

Analyzing Company Segments Through a Series of Time

Alex Cho

Executive Summary

The handheld smartphone is probably the defining factor of the turn of events in the 21st century regarding the sector of information technology. One of the companies at the forefront of this movement is Apple, Inc., which provides the most symbolic of portable electronics, services, and products in the world. Apple is a behemoth of a company that really needs no introduction. Hailing from concentrated technology hub Cupertino, California, it's most famous for its powerful and handheld iPhone that has evolved greatly through the years. Apple has additionally created not just a touchscreen digital phone juggernaut, but an ecosystem that encompasses all of your needs, and has successfully integrated itself into the daily lives of many. One can point out that it is a mature company, but that's because of its smart and calculated business decisions over the many years leading up to that. This includes both decisions regarding product development and investor relations as well. After all, the company needs to take care of the latter in an efficient way to have capital to execute orders of business. It would be interesting to blend a qualitative as well as a mathematical approach to dig deeper into what really goes on behind the scenes regarding these variables, through a series of time.

For this analysis, we worked with four datasets carefully chosen and tidied beforehand that represent the health of Apple, Inc. A Box-Jenkins methodology approach was used with the general steps as the following: Model identification and selection by making sure that the variables are stationary, identifying seasonality and differencing as needed, and using the ACF and PACF plots and functions of the time series to decide which MA and AR components should be included in the model. We read in the four datasets and plotted the time series for each of them. Upon realizing that there was a lot of variability and volatility in the data, we performed Box-Cox transformations on them to reduce the variance and make it stationary. We then chose among the datasets that fit the logarithmic nature of the corresponding transformation well and then determined the correlation between them and the lagged time values to ultimately help us choose which model to run for analysis and forecasting. Then we used parameter estimation using techniques and methods such as maximum likelihood estimation to select the coefficients that best fit the chosen SARIMA models. Afterwards, we computed the mean and variance of the residuals for each model and performed various hypothesis tests on them and plotting the ACF and PACF plots once again to identify any serial correlations as well as deviations in linearity or normality. Finally we undertook a forecasting analysis with our models and made suitable assumptions while tying in macro properties regarding the underlying business structure. The model forecasting the company's share buybacks showed the least variability out of all tested and accurately predicted 10 quarters into the future based on past data.

Introduction

The question I would like to propose is “How do different revenue factors and their consequences, and product development and research expenses relate to the net income of Apple, Inc.?”. This analysis will use not just a qualitative analysis of factors affecting the main idea, but a time series approach as well using comprehensive modeling techniques in order to dive deeper into the data. By studying this data, we will be able to see some of the factors that drive this profit machine. After all, the growth and state of Apple can tell us a lot about the future of the entire industry as it plays a massive role in shaping it. This is one of the main motivations for studying Apple, Inc.’s impact on the global tech sector as its leader. From a surface-level standpoint, we do know a few things surrounding this question. For example, we know that Apple releases its newest generation iPhone in September and makes it available for purchase soon after, therefore we can predict that fourth quarter results will be stronger. We will analyze the consequences of this to check to see how much seasonality in the data there is and if it explains and confirms similar fluctuations in the total net profit as well. It is of great interest for us how Apple’s business structure regarding cash flow and its use of it can be a model for many other large companies while they are in different stages of their life cycles. Studying Apple’s financial health and business structure can provide valuable insights into the performance and prospects of one of the world’s most successful companies. Apple is one of the largest and most profitable companies in the world, with a market capitalization that consistently ranks among the highest of any publicly traded company. Understanding how Apple generates revenue and profits can provide insights into the global technology market, consumer trends, and economic conditions. Apple’s financials provide a wealth of information about the company’s performance, including revenue growth, profitability, and cash flow metrics. Analyzing these statements can help investors and analysts make informed decisions about buying, selling, or holding Apple shares as well. Its business structure and strategy can also provide insights into the company’s future prospects. For example, studying Apple’s product development and marketing strategies can help investors and analysts assess the company’s ability to innovate and compete in a rapidly changing industry. It can also greatly impact the broader economy. As a major player in the technology industry, Apple’s performance can have ripple effects on suppliers, competitors, and other stakeholders. These are the main reasons why we will undertake this extensive time series analysis.

Data Analysis and Experimental Design

We'll be loading all of the necessary packages and the raw datasets.

```
library(readxl)
library(MASS)
library(ggplot2)
library(tidyverse)
library(forecast)
require(gridExtra)
```

The following datasets that we'll be considering are quarterly net income, quarterly dividend yield, quarterly share buybacks, and quarterly iPhone sales. This is all quarterly data- one of the challenges faced during the collection of this is that monthly data could possibly have been a little bit more helpful in finding trends and properly fitting the data since we would have more to work with, however, as a publicly traded company, Apple's quarterly financials were the most accessible- and still a standard way of measuring health so it was a good choice. Other data was often behind a pay wall which also made it not readily available. I chose to let the time periods vary a little bit for each data set, some starting sooner- and this is to account for what I think is significant to undertake the study. For example, although Apple has been doing share buybacks for a lot of its existence, I chose to start at FY 2013 as this is soon after Tim Cook took office and started the massive buyback program for its investors. I decided to start the iPhone sales data before that to account for the first release date. Regarding the 5 V's, as mentioned before since it is quarterly data, the volume and velocity is not much to worry about. Veracity is not of worry either since the company is required to report itself on a regular basis. Variety was the main problem- some of it was in the form of a pdf, a bar chart, etc., however, I solved this by beforehand tidying the data and exporting to an excel file so I could easily load it into R. Lastly, the value of the data I would regard quite high because it was carefully selected so that they would all have great relevance with each other- as this report is to analyze the influence and interaction between the data sets. For extraneous columns, as mentioned before I tidied the data so that only relevant columns were remaining.

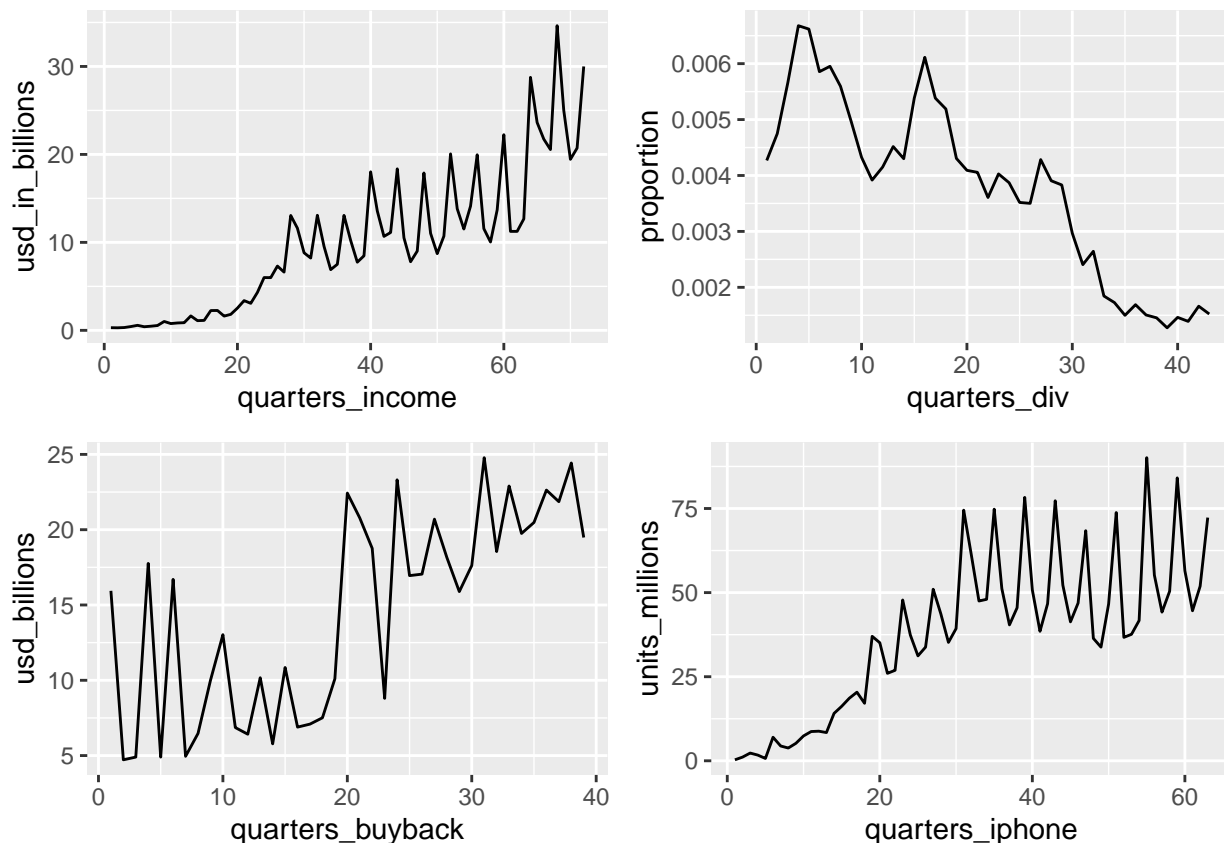
Quarterly Net Income is in simple terms, expenses subtracted from revenues and is widely used as a profitability metric by analysts, and is a great measurement of the company's financial health. Apple is one of the most profitable companies in the world- it is a literal cash cow for both itself and shareholders. The following data is Apple's quarterly net income from 2005-2022.

Dividends are issued by positive cash flow companies to shareholders as a distribution of its earnings. Think of it as sort of a reward for investing in them. Although Apple has paid a cash dividend in its earlier years, we are concerned with its dividend payout since 2012, which is soon after current CEO Tim Cook took office in August 2011.

Share buybacks are done by profitable companies as a form of share remuneration where the company buys existing shares for a higher price- in turn this reduces the float, increases relative shareholder stake and dividends to be received. This is seen as an extremely positive aspect on driver of the share price, which is why we are motivated to analyze these trends. We will be looking at the massive share buyback program that started in 2013.

