



Another indicator for trends are given in the descriptive statistics given in Table 1. For both series, minimum and maximum values are at the extreme points of the training sets. Additionally, 1st and 3rd quantiles and the median are fairly evenly distributed in between. Putting these information together, we can assume that no extreme outline disturbs the growth trends for these time series.

Table 1: Summarize statistics of quarterly real values

		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Consumption	Date	1980-Q1	1989-Q1	1997-Q2	calculated	2006-Q3	2014-Q4
	Value in Million €	57.85	132.04	174.95	184.67	247.83	299.96
Disposable Income	Date	1980-Q1	1989-Q1	1996-Q4	calculated	2005-Q4	2014-Q2
	Value in Million €	91.81	201.79	272.86	281.64	379.63	456.22

Tests for stationarity and seasonality

In order to test the trend and seasonality of *Consumption* and *Disposable Income*, we use a simple regression with them as endogenous variables:

$$Y_{i,t} = \beta_{i,0} + \beta_{i,1} * trend_{i,t} + \beta_{i,2} * Quarter_2 + \beta_{i,3} * Quarter_3 + \beta_{i,3} * Quarter_4 + \varepsilon_{i,t}$$

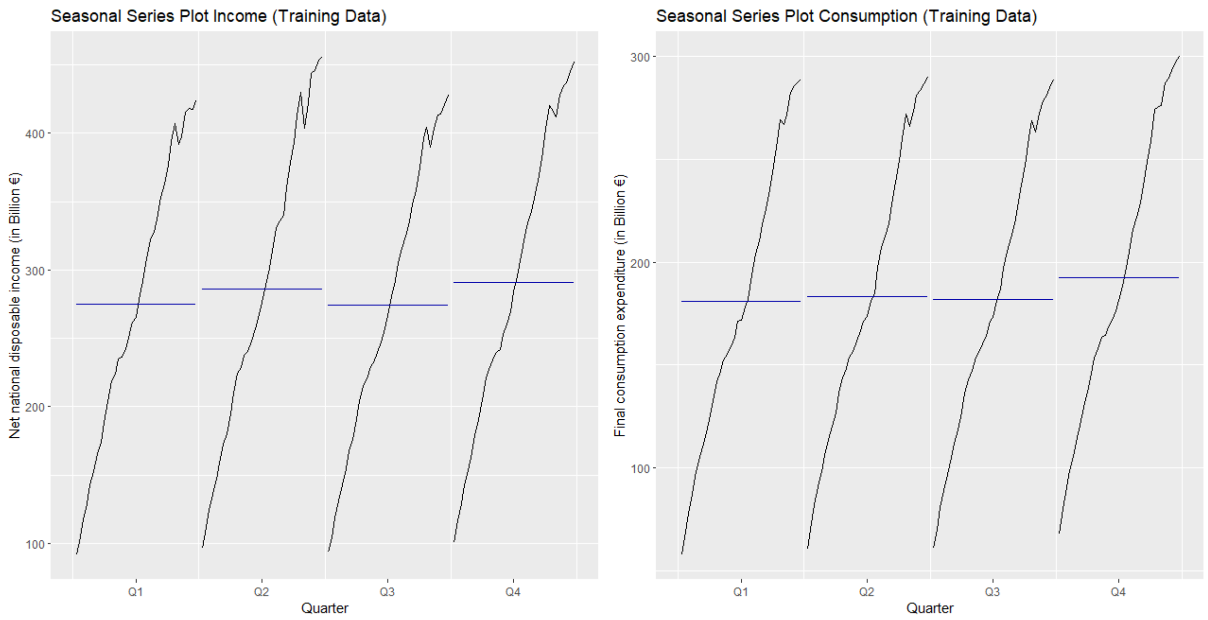
The regression results are presented in Table 2. The trend is very significant for both consumption and disposable income. For disposable income there is a significant seasonality between the first and all other quarters. For the time series of consumption we have significant seasonality only between the first and the last quarter. Thus, the time series of disposable income has a greater volatility throughout the year.

Table 2: Regression on trend and quarterly dummies

	Dependent variable:	
	Consumption	*Disposable Income*
	(1)	(2)
Constant	64.71829***	96.98645***
	(1.33116)	(2.11948)
Trend	1.68641***	2.58290***
	(0.01253)	(0.01995)
2nd Quarter Dummy	0.44033	8.41444***
	(1.43161)	(2.27941)
3th Quarter Dummy	−2.32802	−6.09254**
	(1.43178)	(2.27968)
4th Quarter Dummy	6.13162***	7.91307***
	(1.43205)	(2.28011)
Observations	140	140
R^2	0.9926	0.992
Adjusted R^2	0.9924	0.9918
Residual Std. Error	5.989 (df = 135)	9.535 (df = 135)
F Statistic	4549*** (df = 4; 135)	4211*** (df = 4; 135)
Note:		*p<0.1; **p<0.05; ***p<0.01

Another indication of the higher seasonality of income throughout the year is shown in Figure 2. The mean values of quarters 2 and 4 are significantly higher than those of the first and third quarters. With regard to consumption, only the fourth quarter differs significantly from the others, which can be explained by Christmas spending.

Figure 2: Development per quarter over time



Conclusions about the stationarity of the time series are obtained by applying the Augmented Dickey Fuller and the KPSS (Level) test. In these tests, it should be noted that the Null-hypotheses are defined oppositely.

Augmented Dickey-Fuller:
H_0: no stationarity

KPSS (Level) test:
H_0: stationarity

Table 3: P-values of stationarity tests

	Augmented Dickey-Fuller	KPSS Level
Consumption	0.5152	0.01
Disposable Income	0.3725	0.01

Given the p-values in Table 3, we can reject the null hypotheses for both KPSS tests, whereas this is not possible for the two augmented Dickey-Fuller tests. Based on this information we can infer that the series for consumption and disposable income are both non-stationary.

Autocorrelation pattern

The partial autocorrelation function in Figure 3 shows that for consumption, the first delayed quarter has a strong direct influence on the next, but not on the following one. Due to the high direct correlation of the first lag, the indirect effects are significant up to 40 delays, as the autocorrelation function shows. We would describe this process as an AR(1) with high persistence.

