Report

Could the Financial crisis of 2007-2008 have been predicted before it happened?

Introduction

The global financial crisis of 2007-2008 has been called by many economists as the worst financial crisis since the Great Depression of the 1930s. It started in 2007 with a crisis in the subprime mortgage market in the US. The collapse of Lehman Brothers in Sept 2008 led to this becoming a global financial crisis. Excessive risk taking by investment banks like Lehman Brothers, Bear Stearns, banks, American Insurance Group (AIG) and other financial institutions almost led to a collapse of the world's financial industry.
The costs of the financial crisis were immense.

1. Millions of people lost their homes due to foreclosures, short sales, etc.
2. Massive bailouts of the worlds largest banks and AIG were necessary to prevent the world's financial system from grinding to a halt.
3. The ensuing credit crunch led to the worst recession since the great depression.

Events leading upto the financial crisis.

The Federal Reserve reduced the federal funds rate between 2000 to 2003 from 6.5% to 1%. This was done to spur economic growth as the country was attempting to recover from the dot com bubble. Long term US treasury bond yields dropped as a result. Investors were eager to invest in products which offered higher returns. Wall street answered this demand with products like the Mortgage Backed Securities and Collateralized Debt Obligations (CDO). The Mortgage backed securities and CDOs were given AAA ratings by agencies like S&P and Moody's. This in turn created a demand for more home buyers. To answer this demand, banks and mortgage companies started looking beyond the pool of people with good credit history.

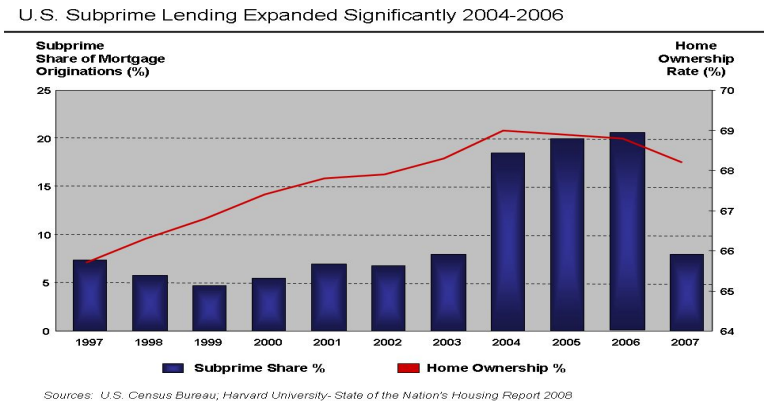
While lending money to borrowers with poor credit histories (Subprime buyers) is not always bad, historically these loans, were under 8 to 10% of the total home loans. Between 2002 and 2007 the share of subprime home loans started to steadily increase upto 20% and in some states in the US to 40% or more. Predatory lending practices were employed, i.e buyers weren't fully educated on the decisions they were taking. Adjustable Rate Mortgages (ARM's), Interest Only Loans, etc were offered to buyers. These loans have a low rate for an initial period, which would reset after that.

Investors required assurances that the mortgage backed securities and CDOs would offer the promised returns. Hence the concept of Credit Default Swaps came about. This works like an insurance policy where the buyer of the swap agrees to pay a premium until the conditions of the protection are met. For the sellers of these swaps, it was like getting regular money at very little risk. **Mortgages don't default or do they?** AIG at some point had insured around securities worth 540 Billion dollars.

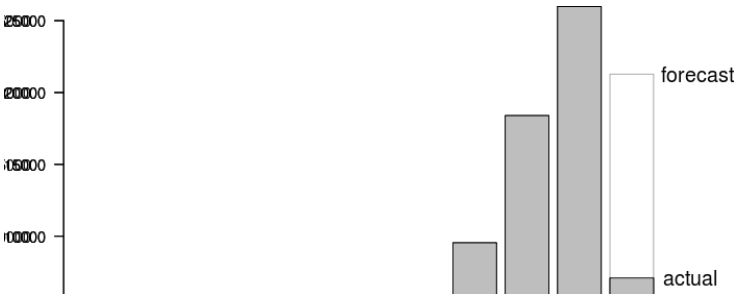
This expansion in home lending led to a dramatic increase in home prices and inflation started to increase around 2004. To curb inflation the Federal Reserve raised interest rates 3 times in 2004 and then gradually after that through 2006.

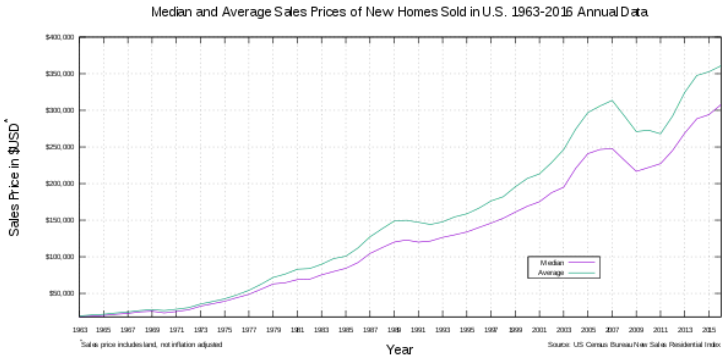
The ARM loans which were sold to buyers started to reset in this period. The prevailing interest rate was too high for these people to afford the monthly payments. Hence mortgage delinquencies started to go up leading to home foreclosures. This put downward pressure on the house prices, thus creating a vicious cycle of events. Construction activity stopped, leading to job losses, which led to more foreclosures and so on.