This is by far Apple's definitive driver in revenue, its flagship product, the iPhone. It's gone through many evolutionary changes since its release in 2007, which has made more and more people fall in love with its easy to use and convenient interface. The release of the iPhone 4 in 2010 for example, was revolutionary. Countless feature updates such as high resolution display, design, software updates steered the product to its course that brought it to what it is today.

First, we will create the time series for the data and plot them to see what is going on.



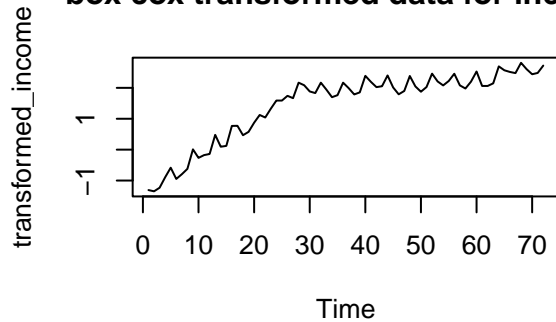
It appears that as expected, there seems to be seasonality in the iPhone sales data. There seems to be much seasonality in the net income time series as well, whereas the buybacks seem to have no visible trend.

Since time series analysis is done under the assumption that the data is normally distributed, as we can see our data is not of that nature, so we will undertake the Box-Cox transformations for them accordingly. We will find the ideal lambda value for each transformation and perform it following. Our goal is to make the data stationary and stabilize the variance as well.

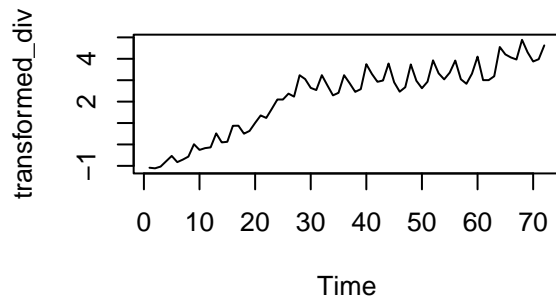
This is the mathematical notation for the transformation used to stabilize the variance, in order to fit towards normality.

$$Y_t = \frac{1}{\lambda}(X_t^\lambda - 1)$$

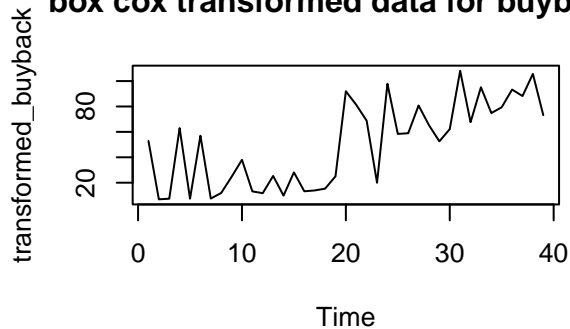
box cox transformed data for income



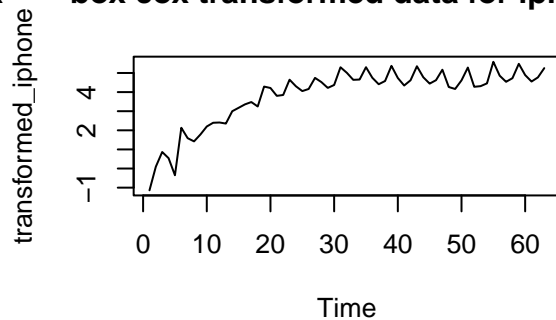
box cox transformed data for div. yield



box cox transformed data for buyback



box cox transformed data for iphone



```
## [1] 66.14034
## [1] 1.314051
## [1] 2.585462e-06
## [1] 2.728384
## [1] 44.5955
## [1] 1097.305
## [1] 551.492
## [1] 2.30647
```

When comparing the variances, the transformed data shows much less variability within, which means that the transformation was necessary and successful.

Looking at the transformed curves- the net income, dividend yield, share buybacks, and iPhone sales seem to match the logarithmic curve accounting for the optimal lambda value therefore I chose to perform the empirical analysis on this data to follow.

To analyze this process underneath the code, the optimal value for lambda can be calculated by using the process of maximum likelihood estimation (MLE), which finds the value of lambda that maximizes the likelihood of observing the transformed data given the original. The function is given by

$$L(\lambda) = -\frac{n}{2} \log(s^2(\lambda)) - (\lambda - 1) \sum (\log(y_i))$$

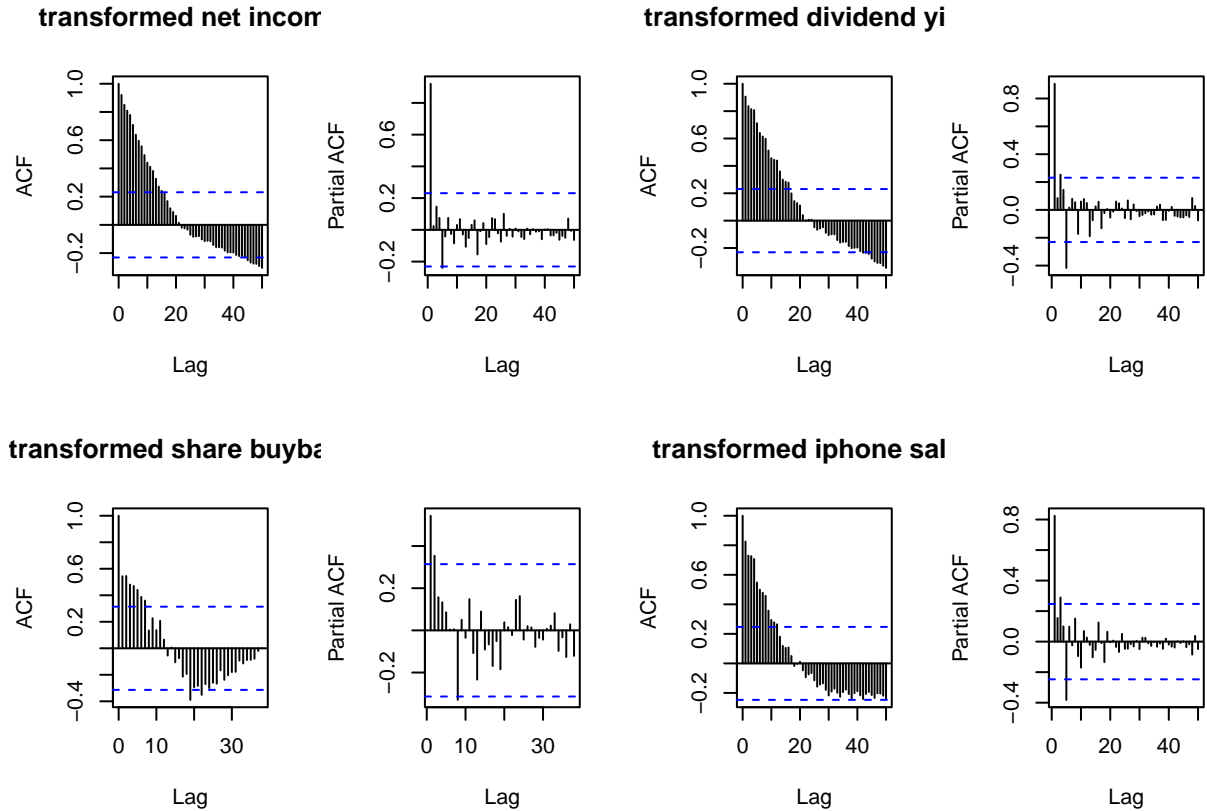
where n is the sample size, $\log s^2$ is the variance of the transformed data, and the summation portion is the sum of the logarithms of the original data.

In relation to our data, the box cox transformation assumes that the data points are non-zero before the transformation, which then involves raising the function to a power (lambda) that varies. It determines the type and strength of the transformation which we will need to eventually figure out which models to use for analysis and forecasting. Since our lambda values are quite close to 0, which means that the transformation will take on a curve of a natural log, and be given by the equation

$$Y_\lambda = \ln(Y)$$

We will look later on to see whether our data is fitting of an ARDL time series regression model in the later steps, as this is what a transformation is most useful for.

Here are the ACF and PACF plots of the transformed data.



In the context of our data, the Autocorrelation and Partial Autocorrelation functions, are explained by tools used to identify the presence of autocorrelation in a time series data. The ACF measures the presence of correlation between a time series process and its lags, or lagged values in time. Although we accounted for 50 lags in the above plots, the equation for the ACF of the net income can be given by:

$$\rho_k = Corr(Y_t, Y_{t-k})$$

The ACF gives a visual representation of how correlated the time series is with its past values, up to a certain lag. If the ACF plot shows a significant correlation at lag k , it indicates that the value of Y at time t

is influenced by the value of Y at time t-k. This is useful for identifying trends, seasonality or other patterns in the data.

The PACF measures the correlation between a time series and its lags, after removing the effect of any intervening lags. The equation for the PACF is given by:

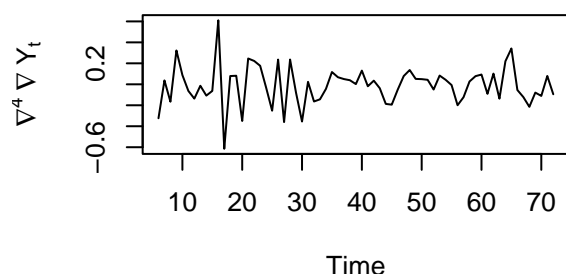
$$\phi_k = Corr(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1})$$

The PACF plots give us a visual representation of the direct influence of the past values on the present value, after removing the effect of any intervening values. If the PACF plot shows a significant correlation at lag k, it indicates that the value of Y at time t is directly influenced by the value of Y at time t-k.