Figure 3: Autocorrelation functions for consumption

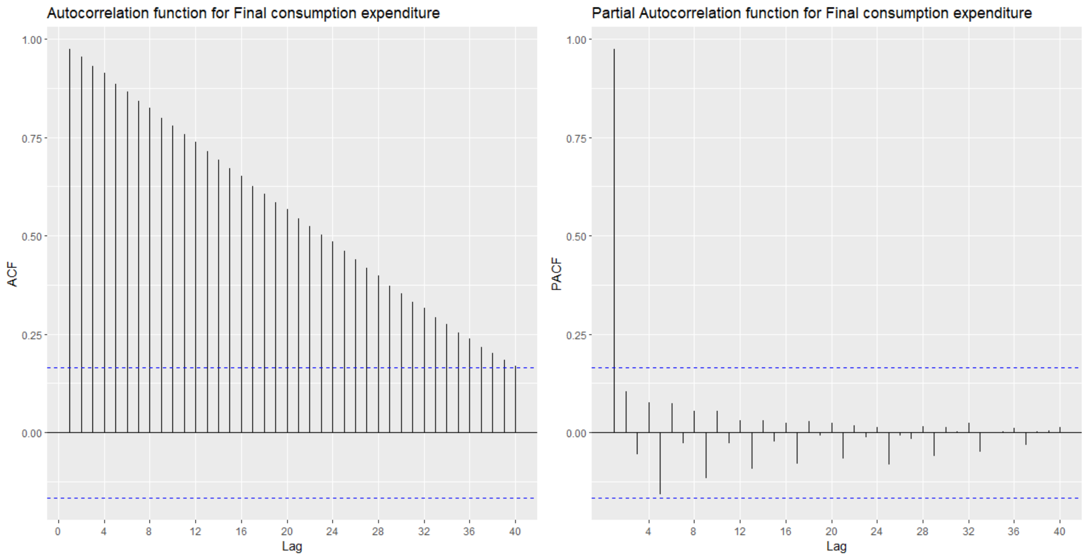
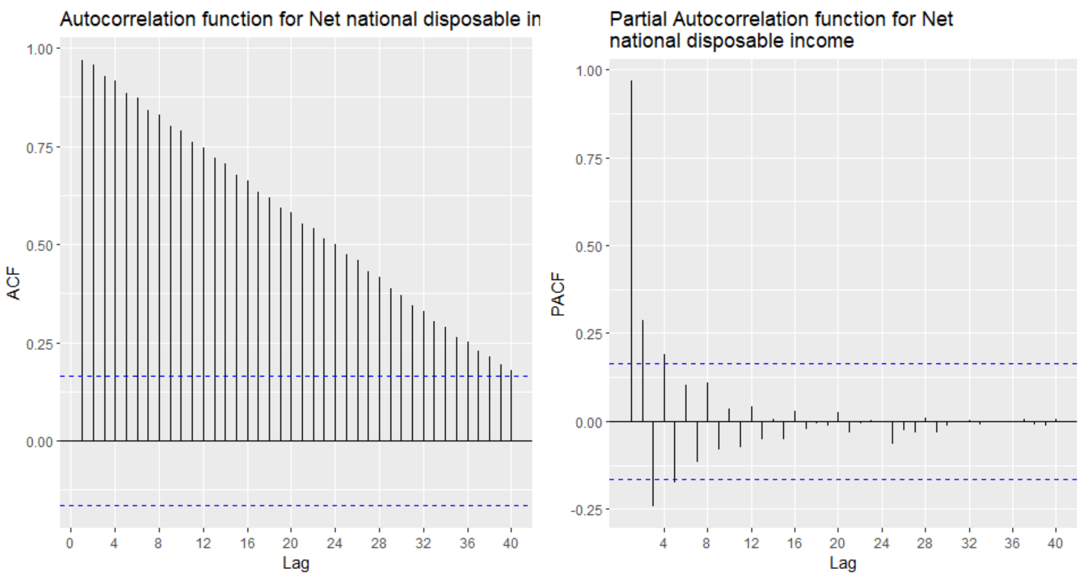


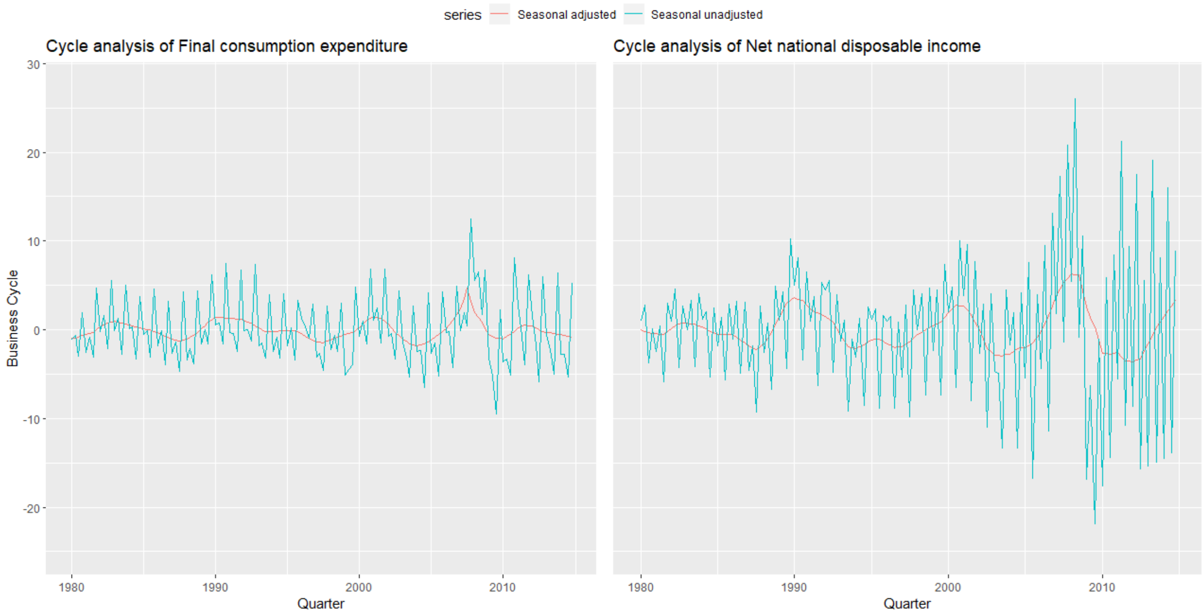
Figure 4 shows that the partial autocorrelation function is significant for disposable income up to the fifth lag, so these five lags have a direct impact on the first quarter. Similar to the autocorrelation function for consumption, the autocorrelation function for income falls very slowly, indicating a persistent pattern.

