TS5

Dylan Hayashi

MSDS 413, Fall 2022, Section 55

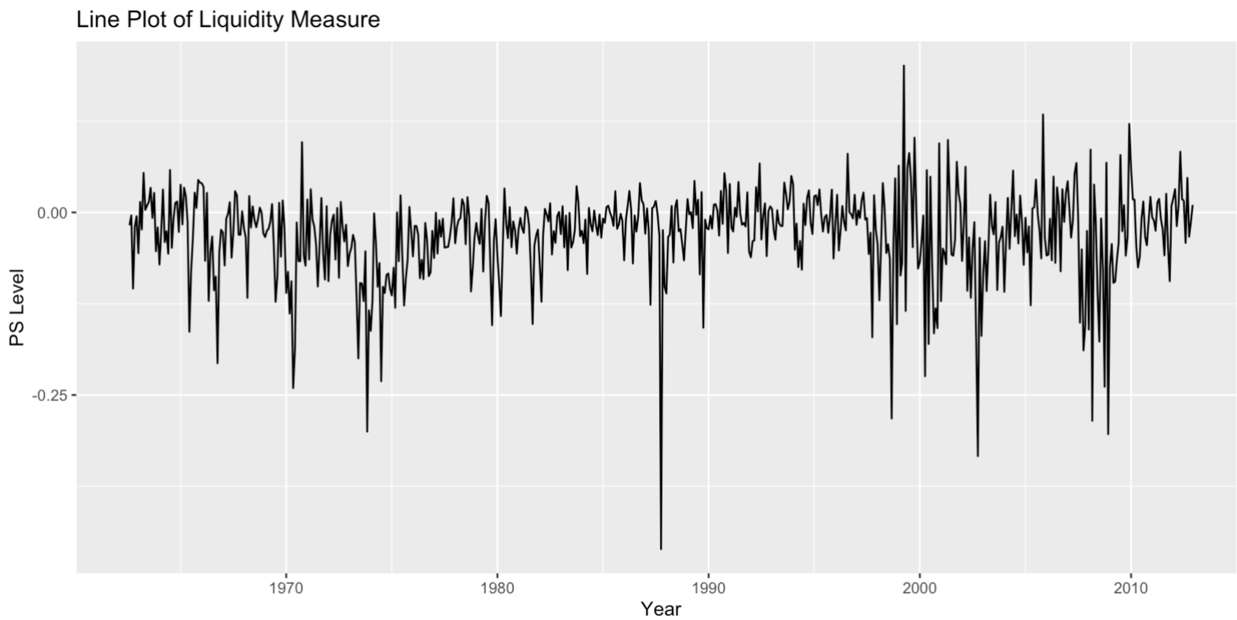
Northwestern University, Time-Series Analysis & Forecasting

October 24, 2022

1. Outlier Management

1.1 EDA

Below is a line plot of the raw liquidity measure data.



Prior to analysis, the data must be confirmed to be time series data. In order to fit the definition, the data must be a time ordered sequence of observations of a stochastic variable over constant time intervals.

To test that this is a time ordered sequence, we can confirm that the data is indexed by unique time periods and that there is an observation of liquidity measure for each time period. The vectors Date, unique(Date), and PS_LEVEL all have 605 observations, confirming that this is the case.

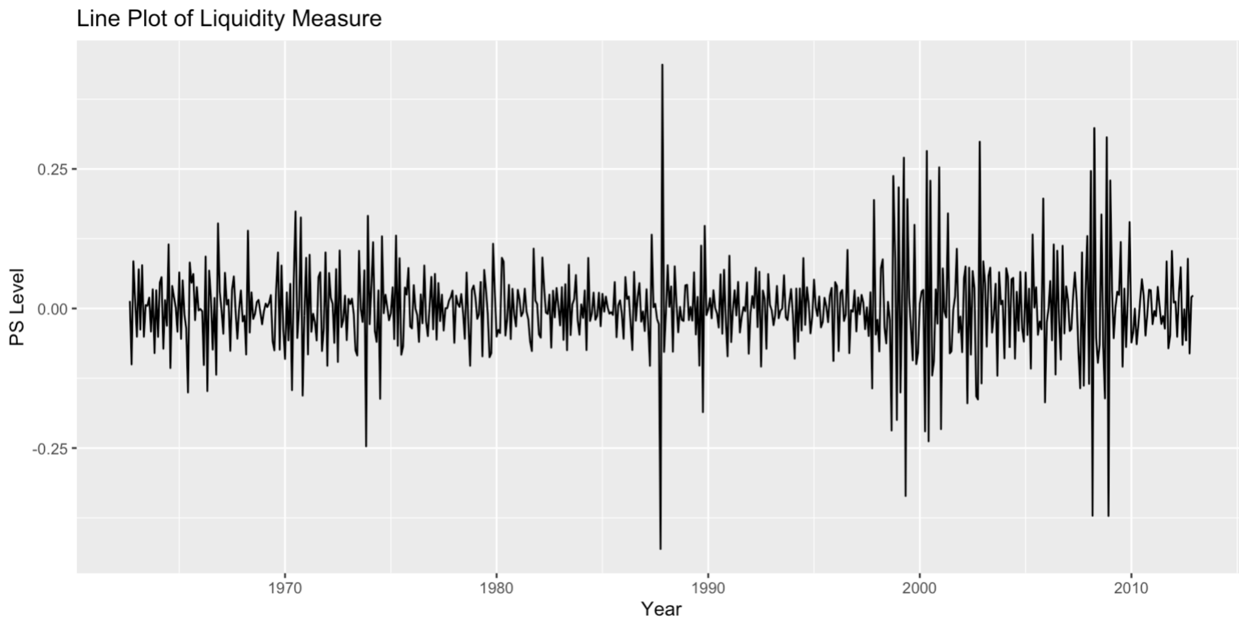
To test for constant intervals between time periods, we can sum the difference between each pair of successive periods, and this should equal one less than the total number of observations. The data are monthly data, so the difference between each date should be expressed as months. When this is done, the sum of the differences between observations is 604, one less than the index length of 605, confirming that the dataset has constant time intervals.

The data summary on the right, which is of the PS_LEVEL variable, shows that the variable has variance. Having variance is indicative of stochasticity. This confirms that the data match the definition of time-series in its entirety.

nobs	605.000000
NAs	0.000000
Minimum	-0.461007
Maximum	0.201015
1. Quartile	-0.056921
3. Quartile	0.007003
Mean	-0.030638
Median	-0.021521
Sum	-18.536186
SE Mean	0.002557
LCL Mean	-0.035661
UCL Mean	-0.025616
Variance	0.003956
Stdev	0.062900
Skewness	-1.529748
Kurtosis	6.265555

The line plot indicates that the data are not stationary. They do not appear to be mean zero, rather they appear to center around -0.02. The summary table of the variable, shown on the previous page, indicates that the mean is -0.031. A t-test of this mean returned a 95% confidence interval of (-0.036 -0.026) and a p-value of 2e-16, allowing us to reject the null hypothesis that the data is mean zero. This is sufficient for us to conclude that the data is not stationary.

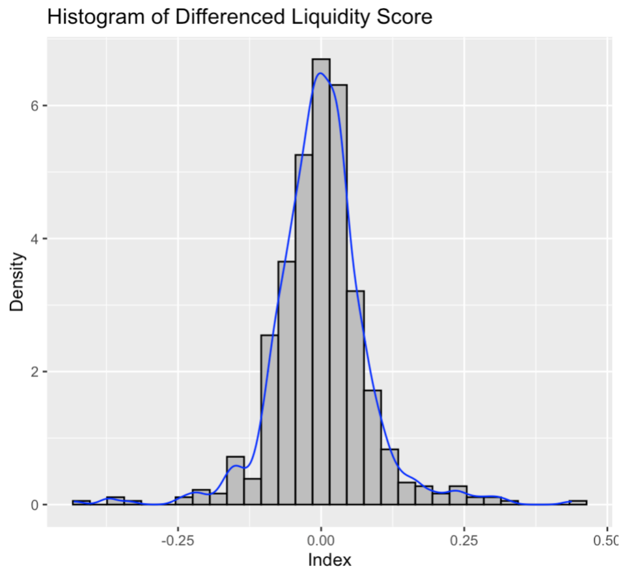
Applying first differencing to the data produces data with a line plot shown below:

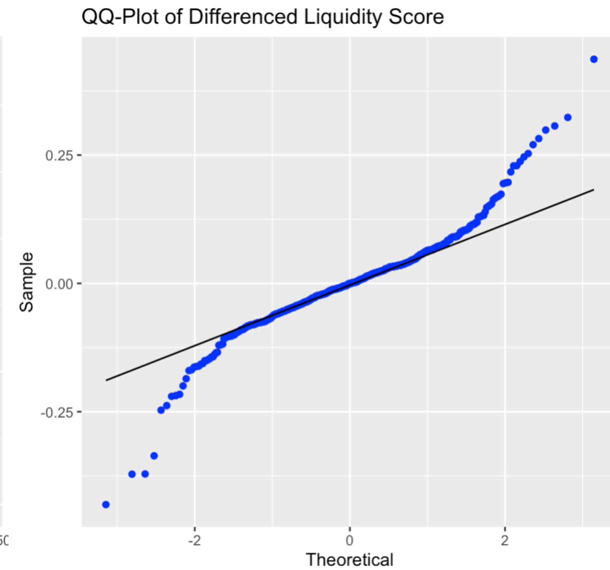


nobs	604.000000
NAs	0.000000
Minimum	-0.430884
Maximum	0.436738
1. Quartile	-0.042995
3. Quartile	0.036713
Mean	0.000046
Median	-0.000194
Sum	0.028018
SE Mean	0.003376
LCL Mean	-0.006584
UCL Mean	0.006676
Variance	0.006884
Stdev	0.082969
Skewness	0.037209
Kurtosis	4.929112

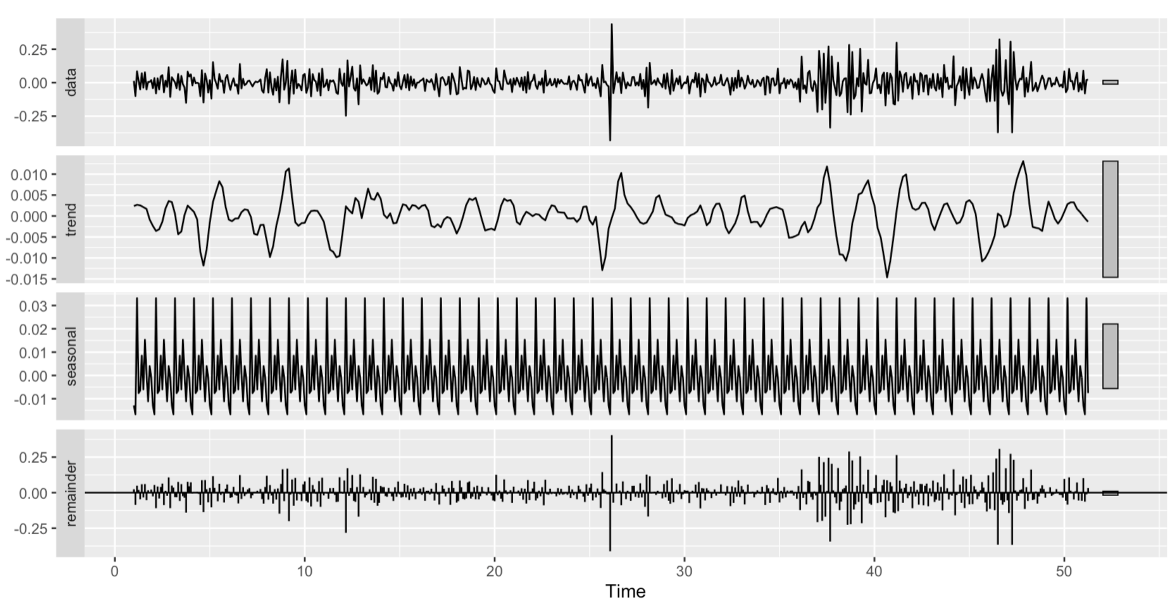
The new data are stationary. As the line plot shows, they are now centered around zero. A summary table of the data, shown to the right, indicates that the new mean is 0.0005, with a t-test 95% confidence interval of (-0.007, 0.007) and a p-value of 0.989, allowing us to accept the null hypothesis that the differenced data are mean zero. While the plot does appear to have non-constant variance, ADF and KPSS tests both substantiate stationarity and the absence of random-walk and drift, as each had a p-value/test statistic beneath the critical value required to assert linear-stationarity (ADF p-value: 0.01 < 0.05, KPSS test statistic: 0.0064 < 0.146.) Thus, I will apply first differencing while making my model.

The differenced data are not normally distributed. While their histogram, shown on the next page, displays a unimodal distribution, there is poor symmetry and almost a lean in the line of best fit. The QQ plot, also shown on the next page, more easily displays non-normality by showing how far the tails deviate from the normal line. Further, neither of the 95% confidence intervals of skew (0.01, 0.058) nor kurtosis (5.005, 5.121) contain zero. Lastly, Shapiro-Wilk, Lilliefors, and Anderson-Darling tests all indicate non-normality; their p-values were all 2.2e-16, allowing us to reject their null hypotheses of normality. Unfortunately, log transformation cannot be applied to the data due to there being negative values.



