Principal Components Analysis of US Treasury Yield Curves

Haci Ibrahim Yalcin

Abstract

An estimation of future yield curves is an enduring problem in the fields of finance and the economy. Several risk factors drive the shape of yield curves. A scaling method is often used to decompose the yield curve into its components. When broken down into factors, the yield curve can then be modelled with their alternating values. The number of factors needed to model the curve depends on the wanted precision and they are highly susceptible to the impact of surrounding factors changing over time, like social issues, timeframes and the periodical economic situation.
The model should be built with the context in mind. A complete solution to predictive modelling of yield curves will be attempted in this work along with appropriate interpretation.

Keywords: Principal Component Analysis, Yield Curve, Driving Factors, Predictive Modelling

1. Introduction

Principal Component Analysis (PCA) is a statistical technique for reducing the dimensions of datasets while preserving the variability of data. All components are linearly independent of each other and can explain the variability of datasets in fewer dimensions.

The yield curve is a curve that plots interest rates or yields across a range of maturity. The slope of the yield curve is an indicator of economic activity and possible change in the future interest rate. There are three types of yield curves which are namely normal, inverted and flat yield curves.

This work is aimed to decompose the yield curve into its uncorrelated components for modelling the future yield curves.

2. Theoretical framework

To understand the principal components of a dataset, one should be well familiar with eigenvalues and eigenvectors. High dimensional data can be presented through matrices. In a simple two-dimensional matrix, both rows and columns can be represented vectors. A column eigenvector $\vec{v}$ with an eigenvalue, λ satisfy the equation

$$M\vec{v} = \lambda\vec{v}$$

where $M$is the square matrix.

Working with matrices and vectors can be computationally expensive in large datasets. An approach of capturing enough information while using less dimension will come in handy in the case of large datasets. Principal Component Analysis is the best method to reduce the dimension while conserving most of the data.

Apart from the high complexity, more dimensions usually lead to higher uncertainties and noise. When two features have a correlation more than zero, it means the first feature can be explained by the dynamics of the second feature to some extent and vice versa. If their correlation equals one (or negative one), one can be transformed to the other through a simple transformation of scalar multiplication. It is safe to say that if one feature can capture information of another feature, then there is redundant data in one of the features. Therefore, we tend to capture only a set of uncorrelated data by using Principal Component Analysis.

Complex systems depending on a large set of parameters may have a large variability. Yield curves usually have more than 10 parameters as each maturity rate can be represented by one parameter. Principal Component Analysis can be used to decompose the data into a set of uncorrelated features. Each component will be pointing in directions of the highest variance in the parameter subspace. For a dataset of n features, up to n principal components can be constructed. The amount of the principal components used in dimensionality reduction is highly dependent on the nature of the problem.

3. Methodology

We can now construct the approach as follows:
Once the principal components are retrieved, we can run tests. One of our main interests is whether the principal components are interpretable.
1. For m observations of n variables represent the data in a m by n-dimensional matrix
2. Calculate the covariance matrix of the dataset
3. Find the eigenvectors of the covariance matrix with their respective eigenvalues
4. Sort the eigenvectors from the largest eigenvalue
5. Check the variability explained by each component

The first three principal components are often expected to imitate the three factors of yield curves, namely level, slope and curvature. To check if a principal component is similar to a factor, we can simply calculate the correlation between the constructed component and the actual data.

We expect the first principal component to be highly correlated to the original data. The second component will resemble the slope of the yield curve and the third component will look like the curvature of the yield curve.

It is expected that the first three components will be sufficient in modelling, at least for smooth yield curves. Higher-order components might prove useful in modelling yield curves with unusual shapes.

3.1. Data

We will be working on the US Treasury yield curve retrieved from the Quandl dataset. Quandl is a free source for financial, economic and alternative datasets. The link [1] will providev15-years of US Treasury yield curve data. Figure 1 shows the first five rows of the Yield curve after cleaning and standardization. Figure 2 is a graphical representation of the Yield curve consisting of 10 different maturity rates.

Date	1 MO	3 MO	6 MO	1 YR	2 YR	3 YR	5 YR	7 YR	10 YR	20 YR
2006-01-03	0.768501	0.801541	0.824859	0.825095	0.817308	0.813953	0.815476	0.813142	0.812236	0.820569
2006-01-04	0.764706	0.807322	0.819209	0.819392	0.811538	0.810078	0.811508	0.811088	0.810127	0.816193
2006-01-05	0.768501	0.809249	0.819209	0.821293	0.813462	0.812016	0.813492	0.811088	0.810127	0.818381
2006-01-06	0.770398	0.813102	0.822976	0.825095	0.821154	0.817829	0.819444	0.815195	0.814346	0.822757
2006-01-09	0.776091	0.815029	0.824859	0.826996	0.821154	0.817829	0.819444	0.817248	0.814346	0.822757

Figure 1 - First Five Rows of the Yield Curve After Standardization

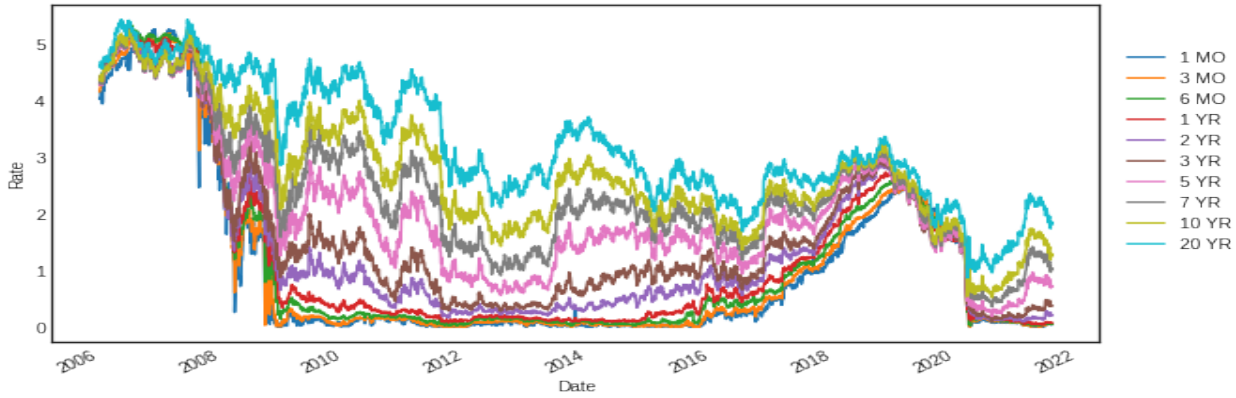


Figure 2 Graphical Representation of the Standardized Yield Curve

3.2. Break Points - Structural Change

We will be looking into the first column to see if there is a structural change in the dataset. The breakpoint function (chow test) from the strucchange package was used to find the structural breaks. Figures 3 and 4 will show us the whereabouts of breakpoints.

```
        Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = ts(df) ~ tt, h = 0.1)

Breakpoints at observation number:

m = 1         678
m = 2         678                          3051
m = 3   404 792                            3051
m = 4   404 792                       2657      3361
m = 5   404 792                       2642 3105 3493
m = 6   404 792                  2293 2681 3105 3493
m = 7   404 792 1297             2293 2681 3105 3493
m = 8   404 792 1299 1734        2293 2681 3105 3493
m = 9   388 776 1164 1552 1941 2329 2717 3105 3493

Corresponding to breakdates:

m = 1         678
m = 2         678                          3051
m = 3   404 792                            3051
m = 4   404 792                       2657      3361
m = 5   404 792                       2642 3105 3493
m = 6   404 792                  2293 2681 3105 3493
m = 7   404 792 1297             2293 2681 3105 3493
m = 8   404 792 1299 1734        2293 2681 3105 3493
m = 9   388 776 1164 1552 1941 2329 2717 3105 3493

Fit:

m          0         1         2         3         4         5         6         7
RSS  7866.45  1670.07   847.02   496.74   235.11   227.79   222.42   221.93
BIC 13780.57  7790.83  5180.82  3134.45   256.25   158.29    90.53   106.69

m          8         9
RSS   221.53   227.77
BIC   124.56   257.10
```