Figure 4: Autocorrelation functions for disposable income



Beyond intra-year seasonality, disposable income also has a higher volatility around its long-term trend than consumption. To determine this statement, which we obtain from Figure 5, we used the Hodrick-Prescott filter. The diagram on the left shows that consumption deviates only slightly from its long-term trend. Even in 2008-2009, during the financial crisis, consumption was almost at the level of its long-term trend, but starting from a strong pre-crisis development. Income, on the other hand, performed significantly weaker than its long-term trend during this period.

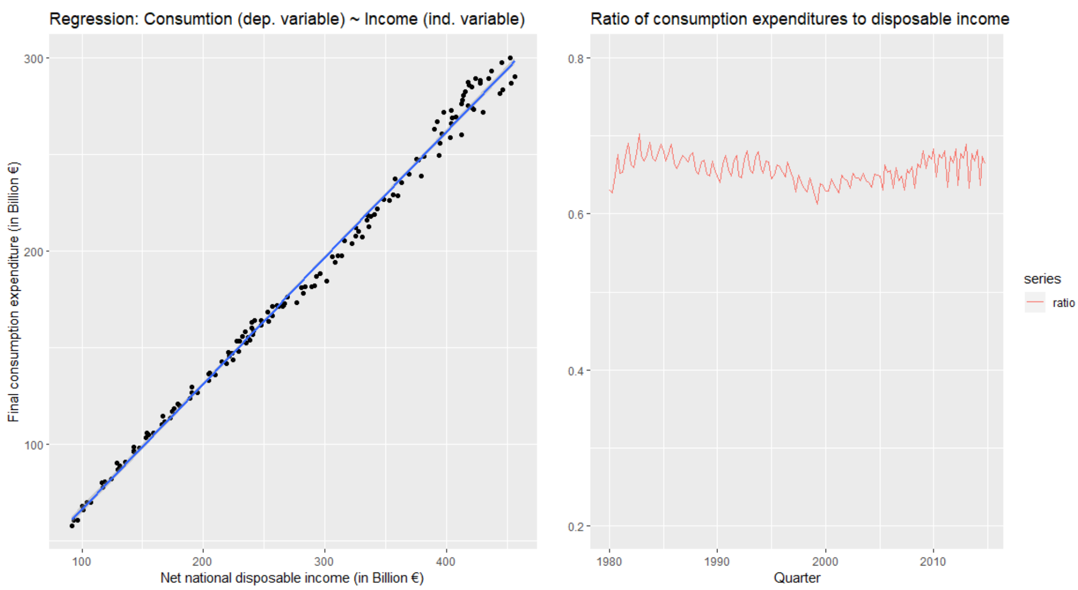
Figure 5: Cycle analysis



2 Long run relationship between consumption expenditure and disposable income

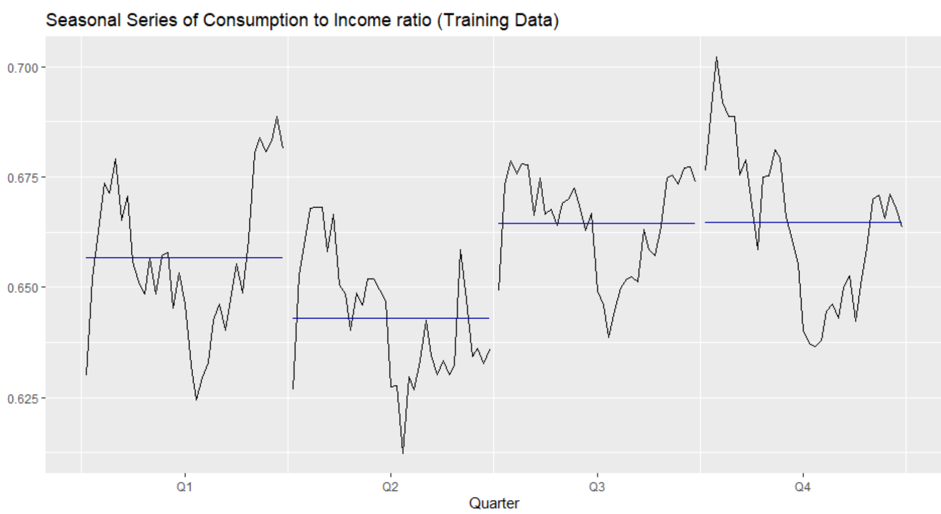
In general, both series have a fairly stable relationship over the entire training period, as one can see in Figure 6. The ratio of consumer spending to disposable income ranges from 61% to 70%. The ratio is highest in the early 1980s and has its weakest phase around the turn of the millennium. The fact that the median and average are almost equal at 65.72% and 65.73% indicates the absence of outliers in the series. In addition, the standard deviation of 1.76% is rather small in consideration of the mean value.

Figure 6: Long-term relationship of consumption expenditure and disposable income



Furthermore, the ratio itself has a seasonality. As Figure 7 shows, the second quarter is considerably weaker than the other ones.

Figure 7: Ratio of consumption expenditure and disposable income per quarter



3 SARIMA model for consumption expenditures

We use the *R* function *auto.arima()* to determine the best ARIMA model. This function uses the model with the lowest information criteria. When looking at the results in Table 4, the function selects an ARIMA(2,0,0)(0,1,1)[4] with a drift, which is a model with a seasonal addition. Such a model is also known as SARIMA.
The seasonal component and drift were expected, as we have already detected seasonality and trend in the previous sections. Seasonality is repeated with a frequency of 4, which is not particularly striking for quarterly data. In contrast, the AR(2) process in the main part is somehow surprising, since the PACF of this series showed only a significant delay[1].
Consumption expenditure is not stationary, so we have to integrate the series, which is carried out in the seasonal part.

