Abstract

The purpose of this project is to determine the best possible model for a time series dataset in order to model the in-sample and forecast the out-of-sample observations of percent change of total consumer credit in the United States. To do so, we entailed into some descriptive analysis to observe how the dataset looked, for the first milestone. Milestone 2 focuses on further analysis of the time series of interest after some adjustments, which helps choose a model to fit the in-sample observations. Finally, on the third milestone we were able to set a prediction of the rest of the observations throughout dynamic and static forecasting and compare the goodness of fit among the models. All in all, after analyzing how our variable of interest behaved, we were able to define several models depending on different criteria and produced forecasts for them, which permitted us to further comprehend how forecasting and time series work.

Milestone 1 | Description of the Time Series

Description and Motivation

The *Percent Change of Total Consumer Credit* statistical release reports outstanding credit extended to individuals for household, family, and other personal expenditures, excluding loans secured by real estate. Total consumer credit comprises two major types: revolving and nonrevolving. Revolving credit plans may be unsecured or secured by collateral and allow a consumer to borrow up to a prearranged limit and repay the debt in one or more installments. Credit card loans comprise most of revolving consumer credit measured in the data, but other types, such as prearranged overdraft plans, are also included. Nonrevolving credit is closed-end credit extended to consumers that is repaid on a prearranged repayment schedule and may be secured or unsecured. To borrow additional funds, the consumer must enter into an additional contract with the lender. Consumer motor vehicle and education loans comprise the majority of nonrevolving credit, but other loan types, such as boat loans, recreational vehicle loans, and personal loans, are also included.

It also reports selected terms of credit, including interest rates on new car loans, personal loans, and credit card plans at commercial banks. Historically, the dataset also included series that measure the terms of credit for motor vehicle loans at finance companies. In the first quarter of 2011, publication of these series was temporarily suspended because of the deterioration of their statistical foundation, but it was eventually resumed and data for that period is currently available.

It has been calculated using the percent change of the seasonally adjusted annual rate, which means that they have removed the predictable seasonal patterns.

This allows us to observe the percentage change of total consumer credit from month to month, which is useful to comprehend the behavior of consumers.

Distributional Properties

If we perform the ADF test at a 95% confidence level, we observe a p-value of 0.01, thus we reject the null hypothesis that all autocorrelations equal 0 and conclude that the time series is stationary. In other words, it doesn't have some time-dependent structure and has constant variance over time.

Through the JarqueBera test, we can say that the time series doesn't follow a normal distribution. Rejecting the null hypothesis of normality, since the value of the test is 0. According to the QQ-plot and Kernel graphics (appendix 1), the time series follows a distribution similar to that of a Laplace's. Where the distribution is over-dispersed in relation to the normal distribution, characterized by having a large number of atypical values, fatter tails (in this case a minimum value of -18.19 and maximum of 26.92, with values for the first and third quartiles being 4 and 10.6, respectively) and having positive kurtosis of 4.09, above the normal kurtosis value of 3. On the other hand, through descriptive statistics, we see that time series has a mean of 7.15 which is higher than the median of 6.66. This defines that the distribution is not completely symmetrical, although it has a fairly low skewness value of 0.11, meaning that it's slightly fat-tailed to the right.

Dynamic Properties

If we look at the ACF (appendix 2) of our data we see a big dependence through several lags, with autocorrelation values of up to 65.8% for the second period. Moreover, the autocorrelogram shows significance for the first 20 lags. It's also quite interesting to see that there is autocorrelation for lots of values decreasing slowly.

The PACF (appendix 2) doesn't hold significance for as many lags as with the ACF, but we can also note a huge autocorrelation on the first lag, even on the second one, although the partial autocorrelations quickly drop and the values are not significant anymore by the fifth lag. If we look at the graph of the data (appendix 1), we are able to see how there is much more variance the further back in time we go, which initially could make us think that reducing our sample would yield even higher values of autocorrelation, but a deeper study of this hypothesis doesn't yield a relevant difference, as the ACF and PACF are relatively similar than the ones using data from 1955.

The ACF and PACF are useful to determine which models we'll be using next to model and forecast our data. Seeing how they both behave it seems that using an AR or ARMA model would be fairly good options, because that the ACF shows a gradual decline in its values, instead of a sudden cutoff, but we can't give a categorical answer to whether the PACF seems to cut off suddenly after *p* lags or if it's more of a gradual decline.

Historical Events Affecting the Data

Next up, we can see the distribution of the percentage change of total consumer credit from 1955 to 2022. It is quite easy to recognize the main ups and downs in the graph and relate them with events we all know happened during this period of time. Generally, positive values mean a percentage increase from one month to the following one in the Total Consumer Credit. Contrarily, negative values explain a monthly decrease. Flat areas don't always mean

0 growth, it depends whether they are over or under the x-axis and when they are over it they translate into constant growth.

For example, during the first few years we see a period of high volatility, induced by the post war instability and growth. Initially the dataset contained data starting from 1943, but because of the instability of this period, we reduced it to the aforementioned interval of time.

Moreover, we can differentiate the recent pandemic recession, financial crisis, petrol crisis, etc. by looking at the negative peaks or gray areas. Furthermore, we can also see some constant growth in the percentage change of total credit which was consumed before 2007 and during the 2010's, after the financial crisis.

Milestone 2 | In-sample analysis

To perform milestone 2, we used 80% of the sample values, as requested. Therefore, we use 644 observations for the in-sample analysis, which correspond to the period from 01-01-1955 to 01/08/2008; and 161 observations for the out-of-sample data to test predictions on. From here, we estimated a number of models and chose the ones that best fit the real time series. Next, we studied the adjusted values, residuals, and other goodness-of-fit metrics of each model.

Choosing the model (Parameter estimates)

In order to choose the model that best fits our data we have fitted multiple autoregressive integrated moving average models (ARIMA) with different combinations of lags for the autoregressive (AR) component, the number of differentiations, and moving average (MA) component. In total, we tested every combination from an ARIMA(p,k,q), where $p \in [0, 4],\ k \in [0, 2],\ and\ q \in [0, 17]$, according to the results yielded from the ACF and PACF analysis. The optimum models have been selected by comparing three different selection criteria: the Log Likelihood, the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC). The best models using each criteria with their attached values can be found in appendix 3 and are ARIMA(4,0,7), ARIMA(0,1,1), ARIMA(3,0,17), respectively. Appendix 5 shows the coefficients for all three models. We have also added the random walk process ARIMA(0,1,0) for comparison since it would be interesting to check (and quantify) whether our predictions would beat it.

1. Following the Log Likelihood criteria ARIMA(3,0,17) has the greatest value which is -1788.475.
2. Using the AIC criteria the best one would be ARIMA(4,0,7) since it has the lowest value. which is 5.609.
3. The BIC criteria chooses ARIMA(0,1,1) because it has the lowest value which is 5.68.
4. The random walk has given a -1948.55 Log Likelihood value, which falls far away from the other models which have a better value.

The random walk ARIMA(0,1,0) provides values for the different criteria that are worse than the previous ones, and in the upcoming sections we are going to see how its predictions are significantly worse.

Fitted values

All models are extremely similar in capturing the data (appendix 4), both from the perspective of the variance of the residuals, as well as the LL values. The best model in this respect being the ARIMA(3,0,17), with a 0.0232 variance of residuals and a LL of -1788.48, but it's important to remark that the difference between an ARIMA(4,0,7) and an ARIMA(3,0,17) are almost negligible. One last important observation is that, even if the differences between the models is not that big, all three of them outperform the random walk by a fairly long measure when we take into account the scale of the data.

Residual diagnostics and Goodness of fit

Another important step to take into the model benchmark is the analysis of the residuals to check that they are normally distributed and not autocorrelated with previous values so that the estimations of these models are consistent.

The Kernel density function plots and the quantile-quantile plots (QQ-plots) can be found in Appendix 6.

