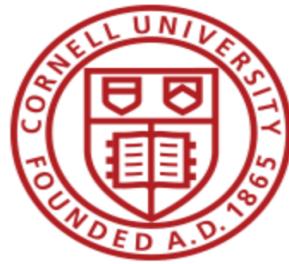


CORNELL UNIVERSITY



ORIE 5640 PROJECT

Statistical Traded Factor Analysis by Regime

Author:
Mingcan Yuan MY463
Supervisor:
Prof. Edward Mehrez
Lihua Dong LD457
Duo Sun DS2324
Runsheng Lin RL738

May 13, 2018



Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Regime Switching Model | 2 |
| 2.1 | Regime Switching Model by GNP | 2 |
| 2.2 | Regime Switching Model By S&P 500 | 3 |
| 3 | PCA on Fama-Bliss Yield | 3 |
| 3.1 | Full Period | 4 |
| 3.2 | Different Regimes | 4 |
| 4 | Regressions | 5 |
| 4.1 | Regressions in the Whole Sample | 6 |
| 4.2 | Regressions in the Normal Period | 6 |
| 4.3 | Regressions in the Crisis Period | 6 |
| 4.4 | Results Analysis | 7 |
| 4.5 | Regressions of the Market Portfolio | 9 |
| 5 | Strategy and Portfolio | 9 |
| 5.1 | AAA bond | 10 |
| 5.2 | BB bond | 11 |
| 5.3 | CCC bond | 11 |
| 5.4 | Portfolio Evaluation | 12 |
| 6 | Summary | 13 |
| 7 | Appendix | 13 |

1 Introduction

In this project, a credit portfolio was developed by statistical traded factor analysis with regime switching model. As the first step, we trained a regime switching model, separating market into 'normal' state and 'crisis' state. Then, principle component analysis was applied on segmented market, and statistical factor regression was built on result of PCA. Based on residuals, we constructed credit portfolio and back test the performance.

2 Regime Switching Model

Markov regime switching models are a specification in which the selling point is the flexibility in handling processes driven by heterogeneous states of the world (J. Hamilton 1994). In our project, we take two different regime for country economy and financial market, that one regime is with high volatility and another in low volatility. In other word, we can identify the states of the market with:

$$Y_t = \mu_1 + \epsilon_t \text{ For State 1}$$

$$Y_t = \mu_2 + \epsilon_t \text{ For State 2}$$

Where:

$$\epsilon_t \sim (0, \sigma_1^2) \text{ For State 1}$$

$$\epsilon_t \sim (0, \sigma_2^2) \text{ For State 2}$$

This represent two processes for our market indicator. When the state of market is 1 at time i , the expectation of the indicator is μ_1 with volatility at σ_1^2 . Likewise, when the state of market is 1, the mean and volatility take other values. In our project, the dependent Y_t can represent a vector of return. The μ_1 is the expected return on a bull market state, generally positive. The lower, and possibly negative value of μ_2 is the expected return on the bear market state, where asset price tends to go down. The difference in volatility suggests one with higher uncertainty than the other. In bear market, the volatility is expected to be greater than in the bull market possibly because of reasons such as risk averse and stop loss orders. This means we can expect the volatility in Crisis State to be higher than the volatility in Normal State.

In this section, we apply regime switching model on Gross National Product(GNP) and S&P 500 Index, to find the crisis predicted by regime switching model based on Macro and Market data. Here, we use the GNP data for it perfectly reflect the development of economy, and S&P 500 Index is one of the best indicator of the market.

2.1 Regime Switching Model by GNP

We implemented this Markov switching process using R's 'fMarkovSwitching' package and fit it with GNP data. We can notice that GNP data is quarterly, and generally GNP grows in a steady speed.

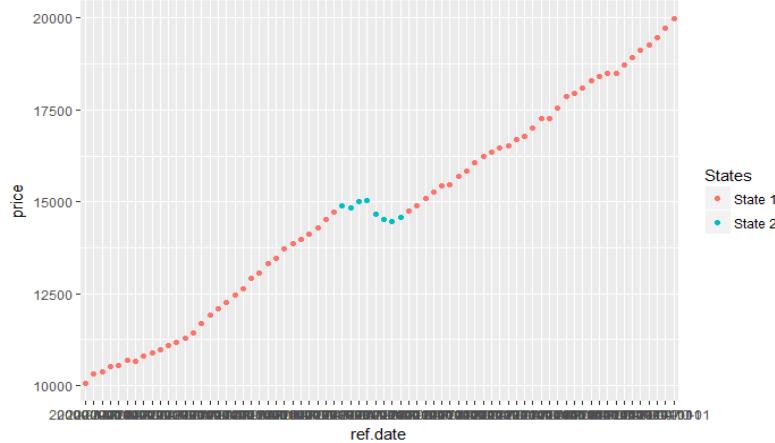


Figure 1: Variance explanation

Obviously, using GNP data, the regime switching model can only detect the financial crisis in 2008, with green dots shown on the diagram. For GNP represent the American economy instead of market trend, it is hard to impact the country economy for normal market crisis, which means the GNP is not sensitive to market crash. So we can hardly use the GNP to detect the crisis in market and give us a good insight about the market state.

2.2 Regime Switching Model By S&P 500

We also implemented the regime switching model with S&P 500 Index, that could offer a better insight on market states. Actually we need to use monthly data in S&P 500, to match with our bond yield in next section.

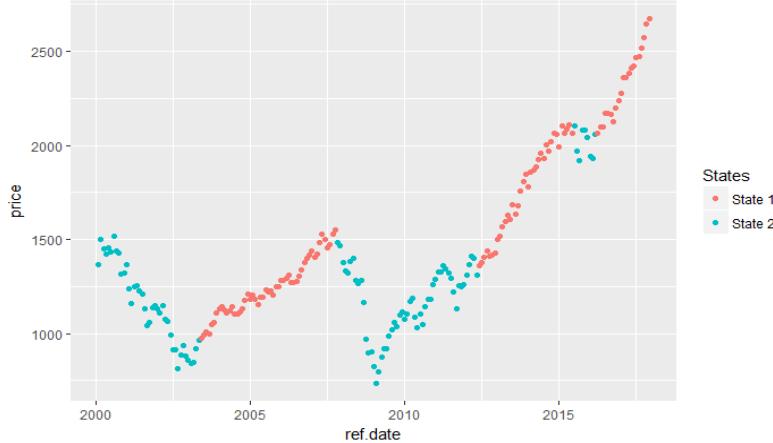


Figure 2: Variance explanation

As what we expected, the S&P 500 data works as a good indicator for market crisis. From the green dot, we can identify three period of crisis in financial market, the collapse of dot-com bubble at beginning of new century, the financial crisis in 2008 and oil crisis in 2015. the partition is good for our next work.

For our S&P 500 monthly return model, the mean return during Normal State is 1.14% while the model standard deviation is 2.21%. In Crisis State, the mean return is -0.4% and the model standard deviation is 5.5%. The key takeaway we found in almost all of our model is that S&P 500 monthly return goes up faster in the Normal State than in the Crisis State. Also, The return in Crisis State is generally negative.

Overall, regimes decided by the S&P 500 data are more reasonable compared to the GNP data, so in the next few sections, our analysis will focus more on the S&P 500 market regimes.