Table 4: SARIMA model for consumption expenditures

	ARIMA(2,0,0)(0,1,1)[4] with drift	
	Coefficients	*Standard Errors*
AR(1)	1.1120	0.0841
AR(2)	-0.1647	0.0861
SMA(1)	-0.6273	0.0921
Drift	1.6779	0.2202
σ^2	2.435	
Log likelihood	-253.05	
AIC	516.1	
AICc	516.56	
BIC	530.66	

Table 5 shows tests for autocorrelation of the residuals of this model.

Table 5: Autocorrelation tests of SARIMA model residuals

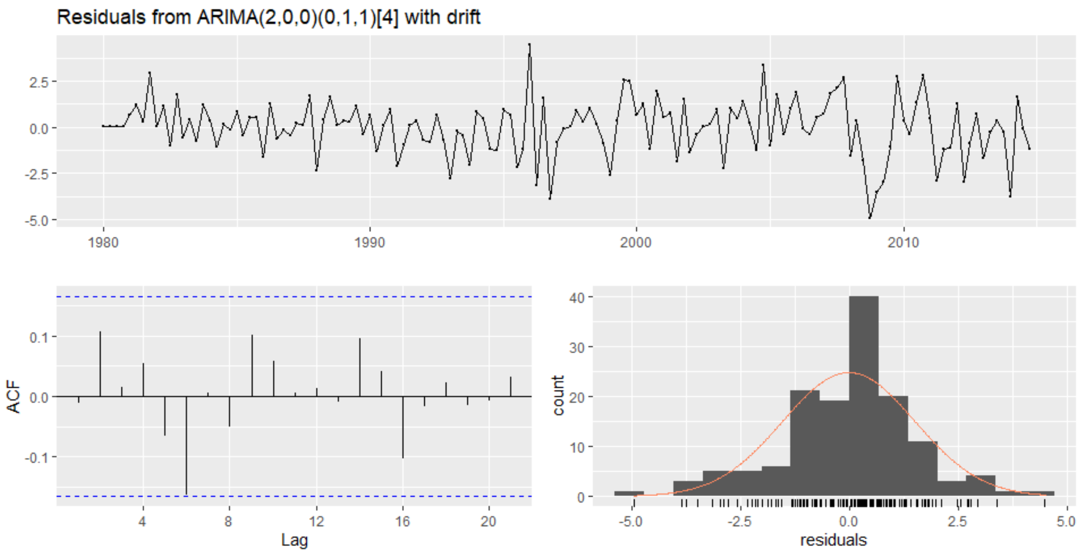
	χ^2	Degrees of freedom	p-value
Box-Pierce test	8.6492	10	0.5657
Box-Ljung test	9.1597	10	0.517

Since both test have p-values above any critical threshold, we can keep the Null-Hypothesis of independent distribution, hence no autocorrelation. This is supported by the ACF

[1] However the second lag is only to the 10% and not the 5% level significant

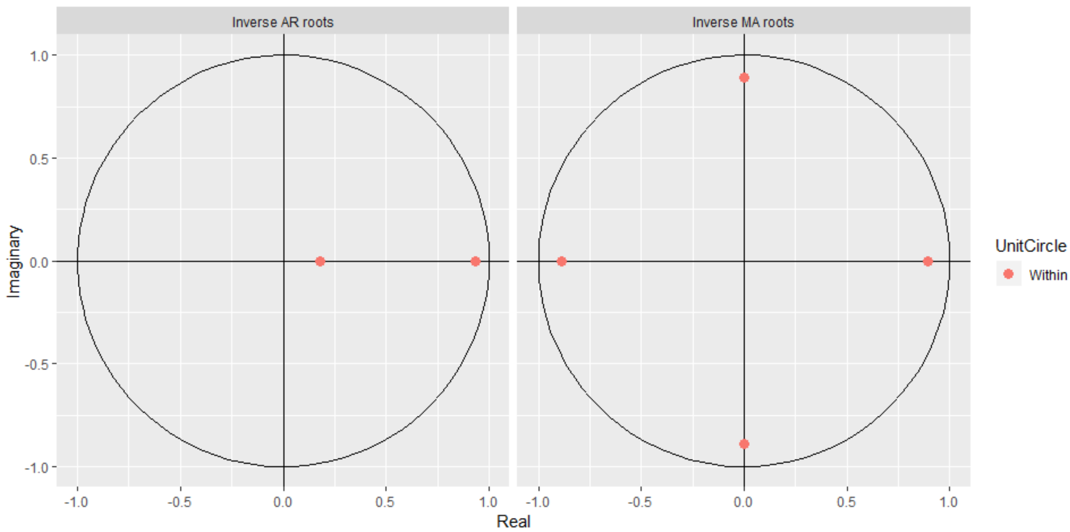
function shown in Figure 8, which does not show a recognizable pattern. The density plot seems slightly left skewed. However, the Shapiro-Wilk test for normality has a p-value of 0.06953. Consequently, we can just about maintain H0 for normally distributed residuals. Based on the ACF and the normality result, we conclude white noise for the residuals.

Figure 8: Residual inspection of SARIMA model for consumption



All the eigenvalues are inside the unit circle in Figure 9, therefore the model is stable. There are no common roots. Due to the eigenvalues inside the circle, the model is invertible.

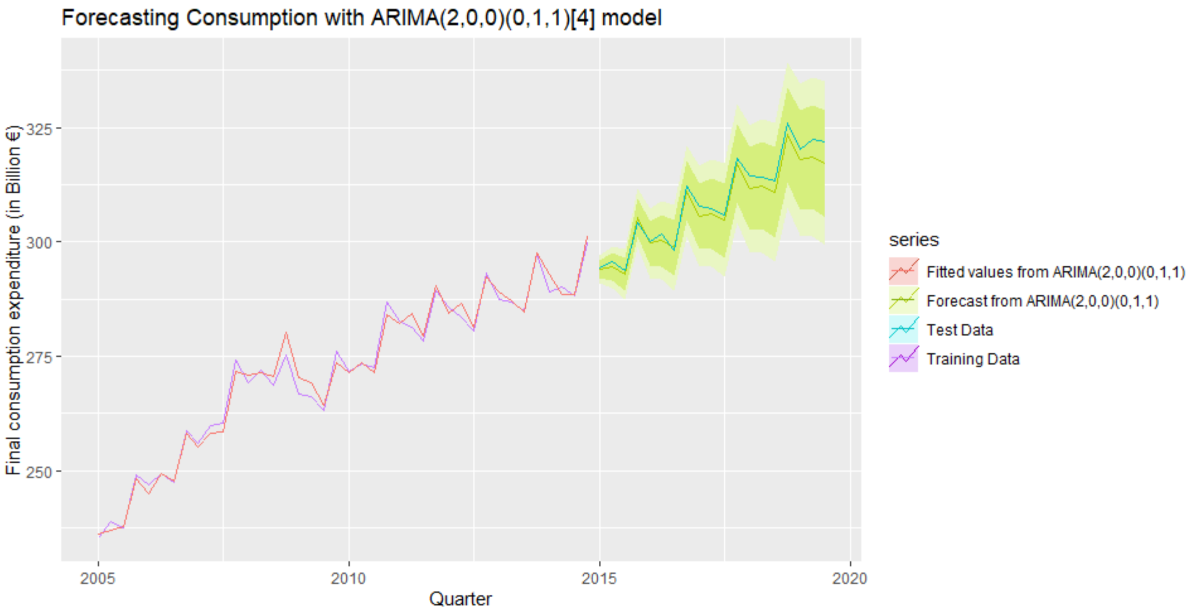
Figure 9: Unit root inspection of SARIMA model