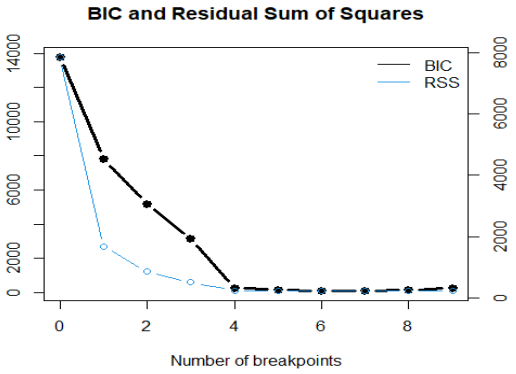
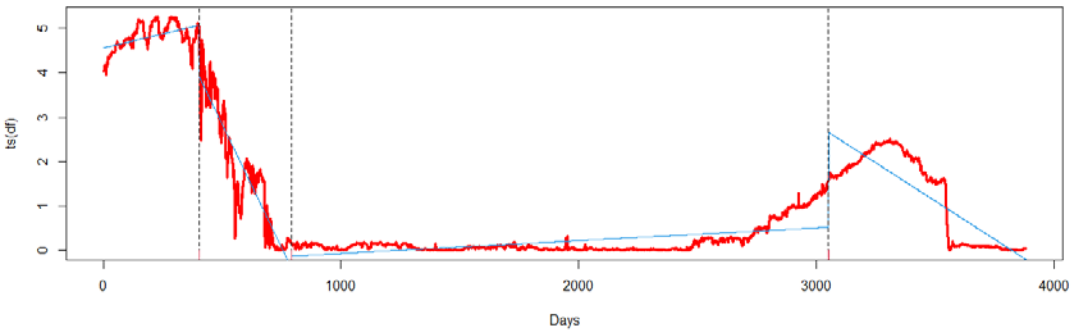


Figure 3



"2007-08-09" "2009-03-02" "2018-03-09"

Figure 4

It appears we have three breakpoints which are located at "2007-08-09", "2009-03-02", "2018-03-09".

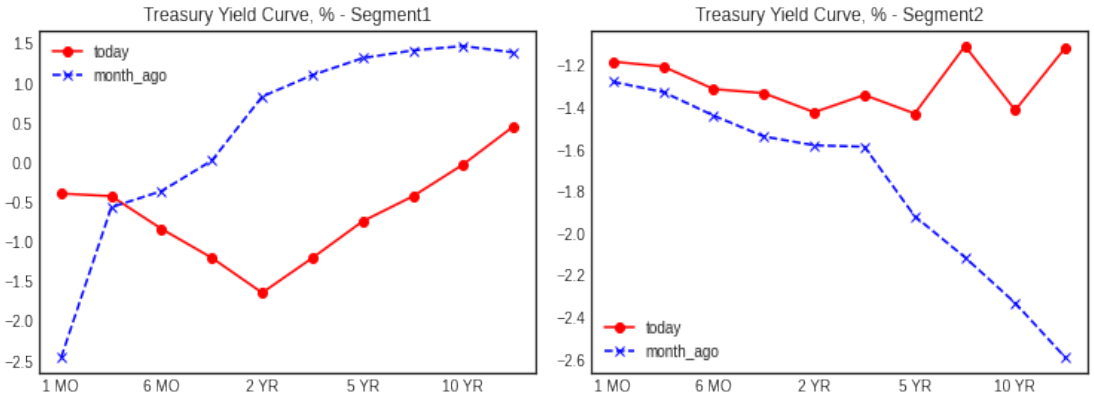
3.3. Segmenting the Data Set into Four Subsets Using Breakpoints

Figure 5 is the graphical representation of the segmented dataset.

Figure 5 Standardized Segment 1,2,3,4

3.4. Autocorrelation

Let's inspect the behavior of the nth month by comparing it to the n-1st month. If the behavior is similar, we can suspect there is autocorrelation. Figure 6 is a graphical representation of nth month data and n-1st month data for each segment.



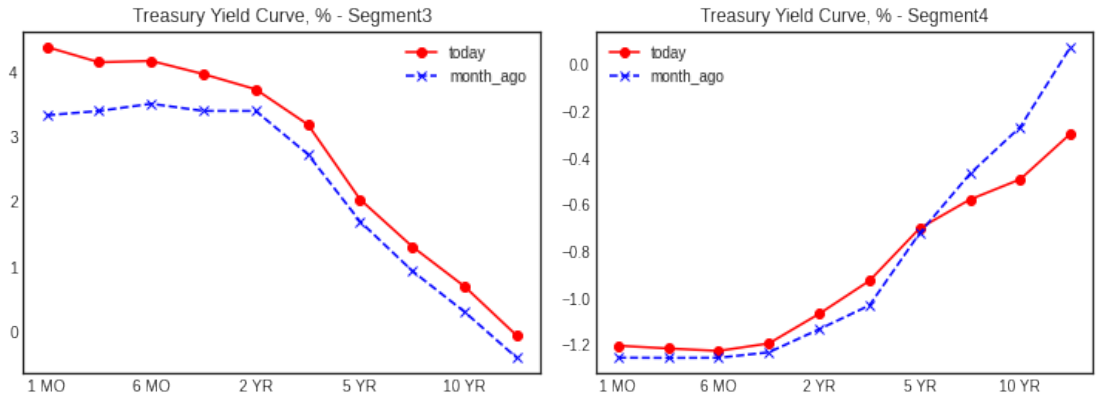


Figure 6 nth and n-1st month comparison

3.5. Forecast Next 500 Days of The Yield Curve for 1 Month, 2 Year and 30 Year Yield Rates Using AR(1) Model on The Full Dataset.