In all models we can conclude that the residuals are not normally distributed since they have positive kurtosis. This can be corroborated by the Jarque-Bera test, in which we reject the null of normality since the p-values are all close to zero and therefore below any significance level. Even though the residuals for all the models are not normal, we can still conclude that the best model in this sense is ARIMA(3,0,17) since it has the lowest statistic. Following this are ARIMA(4,0,7) and ARIMA(0,1,1) and the random walk.

Observing the QQ-plots, the tails of the distributions turn clockwise which indicate that the values are not normal. This means that the distribution is heavily-tailed, or in other words we should see a negative kurtosis and this contradicts the density function plots.

An explanation of this event would be that the QQ-plot shows the ranking in the empirical distribution compared to the normal distribution and because our data is symmetrical, near the middle the ranks will be similar. With respect to the contradiction on the tails, it could be that the software used for this plots (R), computes its fit to the QQ-plots based on some moderate percentiles, such as quartiles, while the fit to the histogram is based on matching moments. Nevertheless, the error terms are not normally distributed.

What is left is to check for autocorrelation. For this we can make use of the autoregressive function (ACF) and the partial autoregressive function (PACF) on the residuals of each model (appendix 7).

For ARIMA(3,0,17) the residuals are not autocorrelated up until lag 24, even though such large lags are often imprecise. Same applies for ARIMA(4,0,7). For ARIMA(0,1,1) we have autocorrelation at lag 6 even though it is minimal.

Another test is the one of Box-Ljung. If the p-value is below a confidence level of 5%, we accept the null that the residuals are correlated. In this case, we can state that the residuals are not autocorrelated for all the models.

Overall, we can conclude that the errors are not autocorrelated for any model although we should consider two facts which seem worth commenting on:

1. We can see that in all our autocorrelograms we have some significant values for lag 24, which may be due to the fact that it is the correlation between two years of distance. This may be due to the fact that there exists correlation between data every two years although it does not change the outcome of our study, except for the case of the Ljung-Box for the ARIMA(0,1,1) depending on the number of lags used (Appendix 7).
2. The random walk does present autocorrelated lags at every level of significance on a Ljung-Box test. In fact, this is proof that our models beat the random walk process, because they are unbiased.

Milestone 3 | Out-of-Sample Prediction

In this third milestone we were asked to compute the prediction of the out-of-sample observations by forecasting the models obtained in the in-sample analysis on milestone 2. Taking the remaining 20% of the sample (using 161 data from the sample, which correspond in the period from 01/09/2008 to 01/01/2022). To do so, we carried out the dynamic and static forecasts for the rest of the observations and computed the Root Mean Square Errors and R-Squared values, respectively, so that we could establish which models forecasted best the rest of the time series observations:

Dynamic Forecasting

In this forecast we can observe how most of the predicted values for the out-of-sample converge and oscillate around the sample mean throughout the period analyzed (appendix 8). That is because computing a high order model introduces complex numbers, which behave differently than we would expect for any model, not only the models we tried to estimate. On the other hand, all models converge after a number of lags, which is also common for any model dynamically forecasted; what we mean by this is that, even if we had tried to calculate the same quantity of step ahead values with the code used in class for an AR(1), MA(1) or ARMA(1,1) we would have obtained similar forecasts around the sample mean.

So, when computing the dynamic forecasts for the models we chose to analyze we were able to calculate their Root Mean Square Error values to compare them and choose which model fitted best for the various out-of-sample predictions produced. The results provided can be found shocking enough because the smallest loss function obtained is the one for the ARIMA(0,1,1) dynamic forecast, where RMSE times 100 was 503.738. The second best model evaluation is the one for the ARIMA(3,0,17) with RMSE times 100 was 522.229. For the last two spots in the ranking of the best two dynamic forecasts we can find the ARIMA(4,0,7) and ARIMA(0,1,0), with loss function values of 574.954 and 670.930, respectively.

Static Forecasting

For the following forecast, the sexier one if you let us, we computed the static forecast with the code provided in class. This kind of forecasting produces a prediction of the 1-step ahead value calculated from the last value observed while it re-estimates the model each step, so that the estimation is done for one value while the model for the values observed is recalculated each time. We will see that these processes are much more precise and are able to construct better graphs, as everybody knows that any paper's quality is directly related to how good the data is presented and how good the graphs look. Furthermore, we will use the R-Squared values for the models predicted as loss functions in order to evaluate how good the out-of-sample forecasts fit the data.

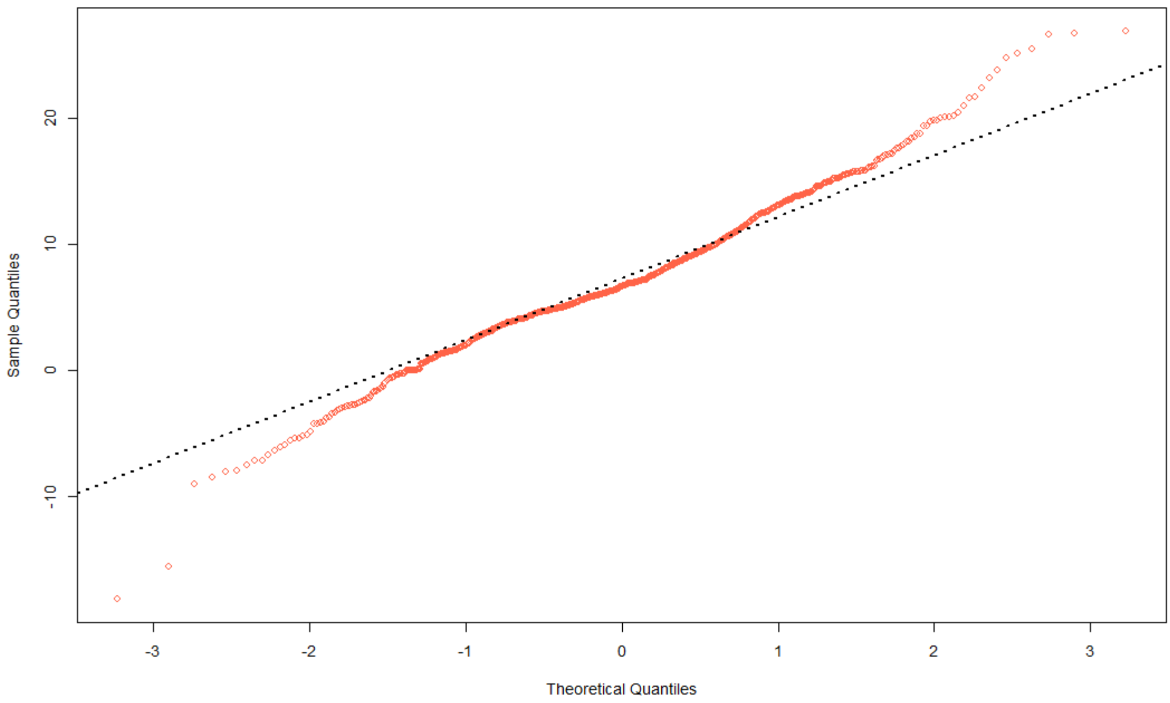
Therefore, we can conclude that by comparing the various predictions obtained for the different models estimated (appendix 9) we might find that the best model, in terms of R-Squared times 100 is the ARIMA(4,0,7) with a value of 72.865. The second best fit would be the one from ARIMA(0,1,1) with a value of 72.732. Third best fit would have a fitting percentage of 72.099, obtained from the ARIMA(3,0,17). Last but not least, we can find the random walk ARIMA(0,1,0), with a value of 60.569, which is quite a high percentage fit. Lastly, we would have to consider that using integrated data of our sample and estimating an MA(1) would be a good choice because we obtained nearly the same evaluation fitness with less coefficients than the other models.