AIG and other banks which had sold Credit Default Swaps now were on the hook to pay. They did not have the cash reserves and had to be bailed out. Investment banks like Lehman Brothers, Bear Stearns folded under the sheer volume of the risk they had undertaken.



Growth in mortgage loan fraud based upon US Department of the Treasury Suspicious Activity Report Analysis (https://en.wikipedia.org/wiki/Subprime_mortgage_crisis)





Graphs showing median and average prices of homes in the US (https://en.wikipedia.org/wiki/Financial_crisis_of_2007%E2%80%932008)

Problem Formulation

Based on the analysis above, an interesting question is, could the 2007-2008 crisis be explained and possibly predicted earlier? This could potentially have softened the blow to the US and worldwide economy in general, and may have helped millions of homes from being foreclosed, jobs, etc. Our approach was to try to frame this as a linear regression problem. We now know that AIG, Banks, Fannie Mae, Freddie Mac, Investment banks like Lehman brothers, Bear Stearns, etc played a central role in creating this crisis. We asked the question, could we have used the stock performance of the above entities to explain and possibly predict the crisis earlier than it happened. The stock tickers we looked at are as below:

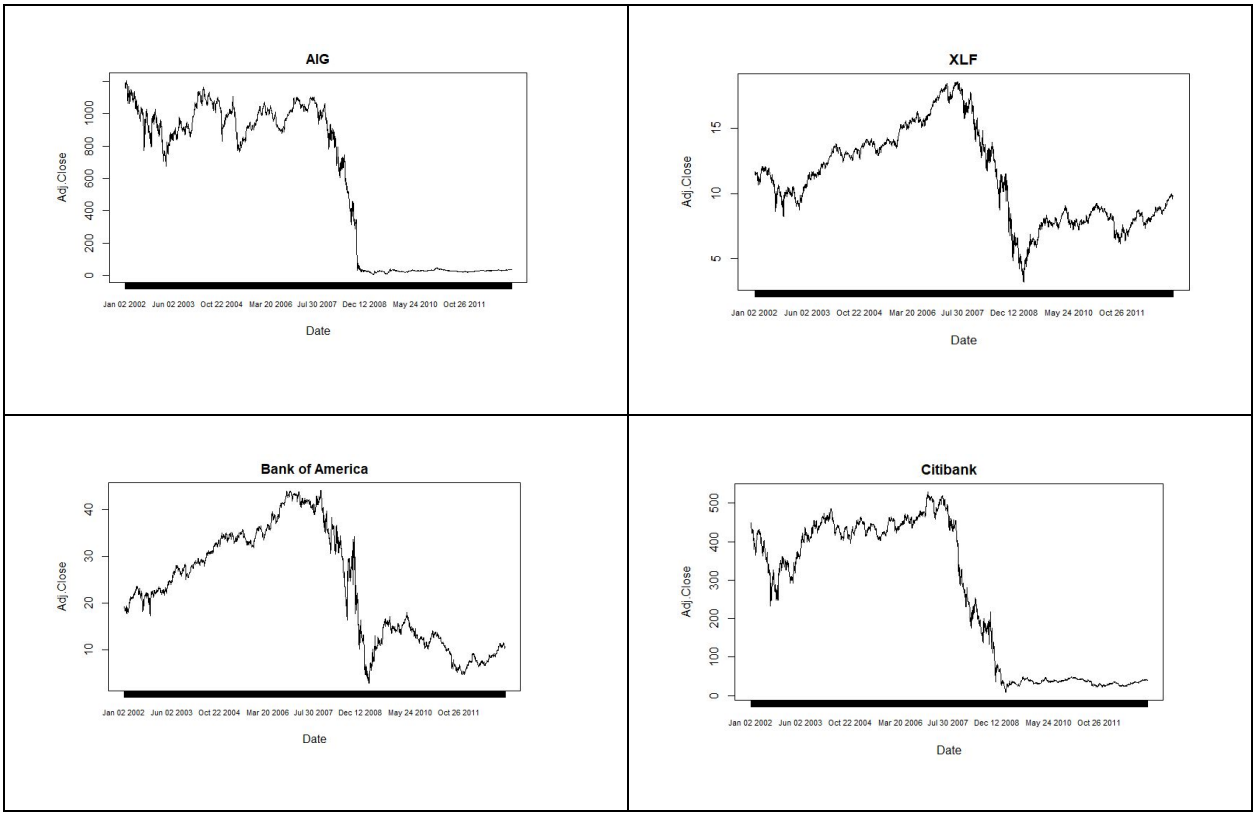
1. AIG
2. XLF
3. Citibank
4. Bank of America
5. Fannie Mae
6. Freddie Mac
7. S&P 500

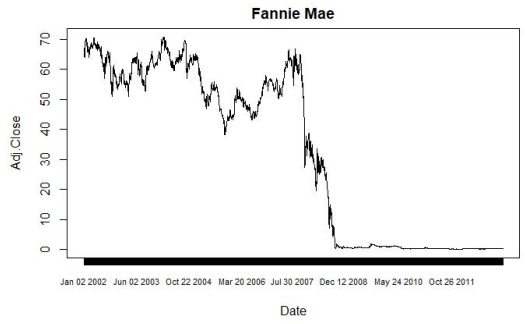
Stock prices and the economy in general is affected greatly by Interest Rates. In general lower interest rates help spur lending and economic activity, which in turn helps the stock market. Higher interest rates on the other hand could indicate inflationary pressure on the economy and could trigger a spending crunch.

We looked at the following plots to help figure out which regression models would make sense.