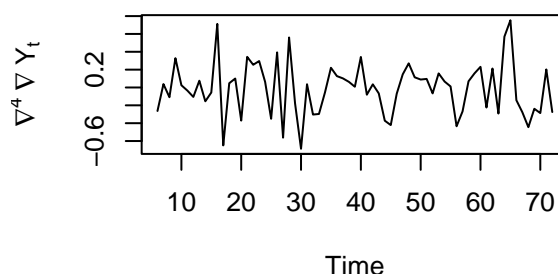
In the context of our datasets, looking at the plots, we observe that there seem to be significant correlations every four lags, which implies that the period of the seasonal component is 4, where d=4. This makes sense because there are four periods when accounting for quarterly data.

Now we will be using this information for de-trending and de-seasonalizing the data in order to find the p and q values from the PACF plots of those datasets for our models.

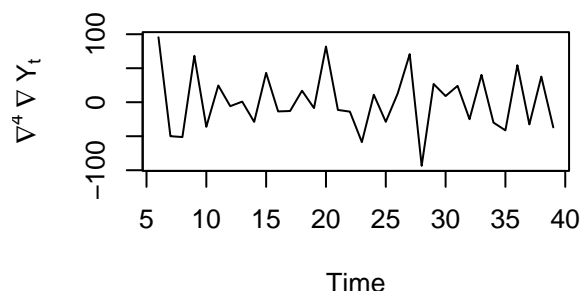
De-trended/seasonalized net income



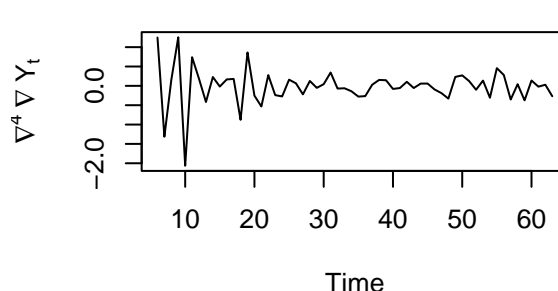
De-trended/seasonalized dividend yield

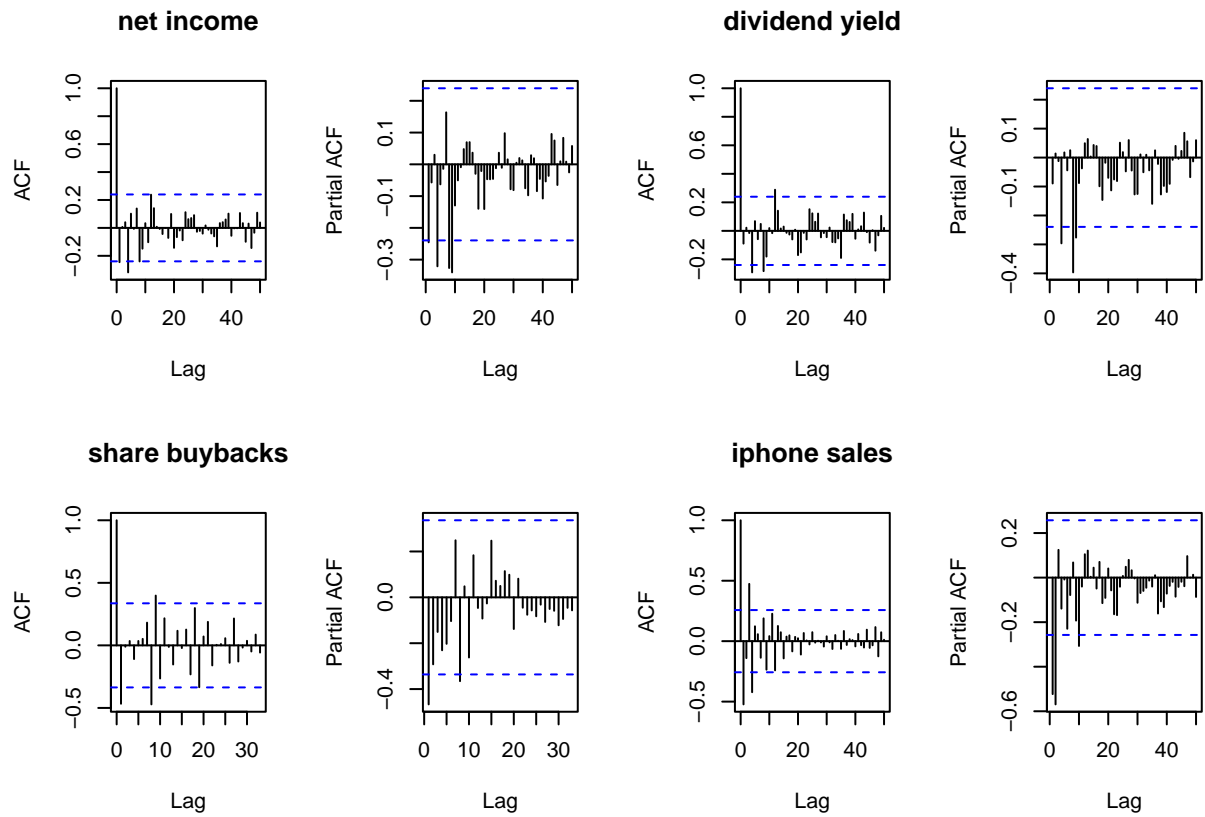


De-trended/seasonalized share buyback



De-trended/seasonalized iphone sale:





As seen in the above plots, things seem to be all over the place when concerning the PACF, where for the most part the lags seem to be within the limits of the significance levels but further research will have to be done in order to analyze its potential white noise properties.

Time Series Modelling and Forecasting

In the above analysis, we used differencing at lag 1 to remove the trend component and then at lag 4 as well to remove the seasonal component. After using a Box-Cox transformation as performed above and differencing at lags 1 and 4, the time series for the datasets are now stationary. We will be looking for the most optimal model for each time series, by using the respective properties of its MA and AR order polynomials and ultimately the ARMA coefficients.

For the net income time series data, we see that there is a seasonal component since there are spikes at lags $l = 4n$. The ACF cuts off after lag 4 so we can assume a simple moving average model of order 1, SMA(1). The PACF cuts off after lag 4 so we can assume SAR(4). For the dividend yield time series data, we see that there is a seasonal component since there are spikes at lags $l = 4n$. The ACF cuts off after lag 4 so we can assume a simple moving average model of order 1, SMA(1). The PACF cuts off after lag 4 so we can assume SAR(4). For the share buyback time series data, we see that there is a seasonal component since there are spikes at lags $l = 4n$. The ACF cuts off after lag 3 so we can assume a simple moving average model of order 1, SMA(1). The PACF cuts off after lag 4 so we can assume SAR(3). For the iPhone sales time series data, we see that there is a seasonal component since there are spikes at lags $l = 4n$. The ACF cuts off after lag 4 so we can assume a simple moving average model of order 1, SMA(1). The PACF cuts off after lag 4 so we can assume SAR(4). We also consider SAR(2) since the lags observed after lag 2 are not as large as the lags prior meaning that they may just be from noise.

Next we will look at lags 2 and 3 to determine the order of MA and AR polynomials. The ACF cuts off immediately therefore we must have MA(0). The PACF decays quite fast as well, so we will consider AR(0).

Using the findings from above, we will consider the following models now to do analysis and forecasting:

- Net Income SARIMA (0,0,0) x (2,0,1)₄
 - AIC = 31.65
 - $\nabla_4 \nabla X_t = (X_t - X_{t-4})(X_t - X_{t-1})Z_t$
 - $(X_t + 1.9445X_{t-4} - 0.963X_{t-8})\nabla_{12} \nabla X_t = (Z_t - 0.7104Z_{t-4})$
- Dividend Yield SARIMA (0,0,0) x (1,0,1)₄
 - AIC = 29.52
 - $\nabla_4 \nabla X_t = (X_t - X_{t-4})(X_t - X_{t-1})Z_t$
 - $(X_t + 0.9421X_{t-4})\nabla_4 \nabla X_t = (Z_t - 0.3826Z_{t-4})$
- Share Buybacks SARIMA (2,0,0) x (2,0,1)₄
 - AIC = 372.08
 - $\nabla_4 \nabla X_t = (X_t - X_{t-4})(X_t - X_{t-1})Z_t$
 - $(X_t + 0.4043X_{t-1})(X_t + 0.3799X_{t-4})\nabla_4 \nabla X_t = (Z_t + 0.0213Z_{t-4})$
- iPhone Sales SARIMA (2,0,0) x (3,0,1)₄
 - AIC = 100.82
 - $\nabla_4 \nabla X_t = (X_t - X_{t-4})(X_t - X_{t-1})Z_t$
 - $(X_t + 1.1247X_{t-1} + 0.4482X_{t-2})(X_t - 0.6333X_{t-4} + 0.1454X_{t-8})\nabla_4 \nabla X_t = (Z_t + 0.0548Z_{t-4})$

Next, we check to see whether our models are casual and invertible by seeing if the modulus of the roots of the polynomials are greater than one, or outside of the unit circle.

```
## [1] -0.4248717+0i  2.4440825-0i
```

```
## [1] 1.407658+0i
```

One of the SAR polynomial roots lie inside of the unit circle, therefore not casual but invertible

```
## [1] -1.088613+0i
```

```
## [1] -1.061458+0i
```

```
## [1] 2.613696+0i
```

All of the roots lie outside of the unit circle, therefore casual and invertible

```
## [1] -0.532114+1.532686i -0.532114-1.532686i
```

```
## [1] 1.602738+0i -1.721667-0i
```

```
## [1] -4.677268+0i
```

All of the roots lie outside of the unit circle, therefore casual and invertible

```
## [1] -0.6435453+0.6143435i -0.6435453-0.6143435i 1.9948121+0.0000000i
```

```
## [1] -6.877579+0i
```

Some of the roots lie inside of the unit circle, therefore not casual but invertible

Now for each of the four proposed models, we will analyze the residuals and perform diagnostic checks including the corresponding hypothesis tests for them. For each of the SARIMA models, we will be conducting Box-Pierce and Ljung-Box test for serial correlations, McLeod-Li/Quadratic correlation test for linearity, and Shapiro-Wilk test for normality, then compare p-values with a significance level of 0.05.