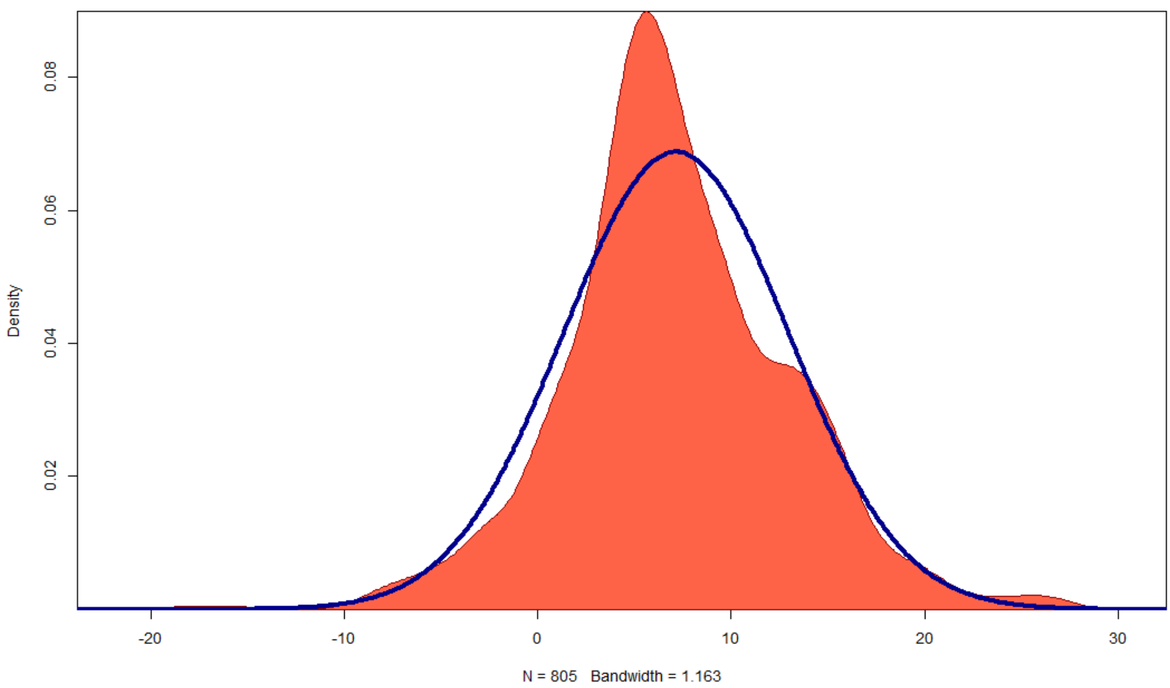
Conclusions

Surprisingly or not, one of the best forecasted models is the one we can obtain with an MA(1) after integrating the data analyzed, because it is the ARIMA(0,1,1); and a volatility clustering model with integrated data might also be appropriate for our variable. Now, considering the log-likelihood criteria to choose which models fitted best our in-sample data, we were able find that the integrated MA(1), ARIMA(4,0,7) and ARIMA(3,0,17), underperforms the other two models (ARIMA(4,0,7) and ARIMA(3,0,17)) log-likelihood value. However, we can conclude that integrating our data and using an MA(1) would be the better model, because even if it's outperformed by both R^2 and LL criteria, the differences are negligible, and we will always prefer to define a model with less coefficients. Regardless of the model used, however, all of them considerably outperform the random walk process used for benchmark.

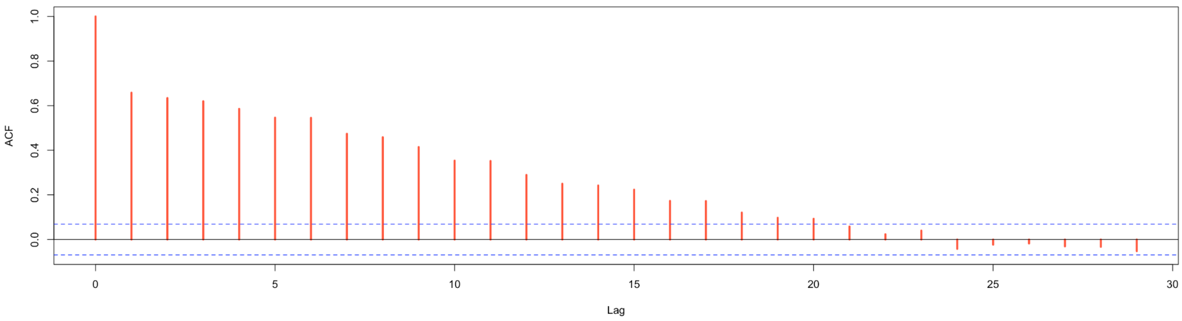
Moreover, this study has helped us to understand how this time series might be useful in order to predict macroeconomic events when analyzing consumer credit. Even if it might be obvious that consequences of unpredicted shocks in the economy need to be complemented with the usual factors (such as GDP, inflation, wages…) in order to find significant and more or less coherent conclusions.

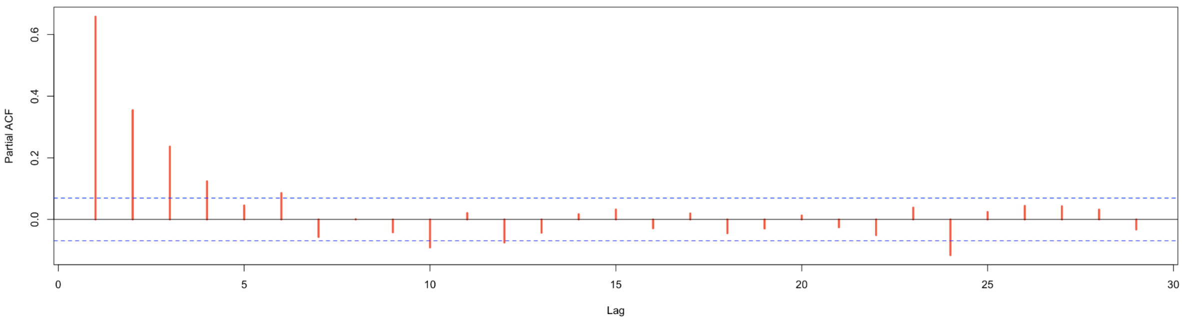
Appendix 1 | Description of the data





Appendix 2 | Dynamic properties of the data





Appendix 3 | Model selection table and criteria

Top three ranking models by criteria, higher in each category is better

By Log Likelihood

tracking	Ljung-Box 22 lags p.value	LL	aic	bic	MSE of 10-step ahead dynamic forecast
arma(3,0,17)	0	-1788.47532358312	5.61948858255627	5.76517441058141	5.96827893005451
arma(3,0,15)	0	-1791.58089696103	5.62292204025164	5.75473302751248	5.90260412985804
arma(3,0,16)	0	-1792.21348615723	5.62799219303489	5.76674060067788	5.96269253628338

By AIC

arma(4,0,7)	0	-1794.25463878793	5.60948645586313	5.69273550044893	6.71454523401226
arma(3,1,8)	0	-1797.81932069865	5.61745130651755	5.6937629307212	5.69683857217813
arma(3,1,9)	0	-1797.31333302374	5.61898550628489	5.70223455087069	5.71374711584748

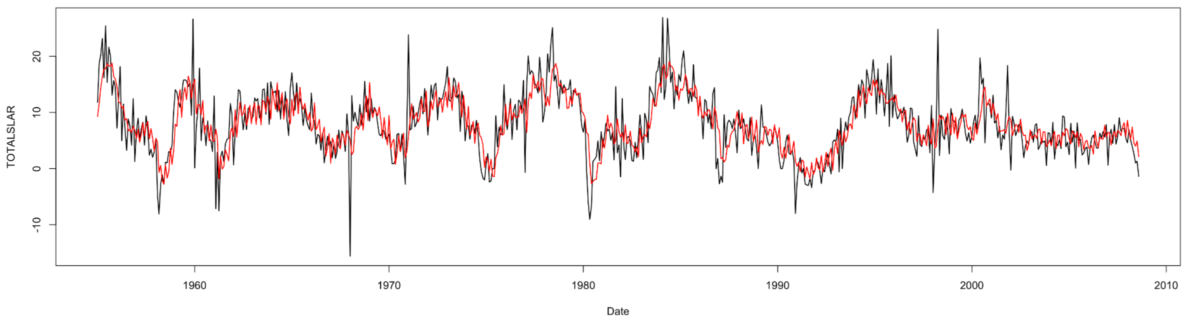
By BIC

arma(0,1,1)	0	-1825.96157033284	5.67379369668585	5.680731117068	4.62393907365134
arma(3,0,3)	0	-1806.86951747968	5.63313514745243	5.68169709012748	6.42937955525433
arma(1,0,1)	0	-1821.54355273161	5.66628432525344	5.68709658639989	6.63224501183884