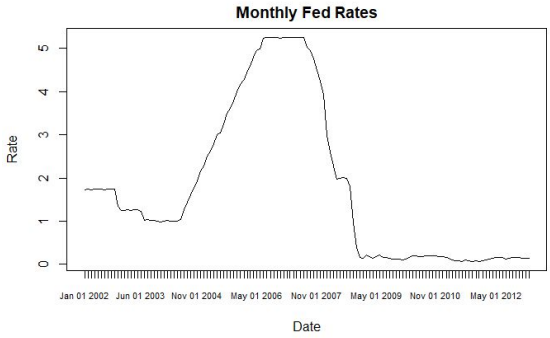
1. Daily stock values of AIG, XLF, Citibank, Bank of America, Fannie Mae, Freddie Mac, S&P 500.
2. Monthly plots of Home prices, Fed Interest Rates, Treasury bonds yields and Quarterly Bank Asset Ratios. The Treasury bonds yield curve was plotted to understand how the long term returns on these bonds were for the 2002-2008 period. Typically when the economy is doing well, the yields are higher to attract investment. In the above period the bonds were offering lower yields, which led to a lot of risk taking from banks and other financial institutions trying to attract investment. The Quarterly bank asset ratio curve was plotted to show how much leverage the banks had undertaken. Leverage is ratio of Assets to Capital. Higher values for this ratio indicates risk.

The plots are as below. We can see from these stock plots that the stock values followed a similar pattern leading upto the 2007-2008 period. They started going down sometime around late 2005 or early 2006.

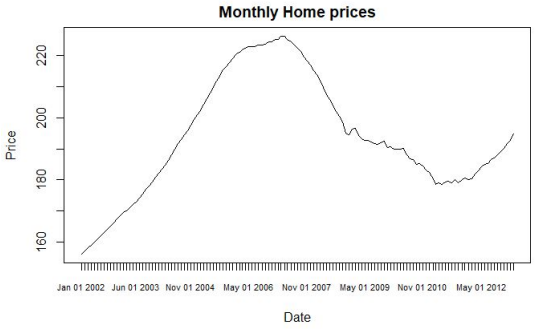




The plots for monthly home price values, Federal interest rates, Treasury bonds yield and quarterly bank asset ratios are below.

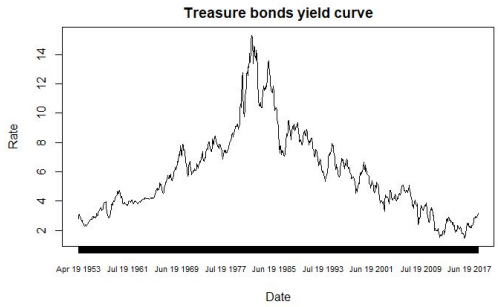


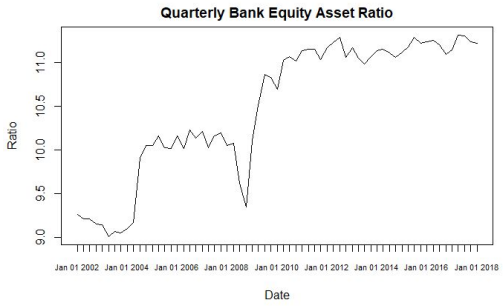
.



We can see that interest rates were low from the early part of 2002 and started steadily going up from 2004 onwards. This seems to coincide with the home price values flatlining or showing a downward trend.

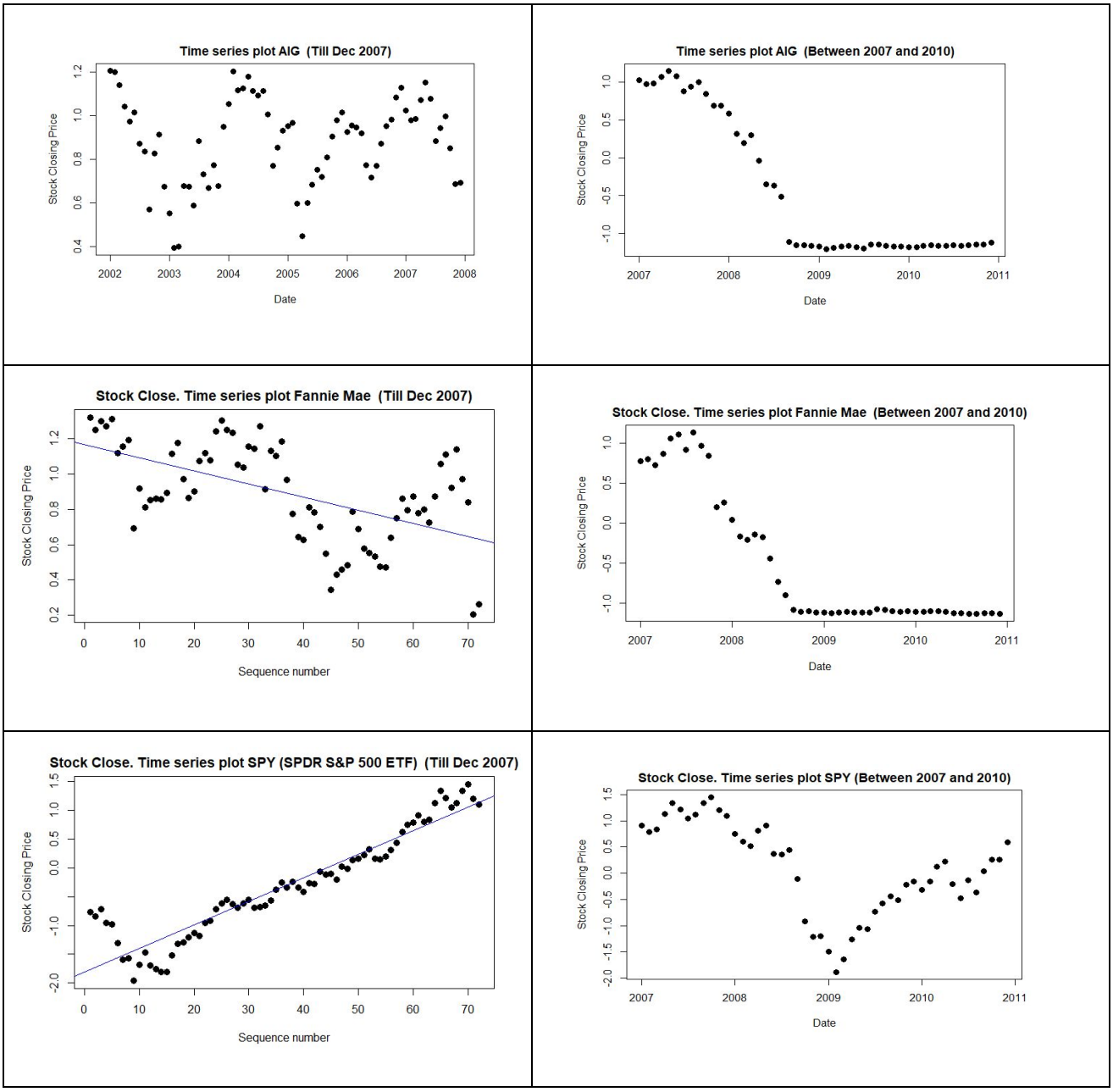
The treasury bond yield curve below shows decreasing yields 2001 onwards.