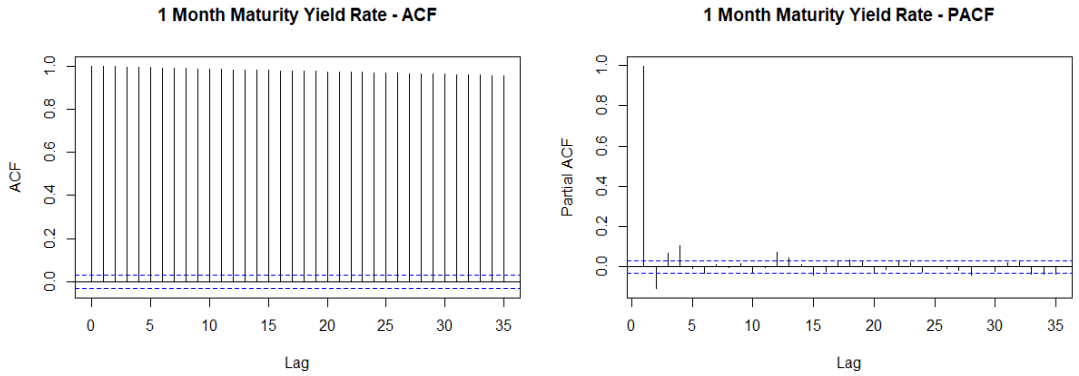


Figure 7 1 Month Yield Maturity Rate ACF and PACF

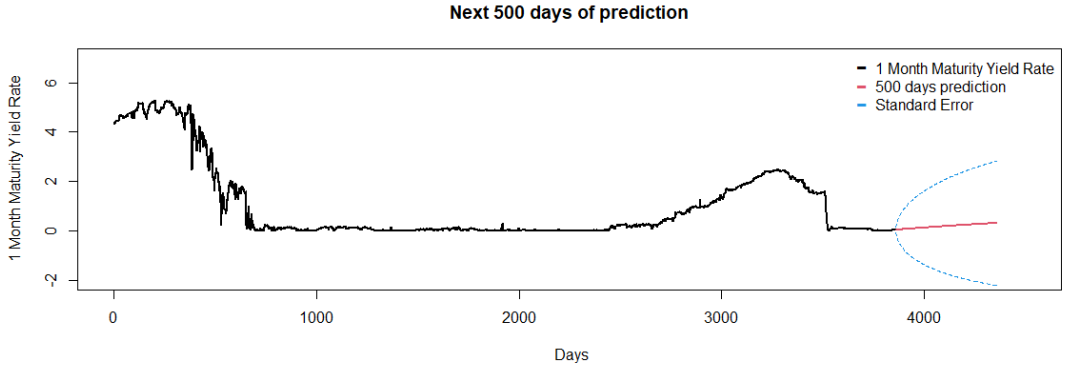


Figure 8 1 Month Yield Maturity Rate Next 500 Days Forecast

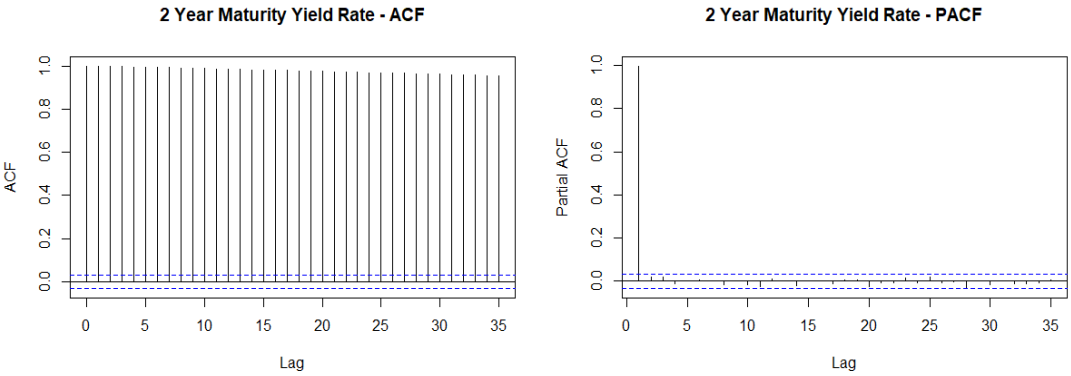


Figure 9 2 Year Yield Maturity Rate ACF and PACF

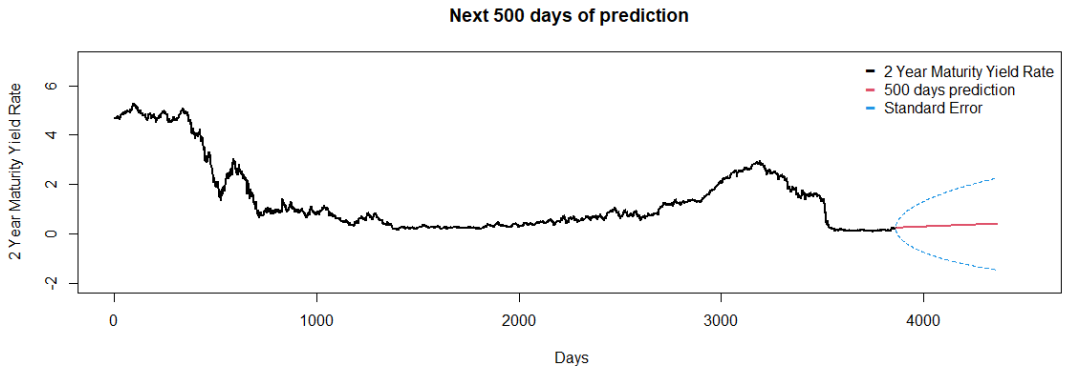


Figure 10 2 Year Yield Maturity Rate Next 500 Days Forecast

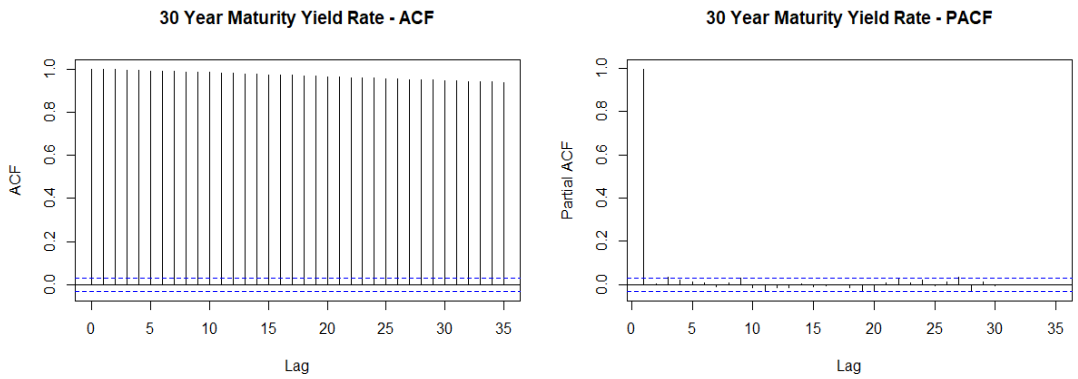


Figure 11 30 Year Yield Maturity Rate ACF and PACF

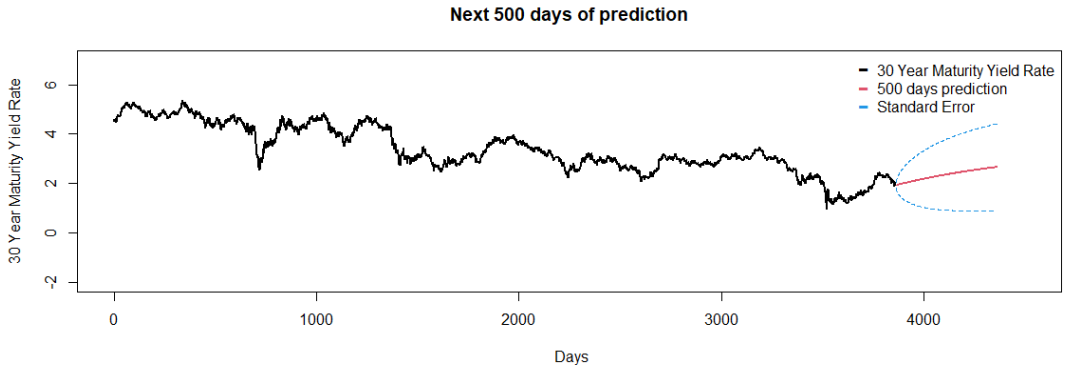


Figure 12 30 Year Yield Maturity Rate Next 500 Days Forecast

3.6. Implement Technical Indicators (Moving Average, Exponentially Weighted Moving Average on a 1-Month Yield Rate)

To see the trends of the Segmented datasets, we will use the moving average and exponentially weighted moving average on a 1-month yield rate.