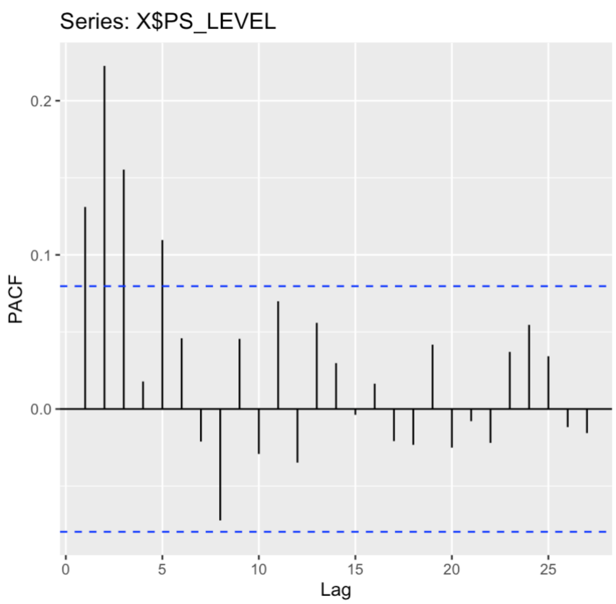


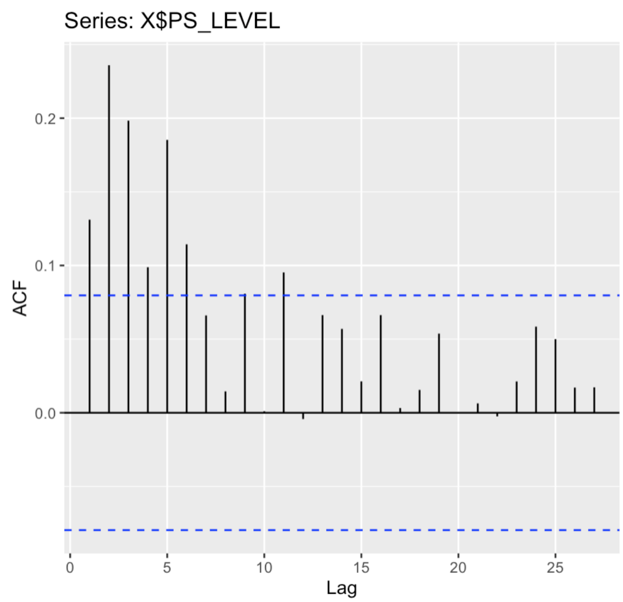
Below is a decomposition of the data. The trend line is a moderately smoothed version of the time plot that keeps the pattern of large variance changes. Based on the time plot, I imagine the proper trend line would be more or less flat and falling along the line y = 0. That is not the case, which indicates that the decomposition failed to explain variance in a manner sufficient to create this ideal trend line. The decomposition indicates that there is a seasonal pattern to the data. It appears to have a large spike in the beginning of the pattern, followed by a number of smaller and very frequent fluctuations. The remainder section of the decomposition looks quite similar to the time plot, indicating that the decomposition didn't explain much of the data.



1.2 Model Construction

Below are an ACF and PACF plot of the raw data. The PACF plot shows three significant lags, indicating that an AR(3) component is appropriate. The ACF plot shows six significant lags, indicating that an MA(6) component is appropriate.





As previously discussed, I will also be applying first differencing in my model. Thus, this would be an ARIMA(3,1,6) model, which would take the general form:

$$x_t = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \theta_3 z_{t-3} + \theta_4 z_{t-4} + \theta_5 z_{t-5} + \theta_6 z_{t-6}$$

The specific form that this model takes is shown in the output below. It has eleven degrees of freedom.

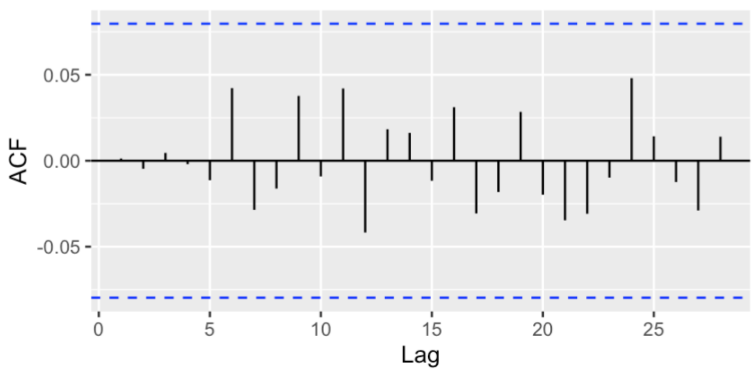
```
ARIMA(3,1,6)

Coefficients:
         ar1     ar2      ar3      ma1     ma2     ma3      ma4     ma5      ma6
      0.3734  0.3347  -0.3340  -1.3082  0.1485  0.5616  -0.4470  0.2407  -0.1733
s.e.  0.2540  0.2058   0.2123   0.2506  0.3313  0.3161   0.1978  0.0833   0.0546

sigma^2 = 0.0036:  log likelihood = 845.57
AIC=-1671.14   AICc=-1670.76   BIC=-1627.1
```

Overall, the residuals of this model indicate that it is a decent fit of the data: they are stationary and do not feature auto-correlation. However, they are not normally distributed, and they feature business cycles.

Stationarity of the residuals is demonstrated by both their summary statistics and statistical tests. The mean, 0.000167, had a t-test 95% confidence interval of (-0.005, 0.005) and an accompanying p-value of 0.94, allowing us to accept the null hypothesis that the residuals are mean zero. The residuals also show constant variance, demonstrated by results of a McLeod-Li test being an empty set, indicating that no lags showed non-constant variance. Thus, the residuals are mean zero and have constant variance, therefore fitting the definition of stationary. This notion, as well as the absence of random-walk and drift, are corroborated by ADF and KPSS tests, as each had a p-value/test statistic beneath the critical value required to assert stationarity (ADF p-value: 0.01 < 0.05, KPSS test statistic: 0.0802 < 0.146.).

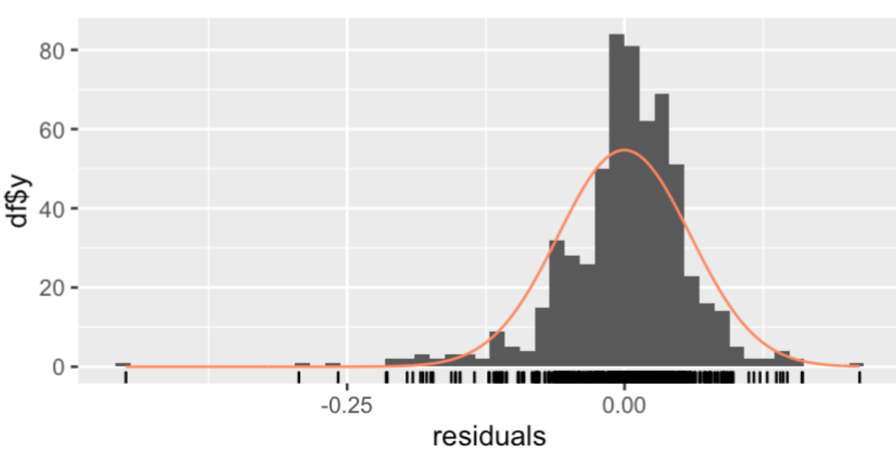
nobs	605.000000
NAs	0.000000
Minimum	-0.449432
Maximum	0.211874
1. Quartile	-0.023562
3. Quartile	0.034129
Mean	-0.000167
Median	0.005773
Sum	-0.101281
SE Mean	0.002421
LCL Mean	-0.004922
UCL Mean	0.004587
Variance	0.003546
Stdev	0.059548
Skewness	-1.458690
Kurtosis	7.464408

Lack of auto-correlation in the residuals is demonstrated in their ACF plot, as well as by a Box-Ljung. The ACF, shown to the right, indicates that there are no lags with significant auto-correlation. Similarly, a Box-Ljung test returned a p-value of 0.9454, allowing the acceptance of the null hypothesis of independence.



Lack of normality in the residuals is demonstrated both visually and by statistical tests. Their histogram, shown below, demonstrates a skewed distribution (actual value -1.458.) Neither of the 95% confidence intervals for skew (-1.527, -1.482) nor excess kurtosis (7.684, 7.994) contain zero. Lastly, the p-values of Shapiro-Wilk, Lilliefors, and Anderson-Darling were all 2e-16, far below the required level to reject their null hypotheses of normality at any commonly used confidence level.

As previously mentioned, the residuals have business cycles. Their respective frequencies are 8.387, 2.698, 12.708, and 3.873 months. This indicates that the model failed to explain patterns in the data.



1.3 Outlier Identification

The largest outlier of the time series occurs at t = 303. Adding an outlier vector to the model to assess the effect of this outlier on the liquidity measure produces a model that takes the following general form:

$$x_t = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \theta_3 z_{t-3} + \theta_4 z_{t-4} + \theta_5 z_{t-5} + \theta_6 z_{t-6} + \beta x_t$$

The precise form of this model is shown below. It has twelve degrees of freedom.

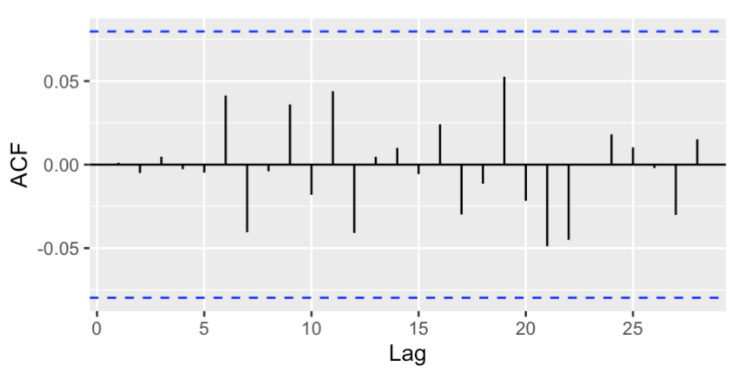
```
Series: X[, 2]
Regression with ARIMA(3,1,6) errors

Coefficients:
         ar1     ar2      ar3      ma1      ma2     ma3      ma4     ma5      ma6     xreg
      0.1924  0.3371  -0.2746  -1.1215  -0.0280  0.5300  -0.3918  0.2192  -0.1767  -0.4265
s.e.  0.2303  0.1962   0.1767   0.2264   0.2641  0.2395   0.1631  0.0779   0.0567   0.0545

sigma^2 = 0.003275:  log likelihood = 874.72
AIC=-1727.43   AICc=-1726.99   BIC=-1678.99
```

The residuals of this model indicate it is not a great fit of the data. The residuals are not strictly stationary, nor are they normally distributed, and they also have business cycles. However, they are mean zero and do not feature auto-correlation.

Lack of strict stationarity is indicated by a McLeod-Li test, which returned a set with every tested lag, indicating non-constant variance. However, the residuals are mean zero. As the summary table on the right shows, their mean was -0.0008, which resulted in a t-test of 95% confidence interval of (-0.005, 0.004) with a p-value of 0.9726. Thus, the mean is statistically indistinguishable from zero. Linear-trend stationarity is indicated by ADF and KPSS tests, as each had a p-value/test statistic beneath the critical value required to assert linear-trend stationarity (ADF p-value: 0.01 < 0.05, KPSS test statistic: 0.0796 < 0.146.).

```
nobs         605.000000
NAs            0.000000
Minimum       -0.293288
Maximum        0.209029
1. Quartile   -0.025147
3. Quartile    0.033910
Mean          -0.000079
Median         0.004799
Sum           -0.047967
SE Mean        0.002307
LCL Mean      -0.004610
UCL Mean       0.004452
Variance       0.003220
Stdev          0.056748
Skewness      -0.870960
Kurtosis       3.138526
```

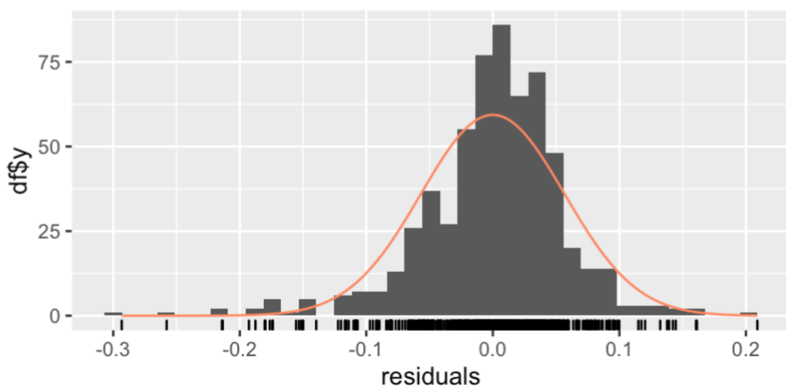
Lack of auto-correlation in the residuals is demonstrated in their ACF plot and a Box-Ljung test. The ACF plot, shown below to the right, shows that there are no lags with significant auto-correlation. A Box-Ljung of the null hypothesis of independence of observations returned a p-value of 0.9454,



allowing us to accept non-auto-correlation.

Lack of normality in the residuals is demonstrated more easily by statistical tests than visually. As the histogram below shows, the residuals of this model are less skewed than for the previous, and do appear to be normally distributed. However, neither the 95% confidence interval for skew (-0.887, -0.855) nor excess kurtosis include zero (3.234, 3.305). Further, the p-values for Shapiro-Wilk, Lilliefors, and Anderson-Darling were all less than 0.001, allowing us to reject their null hypotheses of normality.