We will be using the Box-Pierce and Ljung-Box tests to test the autocorrelation of our time series models. The null hypothesis is that the data is independently distributed, meaning that there is no significant autocorrelation. The alternative hypothesis is that the data distribution is dependent, meaning that there is significant autocorrelation.

The Box-Pierce test statistic is calculated as:

$$Q = n(R_1^2 + R_2^2 + \dots + R_k^2)$$

where n is the sample size, R_1, R_2, \dots, R_k are the autocorrelation coefficients for lags 1, 2, \dots , k , and k is the number of lags being tested. The test statistic follows a chi-square distribution with k degrees of freedom.

In a similar fashion, the Ljung-Box test statistic is calculated as:

$$Q^* = n(n+2)\left(\frac{R_1^2}{n-1} + \frac{R_2^2}{n-2} + \dots + \frac{R_k^2}{n-k}\right)$$

where n is the sample size, R_1, R_2, \dots, R_k are the autocorrelation coefficients for lags 1, 2, \dots , k , and k is the number of lags being tested. The test statistic also follows a chi-square distribution with k degrees of freedom.

We will be using the McLeod-Li test to detect nonlinearity in a time series. The null hypothesis is that the time series is linear, while the alternative hypothesis is that the time series is nonlinear.

The McLeod-Li test statistic is based on the squared autocorrelations of the time series. Specifically, it is calculated as:

$$Q = n(n+2) \sum_{k=1}^h \left(\frac{r_{k+1}^2}{n-k} \right)$$

where n is the sample size, h is the maximum lag being tested, and r_k is the autocorrelation coefficient at lag k . The test statistic follows a chi-squared distribution with h degrees of freedom.

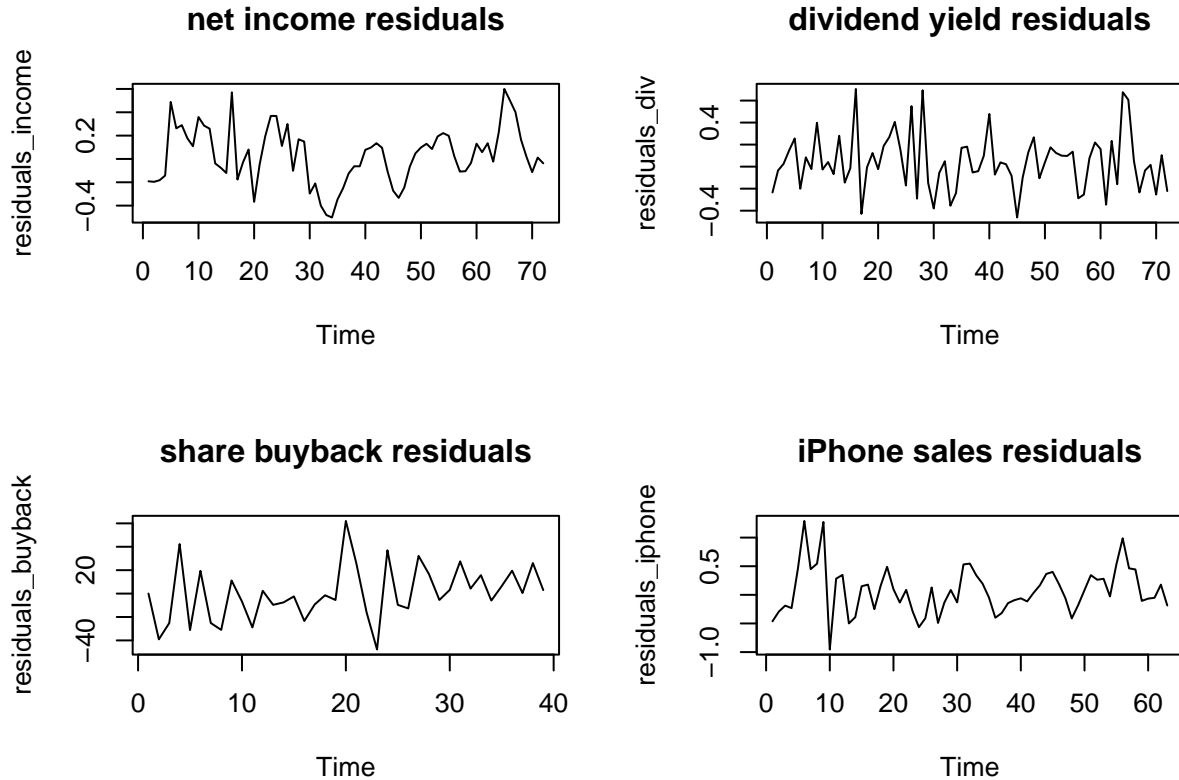
We will use the Shapiro-Wilk test for the normality of a distribution. The null hypothesis is that the data is normally distributed, and the alternative hypothesis is that the data is not normally distributed.

The test statistic is calculated as:

$$W = \frac{\sum_{i=1}^n a_i x_{(i)}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

where n is the sample size, $x_{(i)}$ is the i th order statistic, \bar{x} is the sample mean, and a_i are constants derived from the covariance matrix of the order statistics. The test statistic ranges between 0 and 1, with 1 indicating perfect normality. The critical values for the test depend on the sample size and the significance level chosen.

After all of the hypothesis tests, we will check the ACF and PACF plots to see if the time series resemble that of a white noise.

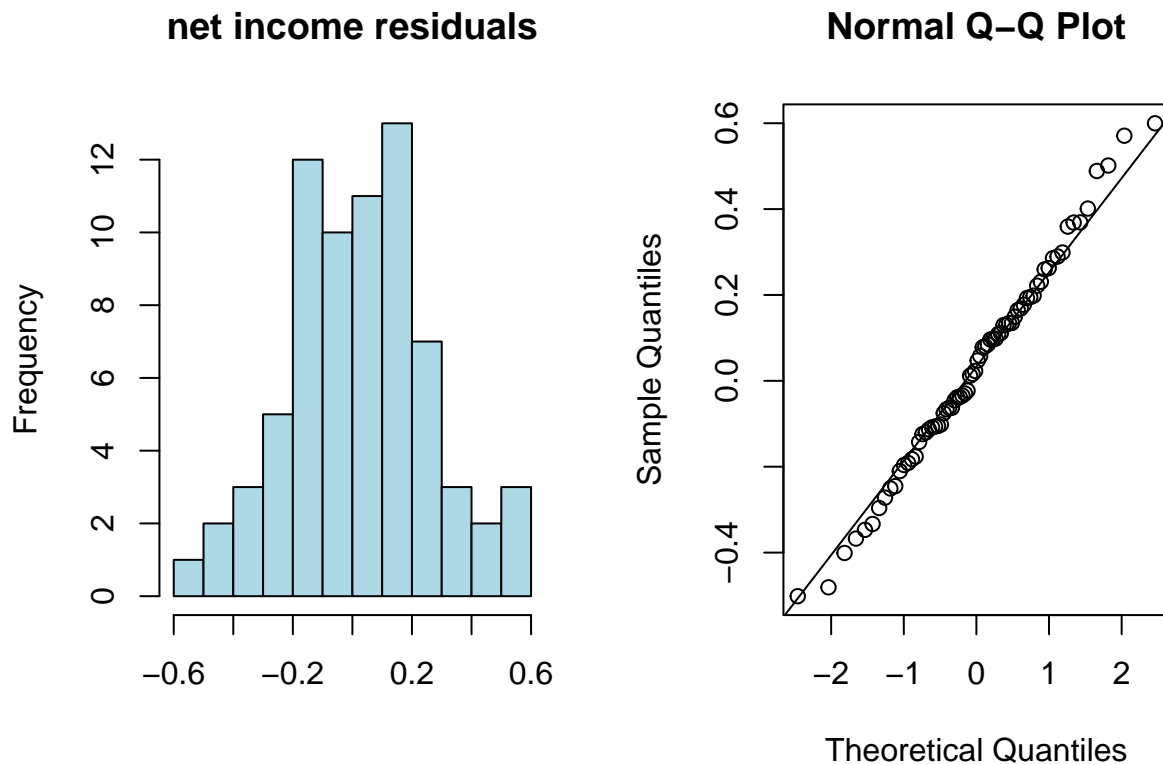


These are the plots for the residuals of each of our models. When considering residuals, we are concerned with the difference between the observed values and the values to be predicted by a model and can be given by the following equation:

$$\varepsilon_t = Y_t - \bar{Y}_t$$

This is helpful in trying to assess the goodness-of-fit of the model. We have plotted the residuals over a period of time to check for patterns or trends. If the residuals exhibit patterns or trends, it may indicate