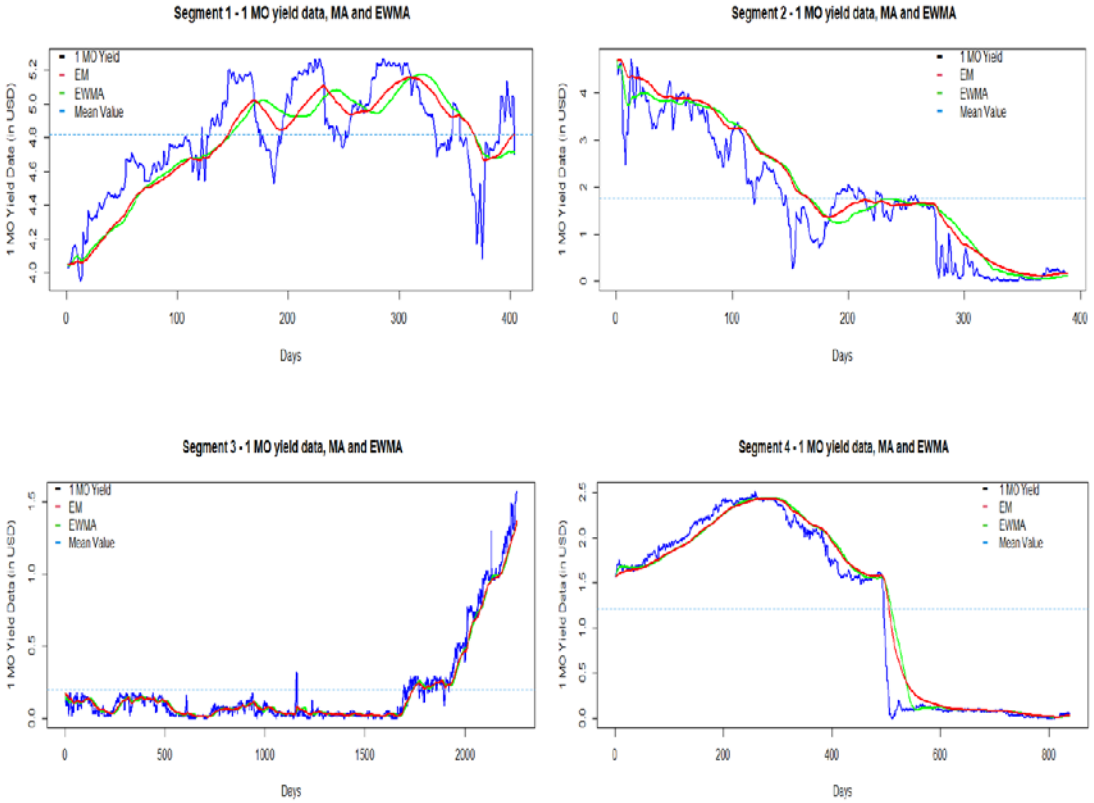
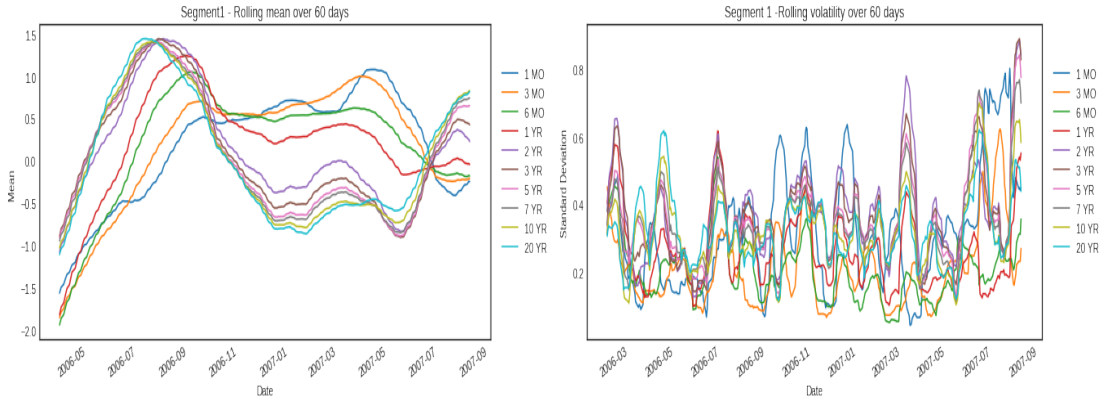


Figure 13 MA and EWMA

By smoothing the graph, we can see that there is an uptrend in segments 1 and 3 and a downtrend in segments 2 and 4.