Forecast with the SARIMA model

Figure 10 shows that the model can predict the data from the test set reasonably well. However, the prediction is systematically somewhat too low. Nevertheless, the test data lies well within the confidence intervals. In previous periods of the training set, the model tended to underestimate the data.

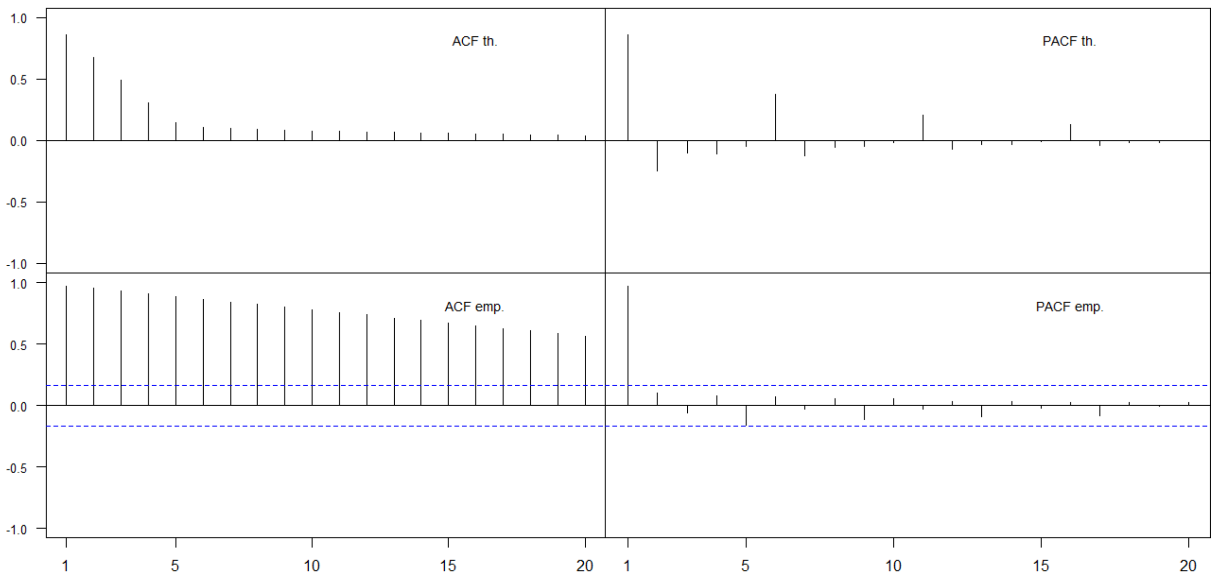
Figure 10: Forecast with SARIMA model



4 Comparison of autocorrelation functions

~~We use the *plotacfthemp()* function - which is based on *ARMAacf()* - of *R* in order to account the seasonal aspect when modeling the Autocorrelation and Partial Autocorrelation function of our SARIMA model.~~ These functions are plotted in the first row of Figure 11. The second row shows the ~~actual~~ ACF and PACF of consumption, similar to Figure 3. By comparing both, one can see that the model is performs reasonable capturing the direct of effects. Both PACF have a clear significant result for the first lag. However, the ACF shows that the SARIMA model does not capture all indirect dynamics of past lags, since the theoretical ACF levels off to fast. ~~Therefore,~~

Figure 11: Autocorrelation functions for consumption and its SARIMA model



5 SARIMAX model for consumption expenditures

In this section we will again use the *R* function *auto.arima()* to determine the best model. In addition to the last time, we include disposable income as an exogenous variable. ARIMA models extended by exogenous variables are also called ARIMAX. Table 6 gives an overview of the model chosen by the function *auto.arima()*. Again, the model contains a seasonal component with a frequency of 4 (quarterly data). This class of models is analogously called SARIMAX.