3 PCA on Fama-Bliss Yield

Principal component analysis (PCA) is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set. We assume the Fama-Bliss Yield data are correlated, hence we use PCA to transform the original variables to the linear combination of these variables which are independent. Then we select the number of Principal Components to be considered, given the percentage of variation that we want to be captured in the abridged data set. In this section, we want to use PCA to find the bonds that affect the interest risk the most and we want to extract those "importance" for later use in the linear regression which can help us to visualize the other risks embedded in the corporate bonds.

We would like to perform PCA by monthly Fama-Bliss data for holding period returns instead of FRED data, because Fama Bliss bonds are fully taxable and non-callable. Also, since there is missing value in the monthly Fama-Bliss data before Year 2000, we only perform PCA on the Fama-Bliss Yield ranging from 2000 to 2017. Also, because Fama Bliss data with long maturity date are more volatile, we only use 10 types of Treasuries with different maturity in the Fama-Bliss Yield data with relative short-term.

3.1 Full Period

We first perform PCA on the Fama-Bliss Treasury yields in the full-period, from 2000 to 2017 and get the 10 PCs.

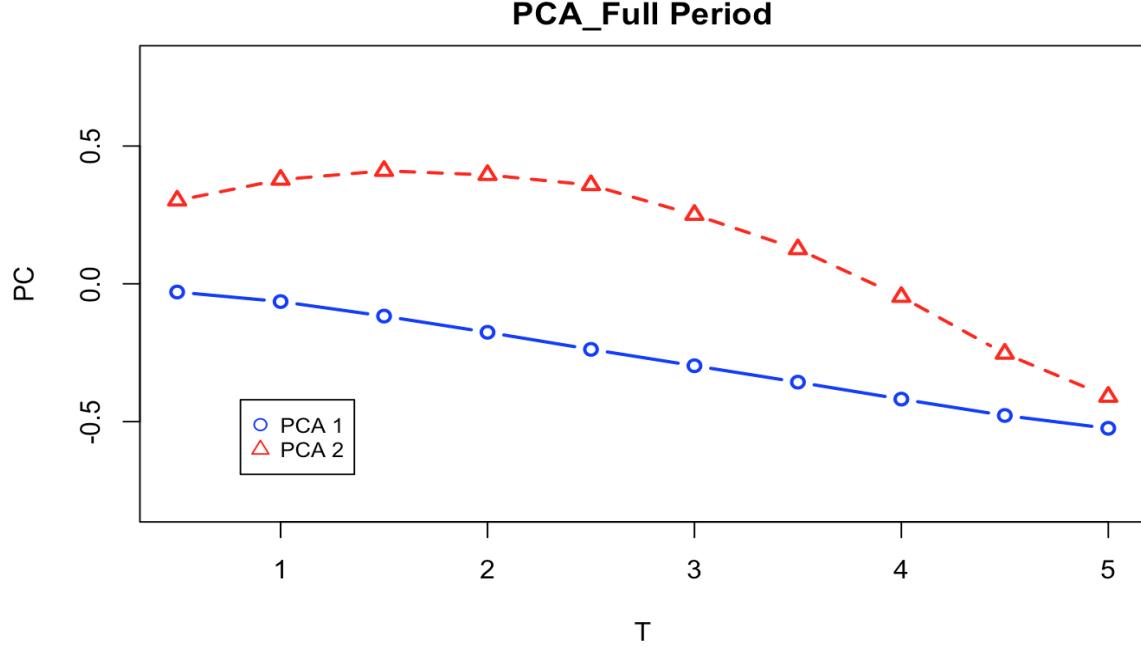


Figure 3: PC1 and PC2 Eigen Vectors

The result of PCA shows that PC1 and PC2 can explain up to 99% of the interest rate risk from 2000 to 2017. Also, from the diagram above, we could assert that bond with maturity 3 years to 5 years are the major component of PC1.

3.2 Different Regimes

Based on the regime switching model introduced in Section 2, we can split the Year 2000 to 2017 into 2 states, the Normal and Crisis State based on S&P 500 Index, for S&P 500 could better reflect the trend of market. We perform PCA on these two state separately to see if there are any difference in the PCs.

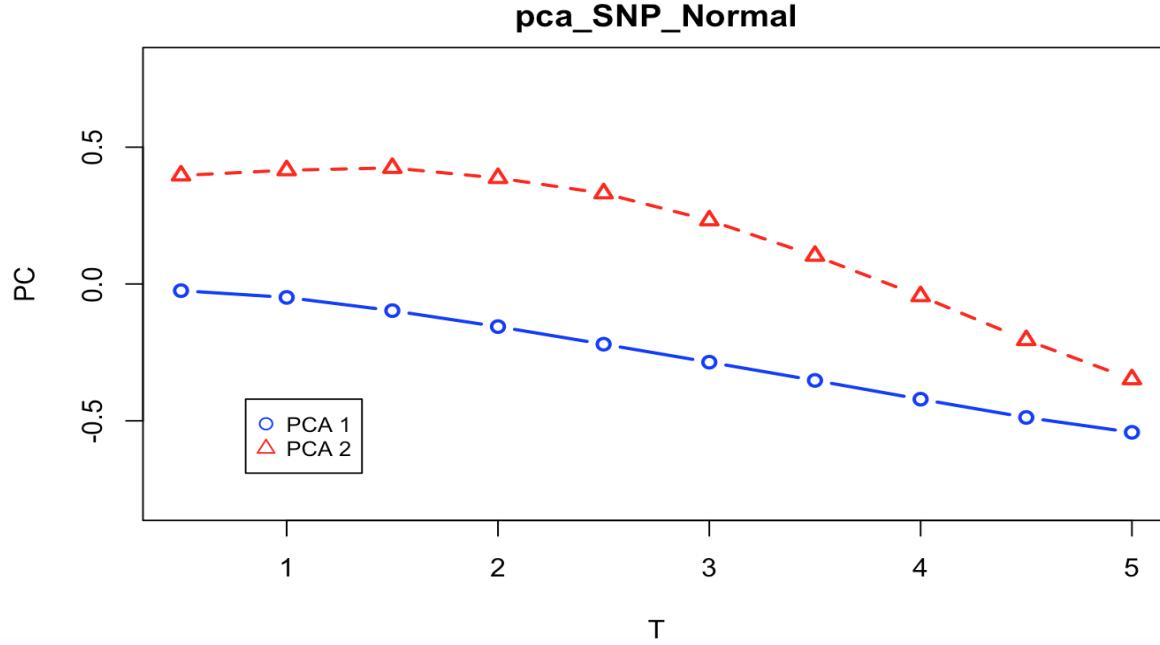


Figure 4: PC1 and PC2 Eigenvectors (Normal State)