that the model is not adequate for representing the time series processes. In the analysis to follow, we will also be calculating summary statistics of the residuals such as the mean and variance- and looking to see that they are close to zero, and have little variance over time, respectively.



```
## [1] 0.03261541

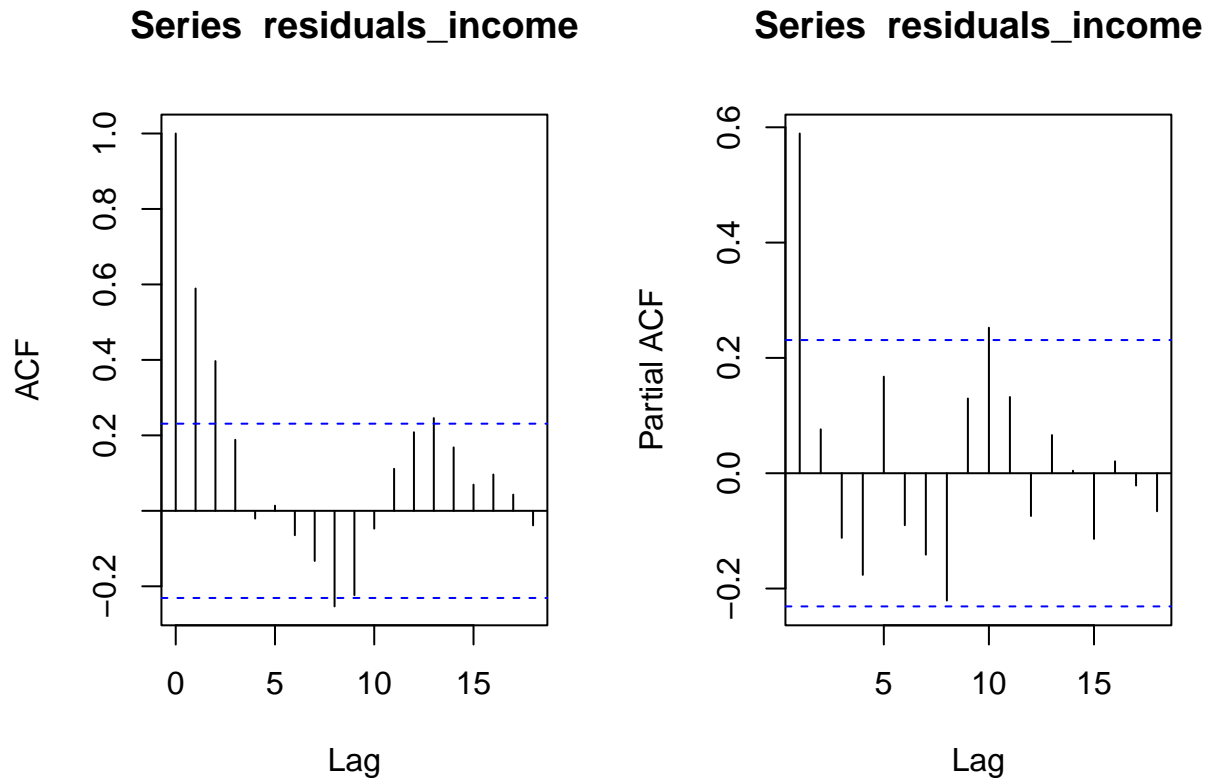
## [1] 0.05830687

##
## Box-Pierce test
##
## data: residuals_income
## X-squared = 52.892, df = 12, p-value = 4.305e-07

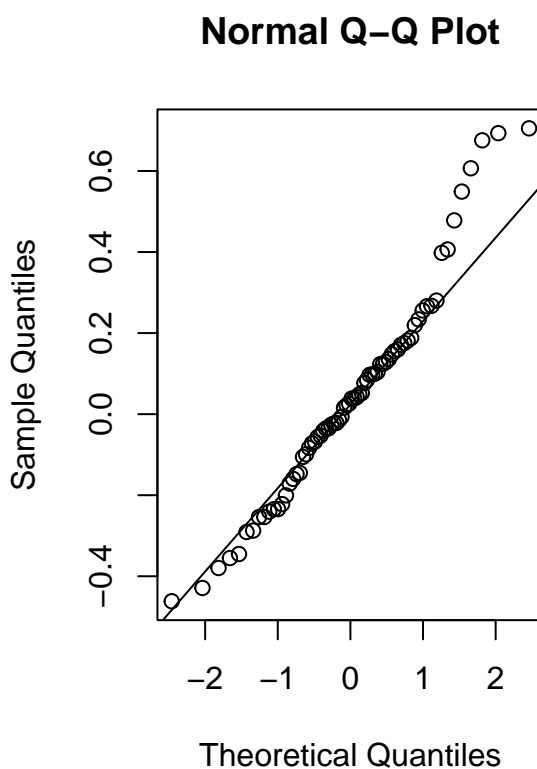
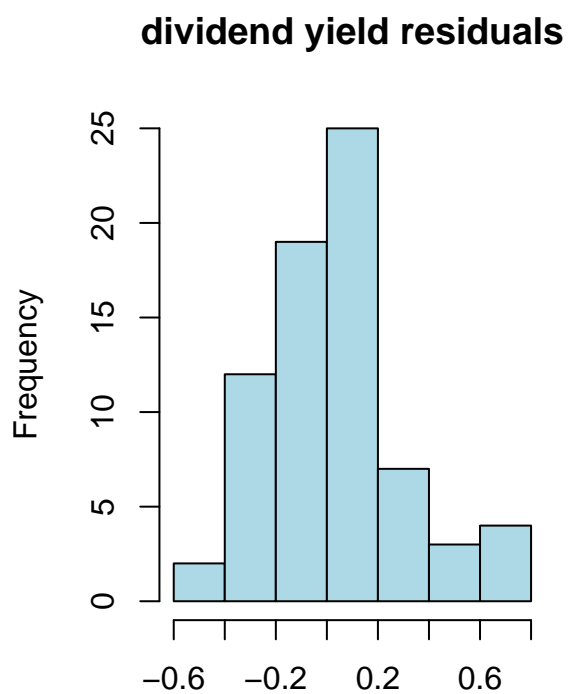
##
## Box-Ljung test
##
## data: residuals_income
## X-squared = 57.297, df = 12, p-value = 6.985e-08

##
## Box-Ljung test
##
## data: residuals_income^2
## X-squared = 20.905, df = 12, p-value = 0.05179
```

```
##
## Shapiro-Wilk normality test
##
## data: residuals_income
## W = 0.99271, p-value = 0.955
```



We see that the plot of the residuals seems to display a bit of a trend, change of variance, and seasonality. The plot of the histogram is fairly symmetric and the QQ plot seems to be normally distributed. The net income model does not pass the Box-Pierce and the Ljung-Box tests but passes the McLeod-Li and Shapiro-Wilk tests when observing a significance level of 0.05. In addition, the ACF and PACF of the residuals are not all quite within the confidence intervals so they do not resemble white noise. Therefore we may have to hold off on this model for the time being as a candidate for our time series forecasting.



```
## [1] 0.04189551

## [1] 0.06670357

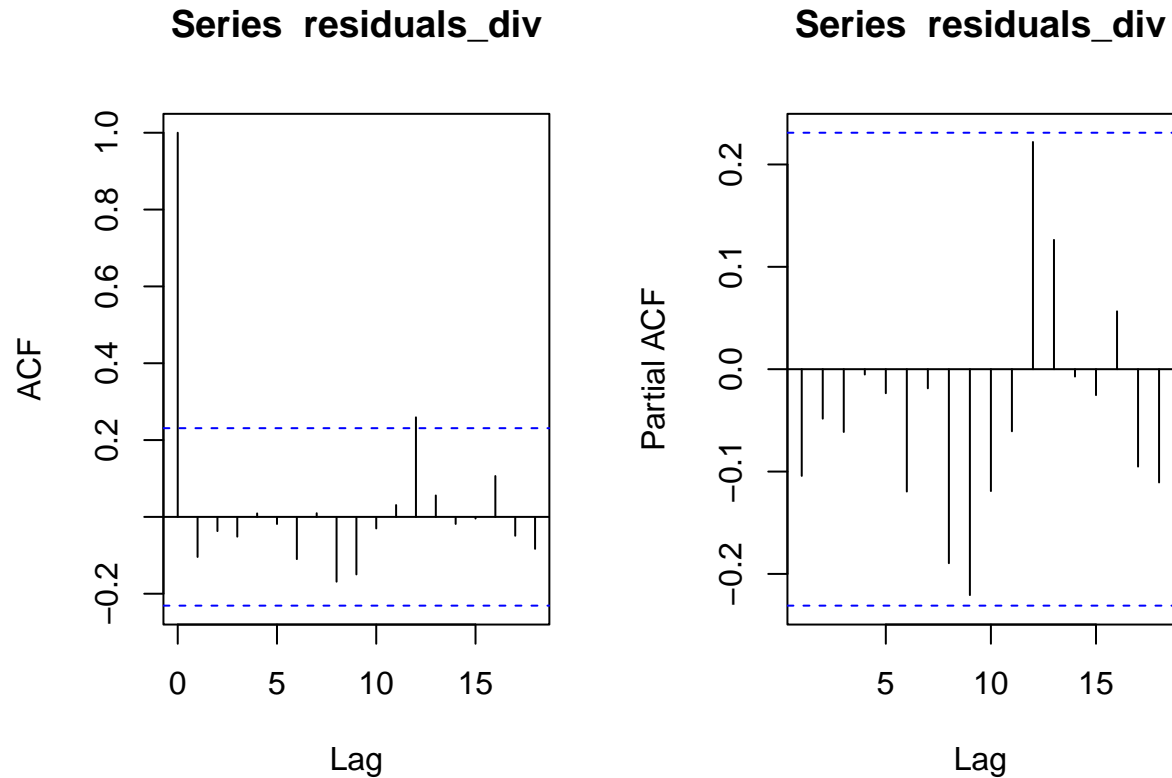
##
## Box-Pierce test
##
## data: residuals_div
## X-squared = 10.611, df = 12, p-value = 0.5625

##
## Box-Ljung test
##
## data: residuals_div
## X-squared = 12.531, df = 12, p-value = 0.404

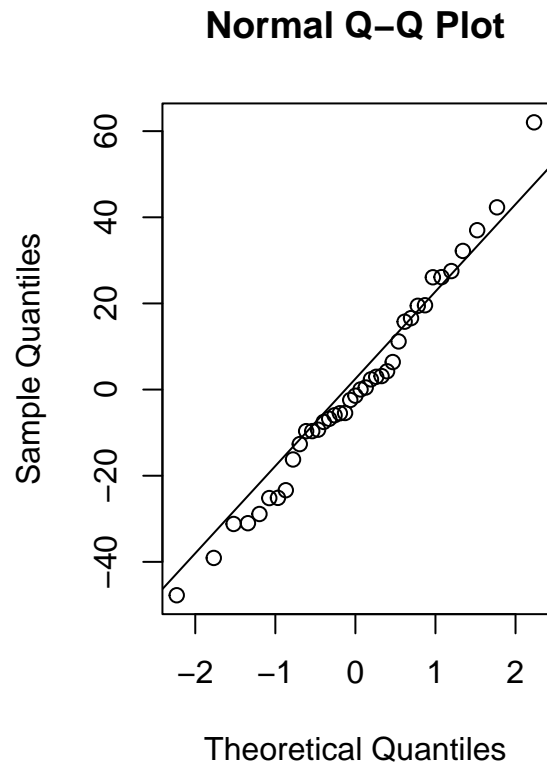
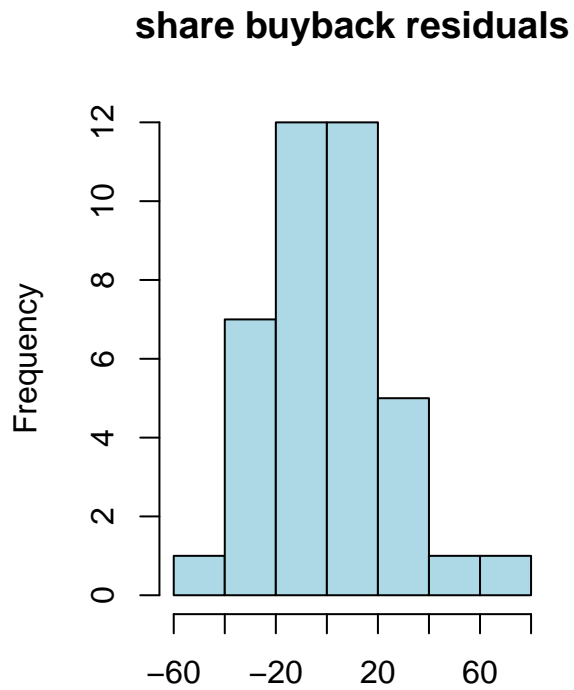
##
## Box-Ljung test
##
## data: residuals_div^2
## X-squared = 8.5677, df = 12, p-value = 0.7393

##
## Shapiro-Wilk normality test
```

```
##  
## data: residuals_div  
## W = 0.96389, p-value = 0.03679
```