As previously mentioned, the residuals have business cycles. Their respective frequencies are 7.730 , 2.728 , 4.085 , 2.053, and 14.755 months. These business cycles indicate patterns in the data that the model was not able to capture.



1.4 Model Reduction

Performing the parameter test function on my model indicated that none of the AR variables were significant, and neither was the MA_3. Removing those variables from my model creates a model of the general form:

$$y_t = c + \theta_1 z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \theta_4 z_{t-4} + \theta_5 z_{t-5} + \theta_5 z_{t-6} + \beta x_t$$

The precise form of this model is shown below. It has seven degrees of freedom.

```
Series: X[, 2]
Regression with ARIMA(0,1,6) errors

Coefficients:
          ma1     ma2  ma3      ma4     ma5      ma6     xreg
      -0.9368  0.1206    0  -0.1308  0.1739  -0.1828  -0.4320
s.e.   0.0398  0.0455    0   0.0499  0.0609   0.0426   0.0545

sigma^2 = 0.003272:  log likelihood = 873
AIC=-1732.01   AICc=-1731.82   BIC=-1701.18
```

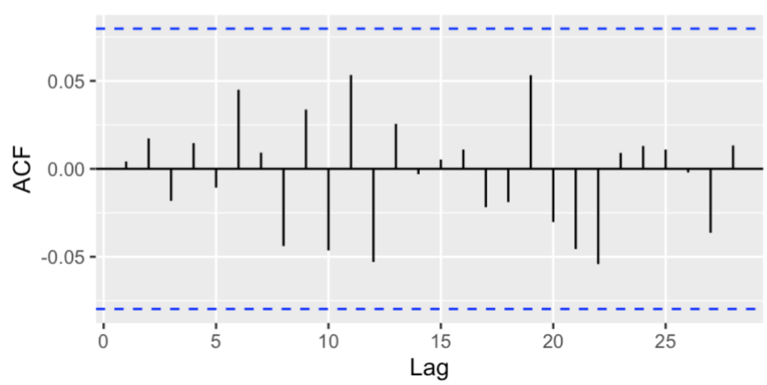
The residuals of this model are quite similar to those of the previous model; they are not strictly stationary, nor normally distributed, have business cycles, but do have mean zero, and do not feature auto-correlation.

As the summary table on the right shows, their mean was -0.0008, which resulted in a t-test of 95% confidence interval of (-0.005, 0.004) with a p-value of 0.9726. Thus, the mean is statistically indistinguishable from zero. Linear-trend stationarity is corroborated by ADF and KPSS tests, each of which had a p-value/test statistic beneath the critical value required to assert linear-trend stationarity. (ADF p-value: 0.01 < 0.05, KPSS test statistic: 0.074 < 0.146) However, a McLeod-Li test returned a set with every tested lag, indicating non-constant variance. Therefore, the residuals are not strictly stationary.

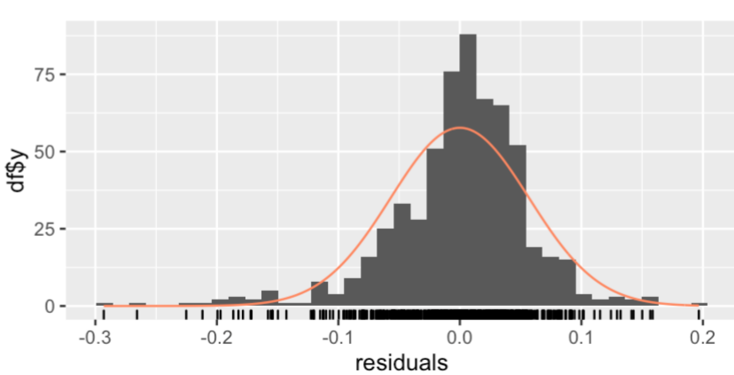
nobs	605.000000
NAs	0.000000
Minimum	-0.293083
Maximum	0.196664
1. Quartile	-0.023346
3. Quartile	0.034388
Mean	-0.000023
Median	0.005413
Sum	-0.013801
SE Mean	0.002314
LCL Mean	-0.004567
UCL Mean	0.004521
Variance	0.003239
Stdev	0.056912
Skewness	-0.942835
Kurtosis	3.270286

Lack of auto-correlation in the residuals is demonstrated in their ACF plot, pictured below to the right, which shows no lags with significant auto-correlation. This is substantiated by a Box-Ljung test, which returned a p-value of 0.7366, allowing us to accept the null hypothesis that the residuals are independent, and therefore non-auto-correlated.

As with the previous model, the residuals' lack of normality is more easily demonstrated by statistical tests than their histogram. Again, their histogram (shown below) appears to have better symmetry than the first model, but also shows somewhat of a left skew. However, neither the 95% confidence interval for skew (-0.952, -0.919) nor excess kurtosis (3.371, 3.487) include zero. Further, the p-values for Shapiro-Wilk, Lilliefors, and Anderson-Darling were all less than 0.001, indicating a rejection of their null hypotheses that the residuals have a normal distribution.



As previously stated, the residuals have business cycles, and their frequencies are 12.812, 2.428, and 4.311 months. Detectable business cycles indicate that the model was not able to capture all patterns in the data's variance.



1.5 Model Comparison

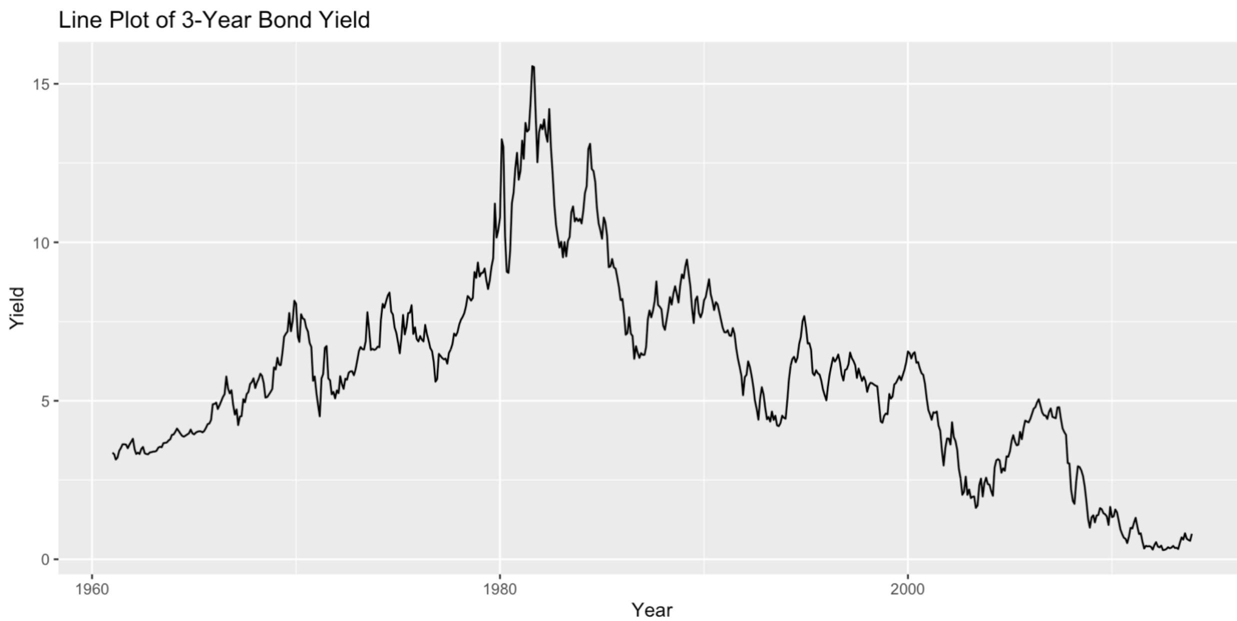
As I have described, the residuals of models 2 and 3 are quite similar. They are stationary, do not feature auto-correlation, have business cycles, and are not normally distributed. As shown below, the results of the models' backtests were also quite similar. Model 2 had better performance for periods 1 and 2, and model 3 performed better for periods 3 and 4. None of these results indicate a clear winner. However, model 3 has four fewer variables than model 2. I find this simplicity preferable, and therefore prefer model 3.

	Model 2				Model 3			
Measurement	t+1	t+2	t+3	t+4	t+1	t+2	t+3	t+4
RMSE	0.034	0.019	0.016	0.031	0.035	0.020	0.015	0.030
MAE	0.027	0.016	0.015	0.031	0.028	0.017	0.015	0.030
Bias	0.015	0.001	0.015	0.031	0.014	0.000	0.015	0.030

2. Box-Jenkins Methodology

2.1. EDA

Below is a time plot of the yield of 3-year bonds:



Prior to analysis, the data must be confirmed to be time series data. This requires the data to be a time ordered sequence of observations of a stochastic variable over constant intervals.

To test that the data are a time ordered sequence, we need to confirm that the data are indexed by unique time periods and that there are an equal number of observations of the yield variable. The vectors qdate, unique(qdate), and yield3 all have 636 observations, confirming that the data are a time-ordered sequence.

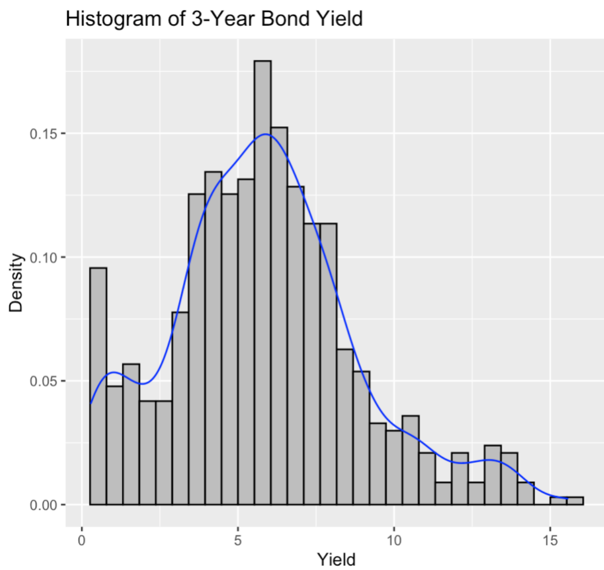
To test for constant intervals, we can sum the difference between each successive pair of periods, and this should equal one less than the total number of observations. These data are monthly, and when these difference in days between periods are converted to months, the sum of the differences is 635, confirming that the dataset has constant time intervals.

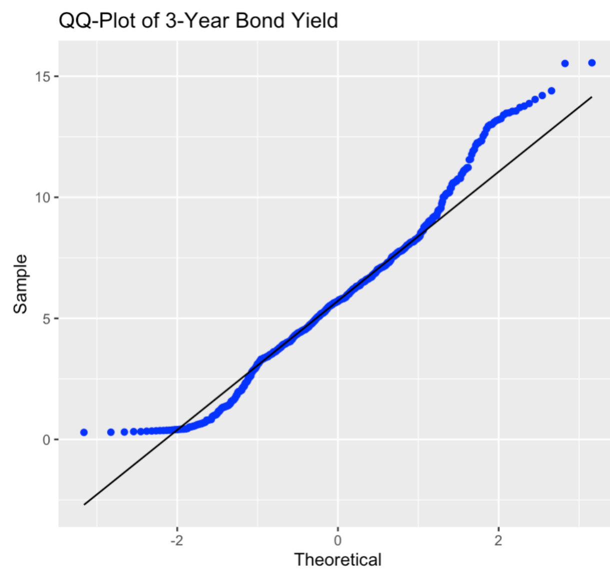
The data summary of yield3, shown on the right, indicates that the variable has variance. Having variance is indicative of stochasticity. Thus, we have confirmed that the data meet the criteria to be considered time series data.