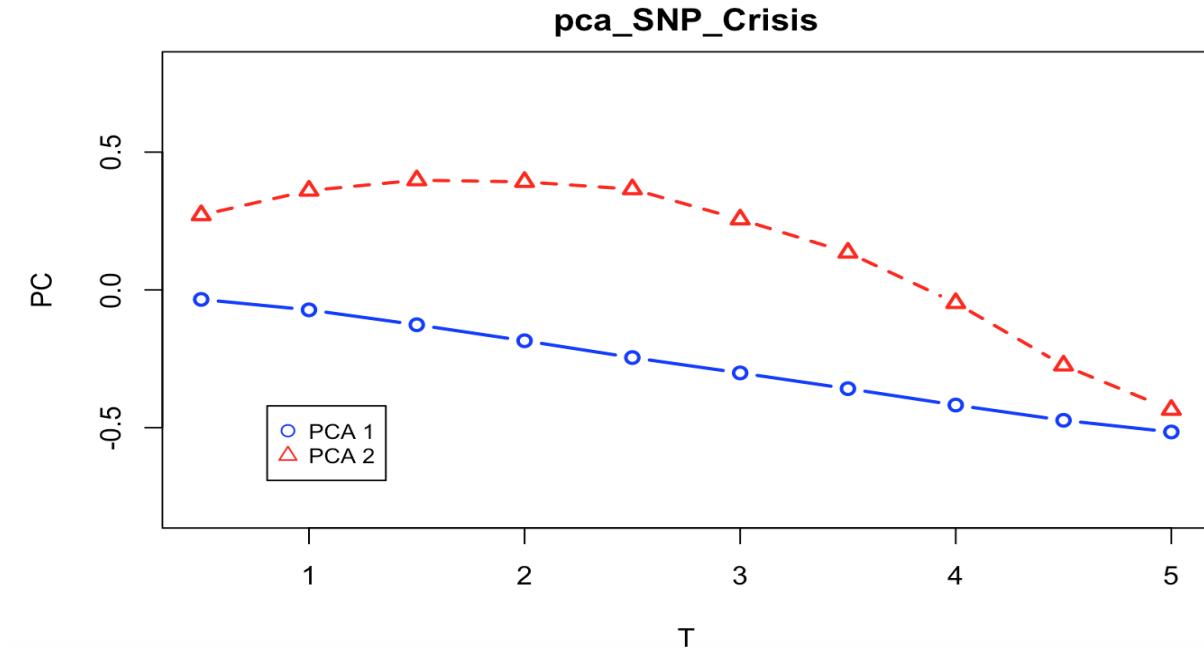


Figure 5: PC1 and PC2 Eigenvectors (Crisis State)

The PCA results show that PC1 and PC2 can explain up to 99% of the interest rate risk from 2000 to 2017 in both the normal and crisis state. Also, Bond with maturity 3 years to 5 years have the highest contributions to PC1.

Table 1: PC1 S&P 500 Regime

| Maturity(year) | 0-0.5 | 0.5-1 | 1-1.5 | 1.5-2 | 2-2.5 | 2.5-3 | 3-3.5 | 3.5-4 | 4-4.5 | 4.5-5 |
|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| PC1 Normal | -0.024 | -0.049 | -0.097 | -0.156 | -0.22 | -0.286 | -0.352 | -0.421 | -0.488 | -0.542 |
| PC1 Crisis | -0.034 | -0.072 | -0.126 | -0.184 | -0.246 | -0.301 | -0.358 | -0.418 | -0.473 | -0.504 |
| Difference | 0.0006 | 0.0028 | 0.006 | 0.010 | 0.012 | 0.009 | 0.004 | -0.003 | -0.014 | -0.029 |

Since each entry of PC1 is the major contribution of that corresponding bond to the interest rate risk, in table 1, last row is to investigate the difference in the contribution of each bond in different states, i.e. Normal and Crisis. The difference is defined as

$$\text{difference}_i = (pc1_normal)_i^2 - (pc1_crisis)_i^2 \quad \text{for } i = 1 \rightarrow 10$$

Since the last few entries in difference are negative while the rest are positive, we can conclude that under the regime switching based on S&P 500 data, the long-term bond could represent more variance in interest rate risk in Normal period and the short-term bond could represent more variance in interest rate risk in Crisis period.

Overall, since there is not much difference in the PCs and the general pattern of variance explained, for the rest of the project, and the regime switching model shows the S&P 500 regimes are more reasonable, we will stick to use the PC1 and PC2 generated from different regimes under S&P 500.

4 Regressions

As we mention above, we performed PCA on Fama-Bliss yields with different maturities and get the principal components. We can see that the first two principal components can explain more than 99% variance in the whole sample, the normal period and the crisis period. Next, we would perform linear regressions on the ICE BofAML US Corporate Effective Yields with different credit ratings (AAA, BB and CCC), and see whether the two principal components of the Fama-Bliss yields can explain the interest rate risk of these corporate bond yields.

We would also analyze the residuals of the regressions because the residuals represent the residual risks of the corporate bond after eliminating the interest rate risk. The residual risks mainly contain the credit risks of the corporate bonds.

4.1 Regressions in the Whole Sample

We first regressed the three corporate bond yields with credit ratings AAA, BB and CCC respectively on the first two principal components of the Fama-Bliss yields in the whole sample. The regression formula is

$$Y_t = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \epsilon_t$$

The table below shows the regression results.

Table 2: Regression results in the whole sample

| Variables | AAA | BB | CCC |
|-----------|----------------------|----------------------|----------------------|
| α | 0.0332*** (36.63) | 0.063*** (39.92) | 0.132*** (28.26) |
| PC1 | -0.138*** (-4.12) | -0.304*** (-5.20) | -0.734*** (-4.23) |
| PC2 | 2.169*** (12.11) | 1.914*** (6.12) | 4.086*** (4.41) |
| R^2 | 0.436 | 0.233 | 0.15 |

t statistics in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.2 Regressions in the Normal Period

We regressed the three corporate bond yields on the first principal components in the normal period that is determined by the Markov switching model on the S&P 500 monthly return. Note that the principal components here are different from those computed in the whole sample. The regression form is

$$Y_t = \alpha + \beta_1 PC1_t^{(normal)} + \beta_2 PC2_t^{(normal)} + \epsilon_t$$

where t is the time in the normal period.

The table below shows the regression results.

Table 3: Regression results in the whole sample

| Variables | AAA | BB | CCC |
|-----------|---------------------|----------------------|---------------------|
| α | 0.029*** (27.70) | 0.051*** (53.17) | 0.110*** (58.40) |
| PC1 | -0.073 (-1.57) | -0.0980** (-2.34) | 0.029 (0.35) |
| PC2 | 2.816*** (10.35) | 2.515*** (10.26) | 0.286 (0.60) |
| R^2 | 0.504 | 0.506 | 0.004 |

t statistics in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.3 Regressions in the Crisis Period

We then regressed the three corporate bond yields on the first two principal components in the crisis period. The regression form is

$$Y_t = \alpha + \beta_1 PC1_t^{(crisis)} + \beta_2 PC2_t^{(crisis)} + \epsilon_t$$

where t is the time in the normal period.

The table below shows the regression results.

Table 4: Regression results in the whole sample

| Variables | AAA | BB | CCC |
|-----------|----------------------|---------------------|---------------------|
| α | 0.037*** (24.71) | 0.078*** (32.22) | 0.171*** (22.91) |
| PC1 | -0.127*** (-2.65) | -0.202** (-2.58) | -0.529** (-2.20) |
| PC2 | 1.926*** (8.07) | 1.419*** (3.64) | 4.596*** (3.84) |
| R^2 | 0.417 | 0.165 | 0.162 |

t statistics in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.4 Results Analysis

From the results of the three regressions in different market regimes, the significant coefficients show that the principal components of Treasury yields can mostly explain the interest rate risk of the three corporate bond yields in the whole sample, normal and crisis period. There is only one exception, that is, for the CCC yields, the coefficients are not significant in the normal period and the R^2 is small.