The Quarterly bank asset plot above shows banks had increasing leverage starting 2002 onwards.

We implemented time series regression models on AIG, XLF, Fannie Mae, SPY, etc. The plots are as below:



The time series plots of AIG, XLF, Fannie Mae and SPY seem to follow similar trends. It seemed like models using these stocks in conjunction with interest rates could offer interesting insights.

We came up with the following regression models.

1. **AIG stock as output with the predictors as XLF, Fannie Mae, Interest Rates.**
   This model allows us to determine the effect of the financial sector which is tracked by the XLF stock, Fannie Mae which tracks the residential mortgage market and Interest Rates. We expect to see a negative relationship between AIG and Interest rates for the period till 2007.

2. **S&P 500 ETF (SPY) as output with the predictors as AIG and Fannie Mae.**
   This model allows us to determine the effect of AIG and Fannie Mae on the S&P 500 index. We now know that these were bell weather stocks which had a significant effect on the financial industry. We expect to see a positive relationship between these predictors and SPY.

3. **AIG stock as output with the predictors as SPY and Interest Rates.**
   This model allows us to determine the effect of the SPY index and Interest Rates on AIG. There is a significant correlation between the financial sector and AIG. This relationship is tracked by the above model. We expect to see a negative relationship between interest rates and AIG for the period till 2007.

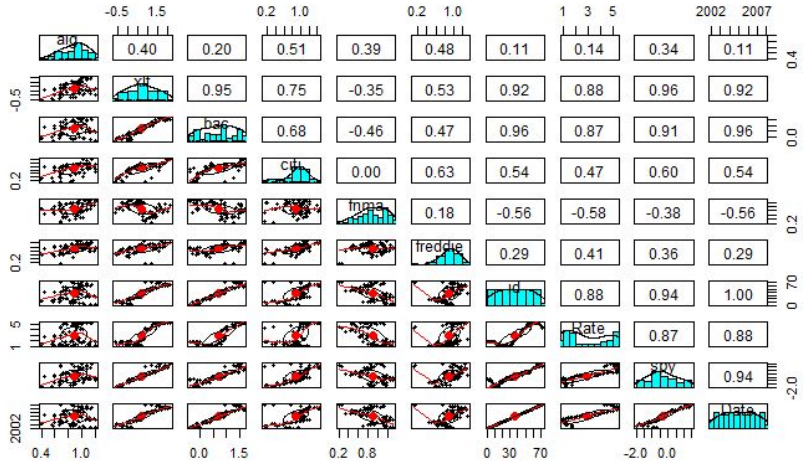
4. **S&P 500 ETF (SPY) as output with Rate as the predictor.**
   We expect in general interest rates to have a positive effect on S&P 500 and stocks in general. Rising interest rates are a good sign of growth in the economy. There should be a positive relationship between SPY and Interest Rates.

We decided to track the monthly stock performance of these entities from the period Jan 1st 2002 to March 1st 2013. With the intent of explaining and predicting the 2007-2008 financial crisis, we looked at regression models for the following periods

1. From Jan 1st 2002 to Dec 31st 2007
2. From Jan 1st 2002 to March 31st 2013
3. From Jan 1st 2007 to Dec 31st 2009
4. From Jan 1st 2009 to March 31st 2013

The above periods were picked as they would help explain how the stock prices of these entities performed leading upto the financial crisis of 2007, in the period of economic decline (2007 - 2009), the whole period and after the financial crisis (after 2009)