These data are not normally distributed. As the histogram on the next page shows, their distribution is roughly unimodal and does not have good symmetry. The tails on the QQ plot,

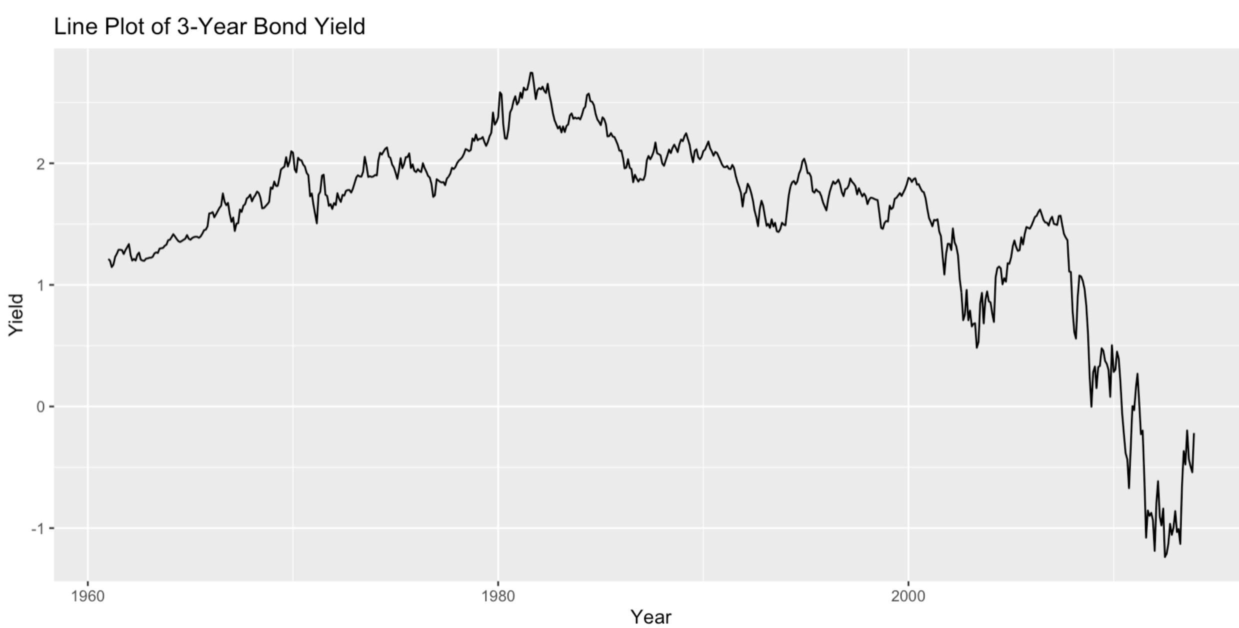
nobs	636.000000
NAs	0.000000
Minimum	0.290000
Maximum	15.557000
1. Quartile	3.926500
3. Quartile	7.523000
Mean	5.825406
Median	5.708000
Sum	3704.958000
SE Mean	0.119152
LCL Mean	5.591426
UCL Mean	6.059385
Variance	9.029418
Stdev	3.004899
Skewness	0.461118
Kurtosis	0.321249

shown below, display how the tails are non-symmetric and deviate from normal. While close to zero, neither of the 95% confidence intervals for skew (0.461, 0.471) or excess kurtosis (0.309, 0.331) include zero. As expected, the p-values for Shapiro-Wilk, Lilliefors, and Anderson-Darling tests were all less than 0.01, allowing the rejection of their null hypotheses of normality. To remediate the distribution, I will apply log transformation to the data.

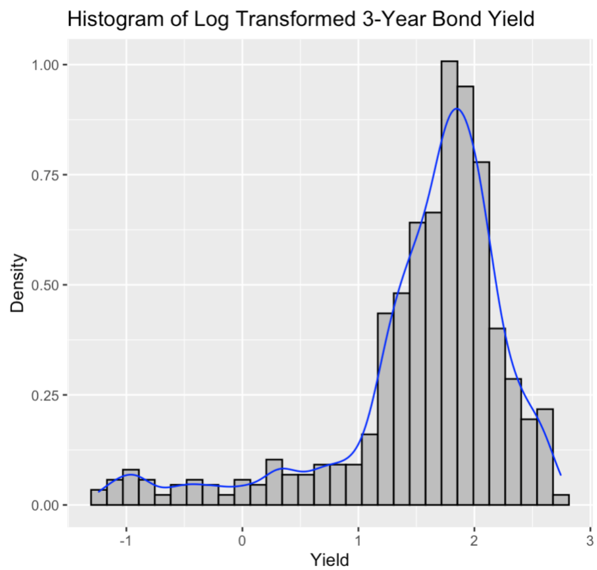


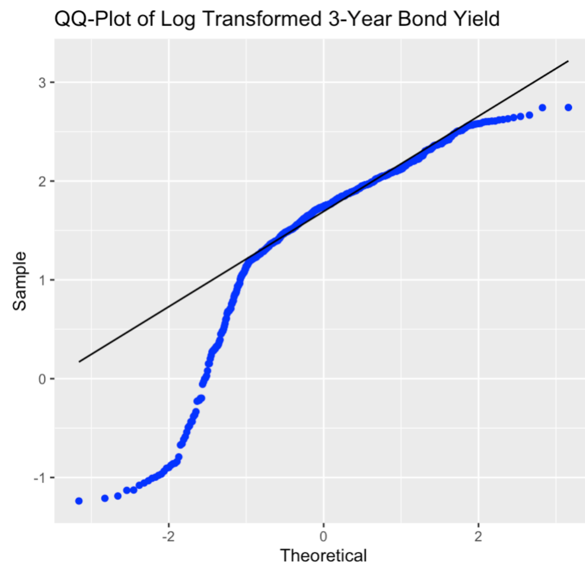


Below is a line plot of the log transformed data:



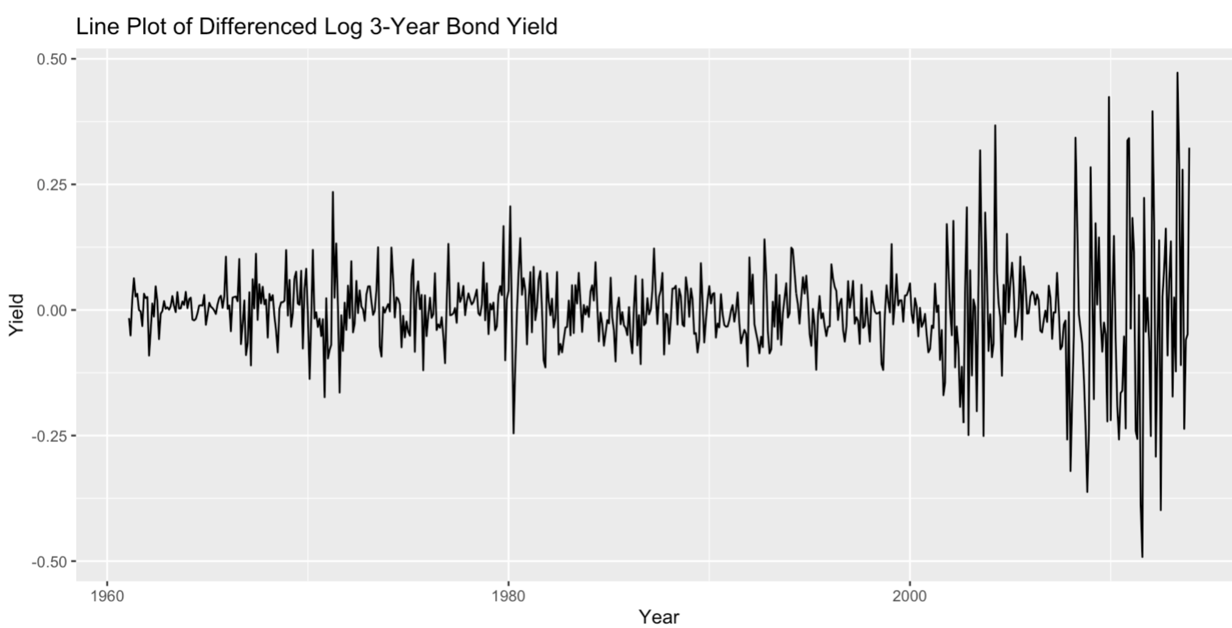
The log transformation did not make the data normally distributed. The histogram below shows that the data looks closer to unimodal, but is still skewed. The QQ Plot shows that the right tail is now closer to the normal line, but the left tail has moved further from it. The 95% confidence interval for skew (0.094, 0.141) moved closer to zero, but excess kurtosis (5.698, 5.807) moved away from zero. Shapiro-Wilk, Lilliefors, and Anderson-Darling tests still produced p-values less than 0.01, indicating that the data are still not normally distributed.





The data are also not stationary. The new mean is 1.56 with a t-test 95% confidence interval of (1.5, 1.619) and a p-value of less than 0.01. Thus, we can conclude that the mean is not zero. A McLeod-Li test of the data returned a set with every lag, indicating the data do not have constant variance. Thus, we can conclude that they are not stationary.

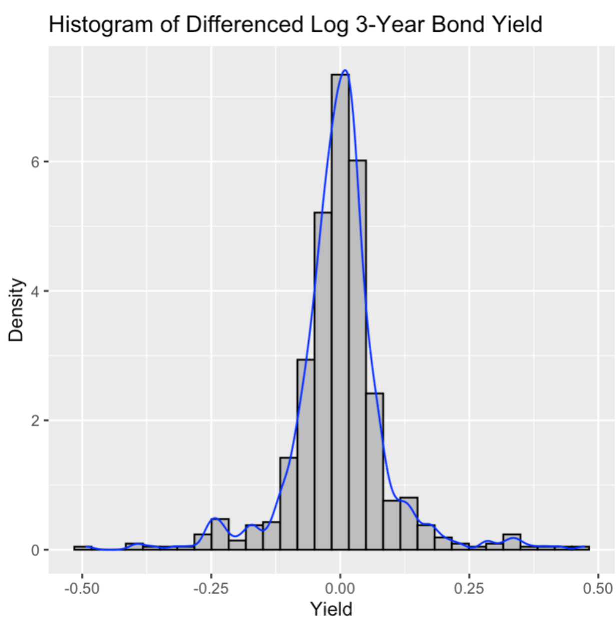
Applying first differencing produces the line plot below:

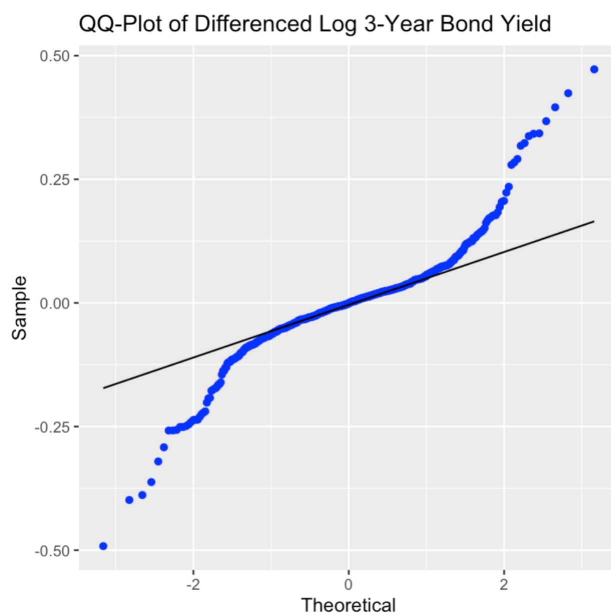


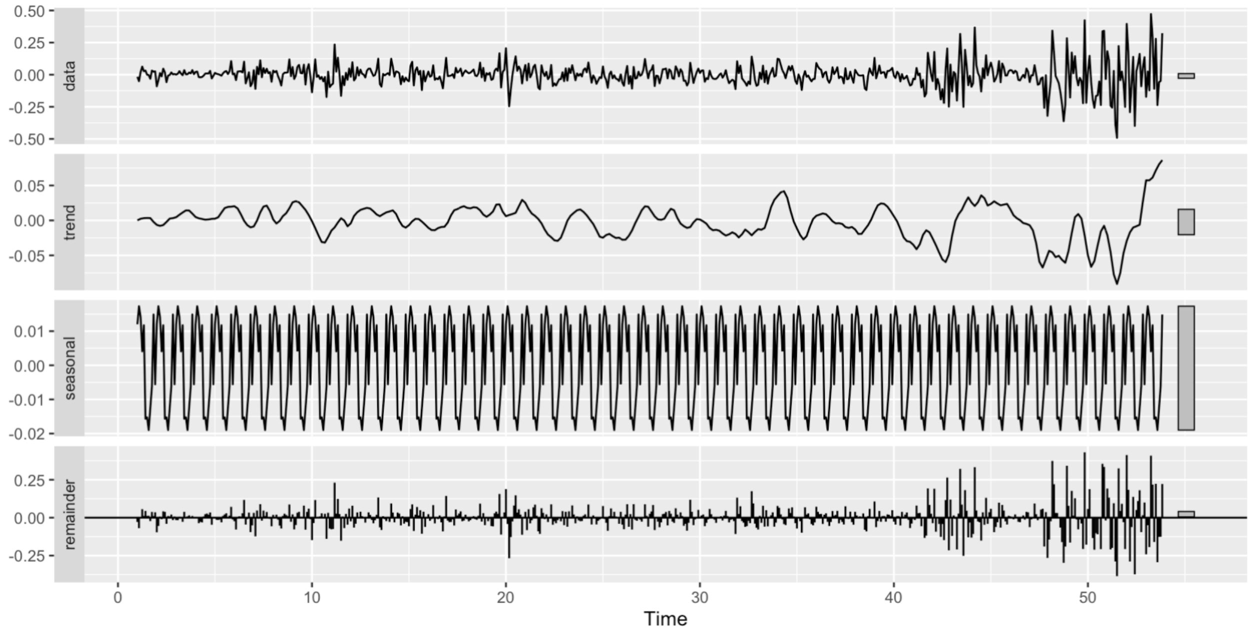
The data are now mean zero. As the summary plot on the right shows, the mean is -0.0022, with a t-test 95% confidence interval of (-0.0096, 0.005) and a p-value of 0.549. Thus, we can conclude that the mean is statistically indistinguishable from zero. Linear-trend stationary is corroborated by ADF and KPSS tests, which each returned a p-value/test-statistic beneath their respective critical values required to assert linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0215 < 0.146$) However, they do not feature constant variance, as a McLeod-Lit test returned a set with every tested lag, indicating non-constant variance. Thus, they are not strictly stationary. However, they are linear-trend stationary, which is still desirable.

nobs	635.000000
NAs	0.000000
Minimum	-0.491823
Maximum	0.472323
1. Quartile	-0.039738
3. Quartile	0.032256
Mean	-0.002254
Median	-0.000743
Sum	-1.431584
SE Mean	0.003760
LCL Mean	-0.009638
UCL Mean	0.005129
Variance	0.008976
Stdev	0.094744
Skewness	0.116144
Kurtosis	5.668950

Applying first differencing also improved the distribution of the data. As the histogram below shows, the symmetry of the distribution has much improved. The QQ Plot shows that the left tail is now closer to the normal line, and the right tail is further. However, their symmetry has improved. Shapiro-Wilk, Lilliefors, and Anderson-Darling tests still produced p-values below 0.01, substantiating the notion that the distribution is still not normal.