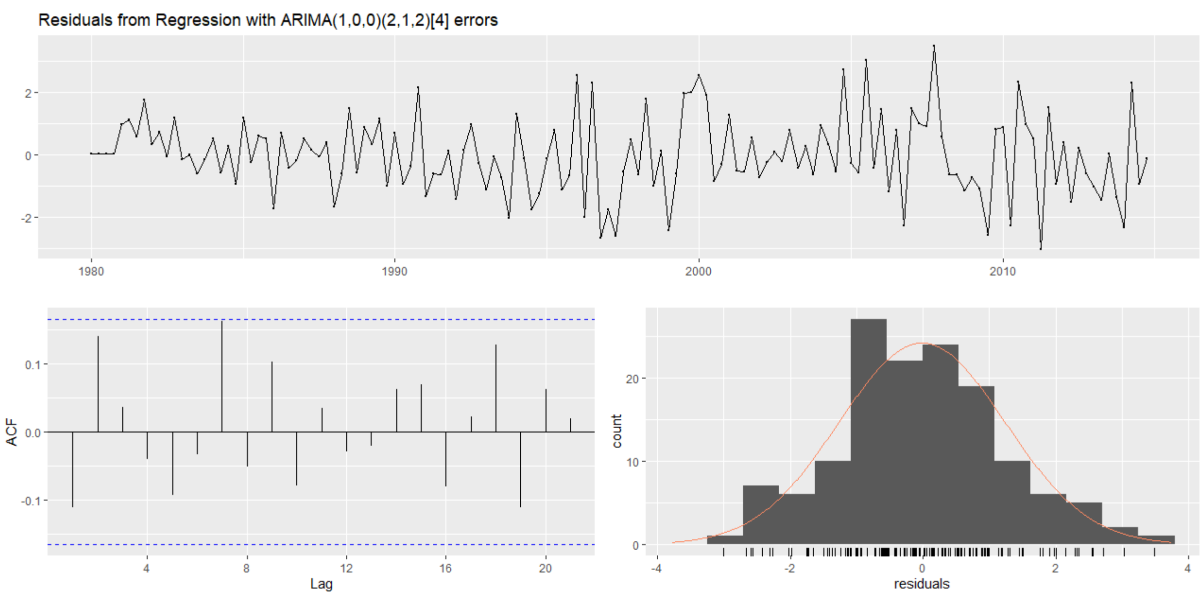
In section 2, we found that the relationship between consumption expenditure and disposable income follows a certain seasonality. Therefore, we have to compensate for additional seasonality in the processes. This may explain why the SARIMAX model has seasonal AR(2) and MA(2) processes in the main part and only one ordinary AR(1) process, whereas the SARIMA model has only one seasonal AR(1) process but an AR(2) process in the main part. The AIC of the SARIMAX model, however, is better at 473 than the AIC of SARIMA, which has a value of 516.

Table 6: SARIMAX model for consumption expenditures

	ARIMA(1,0,0)(2,1,2)[4] errors	
	Coefficients	*Standard Errors*
AR(1)	0.7920	0.0628
SAR(1)	-0.7629	0.0976
SAR(2)	-0.7994	0.1017
SMA(1)	0.5295	0.0886
SMA(2)	0.8997	0.1763
Drift	0.8953	0.1477
Disposable income	0.3136	0.0338
σ^2	1.687	
Log likelihood	-228.5	
AIC	473.01	
AICc	474.14	
BIC	496.31	

In Figure 12, we inspect the residuals of this regression model. The ACF function in this figure has is more persistent than in the SARIMA model, but still without significant lags. The density plot is marginally left skewed.

Figure 12: Residual inspection of SARIMA model for consumption



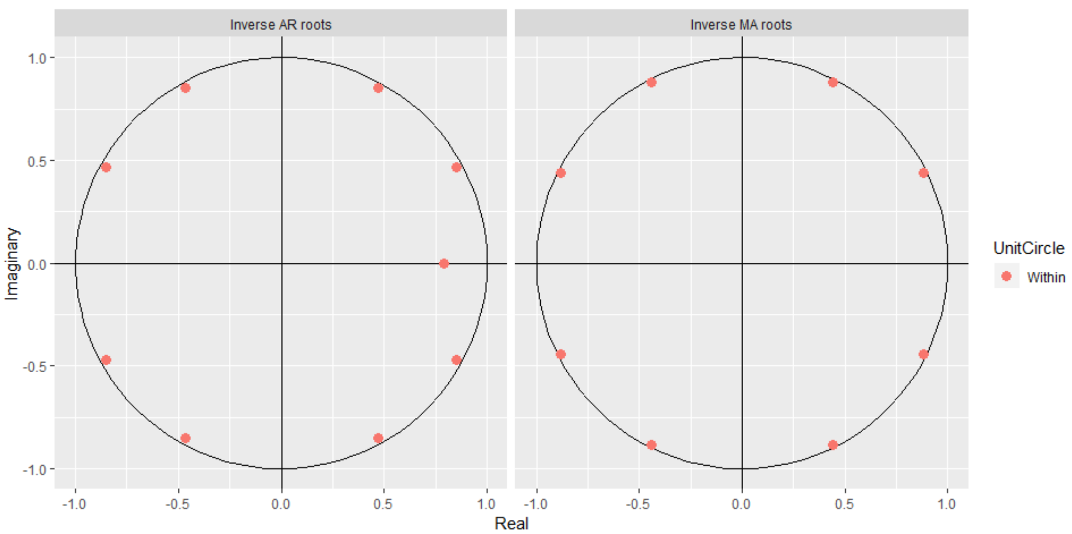
Both test in Table 7 have p-values above the common critical thresholds, hence we can keep the Null-Hypothesis of independent distribution, which implies no serial correlation of the residuals.

Table 7: Autocorrelation tests of SARIMAX residuals

	χ^2	Degrees of freedom	p-value
Box-Pierce test	12.745	10	0.2383
Box-Ljung test	13.438	10	0.2002

The Shapiro-Wilk normality test has a p-value of 0.5779, so we can retain the null hypothesis and assume normality for the residuals. Based on the ACF and the normality result, we conclude white noise for the residuals.

Figure 13: Unit root inspection of SARIMAX model



All the eigenvalues are inside the unit circle shown in Figure 13, therefore the model is stable. There are no common roots. Most roots are at the margin of the circle but still inside, hence they are invertible.