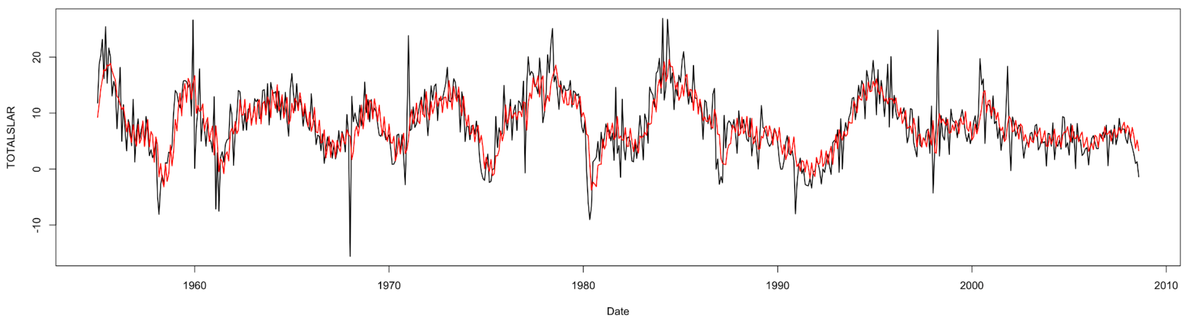
Appendix 4 | Model (fitted) vs data and residuals