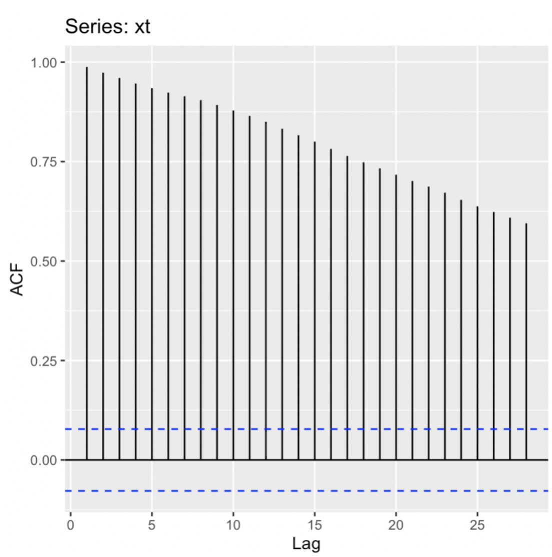


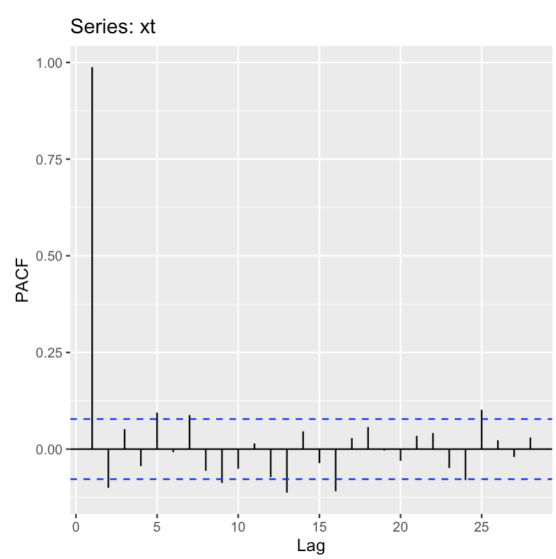


Above is a decomposition of the data. Overall, it does not do a great job of explaining the variance in the data. The trend line is not smooth, and it mirrors much of the variance in the time plot. There does appear to be a well defined seasonal pattern, but it does not account for much of the variance, as it seems to have a min/max range of (-0.02, 0.02), which is small relative to the entire y range of (-0.5, 0.5). The residuals, while mirroring much of the variation in the time plot, do not have a clear sinusoidal type pattern.

2.2. Model 1

Below are the ACF and PACF plots of the log transformed data. Unsurprisingly, the ACF is indicative of a linear trend in the data, which I determined to be the case prior to differencing. This plot would indicate to use a very high number of auto-regressive components. The PACF shows one significant lag, indicating an MA(1) component is appropriate. Given the undesirable ACF, I used the auto arima function to see what was recommended, it was an AR(1) and MA(1). I have already established the need for first differencing, so I will proceed with an ARIMA(1,1,1) model.





Before proceeding, it is worth mentioning that I did not account for outliers in this model for the sake of simplicity. If I were to do so, I would define a threshold for what would be considered an outlier and include a regression component on a vector that indicates the presence of an outlier using a dummy variable. I would then assess the significance of this variable in the model output to determine the effect of outliers on the yield.
The precise form of this model is shown below. It has three degrees of freedom.

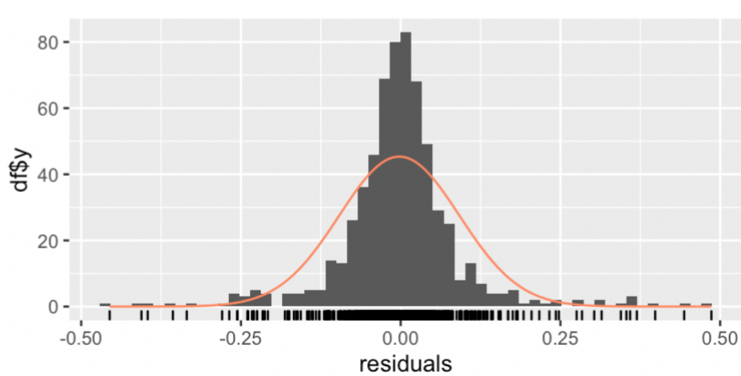
```
Series: xt
ARIMA(1,1,1)

Coefficients:
          ar1     ma1
      -0.7185  0.7969
s.e.   0.1163  0.1000

sigma^2 = 0.008888:  log likelihood = 599.54
AIC=-1193.08   AICc=-1193.04   BIC=-1179.72
```

Overall, the residuals of this model indicate that it is not a good fit of the data. They do not feature strict stationarity, are not normally distributed, and do have auto-correlation. They are, however, linear-trend stationary and do not have business cycles.

Linear trend stationarity is indicated by both summary statistics and ADF and KPSS tests. As the summary table on the right shows, the residuals' mean is -0.002, which produces a t-test 95% confidence interval of (-0.009, 0.005) and a p-value of 0.5713. This p-value and confidence interval allows us to treat the mean as statistically indistinguishable from zero. ADF and KPSS tests both produced a p-value/test-statistic beneath the critical value required to accept linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0215 < 0.146$). Despite this, the residuals are not strictly stationary, as a McLeod-Li test produced a set with every tested lag, indicating non-constant variance.

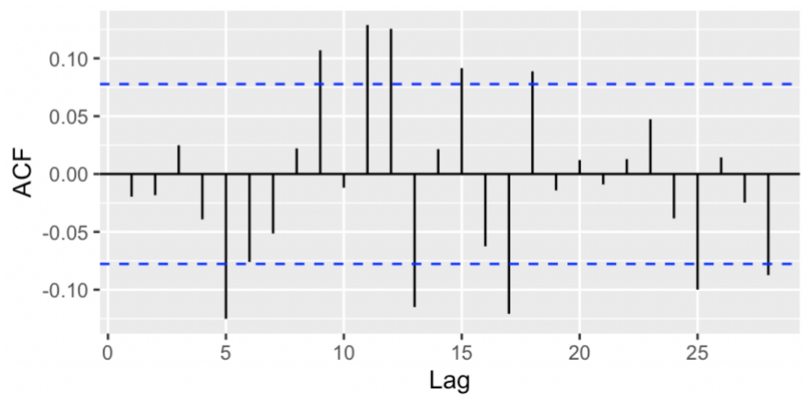
nobs	636.000000
NAs	0.000000
Minimum	-0.455455
Maximum	0.486287
1. Quartile	-0.038960
3. Quartile	0.034196
Mean	-0.002114
Median	-0.001136
Sum	-1.344251
SE Mean	0.003731
LCL Mean	-0.009441
UCL Mean	0.005214
Variance	0.008855
Stdev	0.094102
Skewness	0.216890
Kurtosis	5.824103

Lack of normality in the residuals is most obviously demonstrated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which each returned a p-value of 2e-16, indicating rejection of their null hypotheses that the residuals are normal. 95% confidence intervals for skew and excess kurtosis also do not contain 0: they are (0.198, 0.237) and (5.838, 5.924) respectively. Despite these figures, a histogram of the



residuals (shown below) appears unimodal and with decent symmetry. However, given the absence of 0 in the confidence intervals of skew and excess kurtosis and the extremely small p-values of normality tests, I argue that the residuals are not normally distributed.

Auto-correlation in the residuals is shown both visually and via statistical testing. The ACF plot of the residuals, shown below, includes multiple lags with statistically significant auto-correlation. A Box-Ljung test corroborates this notion, returning a p-value of 0.000006, indicating a rejection of the null hypothesis that the residuals are independent of each other.



2.3. Model 2

Constructing an ARIMA(0,1,1)(0,0,1)$_4$ model produced the following output. This model has three degrees of freedom.

```
Series: xt
ARIMA(0,1,1)(0,0,1)[4]

Coefficients:
         ma1     sma1
      0.0533  -0.0617
s.e.  0.0441   0.0385

sigma^2 = 0.008938:  log likelihood = 597.76
AIC=-1189.52   AICc=-1189.48   BIC=-1176.16
```

The precise form of this model is:

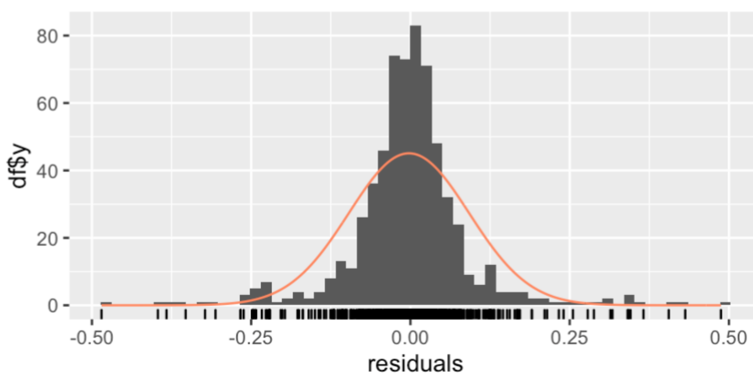
$$y_t = c + z_t + \theta_1 z_{t-1} + \Theta_1 z_{t-4}$$

The residuals of this model are similar to those of the previous, indicating that the model is not a great fit of the data. The residuals do not feature strict stationarity, are not normally distributed, have auto-correlation, and are linear-trend stationary. The one difference is that the residuals of this model do have business cycles.

As the summary table on the right shows, the residuals' mean is -0.002, with a t-test 95% confidence interval of (-0.009 , 0.005) and a p-value of 0.5713. Thus, we can accept the null hypothesis that the residuals are mean zero. ADF and KPSS tests both produced a p-value/test-statistic beneath the critical value required to accept linear-trend stationarity. (ADF p-value: 0.01 < 0.05, KPSS test statistic: 0.0215 < 0.146). However, a McLeod-Li test produced a set with every tested lag, indicating non-constant variance. So, the residuals are linear-trend stationary, but not strictly stationary.

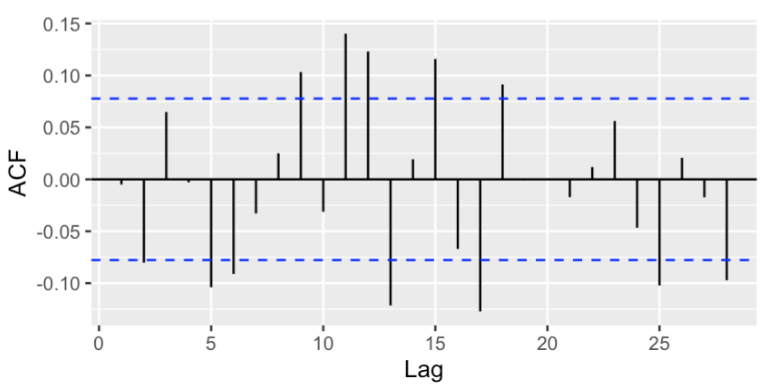
nobs	636.000000
NAs	0.000000
Minimum	-0.484996
Maximum	0.487438
1. Quartile	-0.038743
3. Quartile	0.033736
Mean	-0.002252
Median	-0.000759
Sum	-1.432228
SE Mean	0.003742
LCL Mean	-0.009600
UCL Mean	0.005096
Variance	0.008905
Stdev	0.094364
Skewness	0.163223
Kurtosis	5.845872

The residuals' lack of normality is indicated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which all returned p-values of 2e-16, implying the rejection of their null hypotheses that the residuals are normally distributed. The 95% confidence intervals for skew (0.143, 0.187) and excess kurtosis (6.036, 6.129) do not contain zero. Similarly to the last model, the residuals' histogram (right) appears unimodal and with decent symmetry. However, given the confidence intervals of skew and excess kurtosis and p-values of normality tests, I consider these residuals not normal.



Auto-correlation in the residuals is suggested by both the ACF plot and a Box-Ljung test. The ACF plot, shown below, shows multiple lags with statistically significant auto-correlation. A Box-Ljung test returned a p-value of 0.000001, indicating a rejection of the null hypothesis that the residuals do not feature auto-correlation.

The residuals have a detected business cycle with a frequency of 4.29 months.



I prefer the first model to this model. While the residuals of the two models are similar, the residuals of the first did not contain business cycles. Additionally, the AR_1 variable of the second model had a coefficient of 0.0533 and a standard error of 0.0411, making the significance of the variable questionable. Both of the variables in the first model had coefficients that were at least twice the absolute magnitude of their respective standard errors.

2.4. Model Comparison

As shown below, backtesting showed almost identical results, but the first model did have slightly better results. Given this, and the fact that I prefer the first model to the second based on the reasons in the previous section, I still prefer the first.

	Model 1	Model 2
Measurement	t	t
RMSE	0.34701	0.34712
MAE	0.34701	0.34712
Bias	0.34701	0.34712

3. ARIMA and Regression Errors

3.1. Model 1

A linear regression of 3-year bond yields on 1-year bond yield would take the following general form:

$$y_{3t} = \beta_0 + \beta_1 y_{1t} + \varepsilon_t$$

The precise form this model takes is shown in the output below. This model has two degrees of freedom.