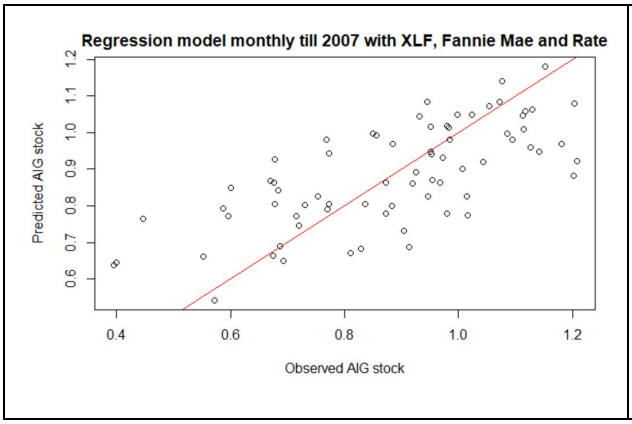
Linear Regression Models for the period from 2002 to 2007

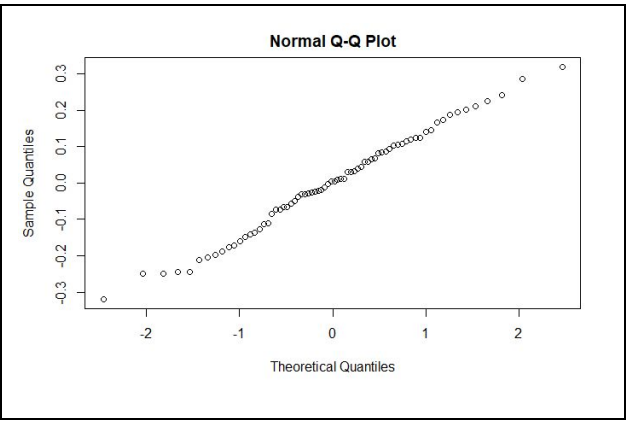


Correlations between the variables which include stock prices and interest rate till 2007.
The correlation plot above shows that there is a positive correlation between AIG and XLF and a negative correlation between interest rates and Fannie Mae, and Fannie Mae and XLF. Fannie Mae purchases mortgages from small banks and packages them as Mortgage securities for investment purposes. A negative correlation between interest rates and Fannie Mae shows that the stock price of Fannie Mae was dependent on the housing market doing well. With rising interest rates, the housing values start to decline leading to lower demand.

Linear regression using monthly stock data between AIG and XLF, Fannie Mae and Interest Rates from 2002 to 2007

```
Call:
lm(formula = aig ~ xlf + fnma + Rate, data = stocks_monthly_till_2007_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.31895 -0.09077  0.00371  0.10434  0.31939

Coefficients:
            Estimate Std. Error t value  Pr(>|t|)
(Intercept)  0.51360    0.11321   4.537 0.0000239 ***
xlf          0.24236    0.05535   4.378 0.0000423 ***
fnma         0.35370    0.08328   4.247 0.0000673 ***
Rate        -0.04109    0.02724  -1.508     0.136
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1447 on 68 degrees of freedom
Multiple R-squared:  0.4932,    Adjusted R-squared:  0.4708
F-statistic: 22.06 on 3 and 68 DF,  p-value: 0.0000000004365
```





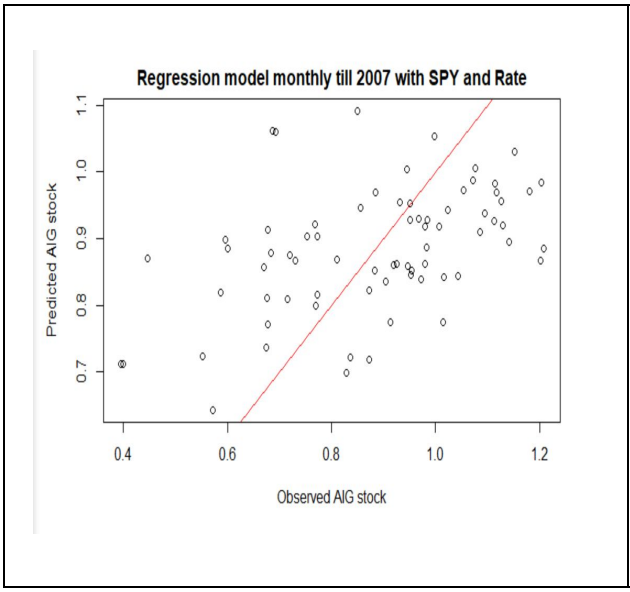
Inference

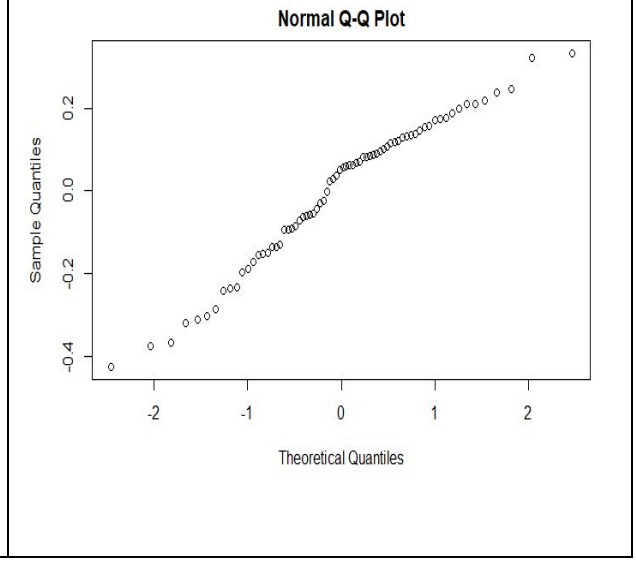
We can see that all predictors in this model are significant at alpha = 0.15, i.e. they have low p values. 47% of the variation in AIG stock appears to be explained by this model. Interestingly enough interest rates have a negative effect on AIG stock price. Rising interest rates are an indication of growth in the economy. However the housing market was negatively impacted by interest rates which negatively impacted AIG due to its exposure to credit default swaps.