We further analyze the residuals of the regressions. The residuals of the regression actually can represent the other risks except the interest rate risk of the corporate bonds. The major risk in the residuals is the credit risk. Thus, we can explore the residuals to analyze the credit risk premium of the bond.

We performed some tests on the residuals. We attached all the test plots in the appendix. First, we tested the auto-correlations of the residuals by the acf plot in R. From the acf plot, we can see that the residuals have significant auto-correlations. These may cause some bias in the coefficients of the regression, but we still kept our analysis because the model has its own economic sense. Then we tested the normality of the residuals. We can see that the AAA bond residuals appear normal in different market regimes. The BB and CCC bond residuals have slightly heavy right tail in the crisis period, which corresponds to high credit risk premium in the crisis period.

Then we plot the residuals of different market regimes in the same figure to compare credit risk premium. We can see a pattern in the figures. The residuals in the crisis period are slightly lower than those in the whole sample, while the residuals in the normal period are slightly higher than those in the whole sample. Our explanation is that people are willing to pay more credit premiums in the crisis period compared to the true credit risk premium, while they pay less credit premiums in the normal period. In the real world, people may find it hard to distinguish the normal and crisis period, so the residuals in the whole sample can represent the credit risk premiums that people actually paid. However, the credit risk premiums in the crisis period may be represented more accurately by the residuals in the crisis period because the residuals are obtained only in the crisis period. This argument can still be applied in the normal period.

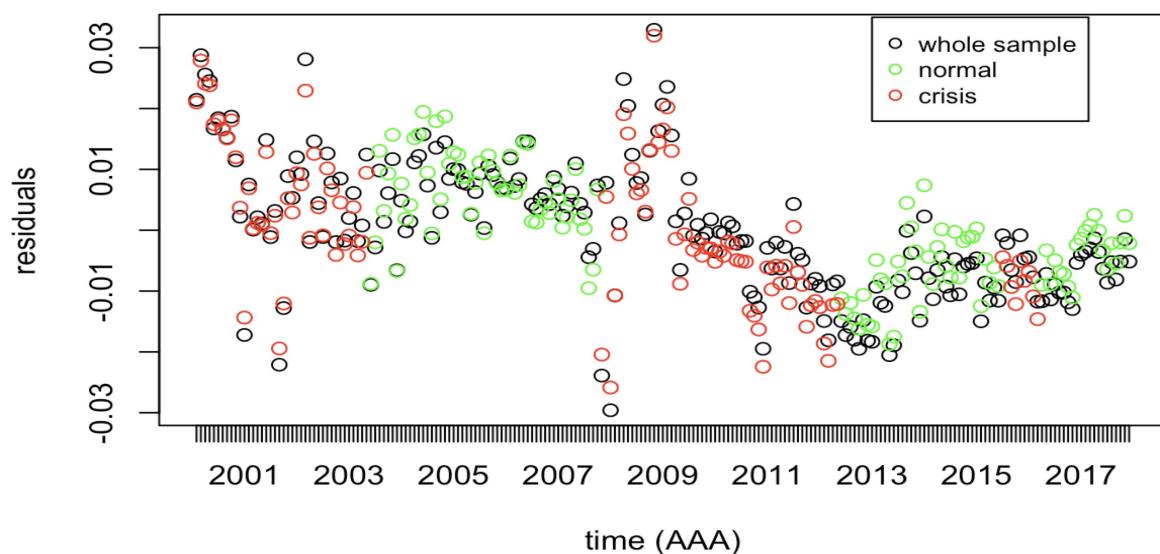


Figure 6: AAA Bond Residuals

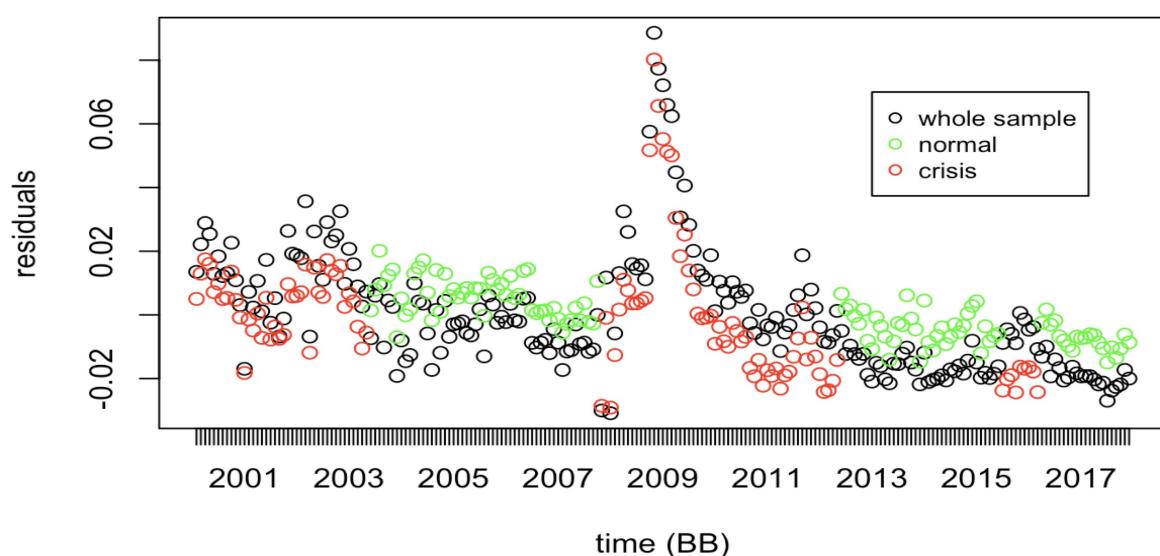


Figure 7: BB Bond Residuals

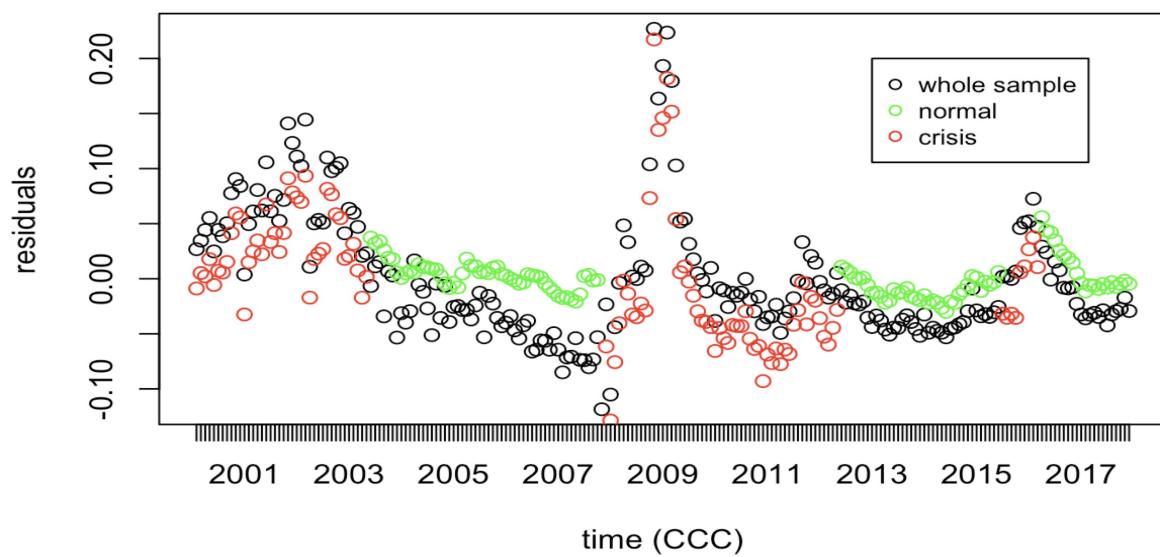


Figure 8: CCC Bond Residuals

4.5 Regressions of the Market Portfolio

We used S&P 500 index as the market portfolio, and projected the market portfolio returns on the first two principal components in the whole sample, the normal and crisis period. The regression forms are

$$R_t = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \epsilon_t$$

$$R_t = \alpha + \beta_1 PC1_t^{(normal)} + \beta_2 PC2_t^{(normal)} + \epsilon_t$$

$$R_t = \alpha + \beta_1 PC1_t^{(crisis)} + \beta_2 PC2_t^{(crisis)} + \epsilon_t$$

The regression results are shown below.