3.7. Rolling Mean and Rolling Volatility



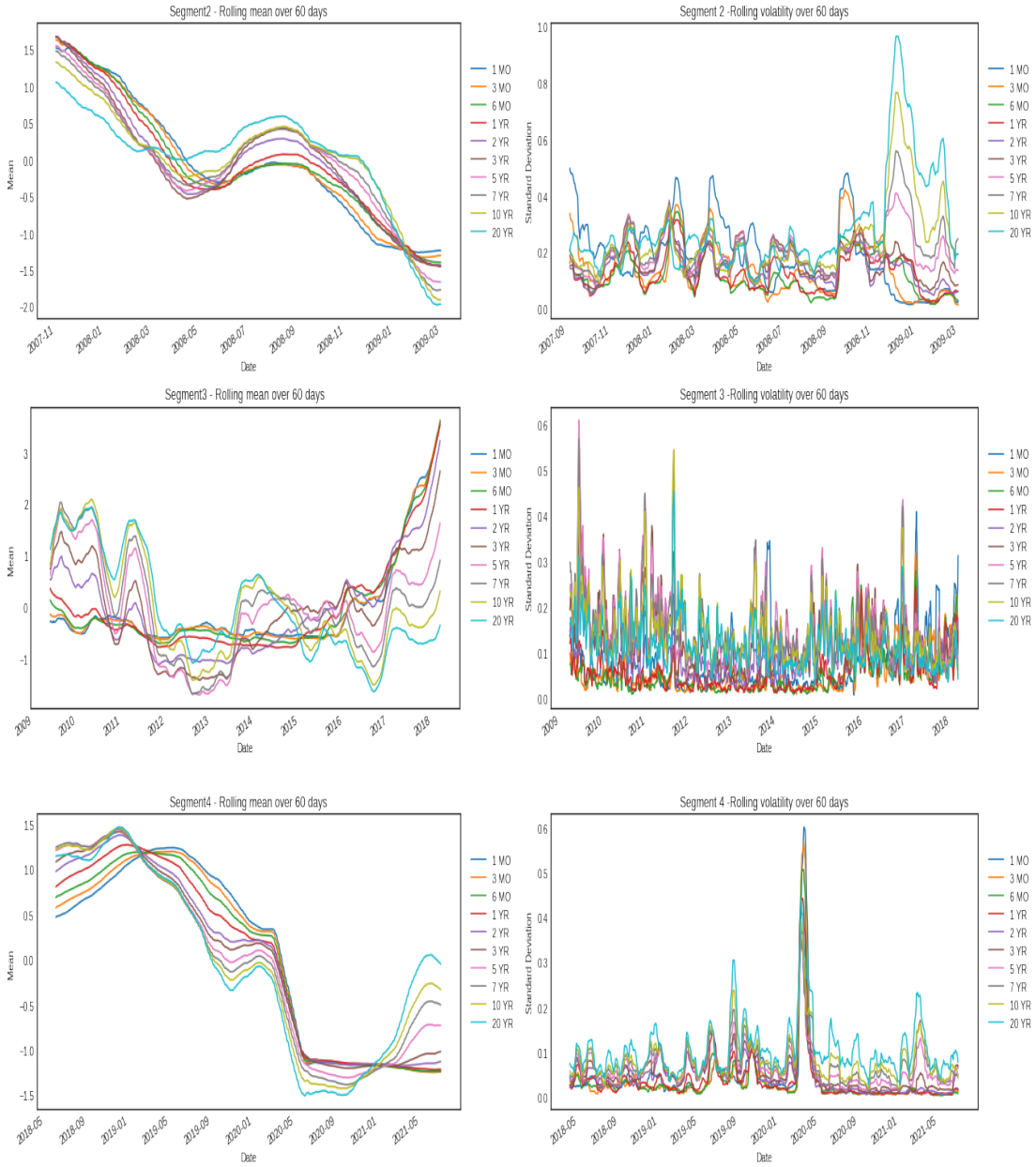


Figure 14 Rolling Mean and Volatility

3.8. Apply Jarque - Bera test for Yield 1 MO

We used the Jarque-Bera test for normality. This test has two components.

1. Skewness

2. Kurtosis

The skewness tells us about the symmetry of a distribution. If skewness is zero in the data, this indicates the data is normally distributed. There are three possible cases: Symmetrical, Positive, and Negative Skewness. The kurtosis tells us about the sharpness of the distribution. If the p-value is less than the critical value, then the null hypothesis will be accepted.

SEGMENT1

Robust Jarque Bera Test

X-squared = 20.48, df = 2, p-value = 3.571e-05

Kurtosis of 1 MO yield data in Segment 1: 11.23377

The skewness of 1 MO yield data in Segment 1: -0.09180424

SEGMENT2

Robust Jarque Bera Test

X-squared = 15.342, df = 2, p-value = 0.0004661

Kurtosis of 1 MO yield data in Segment 2: 1.975493

The skewness of 1 MO yield data in Segment 2: 0.3300049

SEGMENT3

Robust Jarque Bera Test

X-squared = 68460, df = 2, p-value < 2.2e-16

Kurtosis of 1 MO yield data in Segment 3: 7.772561

The skewness of 1 MO yield data in Segment 3: 2.364018

SEGMENT4

Robust Jarque Bera Test

X-squared = 64.747, df = 2, p-value = 8.771e-15

Kurtosis of 1 MO yield data in Segment 4: 1.26494

The skewness of 1 MO yield data in Segment 4: -0.1936394

It appears 1 Month Maturity Rates are normally distributed in all segments as the p values of the test results are less than 5%. Segments 1 and 4 are negatively skewed, Segments 2 and 3 are positively skewed. Additionally, Segments 1 and 3 are Leptokurtic indicating a positive excess kurtosis. This kind of distribution has heavy tails on both sides. This is an indication of large outliers. On the other hand, Segments 2 and 4 are a platykurtic distribution that shows a negative excess kurtosis. This distribution has a flat tail on both sides, and this indicates the small outliers in a distribution.

4. Principal Component Analysis - Covariance Method

Let's look at the correlation using Heat Map

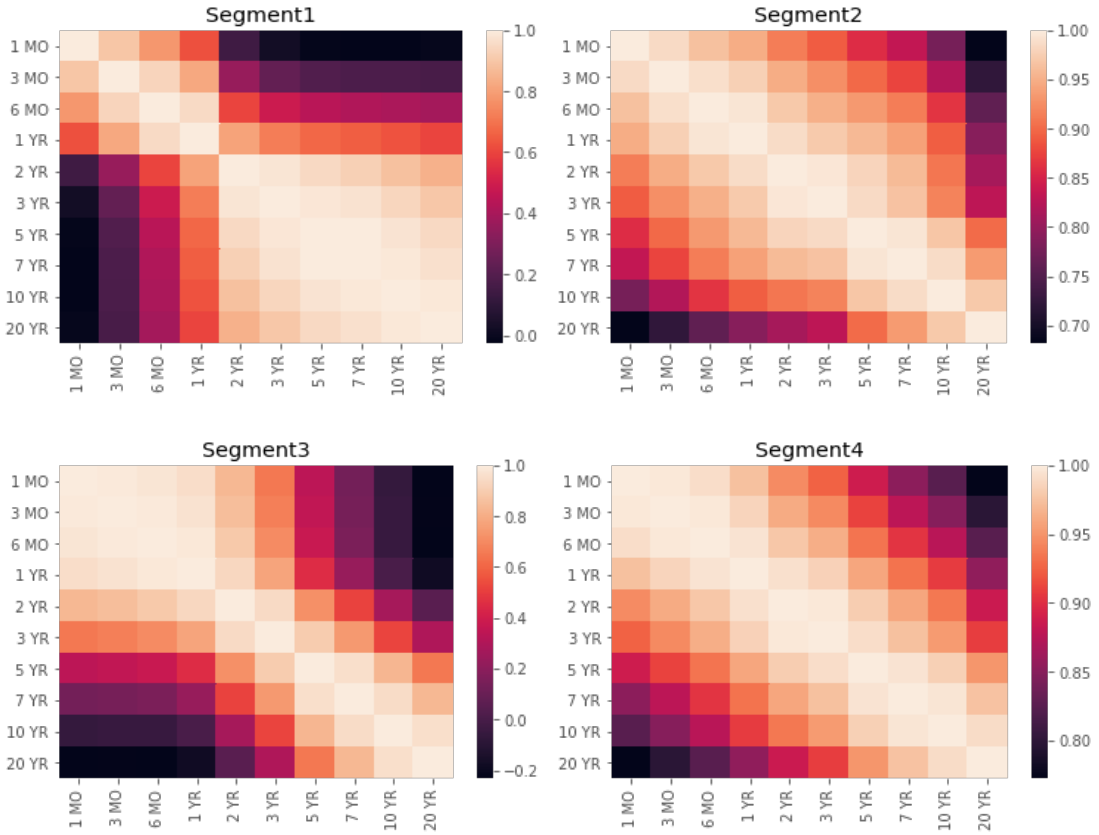


Figure 15 Heat Maps of Each Segment Covariance Matrix

4.1. Explained Proportions

	Eigenvalues	Explained Proportion
1	6.848904	0.684890
2	2.771177	0.277118
3	0.231407	0.023141
4	0.106794	0.010679
5	0.026791	0.002679
6	0.010487	0.001049
7	0.002564	0.000256
8	0.001193	0.000119
9	0.000401	0.000040
10	0.000282	0.000028