The above model appears to offer some insight that things were starting to go wrong in the financial sector.

Linear regression using monthly stock data between AIG, SPY, Rate from 2002 to 2007

```
Call:
lm(formula = aig ~ spy + Rate, data = stocks_monthly_till_2007_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.42453 -0.13171  0.05324  0.12994  0.33359

Coefficients:
            Estimate Std. Error t value             Pr(>|t|)
(Intercept)  1.17597    0.09078  12.955 < 0.0000000000000002 ***
spy          0.20160    0.04769   4.227            0.0000712 ***
Rate        -0.07928    0.02592  -3.058              0.00317 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1781 on 69 degrees of freedom
Multiple R-squared:  0.2212,    Adjusted R-squared:  0.1986
F-statistic: 9.796 on 2 and 69 DF,  p-value: 0.0001799
```





Inference

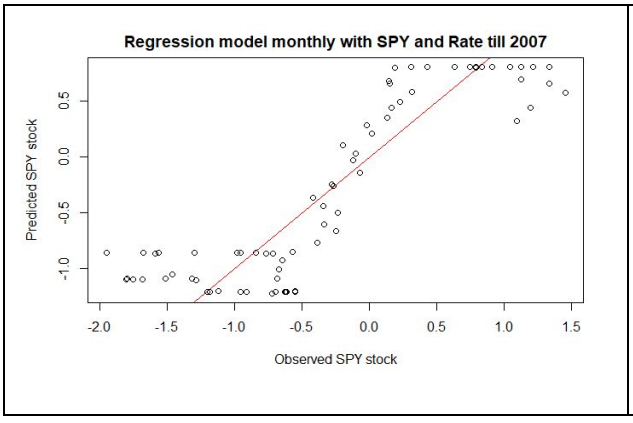
We can see that all predictors in this model are significant at alpha = 0.05, i.e. they have low p values. Around 20% of the variation in AIG’s stock price is explained by the above model.

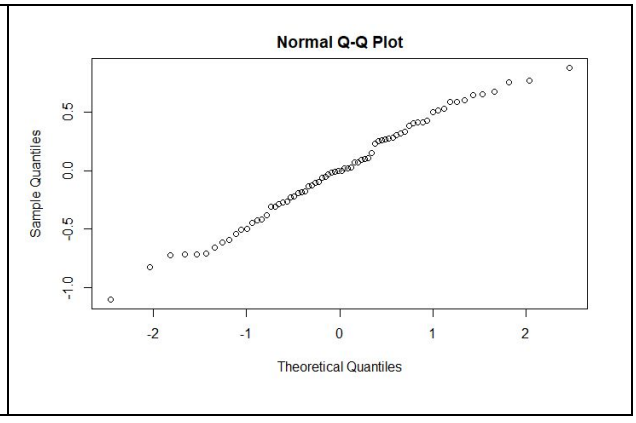
Interestingly rising interest rates show a negative influence on AIG's stock price in this model as well. AIG with its heavy reliance on housing prices going up was starting to see some headwinds on the horizon with its exposure to credit default swaps.

The above model offers some insight that things were starting to go wrong in the financial sector.

Linear regression using monthly stock data between SPY and Interest Rate from 2002 to 2007

```
Call:
lm(formula = spy ~ Rate, data = stocks_monthly_till_2007_df)

Residuals:
     Min       1Q   Median       3Q      Max
-1.09819 -0.28882 -0.00039  0.32349  0.88043

Coefficients:
            Estimate Std. Error t value            Pr(>|t|)
(Intercept) -1.68752    0.10525  -16.03 <0.0000000000000002 ***
Rate         0.47519    0.03154   15.06 <0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4463 on 70 degrees of freedom
Multiple R-squared:  0.7643,  Adjusted R-squared:  0.7609
F-statistic: 226.9 on 1 and 70 DF,  p-value: < 0.00000000000000022
```





Inference

We can see that interest rates overall were having a positive influence on SPY. This makes sense as rising interest rates are a good indicator of economic growth. While the financial sector

was starting to see negative effects due to a slowdown in the housing market, overall the economy was doing well till 2007.

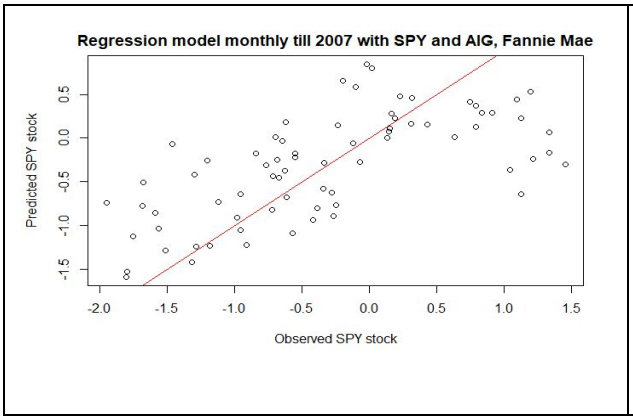
Linear regression using monthly stock data between SPY, AIG and Fannie Mae from 2002 to 2007