Red line denotes the model, black line is the data

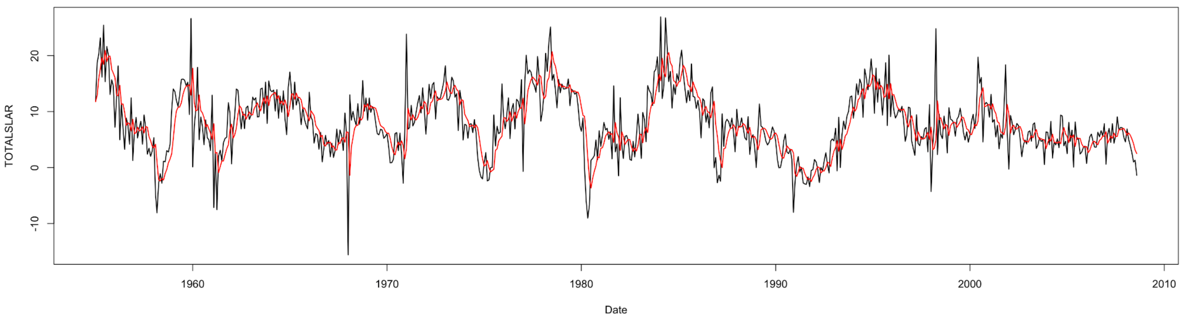
ARIMA(3, 0, 17) | LL: -1788.48



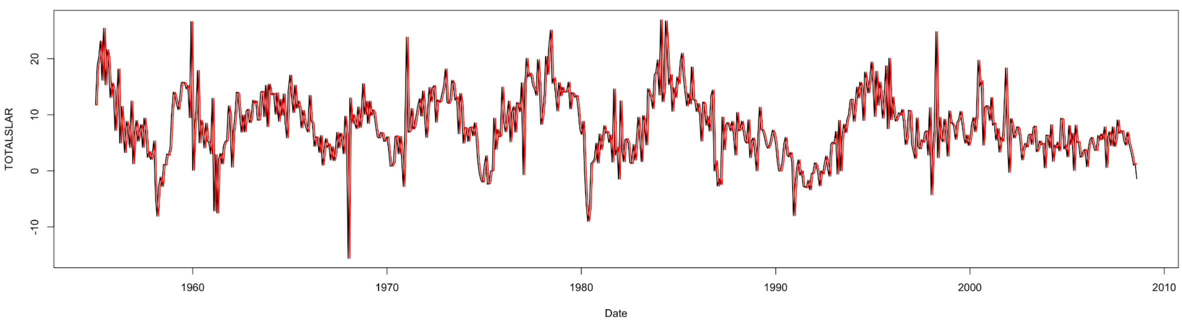
ARIMA(4, 0, 7) | LL: -1794.25

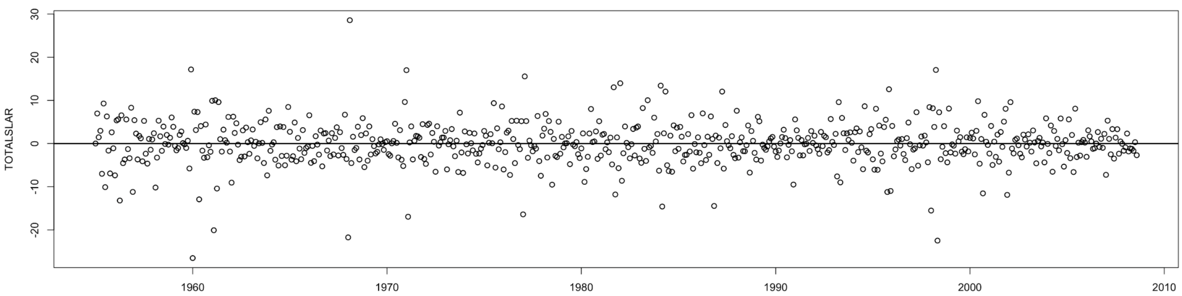


ARIMA(0, 1, 1) | LL: -1825.96



ARIMA(0, 1, 0) | LL: -1948.55



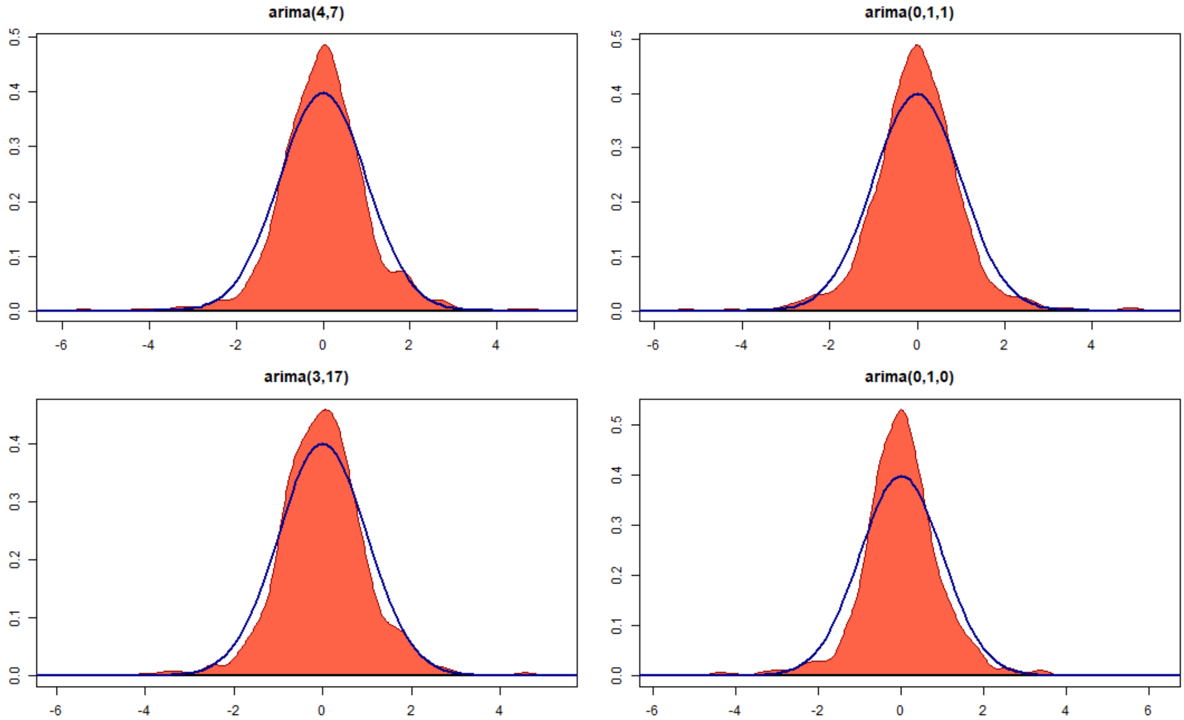


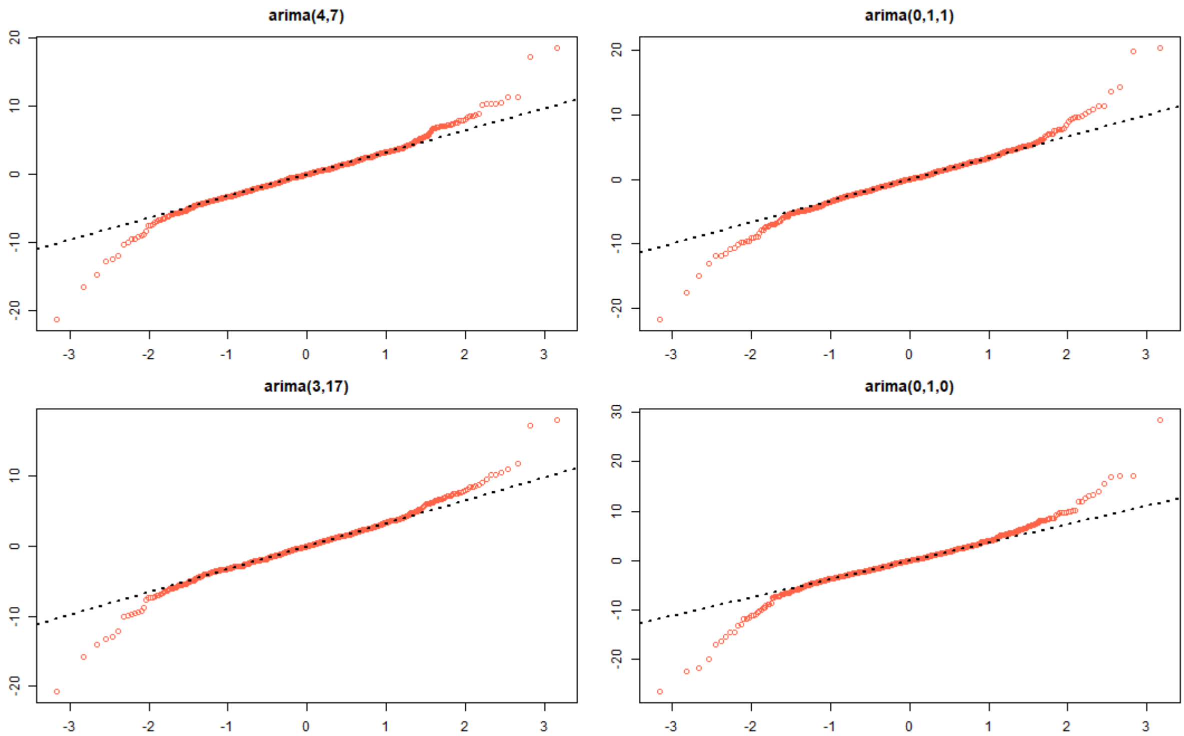
Appendix 5 | Model coefficients

```
stargazer(arima0.1.1,
          arima3.0.17,
          arima4.0.7,
          type = "text",
          column.labels = c("ARIMA(0,1,1)", "ARIMA(3,0,17)", "ARIMA(4,0,7)"),
          dep.var.labels = "Model coefficients (se)",
          align = TRUE,
          model.numbers = FALSE,
          dep.var.caption = "",
          report = "vc*s",
          single.row = TRUE)
```

```
===========================================================================
                                  Model coefficients (se)
                    ARIMA(0,1,1)      ARIMA(3,0,17)         ARIMA(4,0,7)
---------------------------------------------------------------------------
ar1                                 -0.315***  (0.026)  0.761***  (0.032)
ar2                                 -0.022     (0.030)  0.297***  (0.007)
ar3                                  0.842***  (0.026)  0.846***  (0.004)
ar4                                                    -0.925***  (0.030)
ma1               -0.650***  (0.028)  0.638***  (0.033) -0.442***  (0.053)
ma2                                  0.460***  (0.025) -0.214***  (0.046)
ma3                                 -0.475***  (0.037) -0.944***  (0.045)
ma4                                  0.184***  (0.043)  0.708***  (0.055)
ma5                                  0.133***  (0.044)  0.022     (0.047)
ma6                                  0.205***  (0.048)  0.049     (0.042)
ma7                                  0.132**   (0.051) -0.099**   (0.046)
ma8                                  0.206***  (0.053)
ma9                                  0.125**   (0.063)
ma10                                 0.088     (0.063)
ma11                                 0.086     (0.060)
ma12                                -0.035     (0.062)
ma13                                 0.013     (0.065)
ma14                                -0.033     (0.052)
ma15                                 0.170***  (0.059)
ma16                                 0.085     (0.057)
ma17                                 0.112**   (0.044)
intercept                            7.961***  (0.939)  7.959***  (0.569)
---------------------------------------------------------------------------
Observations          643               644                644
Log Likelihood     -1,825.962        -1,788.475         -1,794.255
sigma2              17.129            14.926             15.183
Akaike Inf. Crit.   3,655.923         3,620.951          3,614.509
===========================================================================
Note:                                        *p<0.1; **p<0.05; ***p<0.01
```