Table 5: Regression results in the whole sample

| Variables | Whole Sample | Normal | Crisis |
|-----------|--------------------|--------------------|--------------------|
| α | 0.012*** (3.63) | 0.012*** (4.60) | 0.011* (1.76) |
| PC1 | 0.696*** (5.88) | 0.191* (1.72) | 0.920*** (4.63) |
| PC2 | -0.655 (-1.04) | -0.005 (-0.01) | -0.467 (-0.47) |
| R^2 | 0.144 | 0.0267 | 0.177 |

t statistics in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The results show that only the first principal components of the Treasury yields is significant in different market regimes. The residual plot below shows that there are no significant difference in the residual risks between different market regimes. We think the reason is the stock market is different from corporate bond market. In addition to interest rate risk, there are many other risk sources associated with the stock market.

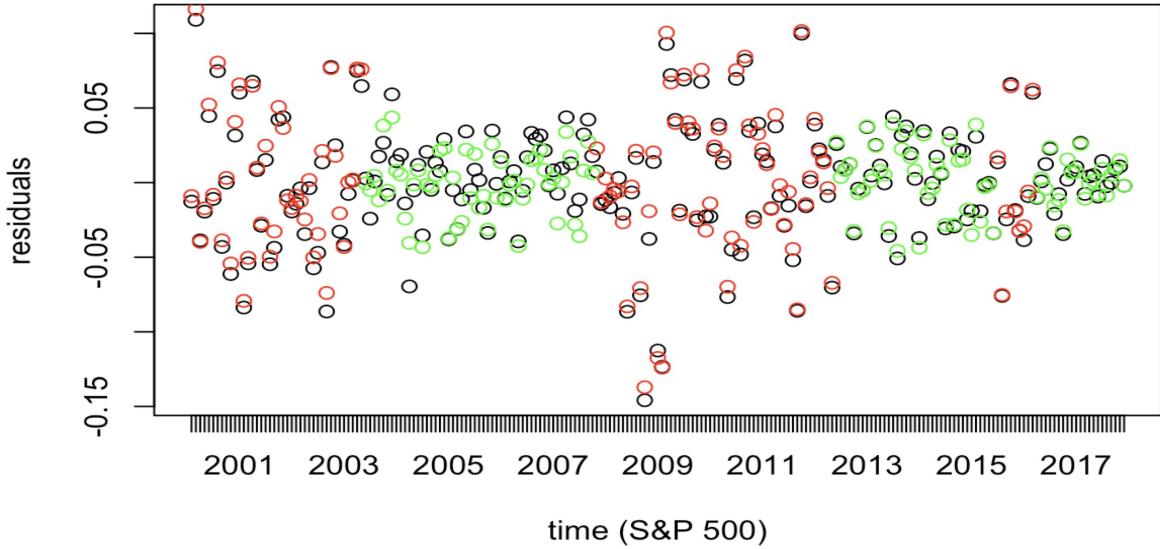


Figure 9: BB Bond Residuals

5 Strategy and Portfolio

In this section, we want to use the results from linear regression and principle component of PCA to construct our strategy and portfolio, to catch the credit risk in different period and get the risk premium to make profit. As we have mentioned above, we can quantify the

credit risk as residual and intercept, when regress bond price, such as AAA bond, against interest rate and Fama-Bliss yield. And we call the portfolio as credit portfolio

In order to construct credit portfolio, we assign weights for each Fama Bliss bond by the sum of weights in PC1 and PC2. As for the bond with maturity at time i , after regressing AAA bonds against Fama Bliss yield, we assign bond with the weight $weight_i$ if we have 'one' AAA bond:

$$weight_i = -(\beta_1 \times (weight\ of\ i\ in\ PC1) + \beta_2 \times (weight\ of\ i\ in\ PC2))$$

When $weight_i$ is positive, it means we need to long $weight_i$ of bond with maturity at time i with every 'one' AAA bond, and when $weight_i$ is negative, it means we need to short $|weight_i|$ of bond with maturity at time i with every 'one' AAA bond. Obviously, in different period, PCA would output different component, and we would construct different portfolio in normal state or crisis state, which make sense that investors should obtain different risk preference in different period.

So we can construct credit portfolio for AAA, BB and CCC bond using regime switching model and PCA, analyze the effect of regime switching model and trend of credit crisis in the past decades.

5.1 AAA bond

For AAA bond, the weight for each bond portfolio with different maturity day, from Fama Bliss yield, is shown below. The blue chart represent weights for each bond in normal state, the red chart represent the weights for each bond in crisis state, and the green chart is for the weights when using whole sample instead of regime switching model.

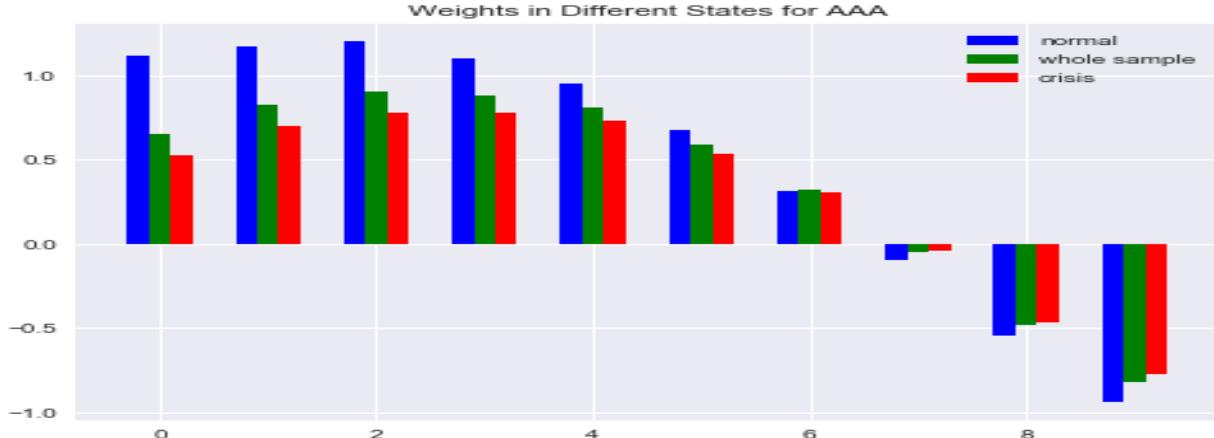


Figure 10: Weights for bond in AAA

As what we have expected, the whole sample weights is between the weights from normal state and crisis state. Investors are always willing to long the short-term yield and short long-term yield if they want to catch interest rate.