	Eigenvalues	Explained Proportion
1	9.220947	0.922095
2	0.610232	0.061023
3	0.120365	0.012036
4	0.028525	0.002852
5	0.010687	0.001069
6	0.004495	0.000450
7	0.001940	0.000194
8	0.001261	0.000126
9	0.000991	0.000099
10	0.000557	0.000056

	Eigenvalues	Explained Proportion
1	6.069410	0.606941
2	3.548766	0.354877
3	0.317633	0.031763
4	0.044494	0.004449
5	0.009403	0.000940
6	0.004478	0.000448
7	0.002281	0.000228
8	0.001745	0.000175
9	0.001116	0.000112
10	0.000675	0.000067

	Eigenvalues	Explained Proportion
1	9.421853	0.942185
2	0.512286	0.051229
3	0.060206	0.006021
4	0.003481	0.000348
5	0.001169	0.000117
6	0.000505	0.000050
7	0.000221	0.000022
8	0.000146	0.000015
9	0.000086	0.000009
10	0.000047	0.000005

Figure 16 Explained Proportion for 10 Component in each Segment

4.2. First Three-Components of Each Segment

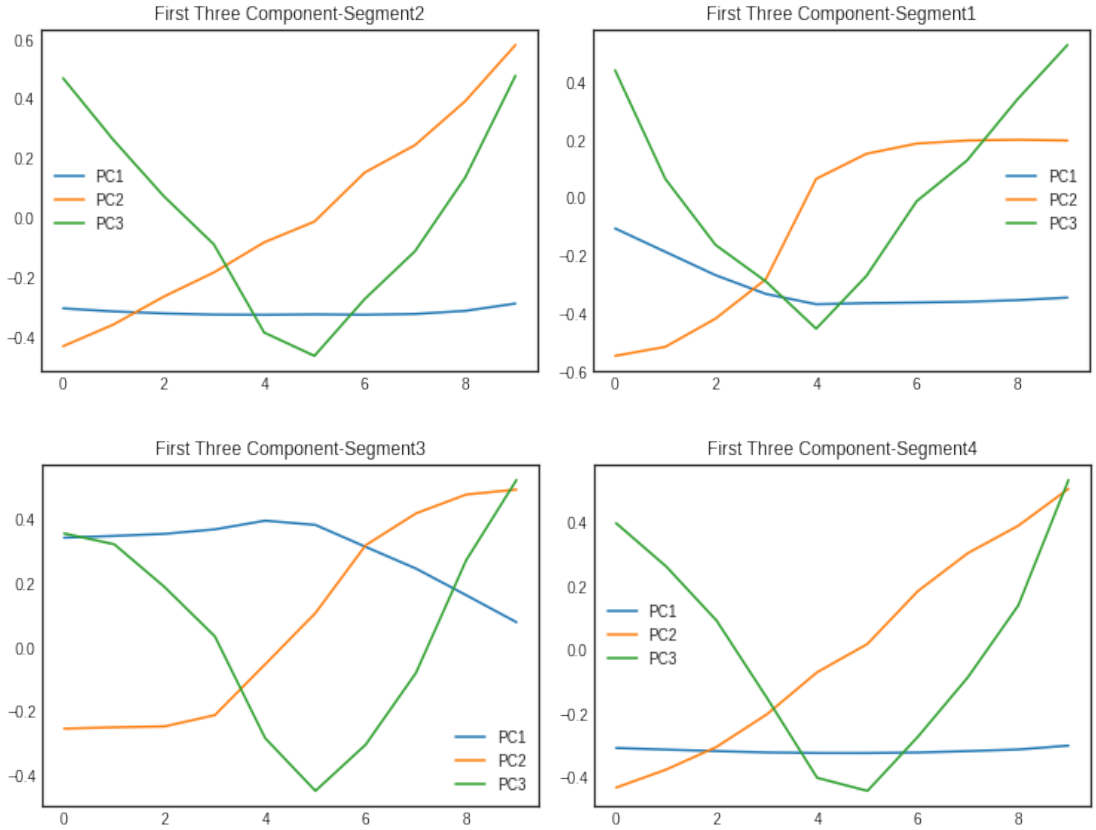
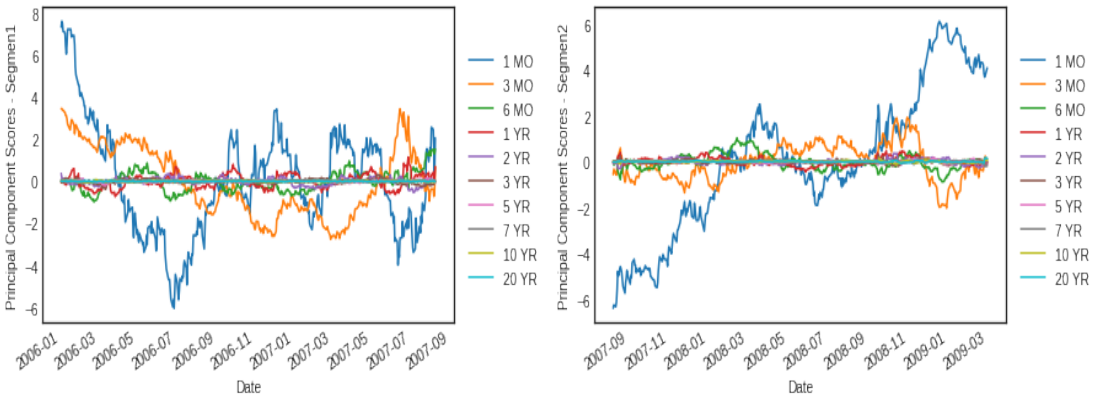