Appendix 6 | Distributional properties of the residuals



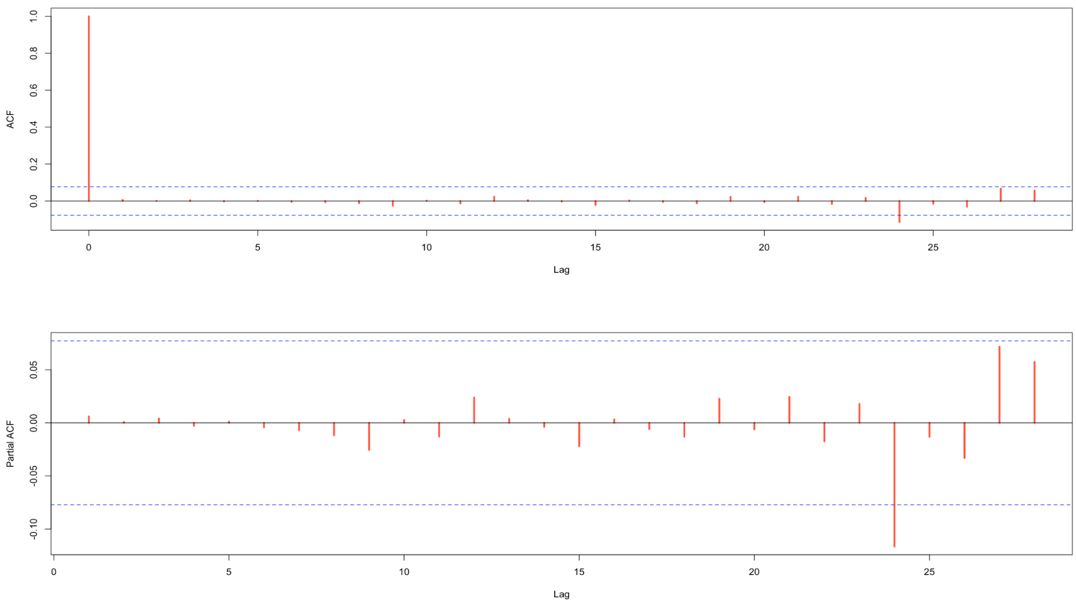


Appendix 7 | Dynamic properties of the residuals

Box-Ljung test

X-squared = 11.478, df = 25, p-value = 0.9903

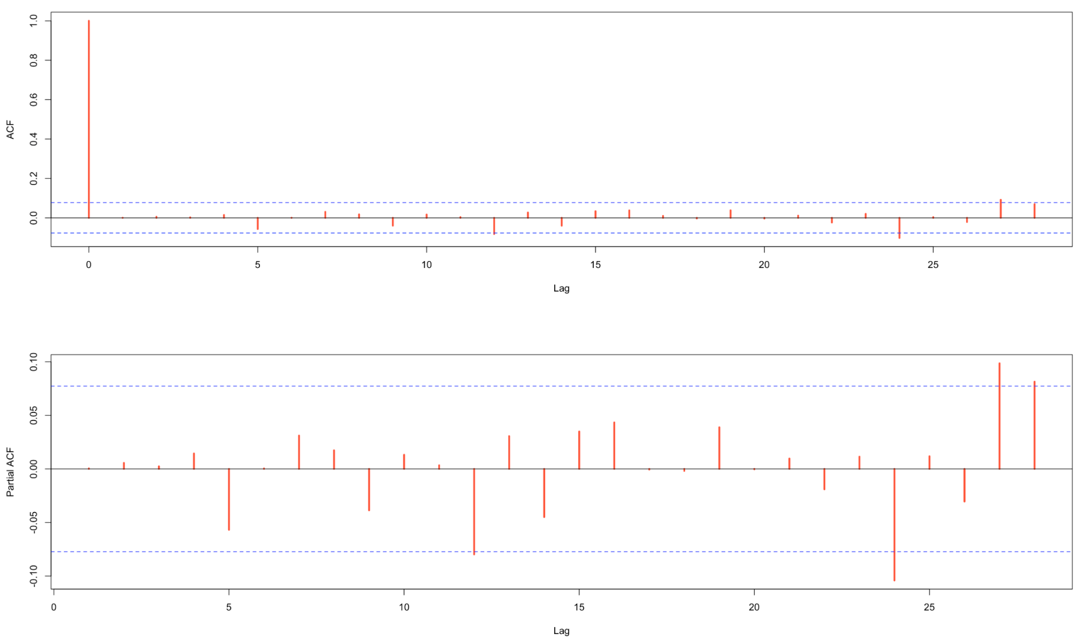
ARIMA(3,0,17)



Box-Ljung test

X-squared = 20.717, df = 25, p-value = 0.7083

ARIMA(4,0,7)

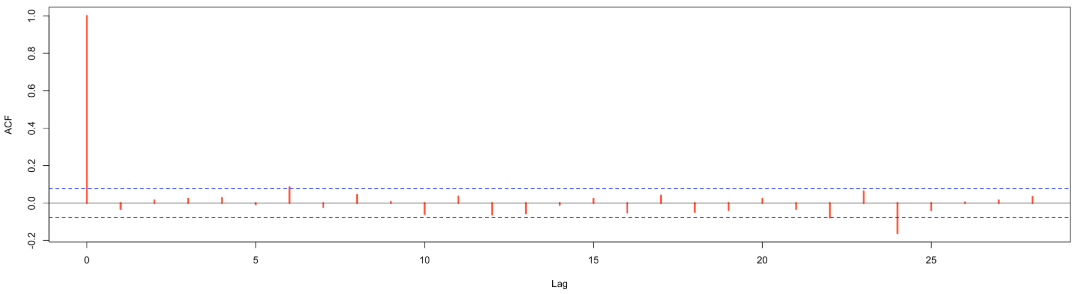


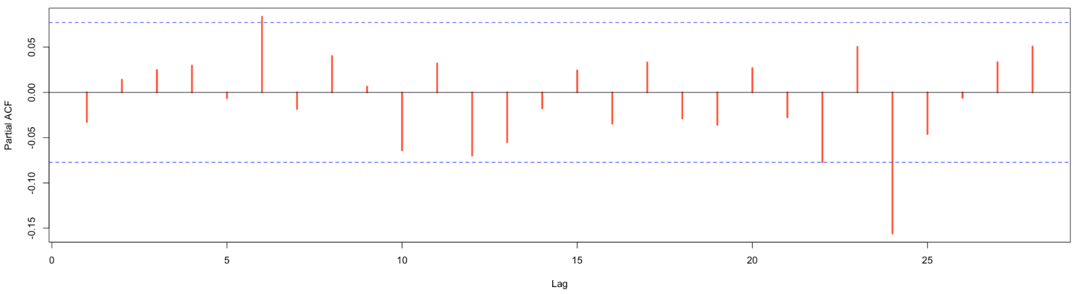
Box-Ljung test

X-squared = 26.633, df = 22, p-value = 0.2255

X-squared = 47.869, df = 25, p-value = 0.003867

ARIMA(0,1,1)