Based on the weight, we could long the AAA bond and short the portfolio above, to obtain the credit premium, which is shown below.



Figure 11: Credit Portfolio for AAA

Using the diagram on the left, we can find that, for AAA bond, regime switching model didn't improve the performance for credit portfolio, which may be explained by the low sensitivity of AAA bond toward market. From the diagram in the center, the return of credit portfolio of AAA bond shows high correlation with the market. We can see that the return of portfolio goes up when market crashed in 2008 and 2015, which means the risk premium increase with crisis and we can catch it with our credit portfolio. The diagram on the right shows the difference of return when using regime switching model, that regime switching didn't work well for AAA bond and sometimes even under performed.

5.2 BB bond

For BB bond, we could regard it as 'Junk Bond'. Intuitively, it would perform differently against the AAA bond. We can see the weights for each bond in order to hedge interest risk in BB bond as below.

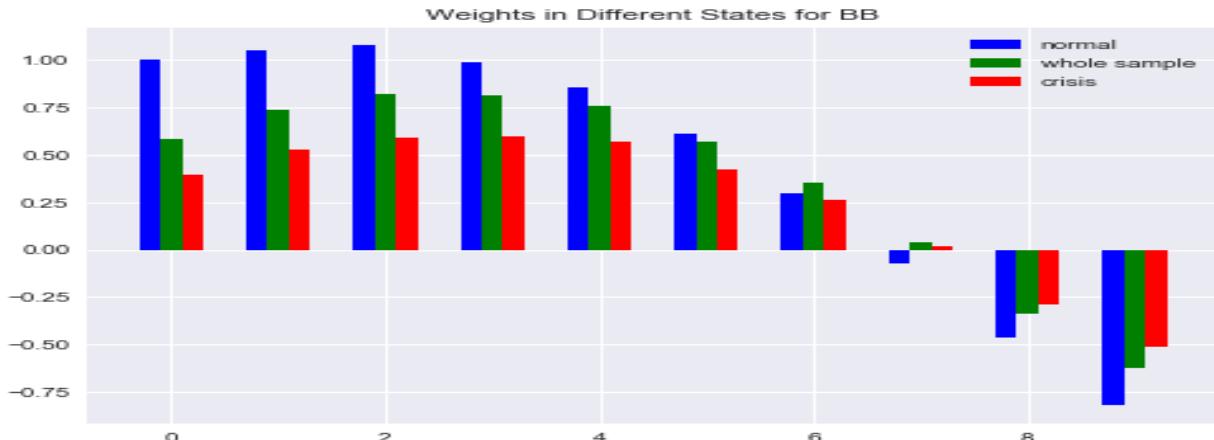


Figure 12: Weights for bond in BB

Obviously, the gap between height of blue bar and red bar becomes larger, which means the BB bond require us to hold the significantly different bond portfolio for normal and crisis period, when hedging the interest rate risk in BB bond. Moreover, we still long the short-term yield and short long-term yield in order to indicate the interest rate risk.

Then, we can check the performance of credit portfolio in BB bond.



Figure 13: Credit Portfolio for BB

Surprisingly, the regime switching model improve the performance for credit portfolio a lot, and the asset process is significantly increasing when separating market into different states. It means for the 'Junk Bond' Investors, they are more sensitive to the market environment, and we could make more profit with credit portfolio when bond are sensitive. Also, the sharp change of the return rate in the center diagram shows the volatile risk premium in BB bond. And the improvement of regime switching model is always positive in left diagram, which means the regime switching model can always improve our credit portfolio.

5.3 CCC bond

For CCC bond, the property of Junk Bond shows more significantly.



Figure 14: Weights for bond in CCC



Figure 15: Credit Portfolio for CCC

It seems that regime switching model could improve more if the bond is more 'junk', in other word, more sensitive to the market and more volatile. We can interpret this phenomenon that regime switching model could provide us with more accurate insight to the market regardless of the states. And it could improve more if the states are more different.

5.4 Portfolio Evaluation

From the credit portfolio constructed by AAA, BB and CCC bond , we have the results in the past decades below. We can assess the credit portfolio by average return and Sharpe Ratio.

Table 6: Performance of Credit Portfolios

| Bond | AAA | BB | CCC |
|--------------|-------|-------|-------|
| Return | 1.23% | 3.15% | 2.93% |
| Sharpe Ratio | 1.85 | 1.60 | 3.36 |

Based on the Sharpe ratio, the return of credit portfolio shows a steady growth and good Sharpe Ratio, which befits from steady yield of the bond and stabilized credit premium. But we need to choose our best portfolio between BB bond and CCC bond, for the reason that credit portfolio built on AAA bond possess relative low annually return, and we can get higher return as well as good Sharpe Ratio in BB bond and CCC bond.

Between BB and CCC, we prefer to choose the credit portfolio built on BB bond, for return dominates Sharpe Ratio when both of them are good enough. For the risk resource of bond portfolio, there are other risk except interest rate risk and credit risk, such as liquidity risk. So we could not perfectly catch the credit risk, and our portfolio may possess other risk. It means we'd better build our portfolio on relatively good bond and BB would be better than CCC.

6 Summary

In this report, we firstly conduct regime switching model on GNP data and S&P 500 Index, to get an insight about different states based on different data. We find that, compared with GNP data, S&P is more sensitive to the market trend and accurate in identifying crisis. The GNP could only show the financial crisis in 2018, which is not enough for us to train our credit portfolio. But S&P could even identify dot-com bubble collapse in 2000 and oil crisis in 2015, which is better for crisis detection.

Also, we apply PCA on the collections of Fama Bliss Yield and analyze the difference between the results of PCA with PCs major component. Then we use the PC1 and PC2 to do the statistical factor regression on high rating bond and low rating bond, checking if regime switching model could offer better understanding of the market. And we analyze the residual in different period to see the changes of credit risk in the past decades.

At last, we construct credit portfolio with the residual and intercept from the result of statistical factor regression. By comparing the effect of regime switching model and credit rating bonds, we find that the lower is credit rating, the better is the effect of regime switching model. And we evaluate different credit portfolio and choose the best credit portfolio built on BB bond.

7 Appendix

Table 7: Maturities of the bond

| | |
|----------------|----------------|
| 0 to 6 Month | 6 to 12 Month |
| 12 to 18 Month | 18 to 24 Month |
| 24 to 30 Month | 30 to 36 Month |
| 36 to 42 Month | 42 to 48 Month |
| 48 to 54 Month | 54 to 60 Month |