We see that the plot of the residuals seems to display a bit of a trend, change of variance, and seasonality. The plot of the histogram is fairly symmetric and the QQ plot seems to be normally distributed. The dividend yield model passes the Box-Pierce, the Ljung-Box, and McLeod-Li tests but does not pass the Shapiro-Wilk test when observing a significance level of 0.05. In addition, the ACF and PACF of the residuals are not all quite within the confidence intervals so they do not resemble white noise. Therefore we may have to hold off on this model for the time being as a candidate for our time series forecasting.



```
## [1] 0.2882012

## [1] 551.184

##
## Box-Pierce test
##
## data: residuals_buyback
## X-squared = 4.6923, df = 12, p-value = 0.9675

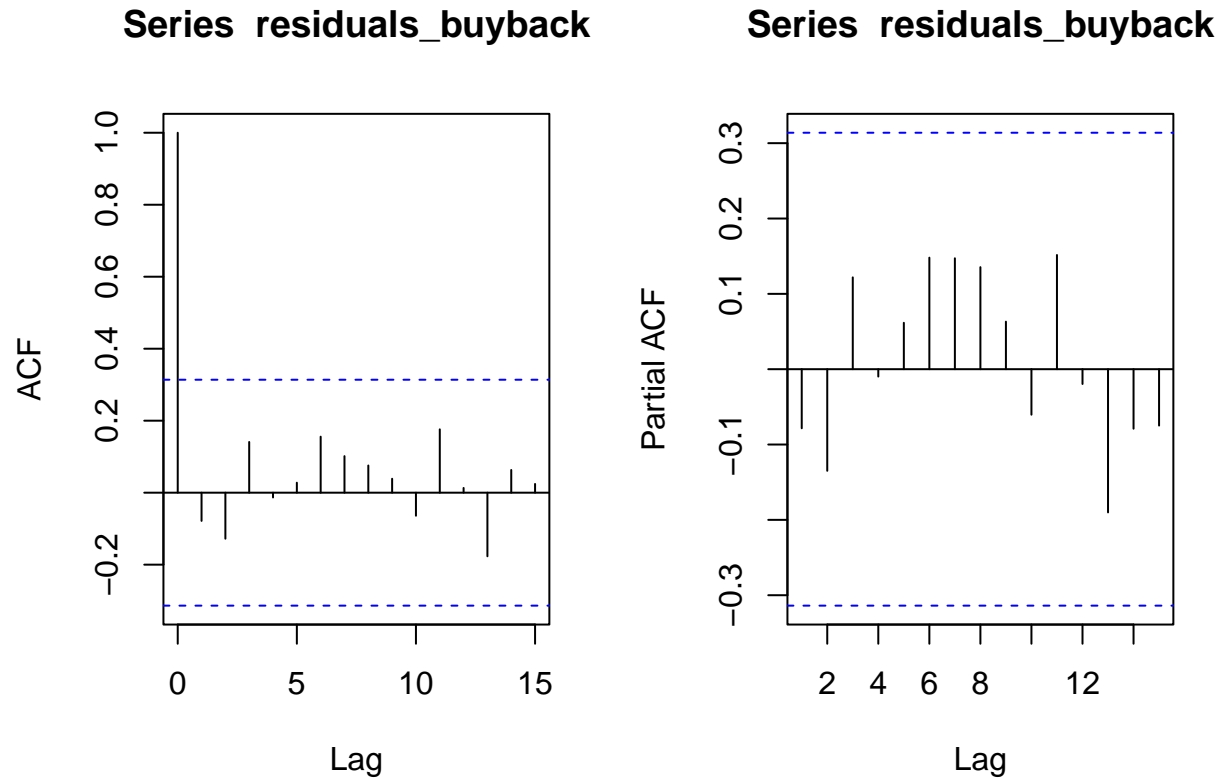
##
## Box-Ljung test
##
## data: residuals_buyback
## X-squared = 5.9612, df = 12, p-value = 0.918

##
## Box-Ljung test
##
## data: residuals_buyback^2
## X-squared = 16.13, df = 12, p-value = 0.1854

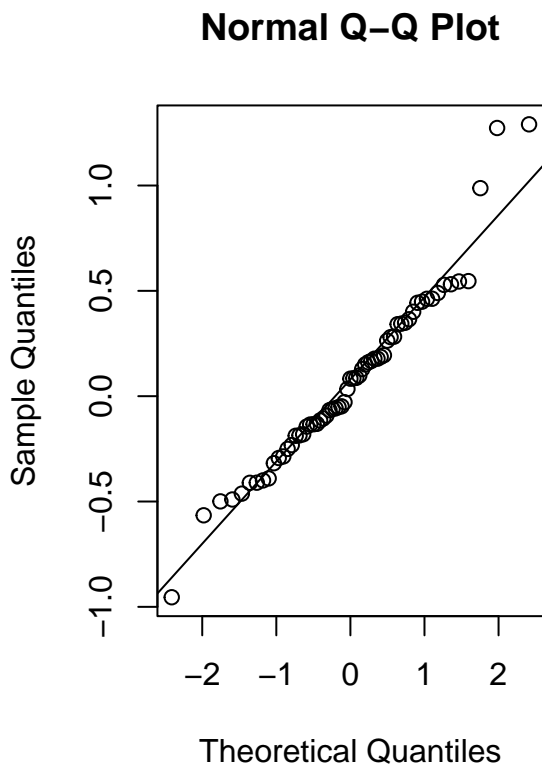
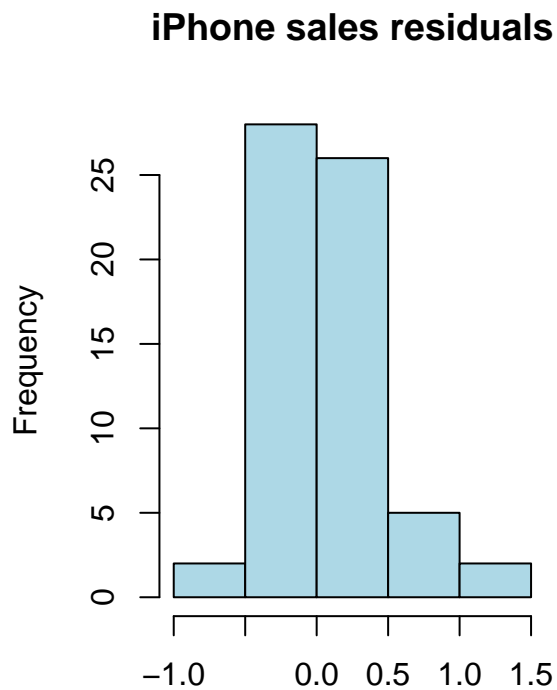
##
## Shapiro-Wilk normality test
```



```
##
## data: residuals_buyback
## W = 0.98518, p-value = 0.8779
```