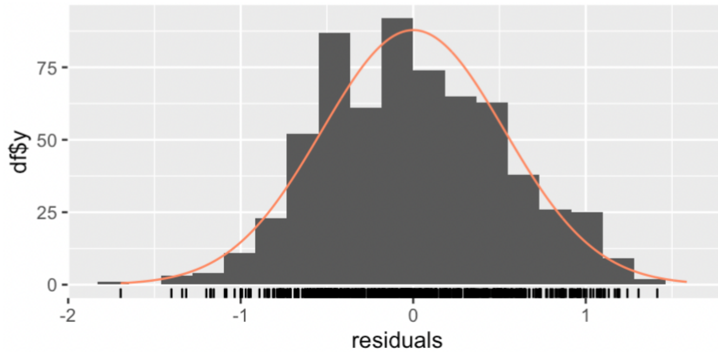
```
Call:
lm(formula = y3t ~ y1t)

Residuals:
     Min       1Q   Median       3Q      Max
-1.69599 -0.42038 -0.03045  0.37993  1.41445

Coefficients:
            Estimate Std. Error t value           Pr(>|t|)
(Intercept) 0.712645   0.041909    17.0 <0.0000000000000002 ***
y1t         0.940816   0.006676   140.9 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.529 on 634 degrees of freedom
Multiple R-squared:  0.9691,    Adjusted R-squared:  0.969
F-statistic: 1.986e+04 on 1 and 634 DF,  p-value: < 0.00000000000000022
```

The adjusted R^2 for this model is 0.969, indicating that almost all of the variance in the 3-year bond yield was explainable by the 1-year bond yield. This is quite high. The p-value associated with the model's F-test was less than 0.01, indicating that this model has more explanatory power than a mean-only model.

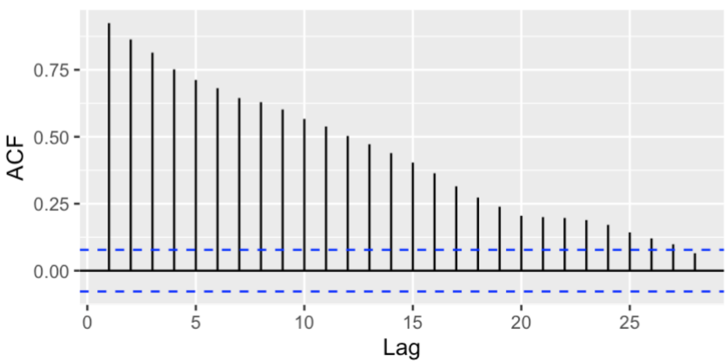
However, the residuals of this model indicate that it is not a great fit of the data: they are not strictly stationary, nor normally distributed, and they do feature auto-correlation. However, they are linear-trend stationary and do not have detectable business cycles.

Linear trend stationarity is indicated by both summary statistics and ADF and KPSS tests. Each produced a p-value/test-statistic beneath the critical value required to accept linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0215 < 0.146$). However, a McLeod-Li test produced a set with every tested lag, indicating non-constant variance. Therefore, they are not strictly stationary.

The residuals' lack of normality is indicated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which all returned p-values less than 0.01, allowing us to reject the null hypothesis of normality. This is further evidenced by neither the 95% confidence interval for skew (0.114, 0.122) nor excess kurtosis (6.036, 6.129) containing zero. Despite this, their distribution is unimodal and with decent symmetry, as shown in the histogram to the right.

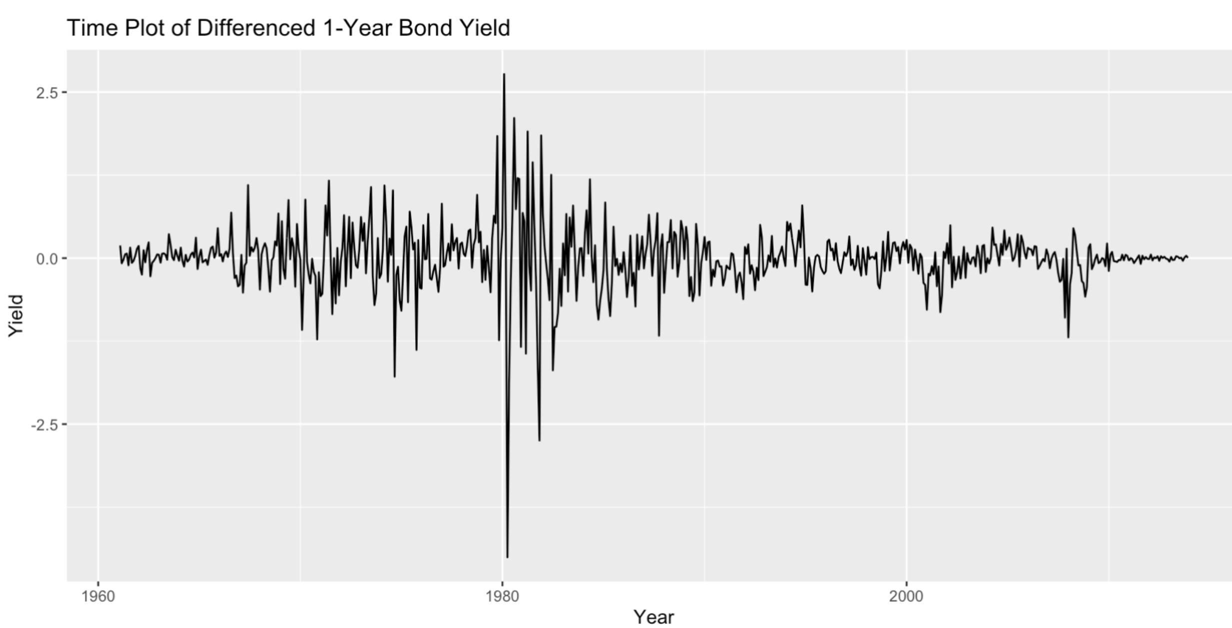


Auto-correlation in the residuals is demonstrated in their ACF plot, pictured below, which shows almost all tested lags having significant auto-correlation. This is substantiated by a Box-Ljung test, which returned a p-value of 2e-16, indicating a rejection of the null hypothesis that the residuals are independent, and therefore they are auto-correlated.



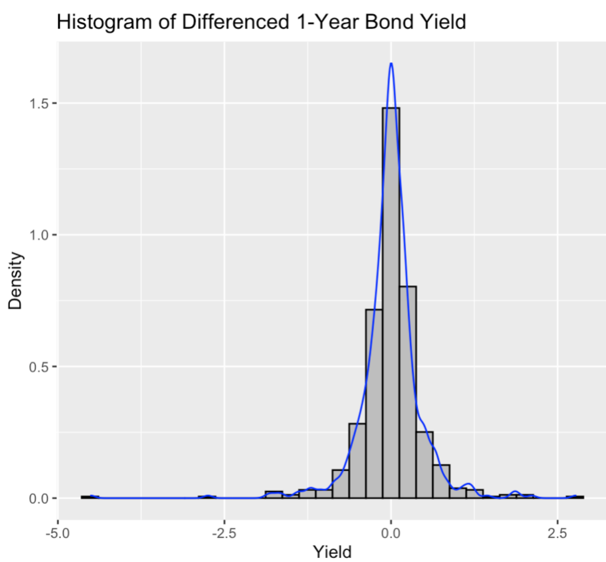
3.2. Model 2

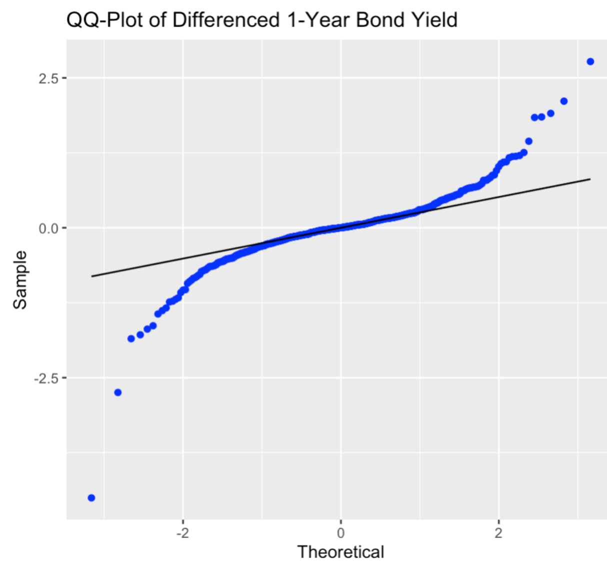
Below is a time plot of the first differenced yield of 1-year bonds:



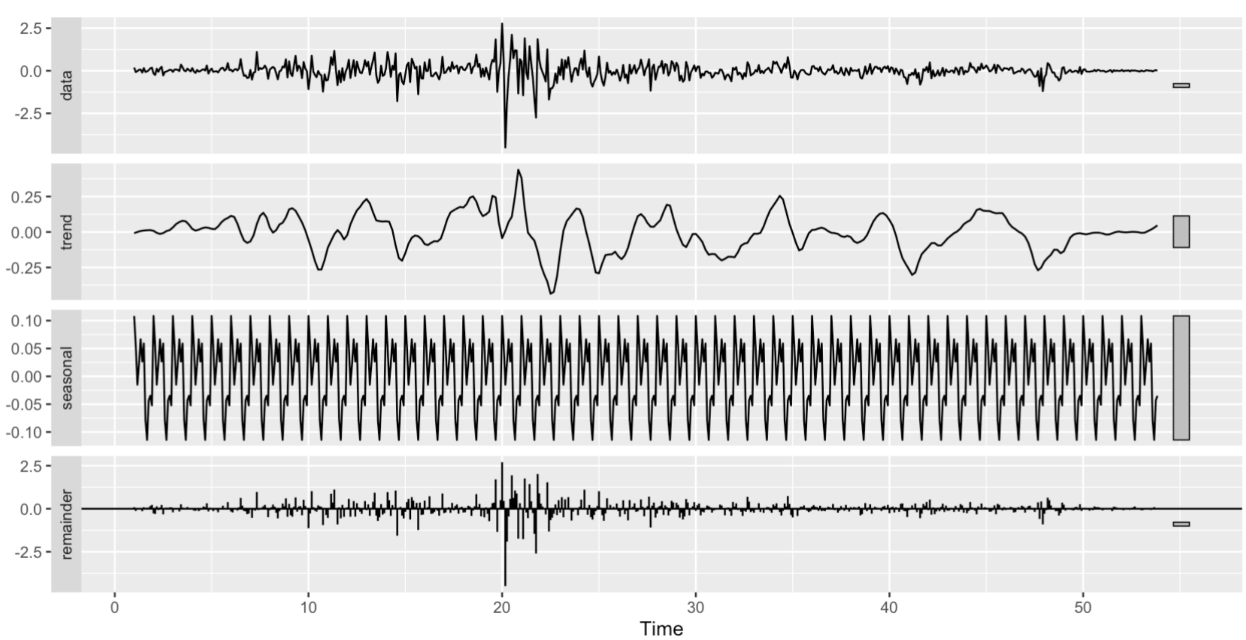
ADF and KPSS tests indicate that the data are linear-trend stationary; each returned a p-value/test-statistic beneath the critical value required to assert linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0273 < 0.146$) However, a McLeod-Li test returned a set of all tested lags, indicating that the data do not feature constant variance. Therefore, they are not strictly stationary.

The data are not normally distributed. While their distribution is unimodal, as shown below in the histogram, the tails do deviate from the normal line, as shown in the QQ plot. Non-normality is also indicated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which all returned p-values of 2e-16, allowing us to reject their null hypotheses of normality.



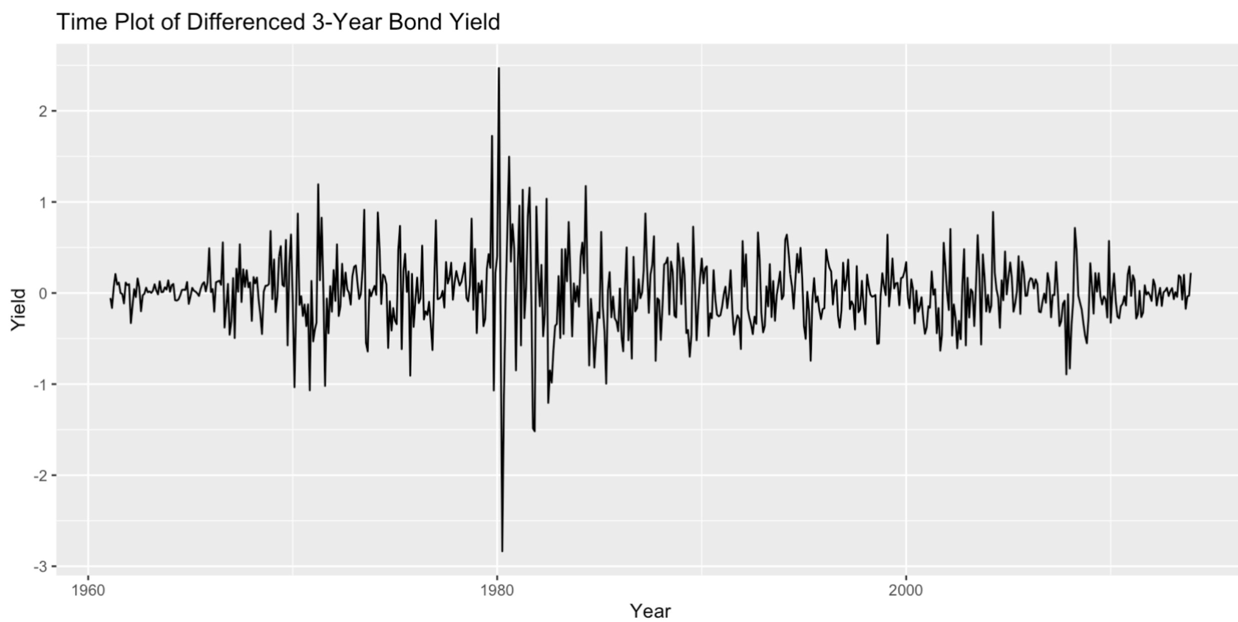


Below is a decomposition of the data:



Overall, it does not do a great job of explaining the variance in the data. The trend line is a moderately smoothed version of the time plot. There is a well defined seasonal pattern, but it doesn't account for much of the variance, as it has a min/max range of (-0.1, 0.1), which is small relative to the entire y range of (-4, 2.5). The residuals, while mirroring much of the variation in the time plot, do not have a clear sinusoidal type pattern.