Variance Explanation in section 3

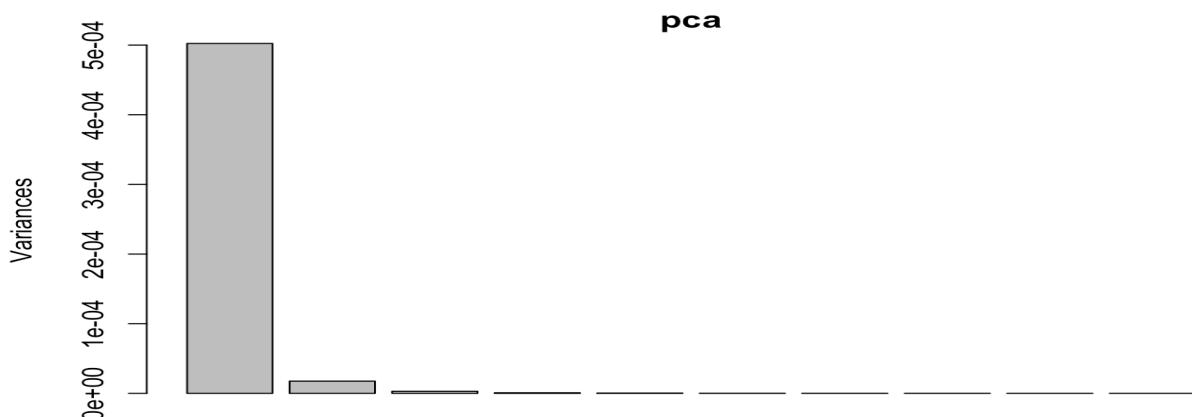




Figure 17: Variance Explanation for Crisis State



Figure 18: Variance Explanation for Normal State

The auto-correlation and normality tests of the residuals in the section 5:

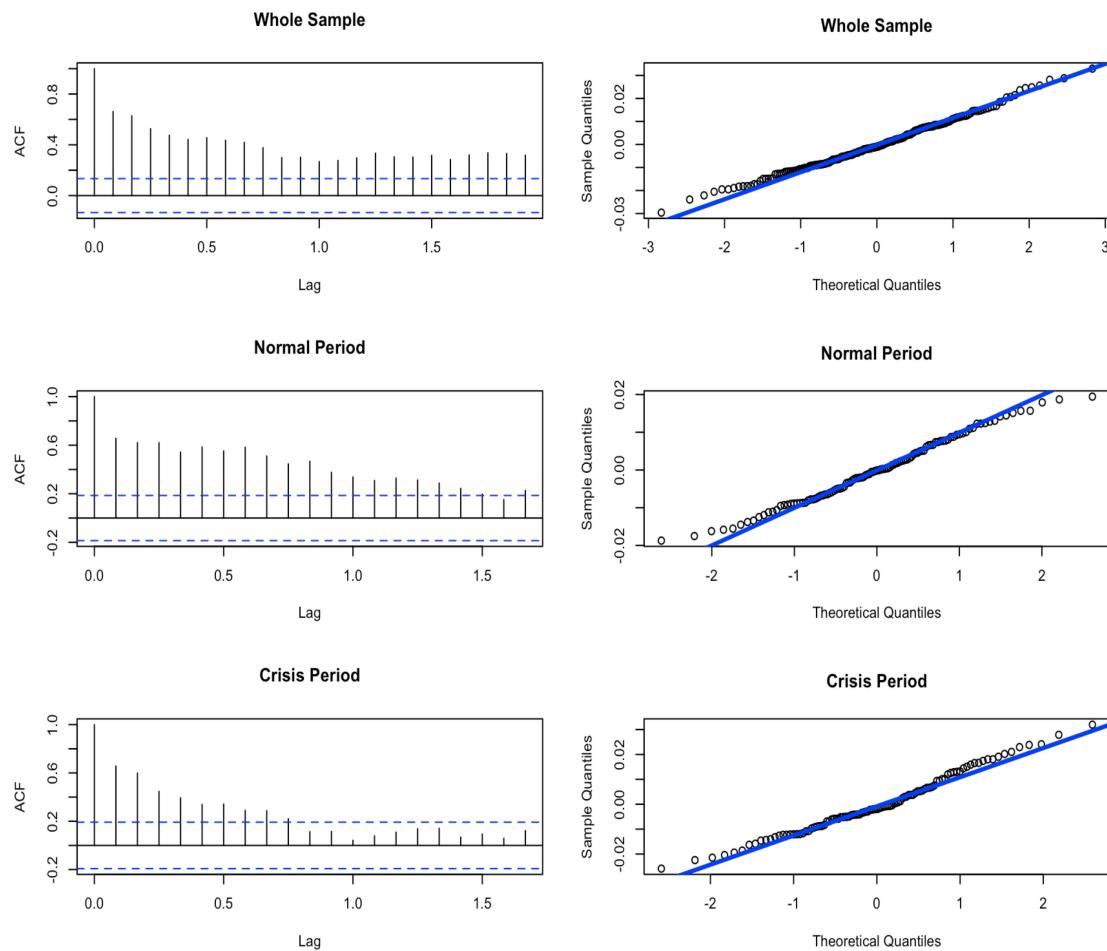


Figure 19: AAA Bond Residual Tests

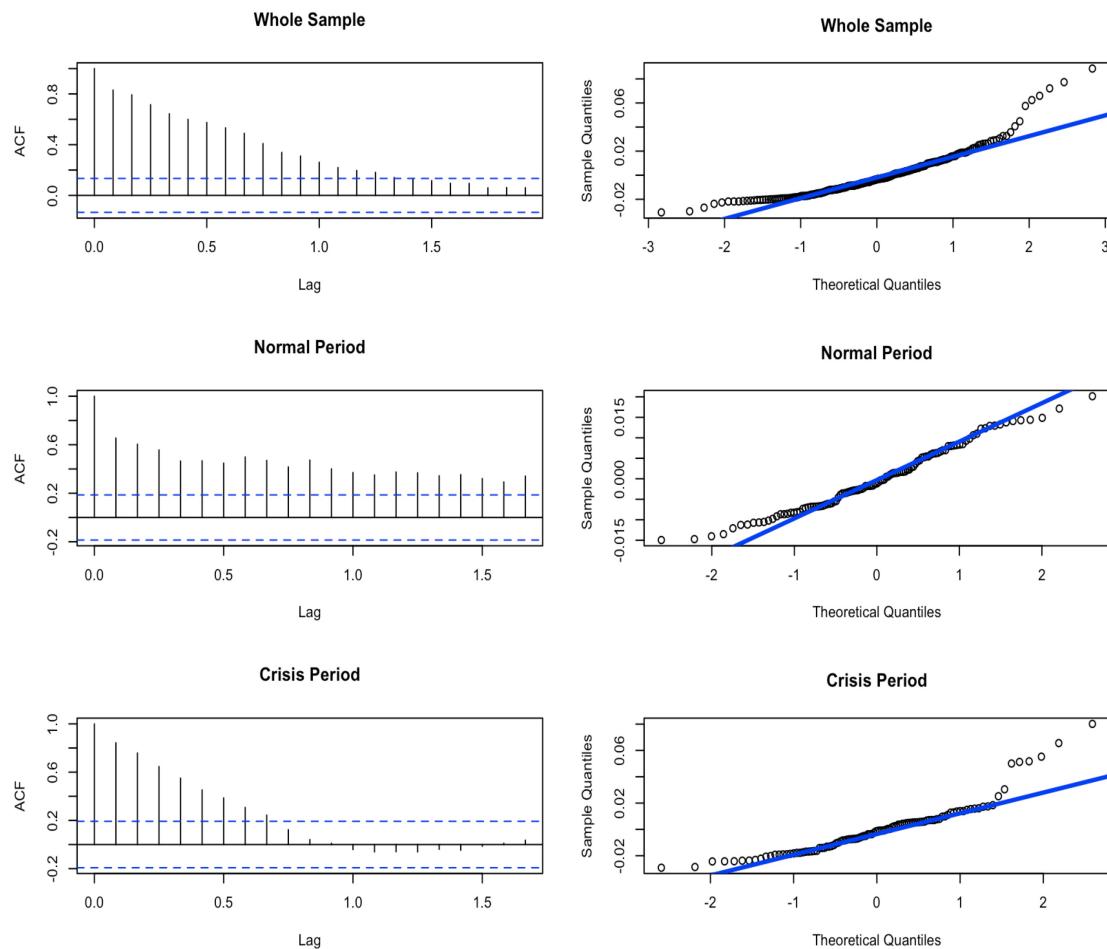


Figure 20: BB Bond Residual Tests

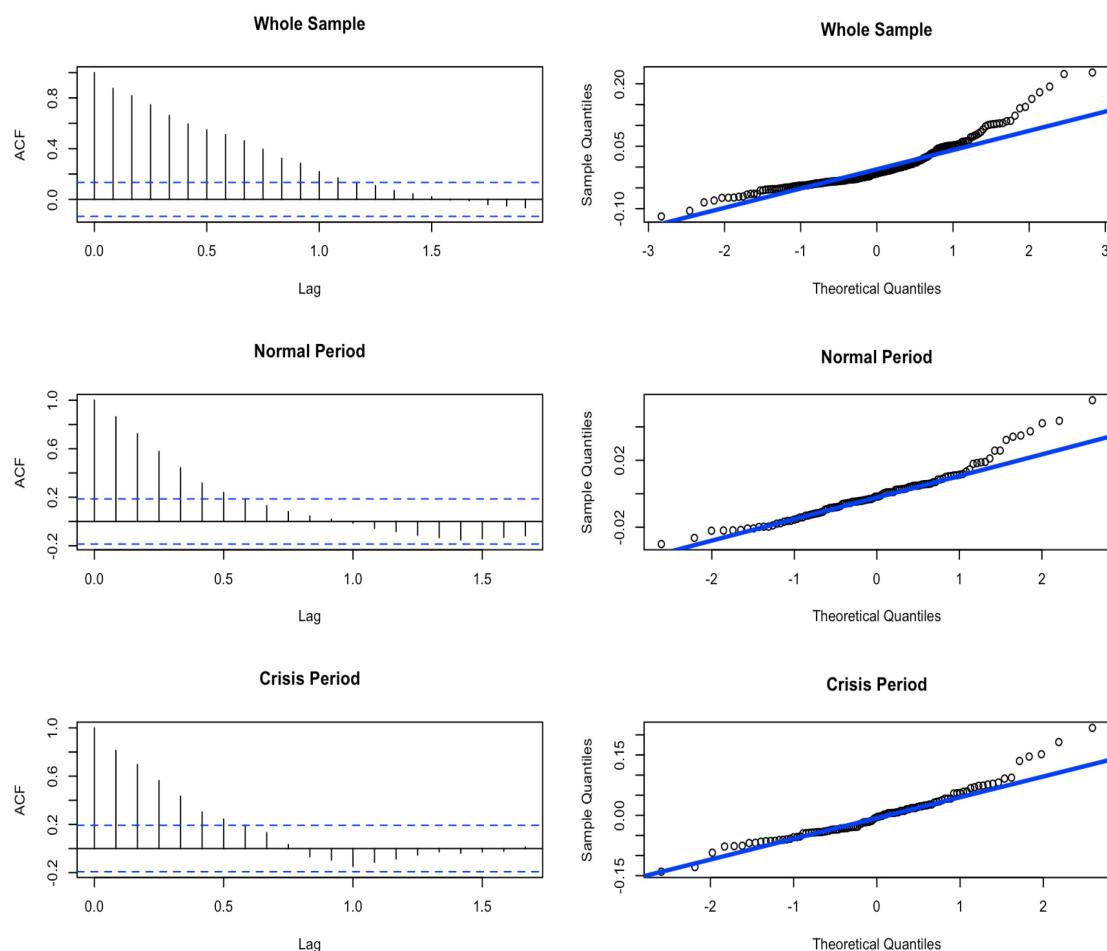


Figure 21: CCC Bond Residual Tests

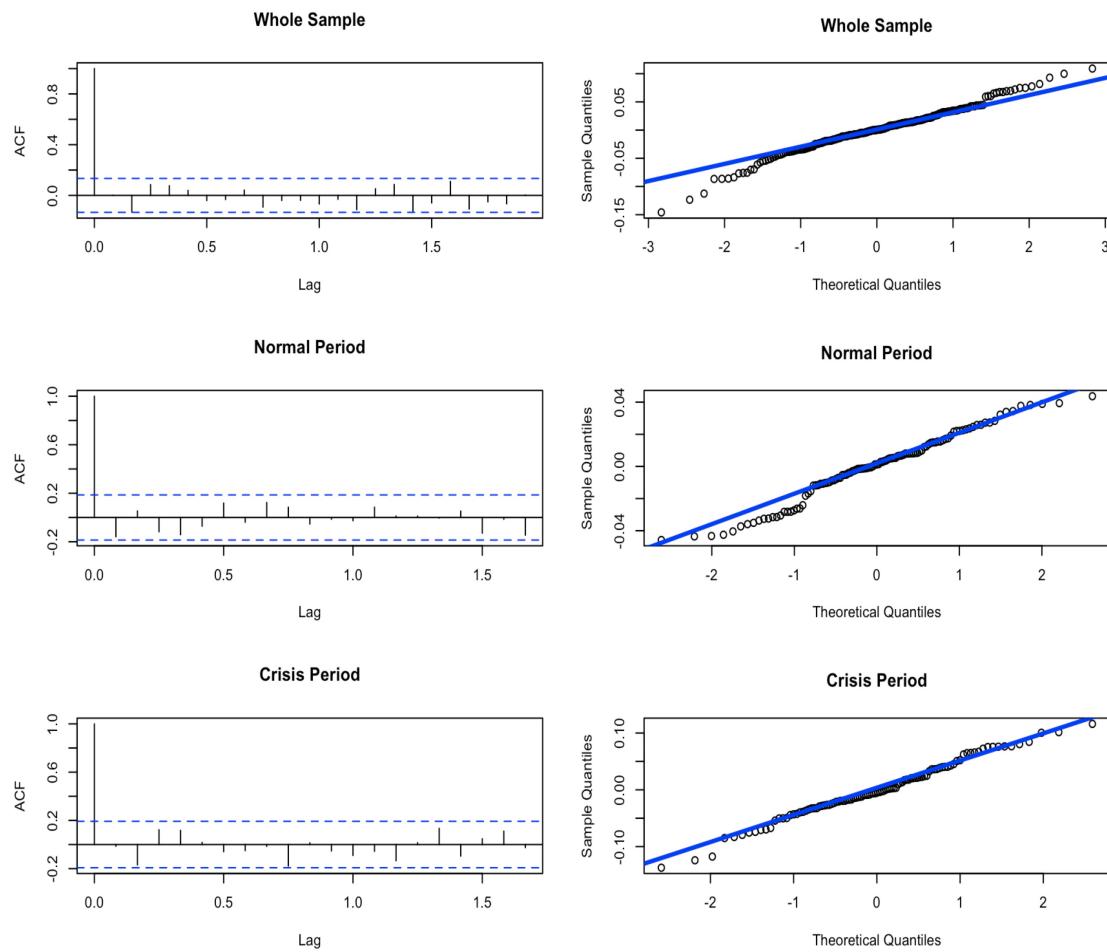


Figure 22: CCC Bond Residual Tests