We see that the plot of the residuals seems to display a bit of a trend, change of variance, and seasonality. The plot of the histogram is fairly symmetric and the QQ plot seems to be normally distributed. The share buyback model passes all of the following Box-Pierce, Ljung-Box, McLeod-Li, and Shapiro-Wilk tests when observing a significance level of 0.05. In addition, the ACF and PACF of the residuals are all within the confidence intervals so they resemble white noise. Therefore, the share buyback time series model is a good candidate for our time series forecasting analysis.



```
## [1] 0.07418759

## [1] 0.1688985

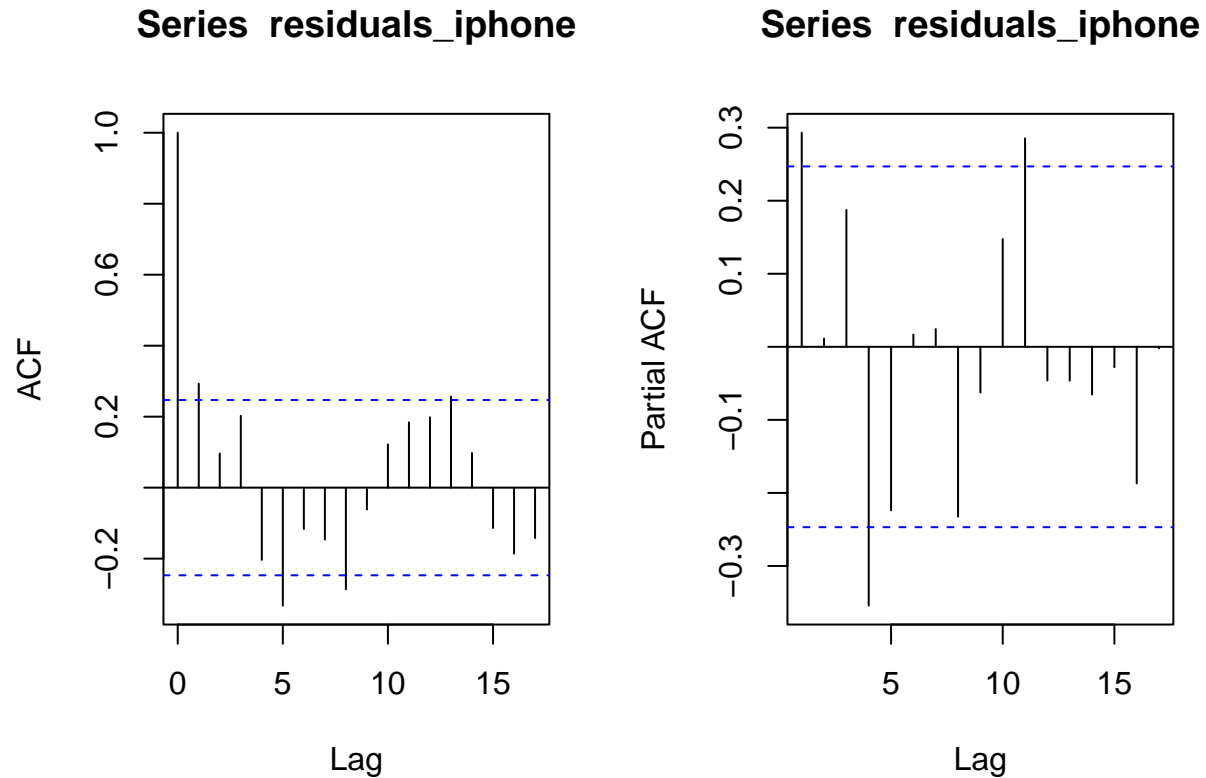
##
## Box-Pierce test
##
## data: residuals_iphone
## X-squared = 31.309, df = 12, p-value = 0.001767

##
## Box-Ljung test
##
## data: residuals_iphone
## X-squared = 35.682, df = 12, p-value = 0.0003644

##
## Box-Ljung test
##
## data: residuals_iphone^2
## X-squared = 16.153, df = 12, p-value = 0.1843

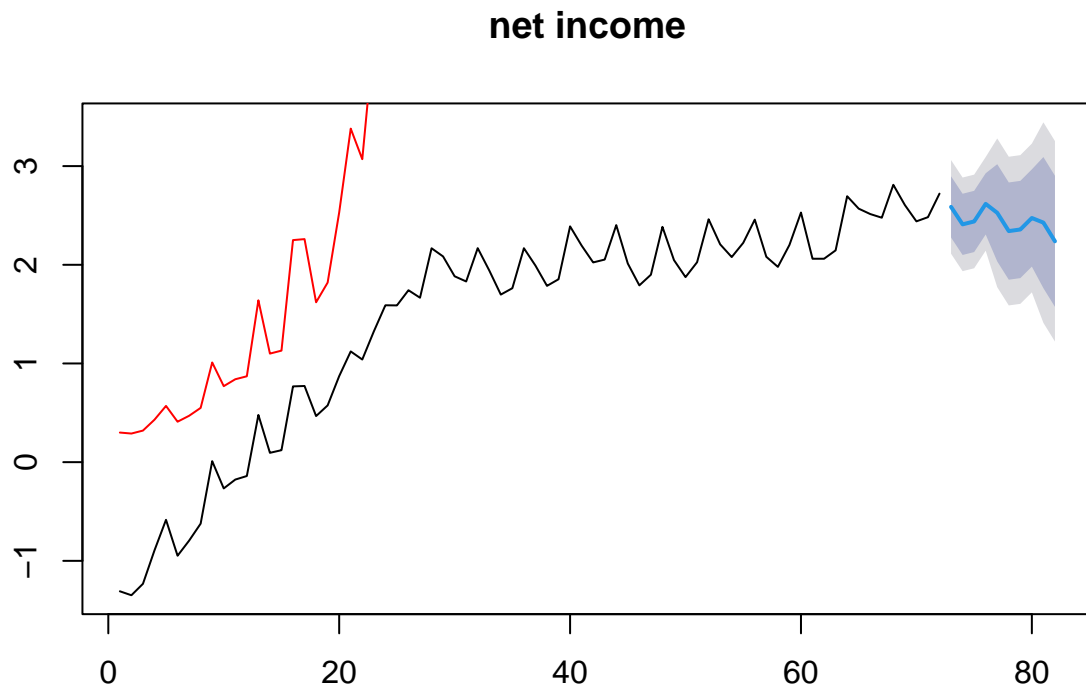
##
## Shapiro-Wilk normality test
```

```
##
## data: residuals_iphone
## W = 0.96258, p-value = 0.05269
```



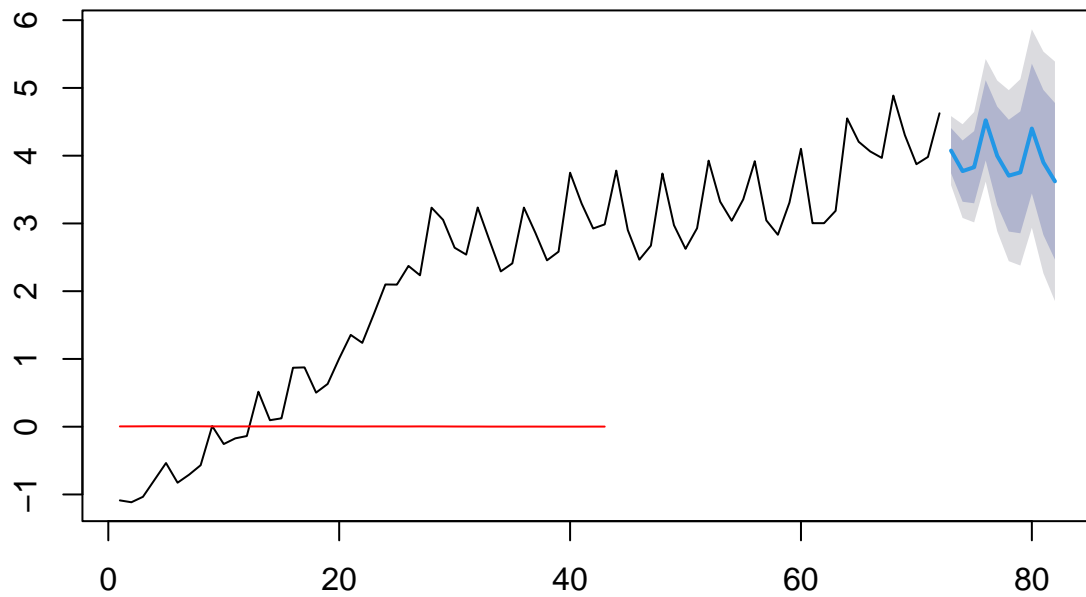
We see that the plot of the residuals seems to display a bit of a trend, change of variance, and seasonality. The plot of the histogram is fairly symmetric and the QQ plot seems to be normally distributed for the most part. The iphone sales model does not pass the Box-Pierce and the Ljung-Box tests but passes the McLeod-Li and Shapiro-Wilk tests when observing a significance level of 0.05. In addition, the ACF and PACF of the residuals are not all quite within the confidence intervals so they do not resemble white noise. Therefore we may have to hold off on this model for the time being as a candidate for our time series modeling.

Through the above analysis for each model, although we found some flaws in three of them including not passing certain tests for correlation, and thereafter the observation in autocorrelations, we will run all of the models for forecasting and see how it turns out.