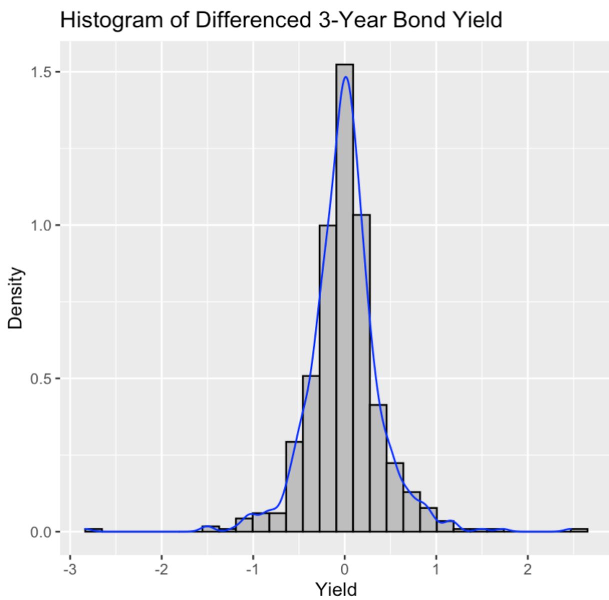
Below is a time plot of the first differenced yield of 3-year bonds:

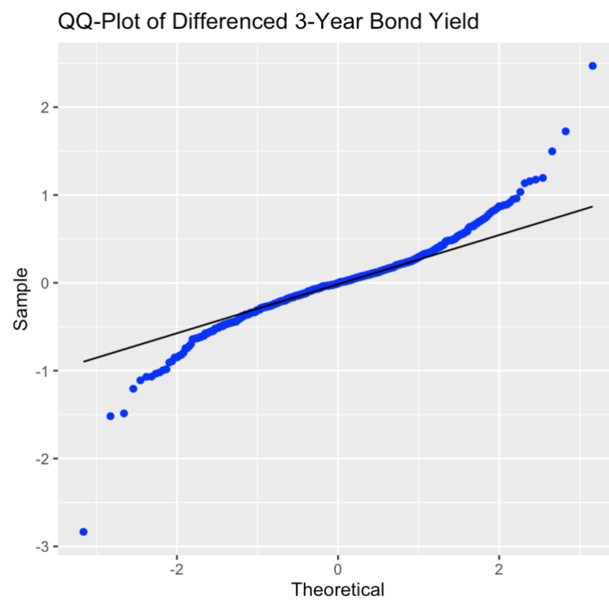


The data very closely resemble the differenced 1-year bond yield data. Unsurprisingly, they have very similar characteristics; they are linear-trend stationary, but not strictly stationary, and they are not normally distributed.

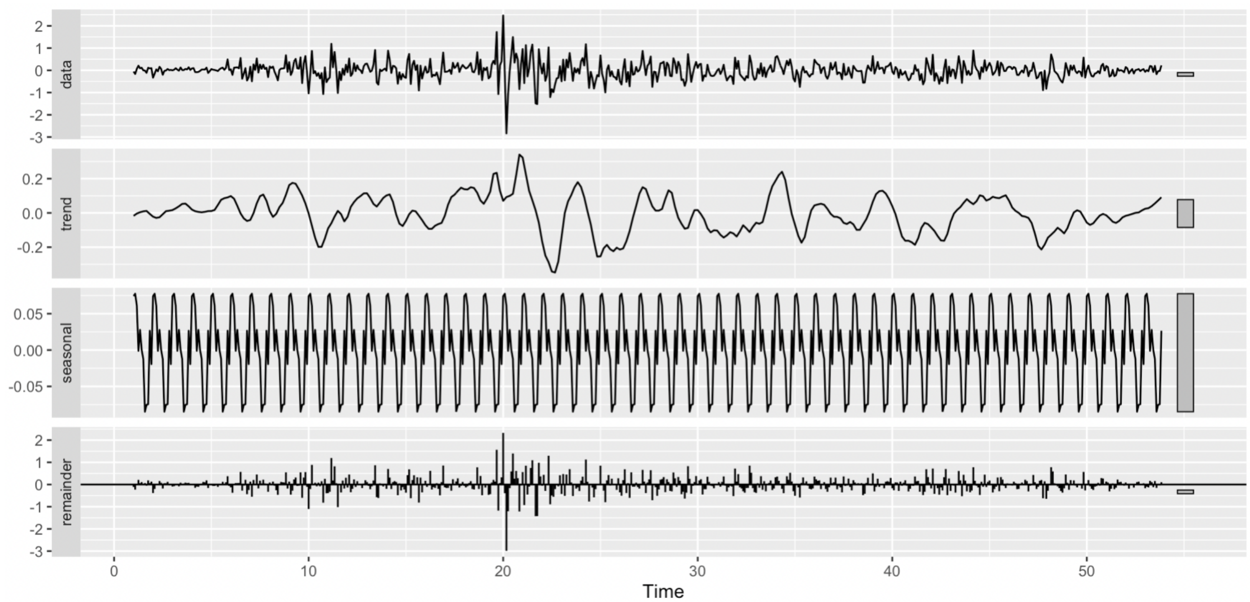
ADF and KPSS tests indicate linear-trend stationary, as both returned a p-value/test-statistic beneath the critical value required to assert linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0316 < 0.146$) A McLeod-Li test returned a set of all tested lags, indicating that the data do not feature constant variance. Therefore, they are not strictly stationary.

The data are not normally distributed. While their distribution is unimodal, as shown in the histogram on the next page, the tails do deviate from the normal line, as shown in the QQ plot. Shapiro-Wilk, Lilliefors, and Anderson-Darling tests all returned p-values of less than 0.01, indicating a rejection of their null hypotheses of normality.





Decomposition of the data produces similar results to those of the 1-year bond yield. Overall, decomposition does not do a great job of explaining the variance. The trend line is not smooth and resembles the time plot. There is a well defined seasonal pattern, but it doesn't account for much of the variance, as it has a min/max range of (-0.1, 0.1), which is small relative to the entire y range of (-4, 2.5). However, this is more of the variance than in the previous decomposition. The residuals mirr much of the variation in the time plot, but they do not have a clear sinusoidal type pattern.



A linear regression of the differenced 3-year bond yield on the differenced 1-year bond yield would take the general form:

$$dy_{3t} = \beta_0 + \beta_1 dy_{1t} + \varepsilon_t$$

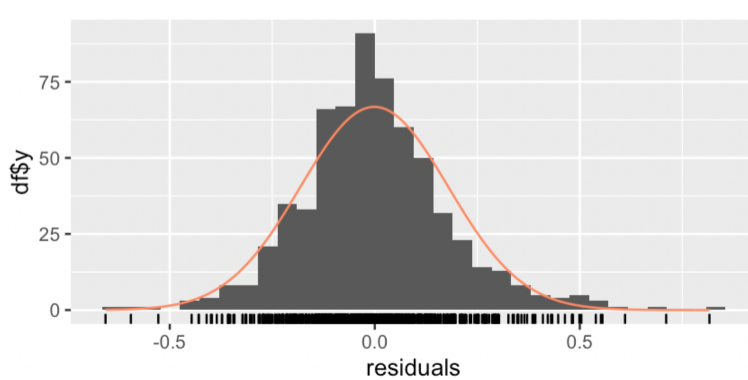
The precise form of the model is shown below in the model output. This model has one degree of freedom.

```
Call:
lm(formula = d3t ~ -1 + d1t)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.65669 -0.11132 -0.01054  0.09621  0.81624 

Coefficients:
    Estimate Std. Error t value            Pr(>|t|)    
d1t  0.73598    0.01478   49.78 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1803 on 634 degrees of freedom
Multiple R-squared:  0.7963,	Adjusted R-squared:  0.796 
F-statistic:  2478 on 1 and 634 DF,  p-value: < 0.00000000000000022
```

The adjusted R^2 for this model is 0.796, indicating that differencing the data resulted in the yield of 1-year bonds explaining less of the variance in the 3-year bonds. Similar to the previous model, the p-value of this model's F-test is less than 0.01, indicating that it has more explanatory power than a mean only model.

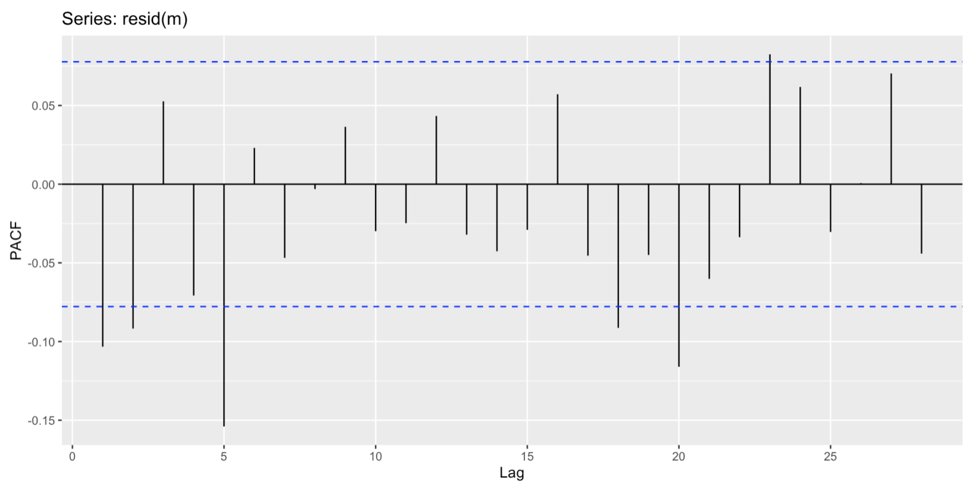
The residuals of this model indicate that it is not a good fit of the data: they are not strictly stationary, nor normally distributed, and they do feature auto-correlation. However, they are linear-trend stationary and do not have detectable business cycles.

Linear trend stationarity is indicated by both summary statistics and ADF and KPSS tests. Each produced a p-value/test-statistic beneath the critical value required to accept linear-trend stationarity. (ADF p-value: 0.01 < 0.05, KPSS test statistic: 0.0219 < 0.146). However, a McLeod-Li test produced a set with every tested lag, indicating non-constant variance. Therefore, they are not strictly stationary.



The residuals' lack of normality is indicated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which all returned p-values less than 0.01, allowing us to reject their null hypotheses of normality. This is further evidenced by neither the 95% confidence interval for skew (0.417, 0.44) nor excess kurtosis (1.604, 1.676) containing zero. Despite this, their distribution is unimodal and with decent symmetry, as shown in the histogram on the previous page.

Auto-correlation in the residuals is demonstrated in their PACF plot, pictured below, which shows that several lags having significant auto-correlation. This is substantiated by a Box-Ljung test, which returned a p-value of 0.00002, indicating a rejection of the null hypothesis that the residuals are independent, and therefore they are auto-correlated.



Overall, I think this model is inadequate relative to the first model. The residuals of the two models are close enough that they do not indicate a clear winner: they are both linear-trend stationary, are not strictly stationary, are not normally distributed, and do feature auto-correlation. However, as discussed in my analysis of the model summary, the adjusted R^2 of this model is lower than the first. (0.969 vs 0.796)

3.3 - Model 1 and 2 Comparison

The first model describes a stronger linear relationship between the yields of the 1 and 3 year bonds than does the second model. This is indicated by the difference in their adjusted R^2 (0.969 vs 0.796). A higher adjusted R^2 indicates that a greater amount of variance in the yield of 3-year bonds can be explained by the yield of 1-year bonds, and thus there is a stronger relationship. When considered in the context of differencing, this would indicate that the yield of the 1-year bond is a greater predictor of the change in overall level of the 3-year bond yield than the month to month change in the 3-year bond yield.