Figure 17 Graphical Representation of First Three-Components of Each Segment

4.3. Principal Component Scores



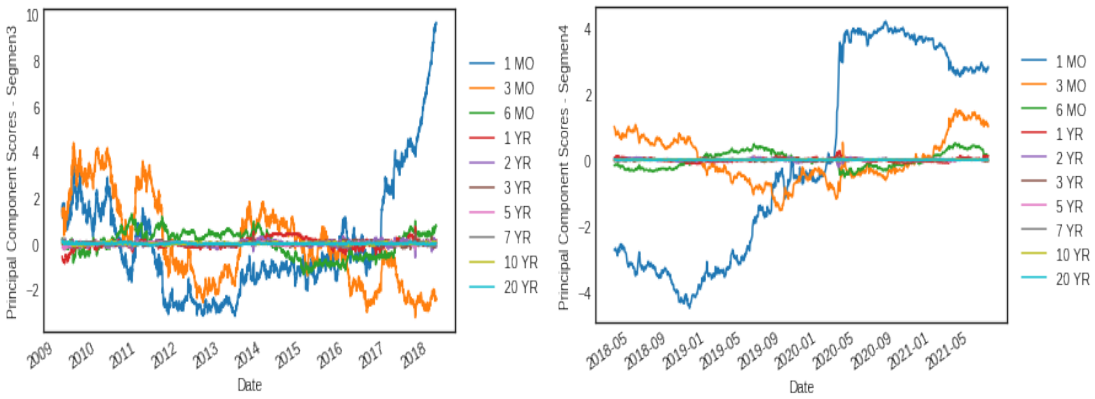
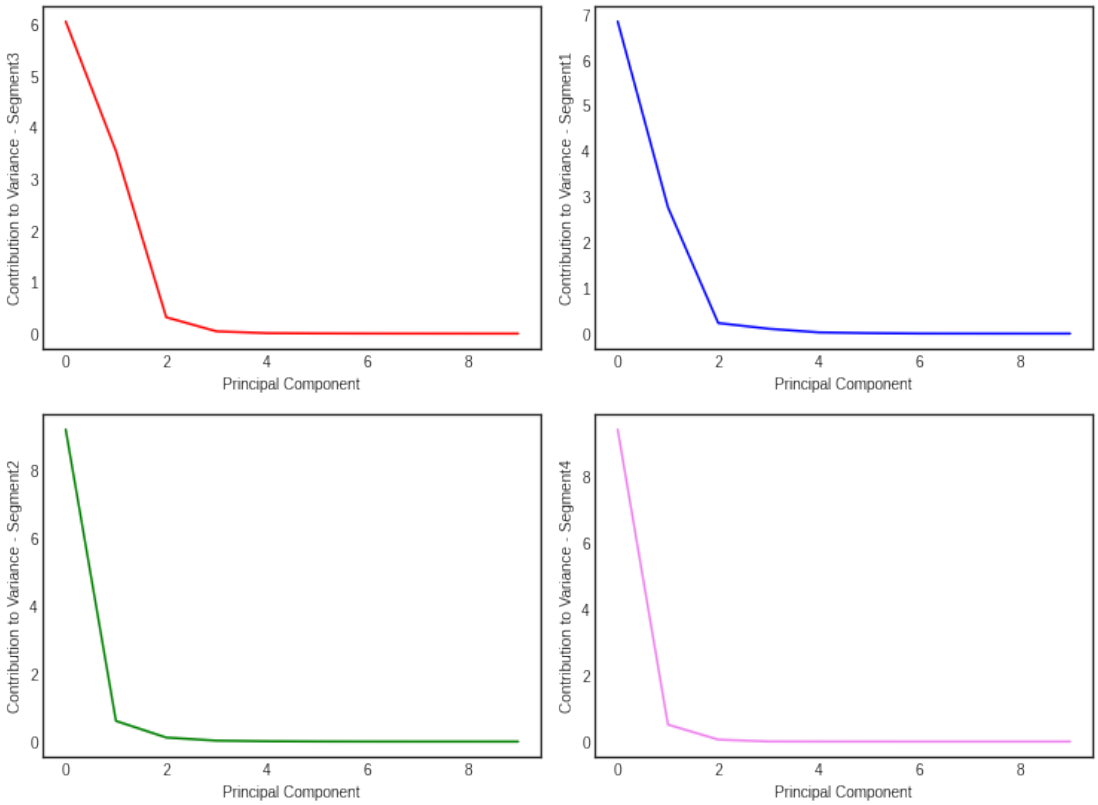