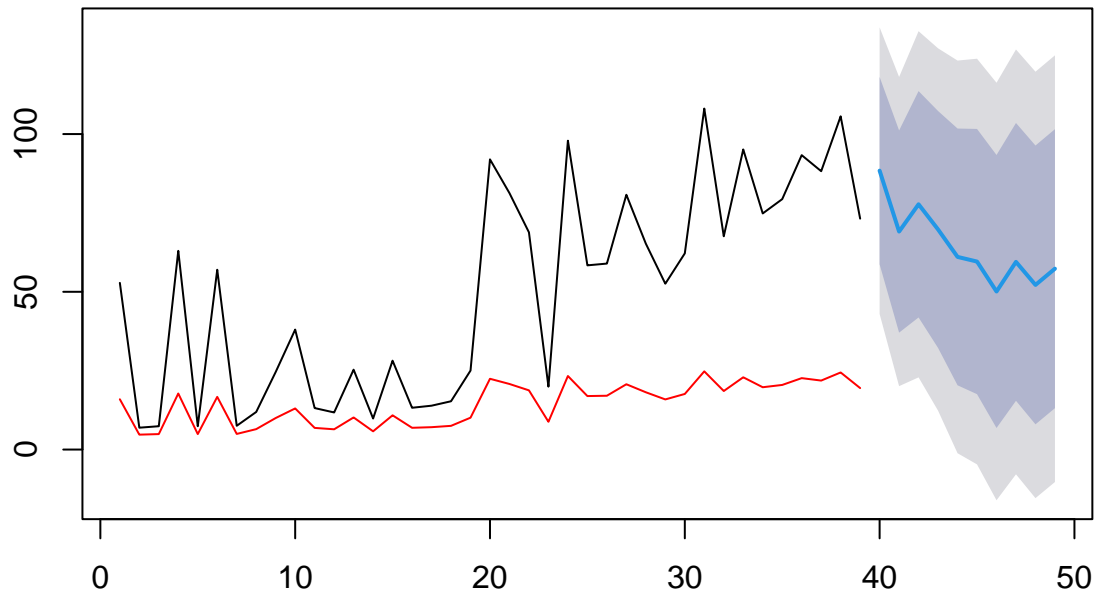
According to the chart displayed above, we can see that the line will stay relatively flat but trend slightly lower over the next 10 quarters, with a 95% level of confidence. It seems like the model accounted fairly well for the seasonality which is explained by Apple releasing a new iPhone every Fall, as it makes up a great portion of its revenues. However, we can only naturally predict that the net income will increasingly show an uptrend which our model forecasting does not exactly display.

dividend yield as a proportion of share price



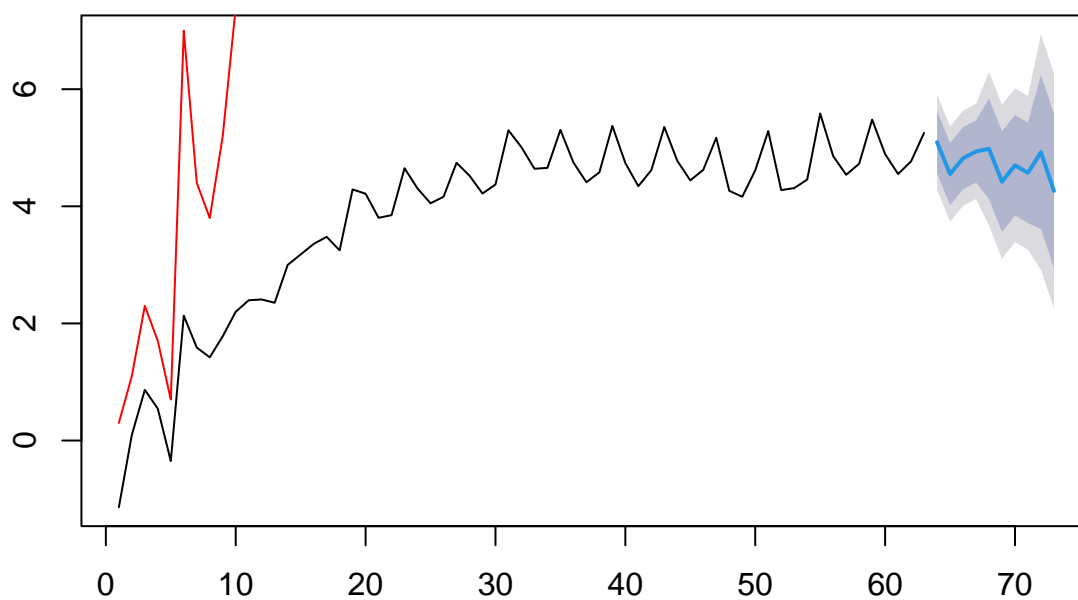
The forecast for the dividend yield as a proportion of the stock price at its declaration date shows a very slight downtrend as well. In the context of our data, this means that as the share price rises, the dividend yield will drop a little bit which implies that revenue reinvestment into segments such as research and development will be more prevalent in the coming 10 quarters.

share buybacks



The forecasting shown for the share buybacks displays a drastically lower trend for the next 10 quarters. In the context of our data, this means that the company will look to reinvest company profits into innovation as mentioned before rather than rewarding shareholders. The forecasting range explained by the 95% confidence interval observed is quite wide as well, so it is important to take note of that.

iPhone sales



iPhone sales are projected to stay stagnant in the next 10 quarters to come, and if we take a look at the previous quarters, we have seen a similar trend while observing seasonality every four quarters which accounts for the timely release of a new product.

Discussion and Conclusions

Previously, we conducted a comprehensive time series modeling analysis on the different segments of Apple, Inc.'s business that make it the technology symbol that it is today. We performed numerous transformations, calculations, interpretations of summary statistics and hypotheses and their corresponding tests, in order to fit the adequate models for forecasting these aforementioned segments to 10 quarters ahead. Variance dependent ARMA models such as the GARCH model, or an extension of the AR model, ARDL time series regression models were briefly considered, but we ended up with SARIMA to properly account for seasonality. The SARIMA models we determined by evaluating for the MA and AR polynomial roots and coefficients gave us an idea of what this would look like mathematically and graphically. Out of these models tested, the share buybacks SARIMA $(0, 0, 0) \times (2, 0, 1)_4$ model was the most promising because we found out that the modulus of its characteristic polynomial roots were outside of the unit circle, thus made it casual and invertible. It also passed all of the hypothesis tests that we decided to perform including the Box tests for serial autocorrelations, McLeod-Li test for linearity, and Shapiro-Wilk test for normality by comparison of p-values with a significance level of 0.05. When looking at the graph of residuals, we saw the least amount of variability in the mean and variance when taken across a series of time. The last straw came when we observed the ACF and PACF plots for the residuals of the share buybacks and saw that none of the lags touched the line indicating the significance level, indicating that we were looking at a white noise process. The twist came when we used the model to forecast future values for share buybacks, and saw that the range indicating the possible values within the 95% and 90% confidence intervals was the widest among all of our models, which raised questions regarding practicality and real life implications that these results would pertain to in terms of Apple's business decisions. We came to the conclusion that the predicted values were to trend lower as a result of the company reinvesting more of its profits back into research and development rather than rewarding shareholders to a similar extent. On an additional note, the other models performed decently, as it was clear that there were trade-offs when observing the range of the confidence intervals as mentioned before. The forecast for net income, dividend yield, and iPhone sales showed that the numbers would stagnate, almost show a decreasing trend for the next 10 quarters, with some seasonality around the third quarter. All in all, the share buyback SARIMA model was used to forecast future values of company profit re-investments having the highest AIC value, however the residuals resembled white noise, and it passed all of the diagnostic checks. We see that it has successfully predicted values close to Wall Street analysts' estimates for future quarters. Thus our model met Box-Jenkins methodology and assumptions for forecasting time series and our goal was achieved.

Bibliography and Appendix

The following are the resources I consulted in the process of undertaking this analysis.

Chatfield C. "The analysis of time series: an introduction." Chapman and hall/CRC; 2003 Jul 29. (7th Ed.)

Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. OTexts; 2018 May 8.

Cruz, M.G., Peters, G.W. and Shevchenko, P.V., 2015. Fundamental aspects of operational risk and insurance analytics: A handbook of operational risk. John Wiley & Sons.

Shumway, R.H. and Stoffer, D.S. (2017). Time series analysis and its applications : with R examples. Cham, Switzerland: Springer.

Appendix

An extra note on SARIMA models

The SARIMA model, short for Seasonal Autoregressive Integrated Moving Average is a powerful time series forecasting model that takes into account for seasonality in data. The SARIMA model has several parameters as listed in the following:

SARIMA(p,d,q)(P,D,Q)m, where

p/P refers to nonseasonal/seasonal AR terms, d/D refers to nonseasonal/seasonal differencing, q/Q refers to nonseasonal/seasonal MA terms, and m refers to the time steps per season. Next, we will denote the time series representations for each term in order to derive the general form for a SARIMA model.

SAR: $\phi_p(B^m)Y_t = (1 - \phi_1 B^m - \phi_2 B^{2m} - \dots - \phi_p B^{pm})Y_t$, where ϕ_k is the seasonal AR coefficient at lag k.

D: $Y_t = (1 - B)^d Y_t$, where B denotes the backshift operator and d represents the number of differences to make the series stationary.

Seasonal D: $Y_t = (1 - B^m)^D Y_t$

SMA: $\theta_q(B)\varepsilon_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)\varepsilon_t$, where θ_k are the MA coefficients at lag k.

Using these, the SARIMA model can be written as:

$$(1 - \phi_1 B^m - \phi_2 B^{2m} - \dots - \phi_P B^{Pm})(1 - B^d)(1 - B^{Dm})(Y_t - \mu_t) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)(1 + \theta_1 B^m + \theta_2 B^{2m} + \dots + \theta_Q B^{Qm})\varepsilon_t$$

where B is the backshift operator and μ_t represents the time series mean.

On the left-hand side is the seasonal AR component, differencing, and seasonal differencing components. The right-hand side of the equation is the seasonal and nonseasonal MA components. We will now divide both sides by $(1 - \phi_1 B^m - \phi_2 B^{2m} - \dots - \phi_P B^{Pm})(1 - B^d)(1 - B^{Dm})$, which will give us:

$$Y_t - \mu_t = \frac{(1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)(1 + \theta_1 B^m + \theta_2 B^{2m} + \dots + \theta_Q B^{Qm})\varepsilon_t}{(1 - \phi_1 B^m - \phi_2 B^{2m} - \dots - \phi_P B^{Pm})(1 - B^d)(1 - B^{Dm})}$$

Expanding the denominator using the geometric series gives us:

$$Y_t - \mu_t = (1 + \theta_1 B^m + \theta_2 B^{2m} + \dots + \theta_Q B^{Qm} + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q + \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p)\varepsilon_t$$

is what is working behind the R code.