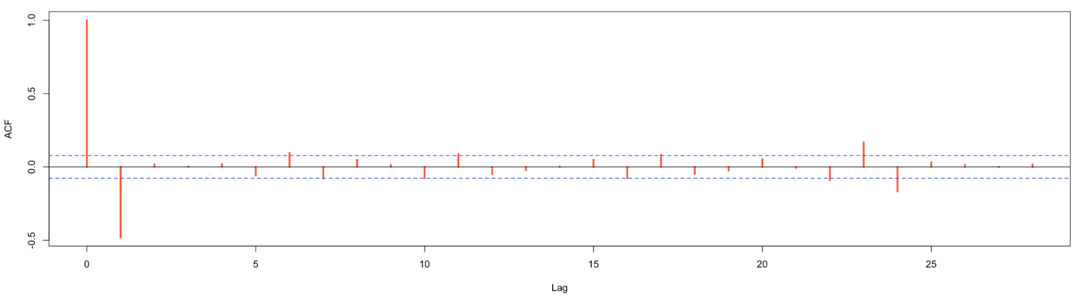


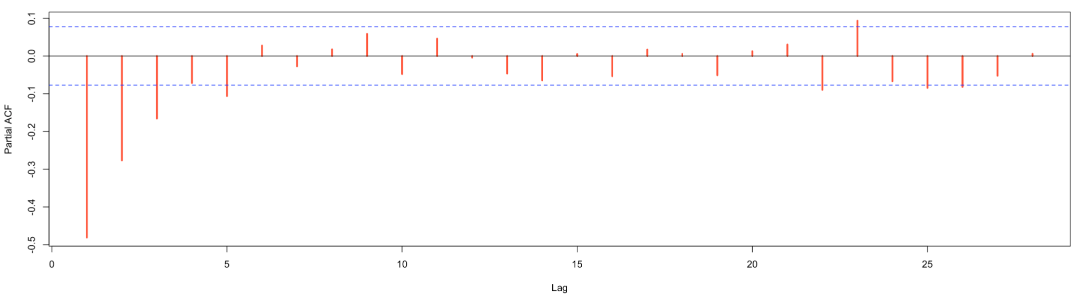


Box-Ljung test

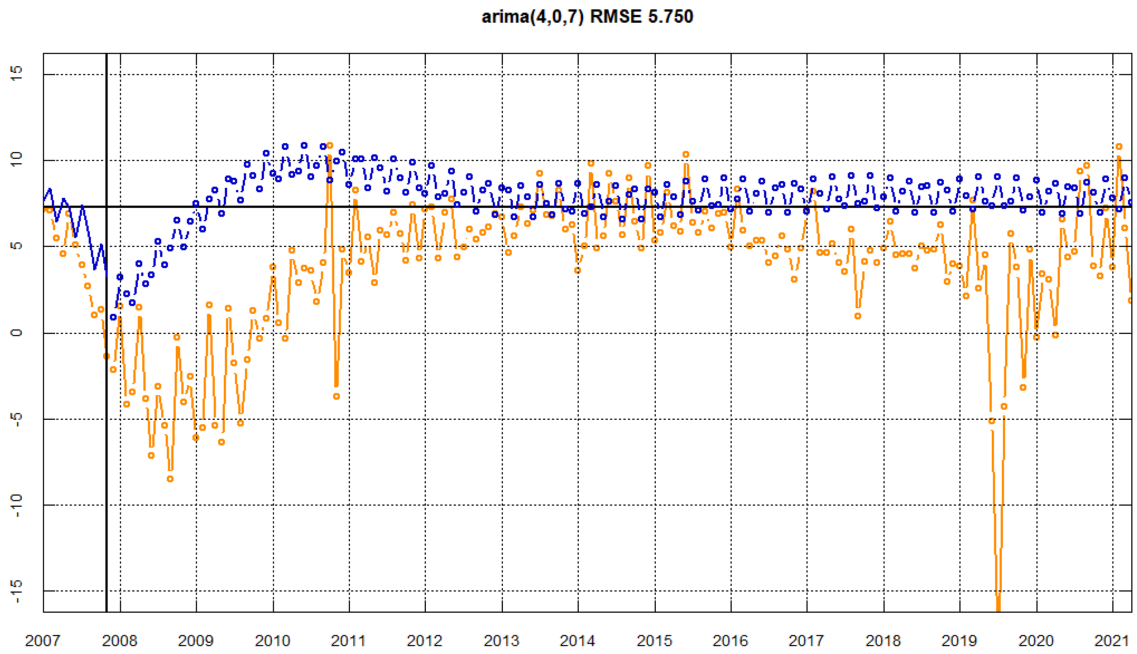
X-squared = 194.86, df = 22, p-value < 2.2e-16

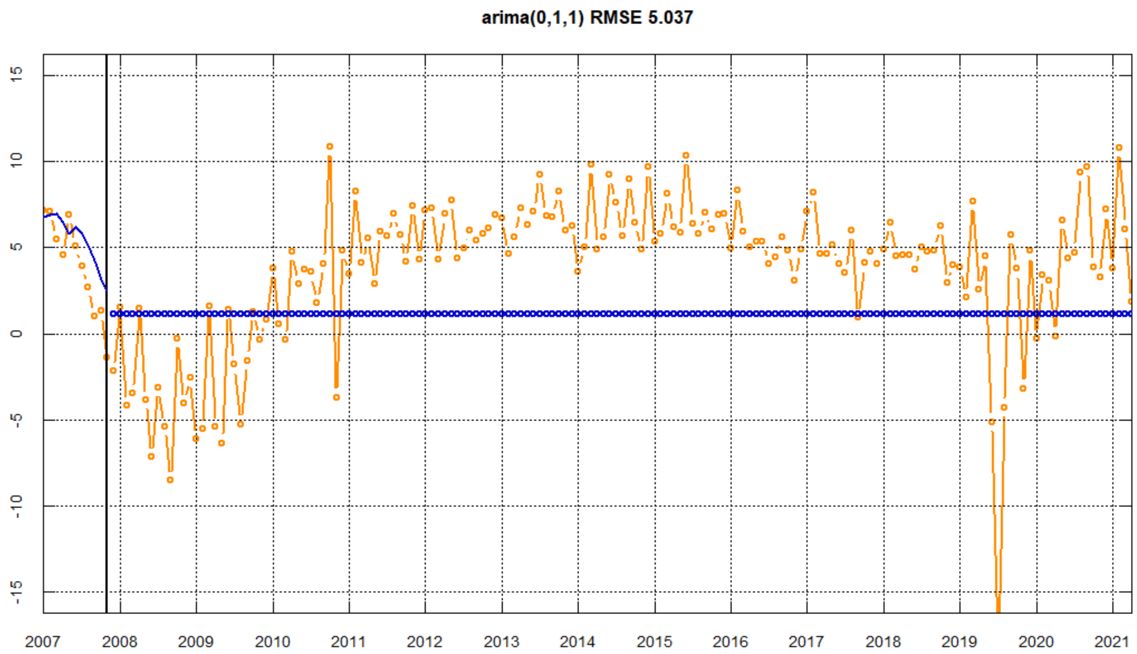
ARIMA(0,1,0) – Random Walk



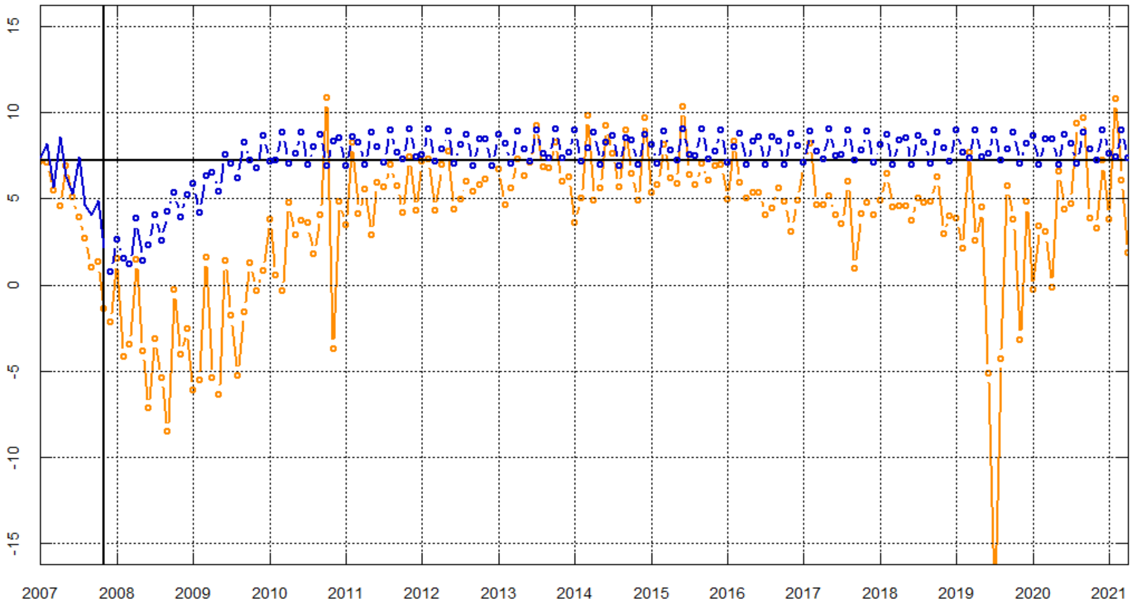


Appendix 8 | Dynamic Forecasts (Out-of-Sample prediction)