Forecast with the SARIMAX model

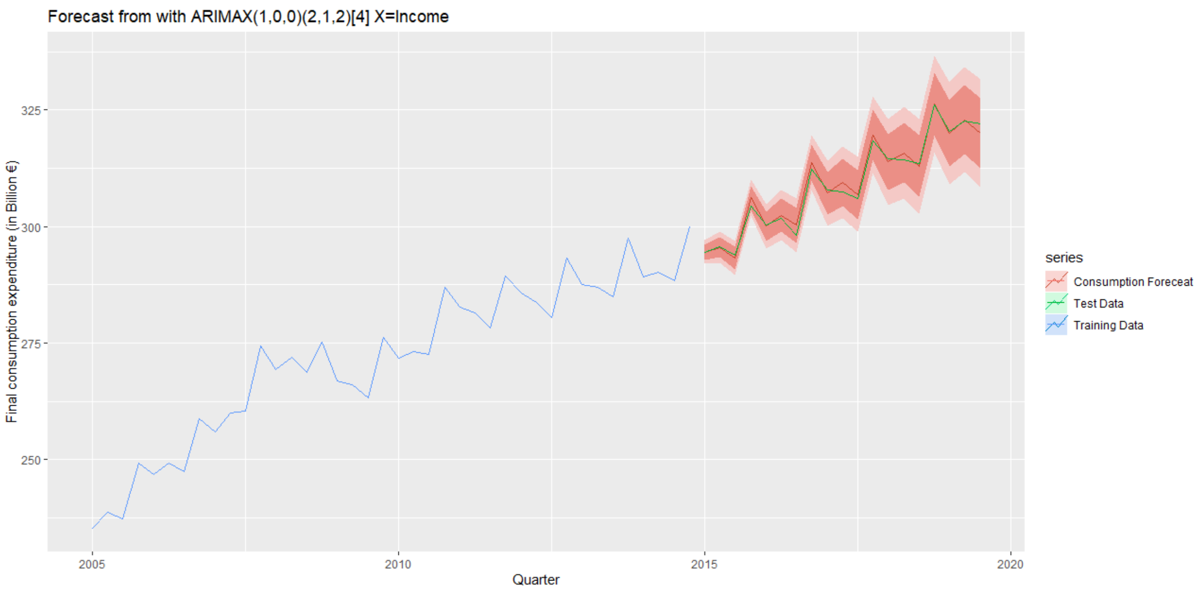
To be consistent, for the SARIMAX equation we should use a forecast rather than the test data of disposable income. This forecast for the income variable requires forecasts ~~at the point t~~ for the period t to $t+h$. To get them, we use the function *forecast()* in *R*. The function selects an ETS model with the best Akine information criteria. Based on our training set of data, the function selects an ETS(M,A,M) model, which is a multiplicative Holt-Winters method with multiplicative errors.

In detail, we get the following estimates for the smoothing parameters:
$\hat{\alpha} = 0.7335, \quad \hat{\beta} = 0.0065, \quad \hat{\gamma} = 0.2665$ which yields following model:

$$y_t = (l_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t) \qquad l_t = (l_{t-1} + b_{t-1})(1 + 0.7335\varepsilon_t)$$
$$b_t = b_{t-1} + 0.0065(l_{t-1} + b_{t_1})\varepsilon_t \qquad s_t = s_{t-m}(1 + 0.2665\varepsilon_t)$$

With this model we obtain a forecast for the the disposable income variable, which in turn is used in the SARIMAX forecast pictured in Figure 14. Our SARIMAX forecast marginally overestimates the actual data, however, this data lies well within the confidence intervals

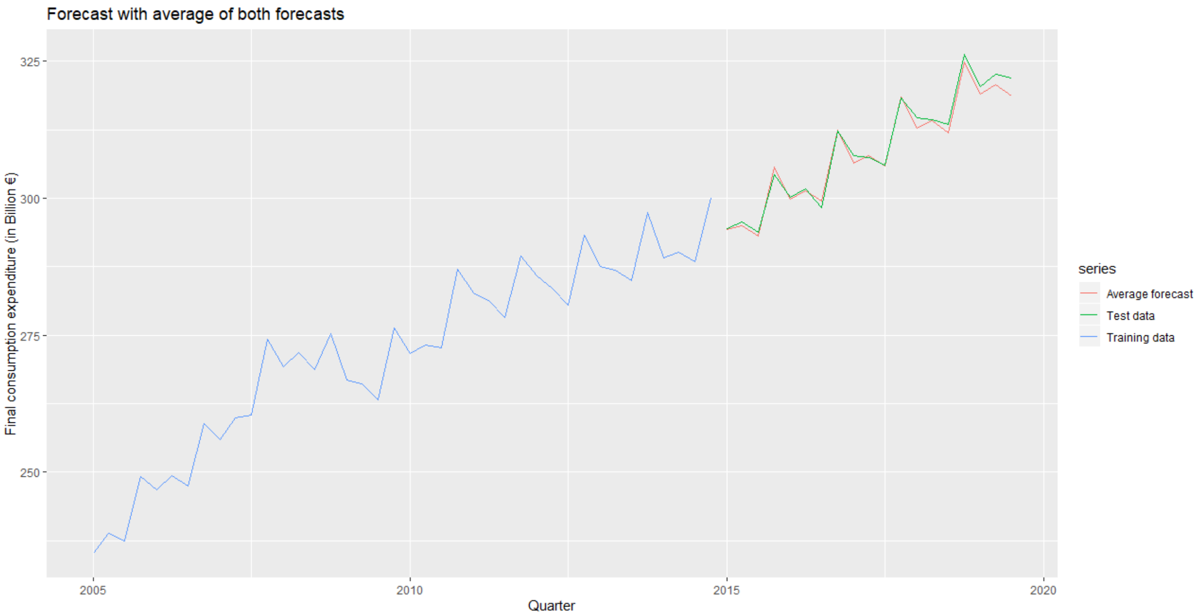
Figure 14: Forecast with SARIMAX model



6 Forecast comparison and merging

When the SARIMA and SARIMAX models are combined equally to predict the data, the combined prediction slightly underestimates the test values, as it can be observed in Figure 15. This is mainly based on the SARIMA model, which underestimates the data more than the SARIMAX model overestimates them.