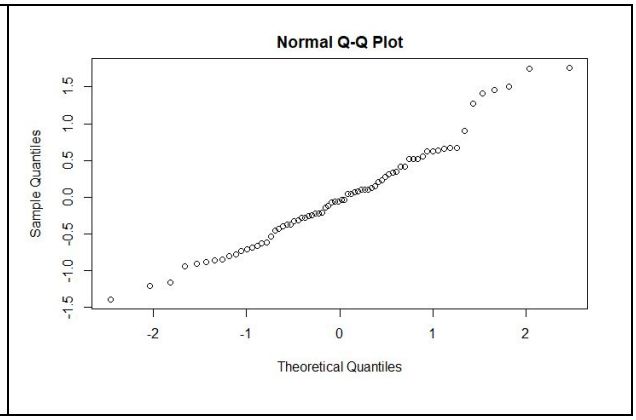
Call:
lm(formula = spy ~ aig + fnma, data = stocks_monthly_till_2007_df)

Residuals:
```
     Min       1Q   Median       3Q      Max
-1.39503 -0.44098 -0.05022  0.41769  1.76429
```

Coefficients:
```
             Estimate Std. Error t value    Pr(>|t|)
(Intercept)  -0.8708     0.4031  -2.160      0.0342 *
aig           2.6341     0.4552   5.787 0.000000192 ***
fnma         -1.9792     0.3260  -6.072 0.000000061 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 0.703 on 69 degrees of freedom
Multiple R-squared: 0.4236, Adjusted R-squared: 0.4069
F-statistic: 25.35 on 2 and 69 DF, p-value: 0.000000005567





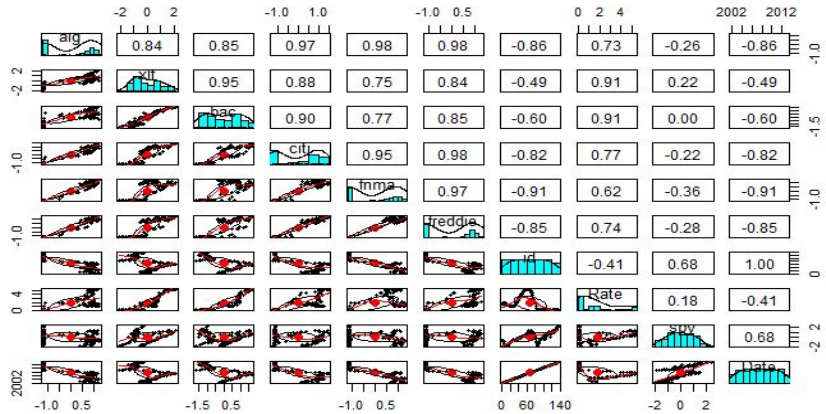
Inference
We can see that this model is significant at alpha = 0.05. Interestingly Fannie Mae has a negative influence on SPY (S&P 500). We would expect a positive relationship between

Fannie Mae and SPY as Fannie Mae buying mortgages provides banks with funds to enable lending.

At this point Fannie Mae was impacted by a slowdown in the housing market due to rising interest rates. It was dragging down the financial sector due to its inability to continue purchasing mortgages from banks. Overall things were starting to sour at this point.

40% of the variance in S&P 500 index is explained by this model.

Linear Regression Models for the period from 2002 to 2013.



Correlations between the variables which include stock prices and interest rate for the period from 2002 to 2013.

We can see that there is significant positive correlation between AIG and XLF and Fannie Mae. SPY has a negative correlation with Fannie Mae in this period. Something to keep in mind. This could be in part due to Fannie Mae being negatively impacted by interest rates.

Linear regression using monthly stock data between AIG and XLF, Fannie Mae and Interest Rates from 2002 to 2013.

Call:
lm(formula = aig ~ xlf + fnma + Rate, data = stocks_monthly_df)

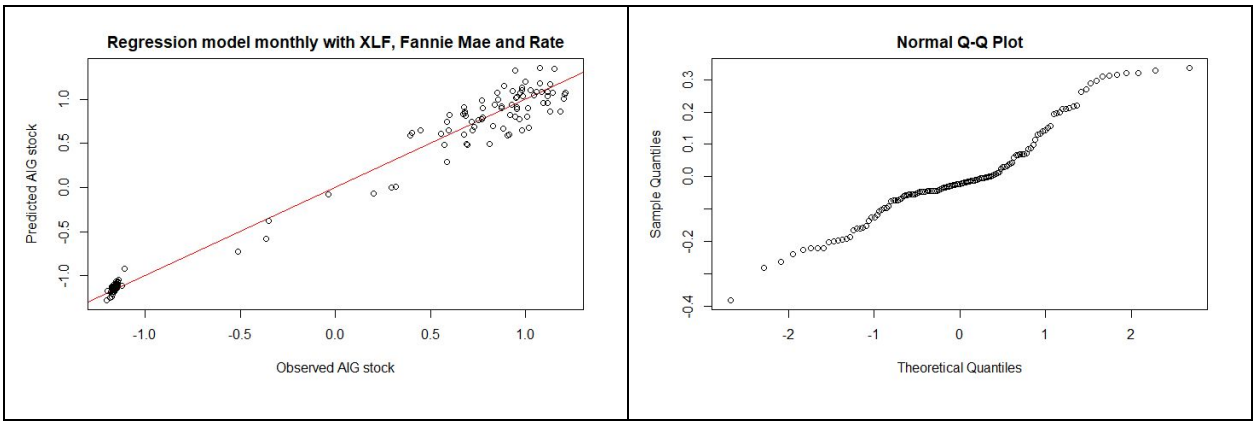
Residuals:
```
     Min       1Q   Median       3Q      Max
-0.38016 -0.06398 -0.02221  0.06715  0.33657
```

Coefficients:
```
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.11089    0.03334  -3.326 0.001144 **
xlf          0.12807    0.03790   3.379 0.000960 ***
fnma         0.80922    0.01950  41.508  < 2e-16 ***
```

Rate 0.06275 0.01753 3.579 0.000485 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1431 on 131 degrees of freedom