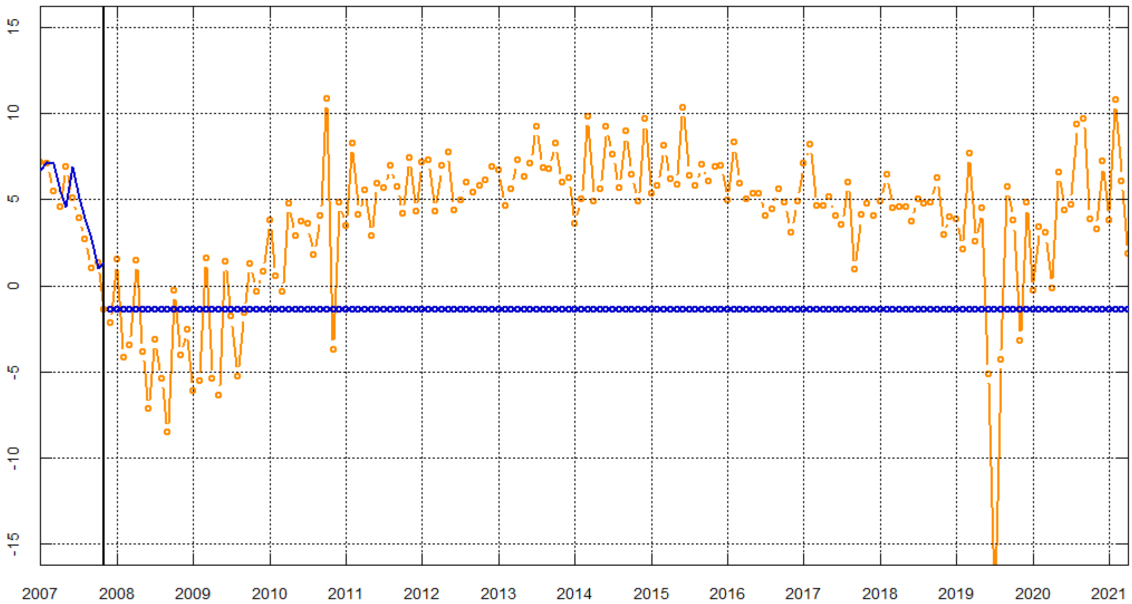




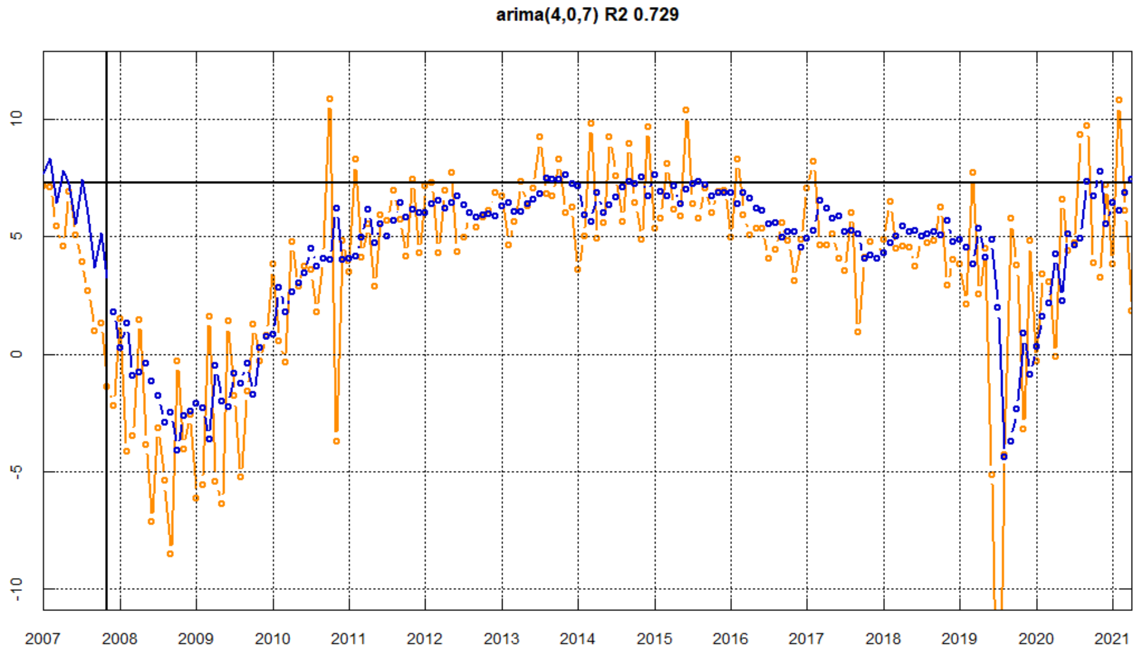
arima(3,0,17) RMSE 5.222

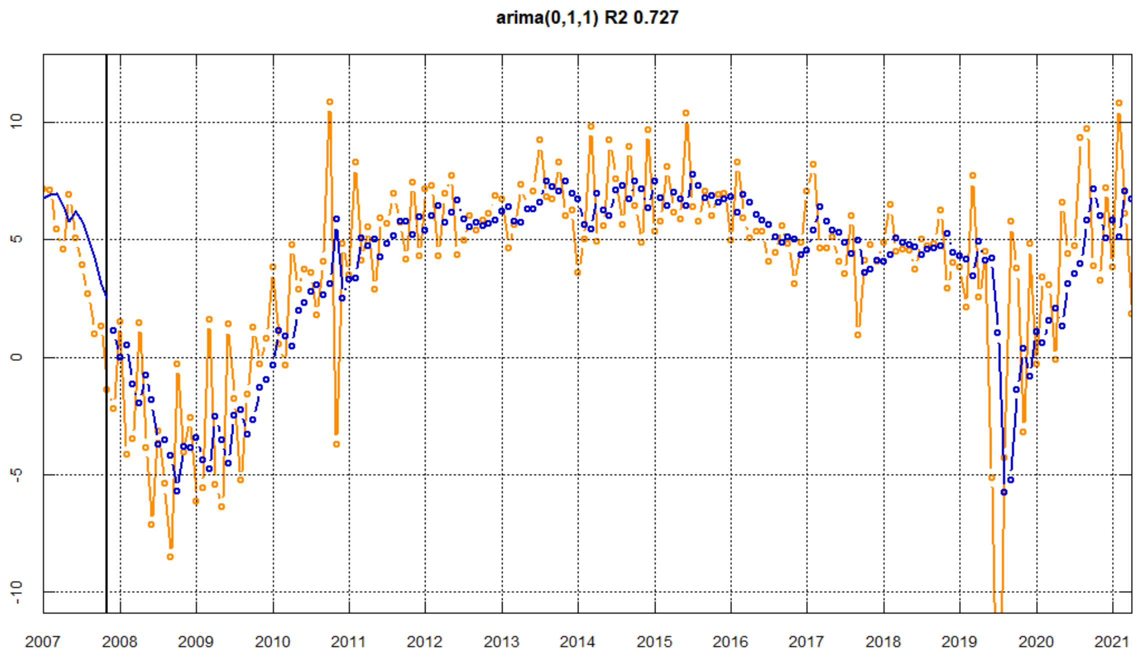


arima(0,1,0) RMSE 6.709

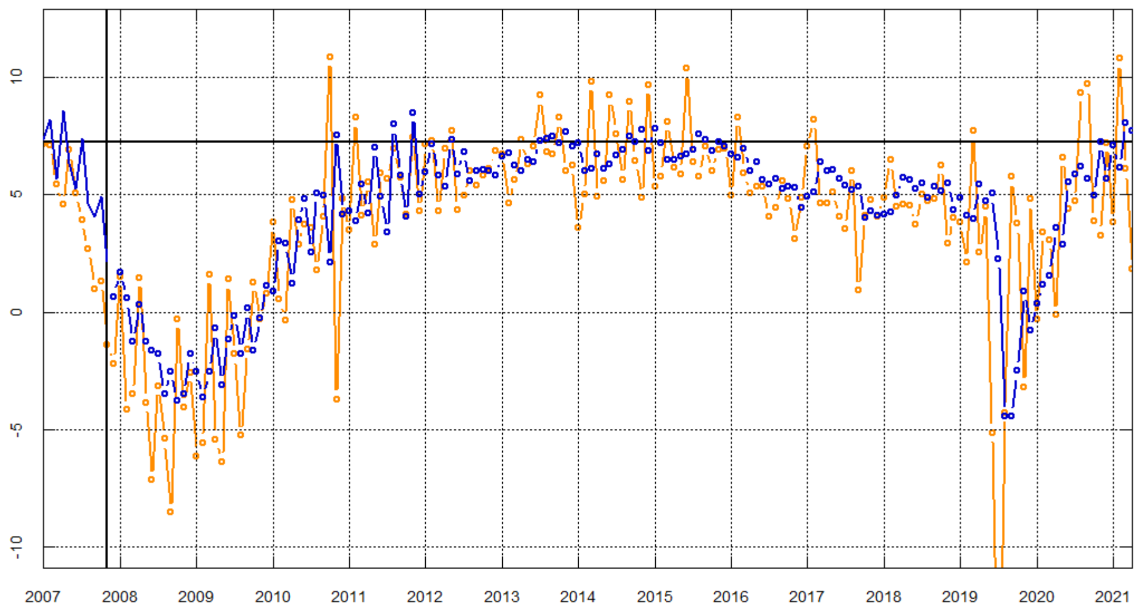


Appendix 9 | One-Step-Ahead Static Forecasts (Out-of-Sample prediction)





arima(3,0,17) R2 0.721



arima(0,1,0) R2 0.606