Figure 15: Forecast with average weights



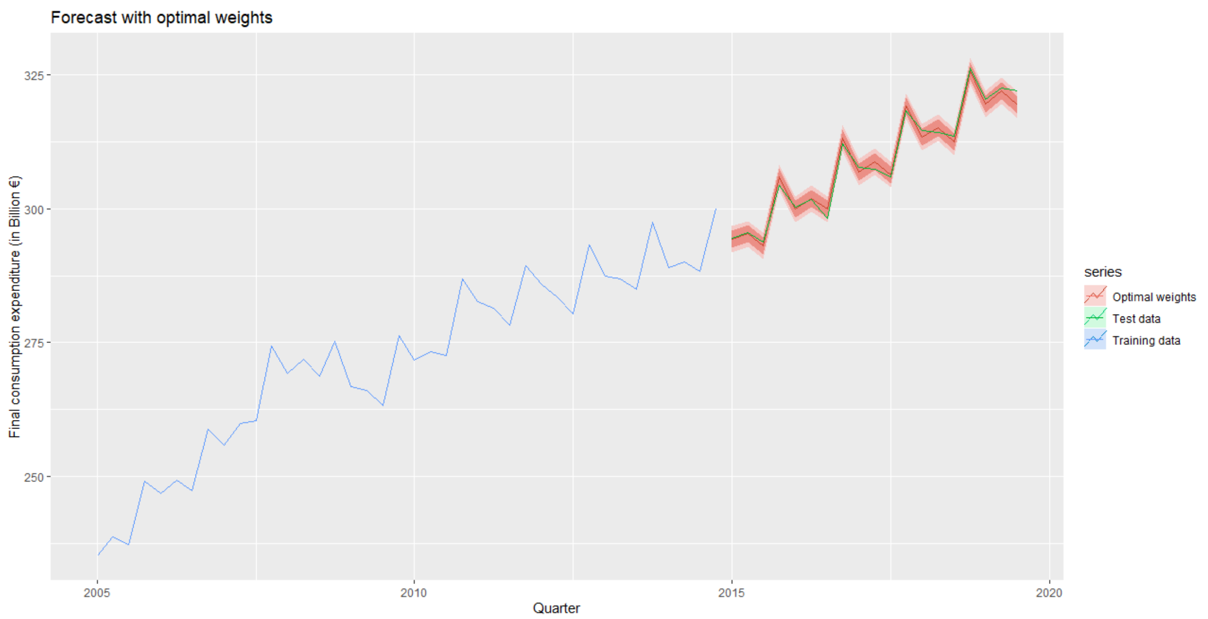
To determine the optimal weights, we regress the consumption expenditures of the trainings set on the fitted values of the SARIMAX and SARIMA for those period. Based on the results in Table 8, we weight the SARIMAX at 81.2% and the SARIMA at 18.8%.

Table 8: Optimal weights regression

	Dependent variable:
	Consumption expenditures trainings set
SARIMAX on traings set	0.812***
	(0.096)
SARIMA on traings set	0.188*
	(0.096)
Observations	140
R^2	1.000
Adjusted R^2	1.000
Residual Std. Error	1.238 (df = 138)
F Statistic	1,771,949.000*** (df = 2; 138)
Note:	*p<0.1; **p<0.05; ***p<0.01

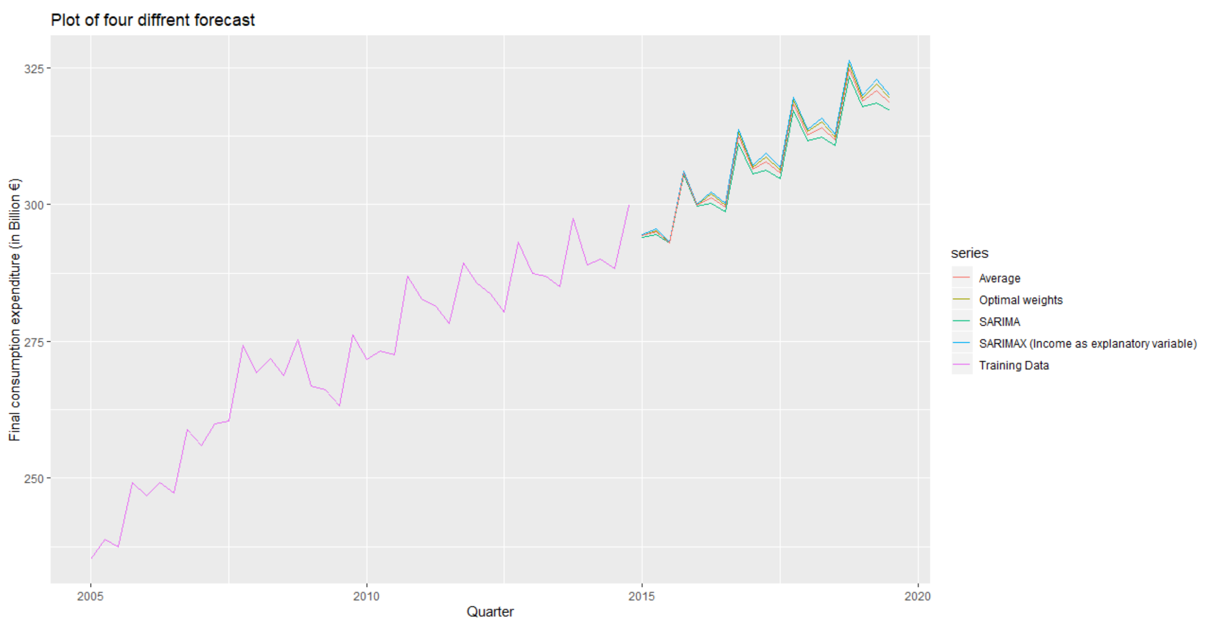
An application of this weighting to the forecast of consumption expenditure is shown in Figure 16. The optimal weighting forecast performs very well for the important fourth quarter, but slightly worse for the quarters in the middle of the year. These, however, lie well between the confidence intervals.

Figure 16: Forecast with optimal weights



Plotting all four forecast variants together, one can observe graphically in Figure 17 that SARIMA is the furthest away from the actual data. The second least accurate forecast is the average weighted forecast. The SARIMAX and optimally weighted forecast are graphically hardly distinguishable from the test series.

Figure 17: Forecasts overview



These visual statements are supported by the out-of-sample error terms in Table 9, which show that the prediction weighted by optimal in-sample performance is also the best prediction for the out-of-sample, given a root mean squard error (RMSE) of 1.0665 . Similarly, the SARIMA model is the one with the weakest performance, since it only achieves a RMSE of 2.1190.

Table 9: Comparison of Residuals

	ME	RMSE	MAE	MPE	MASE	ACF1
SARIMA	1.5153	2.1190	1.5153	2.1190	2.1190	1.5153
SARIMAX	-0.3839	1.1448	0.9186	-0.1261	0.1318	-0.1755
Average weights	1.2756	1.2756	1.2756	1.2756	1.2756	1.2756
Optimal weights	0.0732	1.0665	0.8853	0.0202	0.1270	-0.0457

References

OECD (2020). "Quarterly National Accounts". Paris: Organization for Economic co-operation and Development [referred: 25.3.2020].
Access method: https://stats.oecd.org/Index.aspx?DataSetCode=QNA.

We used following packages in *R*:
"forecast", "ggplot2","ggpubr", "fma", "expsmooth", "fpp2", "seasonal", "mFilter", "tseries", "scales", "stargazer", "urca", "GGally", "caschrono"