3.4. Model 3

Creating an AR model for the 3-year bond yield while using the 1-year yield as a regressor variable would take the general form:

$$dy_{3t} = c + \phi_1 y_{3t-1} + \phi_2 y_{3t-2} + \phi_3 y_{3t-3} + \phi_4 y_{3t-4} + \phi_5 y_{3t-5} + \beta dy_{1t}$$

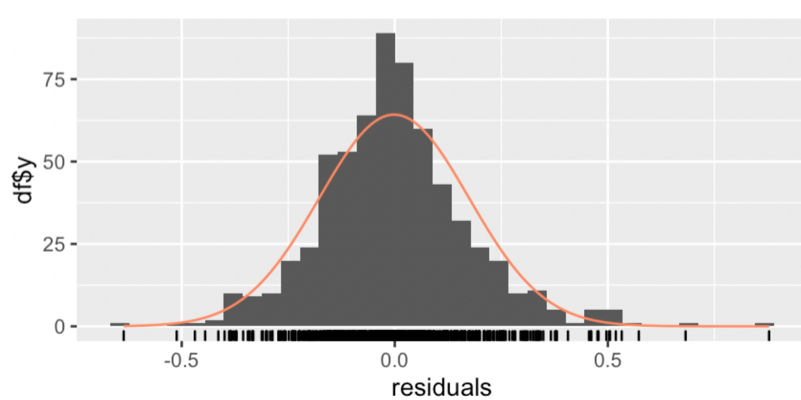
The specific form of this model is shown in the model output below. This model has six degrees of freedom.

```
Series: ts(d3t)
Regression with ARIMA(5,0,0) errors

Coefficients:
          ar1      ar2     ar3      ar4      ar5    xreg
      -0.1178  -0.0845  0.0326  -0.0893  -0.1533  0.7436
s.e.   0.0395   0.0394  0.0396   0.0397   0.0393  0.0142

sigma^2 = 0.03111:  log likelihood = 203.7
AIC=-393.4   AICc=-393.23   BIC=-362.23
```

Overall, the residuals indicate that this model is a decent fit of the data: they are not strictly stationary, nor normally distributed, and have business cycles, but they do not have auto-correlation.

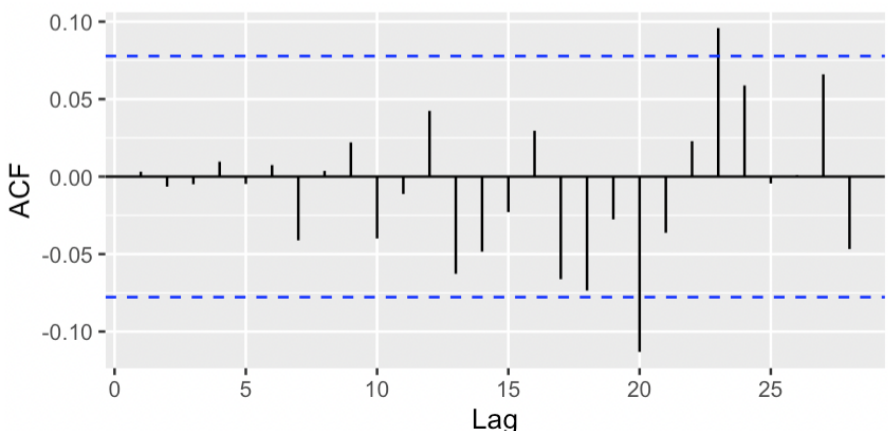
Linear trend stationarity is indicated by both ADF and KPSS tests. Each produced a p-value/test-statistic beneath the critical value required to accept linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0271 < 0.146$). However, a McLeod-Li test produced a set with every tested lag, indicating non-constant variance. Therefore, they are not strictly stationary.

The residuals' lack of normality is indicated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which all returned p-values less than 0.01, allowing us to reject their null hypotheses of normality. This is further evidenced by neither the 95% confidence interval for skew (0.451, 0.472) nor excess kurtosis (1.604, 1.676) containing zero. Despite this, their distribution appears approximately normal in the histogram to the right. However, it does show slight skew.



Lack of auto-correlation in the residuals is demonstrated by a Box-Ljung test, which returned a p-value of 0.9859, indicating the null hypothesis that the residuals are independent, and therefore non-auto-correlated.

The ACF plot, pictured below, shows a few lags having significant auto-correlation. However, given that it is only a few and the p-value of the Box-Ljung test, I believe these residuals do not have auto-correlation.



As stated, the residuals do have business cycles. There are two, of frequencies 6.085 and 3.007 months. These cycles indicate patterns in the variance of the data that the model was not able to capture.

3.5. Model 4

Performing the parameter test function on the previous model indicated that the AR_3 variable was not statistically significant. By removing it, the general form of the reduced model would be:

$$dy_{3t} = c + \phi_1 y_{3t-1} + \phi_2 y_{3t-2} + \phi_4 y_{3t-4} + \phi_5 y_{3t-5} + \beta dy_{1t}$$

The specific form of this model is shown in the model output below. This model has five degrees of freedom.

```
Series: d3t
Regression with ARIMA(5,0,0) errors

Coefficients:
          ar1      ar2  ar3      ar4      ar5    xreg
      -0.1205  -0.0883    0  -0.0929  -0.1563  0.7429
s.e.   0.0394   0.0391    0   0.0395   0.0391  0.0142

sigma^2 = 0.03109:  log likelihood = 203.36
AIC=-394.72   AICc=-394.59   BIC=-368
```

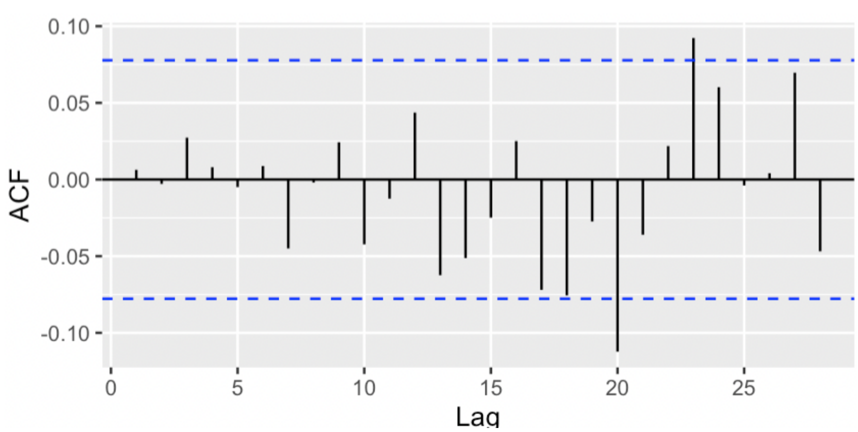
Unsurprisingly, the residuals of this model are quite close to the residuals of the previous model: they are not strictly stationary, nor normally distributed, and have business cycles, but do not have auto-correlation.

Linear trend stationarity is indicated by both ADF and KPSS tests, as each produced a p-value/test-statistic beneath the critical value required to accept linear-trend stationarity. (ADF p-value: $0.01 < 0.05$, KPSS test statistic: $0.0275 < 0.146$). However, a McLeod-Li test produced a set with every tested lag, indicating non-constant variance. Therefore, they are not strictly stationary.

The residuals' lack of normality is indicated by Shapiro-Wilk, Lilliefors, and Anderson-Darling tests, which all returned p-values less than 0.01, allowing us to reject their null hypotheses of normality. Lack of normality is further evidenced by neither the 95% confidence interval for skew (0.45, 0.469) nor excess kurtosis (1.866, 1.931) containing zero.

Lack of auto-correlation in the residuals is demonstrated by a Box-Ljung test, which returned a p-value of 0.9643, indicating the null hypothesis that the residuals are independent, and therefore non-auto-correlated.

The ACF plot, pictured below, indicates two lags having significant auto-correlation. However, given that it is only two and the results of the Box-Ljung test, I assert that these residuals are non-auto-correlated.



As stated, the residuals do have business cycles. There are two, of frequencies 6.082 and 3.008 months.

In comparison to my first linear regression model, the residuals of this model indicate that it is a better fit of the data. For both models, the residuals are linear-trend stationary and do not feature strict stationarity. However, the residuals of this model do not feature auto-correlation, while the residuals of the better linear regression model did.

3.6. Real Solutions to the Characteristic Equation of Model 4

Performing the polyroot function on the reduced ARIMA model produced four solutions, two of which were real. They were 1.152 and -1.031.

```
> #build polynomial (1 + the negative of the coefficients)
> (p2 <- c(1,-m$coef))
                  ar1         ar2         ar3         ar4         ar5        xreg
 1.00000000  0.12053949  0.08831368  0.00000000  0.09287176  0.15630623 -0.74287895
> (s2 <- polyroot(p2))    # if no imaginary part, no cycles
[1]  0.5424517+0.9097942i -0.4977595+0.8730920i -0.4977595-0.8730920i  0.5424517-0.9097942i  1.1520631+0.0000000i
[6] -1.0310416-0.0000000i
> #complex roots are in pairs, so use unique()
> #using round() because there may be very small differences
> (z = unique(round(sapply(all_complex(s2),period),digits=3)))   # lengths of business cycles
[1] 6.082 3.008
```

3.7. Inverse of Absolute Values of the Solutions to the Characteristic Equation of Model 4

Below are the inverses of the absolute values of the solutions of the characteristic equation. The maximum value, 0.707, indicates that the model likely has a unit root, as all roots are less than one and, therefore, fall within the radius of the unit circle.

```
> s1
[1]  1.064082+0.930665i -0.560747+1.298226i -0.560747-1.298226i  1.064082-0.930665i -1.600835-0.000000i
> Mod(s1)
[1] 1.413651 1.414153 1.414153 1.413651 1.600835
> 1/Mod(s1)
[1] 0.7073883 0.7071370 0.7071370 0.7073883 0.6246740
```

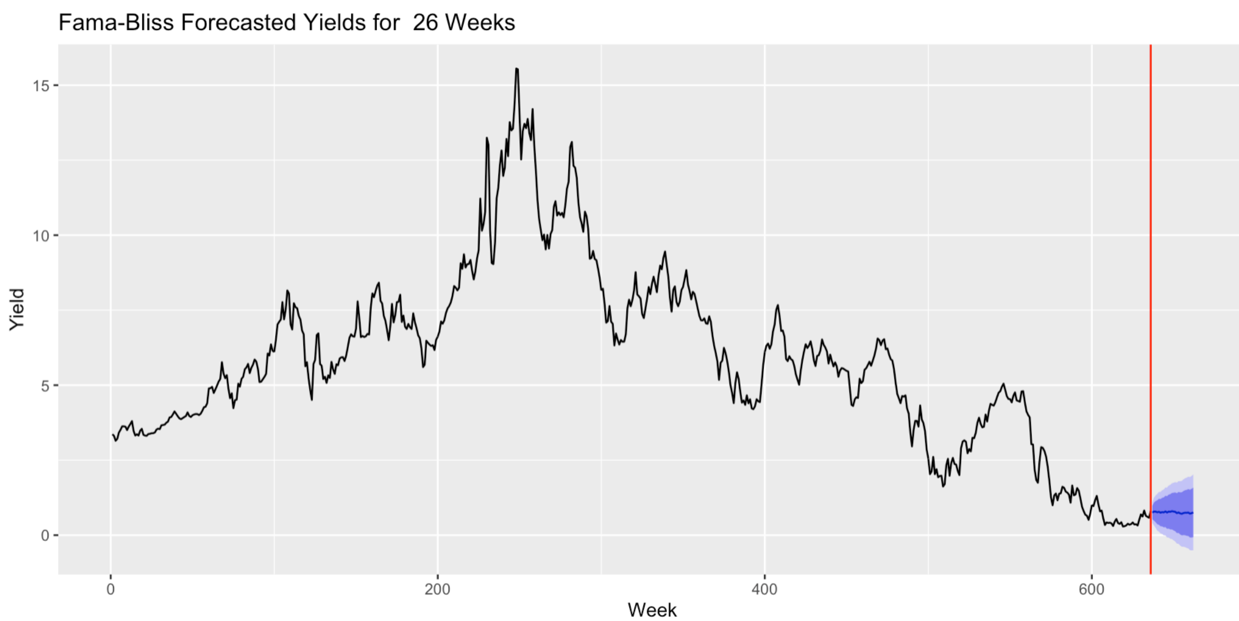
3.8. Model Comparison

As previously stated, I preferred my first linear regression model to my second, and my second ARIMA model to my first. Between the preferred models, the ARIMA model performed best during backtesting. As the results below show, the ARIMA model had better error and bias measurements for all tested time periods. I will use this model to make my final forecasts.

	Model 1				Model 4			
Measurement	t+1	t+2	t+3	t+4	t+1	t+2	t+3	t+4
RMSE	0.667	0.673	0.702	0.804	0.145	0.127	0.154	0.232
MAE	0.661	0.666	0.693	0.804	0.109	0.086	0.114	0.232
Bias	0.661	0.666	0.693	0.804	0.004	0.059	0.103	0.232

4. Report

Below is a plot that shows my predictions for the behavior of the 3-year bond yield over the next 26 weeks. The red line delineates where current data stops and my predictions begin. These are probabilistic predictions, meaning that what is shown below displays possible outcomes and their likelihoods. The dark blue line represents the most likely outcome. Around it are other possible outcomes, with darker shading representing higher probability. The lightest blue regions represent the 5% most extreme outcomes.



As the graph shows, I predict it is most likely that 3-year bond yields will remain stable over the next two quarters. However, there is a very slight decline in the line of expected yield occurring in roughly four months. The slight volatility shown in the line of expected yield embodies the slight volatility in yields over the past half a year. It is notable that my predictions do not see yields exceeding 2%, but do indicate a 2.5% chance that yields turn negative.