Figure 18 Principal Components Score

4.4. Contribution to Variance



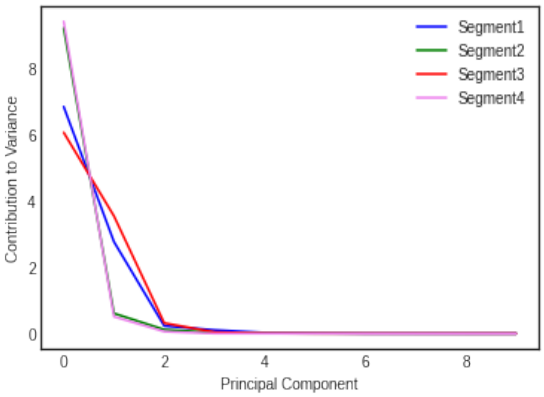


Figure 19 First 10 Components' Contributions to Variance in Each Segment

4.5. Cumulative Explained Variance against Number of Components

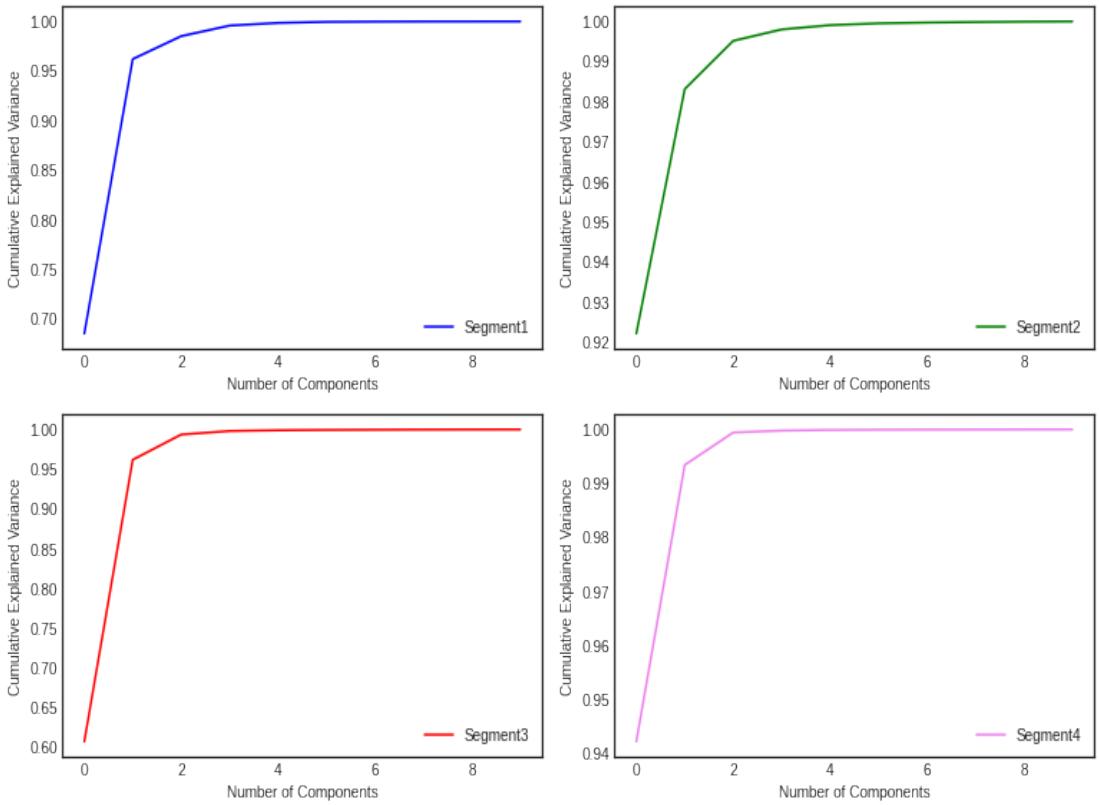


Figure 20 Cumulative Explained Variance of First 10 Components in Each Segment

4.6. Comparing PC 1- and 20-Year Maturity Rate – Level Factor

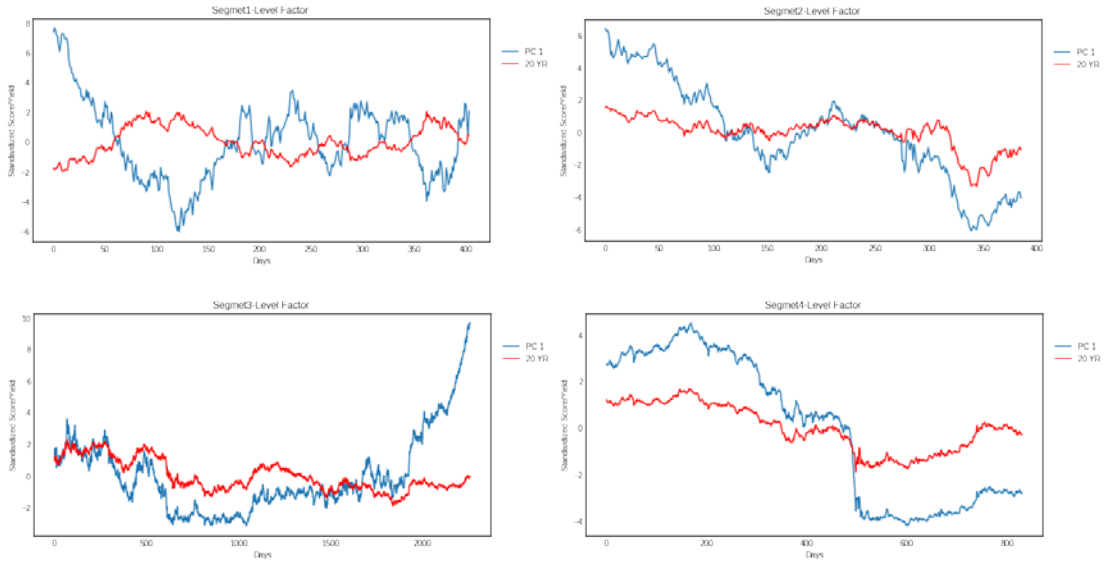


Figure 21 PC 1 - 20 Year Maturity Rate in Segment 1, 2, 3, 4

4.7. Comparing PC 2 and Slope – Slope Factor

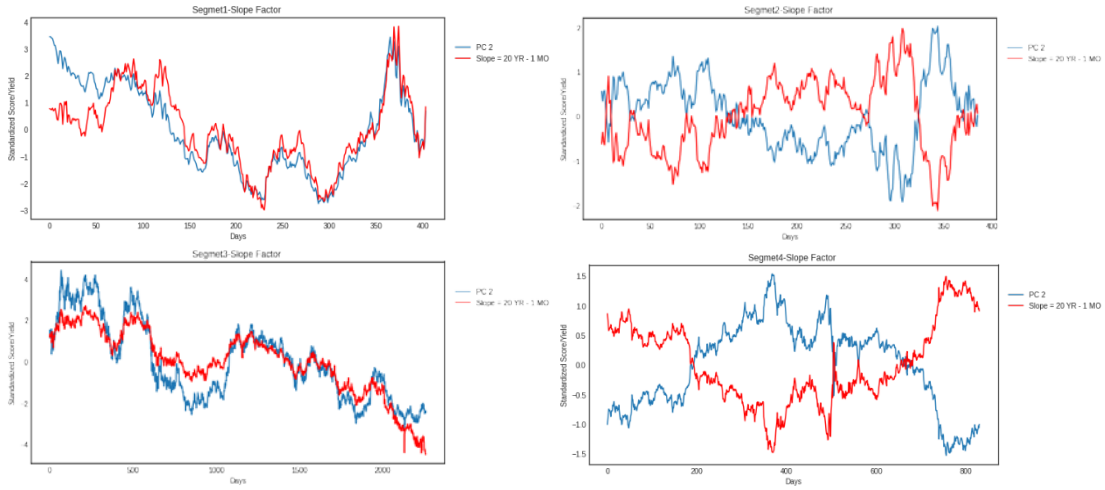


Figure 22 PC 2 – Slope in Segment 1, 2, 3, 4

4.8. Comparing PC 3 and Curvature – Curvature Factor

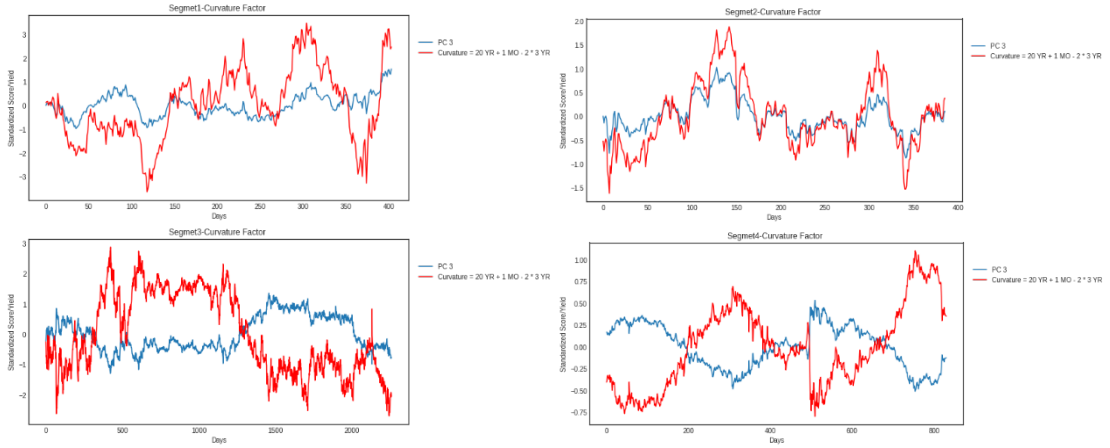


Figure 23 PC 3 – Curvature in Segment 1, 2, 3, 4

5. Conclusion

As it is explained in the Theoretical Framework, we retrieve the first 10 Principal Components of each Segment. Results show that the first two components explain 95 per cent of the variance. This is as expected.

Using 1 component, we can explain 68% of the variability, using 2 components, we can explain 96% of the variability, using 3 components, we can explain 98% of the variability and using 4 components, we can explain 99% of the variability in Segment 1.

Using 1 component, we can explain 92% of the variability, using 2 components, we can explain 98% of the variability, and using 3 components, we can explain 99% of the variability in Segment 2.

Using 1 component, we can explain 61% of the variability, using 2 components, we can explain 96% of the variability, and using 3 components, we can explain 99% of the variability in Segment 3.

Lastly, using 1 component, we can explain 94% of the variability, and using 2 components, we can explain 99% of the variability in Segment 4.

When we plot the level factor, we can see that PC 1 is similar to the 20 YR Yield Rate in Segments 2 and 4. This is because PC 1 can explain more than 90 % of data in those two segments.

In the Slope Factor, a comparison between 20 YR Yield Rate in Segment 1, 3 and PC 2 is expected. Both look similar as PC 2 can explain more than 95% of data in there. But Segments 2 and 4 surprisingly do not look like PC 2. This is not expected.

In the Curvature Factor, according to our theoretical framework, Curvature should look similar to PC 3. We can see that the 20 YR Yield Rate in Segment 1, 2 and PC 3 looks similar. But unexpectedly, Segments 3 and 4 look very different from PC3.

6. References

[1] https://www.quandl.com/data/USTREASURY/YIELD-Treasury-Yield-Curve-Rates

Korean, J. Anethesiol (May 2013). The Prevention and Handling of the Missing data.

Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3668100/

Srivastava, Sonam(October 23, 2019). Regime Shift Models- A Fascinating Use Case of

Time modelling. Retrieved from https://www.analyticsvidhya.com/blog/2019/10/regime-

shift models-time-series-modeling-financial-markets/