Multiple R-squared: 0.98, Adjusted R-squared: 0.9795
F-statistic: 2138 on 3 and 131 DF, p-value: < 2.2e-16



Inference
We can see that this model is significant at alpha = 0.05. Interestingly all coefficients are positive in this model, i.e. interest rates no longer have a negative effect on AIG’s stock price. This can be partially explained by the fact that the interest rates were dropped in 2008 onwards to help kickstart the economy after the financial crisis. The positive relationship between AIG and interest rates for the overall period could also indicate a changing business model at AIG. This helped it return to profitability.

Overall this model explains around 97% of the variation in AIG’s stock price.

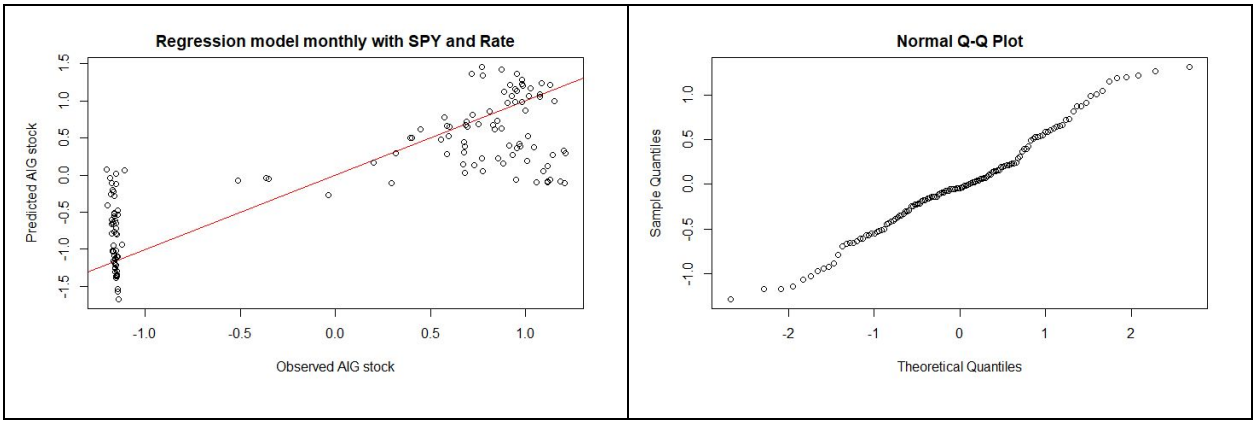
Linear regression using monthly stock data between AIG, SPY and Interest Rates from 2002 to 2013.

Call:
lm(formula = aig ~ spy + Rate, data = stocks_monthly_df)

Residuals:
Min 1Q Median 3Q Max
-1.28162 -0.33607 -0.03981 0.26776 1.31555
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.78268 0.06803 -11.505 < 0.0000000000000002 ***
spy -0.40462 0.04913 -8.237 0.000000000000152 ***

Rate 0.44291 0.02717 16.299 < 0.0000000000000002 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.5599 on 132 degrees of freedom
Multiple R-squared: 0.6912, Adjusted R-squared: 0.6866
F-statistic: 147.8 on 2 and 132 DF, p-value: < 0.00000000000000022



Inference
It is not clear why SPY would have a negative influence on AIG. Perhaps we need more predictors for this model. We decided to drop this model for lack of time required for investigation.

Linear regression using monthly stock data between SPY and Interest Rates from 2002 to 2013.

Call:
lm(formula = spy ~ Rate, data = stocks_monthly_df)

Residuals:
Min 1Q Median 3Q Max
-1.9525 -0.6396 -0.1180 0.7792 2.5039

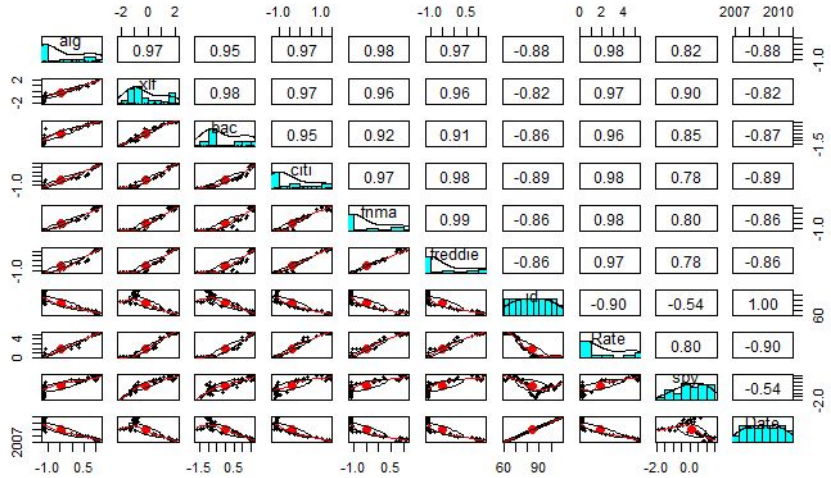
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.17135 0.11915 -1.438 0.153
Rate 0.09697 0.04722 2.053 0.042 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9882 on 133 degrees of freedom
Multiple R-squared: 0.03073, Adjusted R-squared: 0.02344
F-statistic: 4.216 on 1 and 133 DF, p-value: 0.042

Inference
While the interest rate is significant at alpha = 0.05, it appears that this model is missing some predictors. We decided to ignore this model for this period.

Linear Regression Models for the period from 2007 to 2009

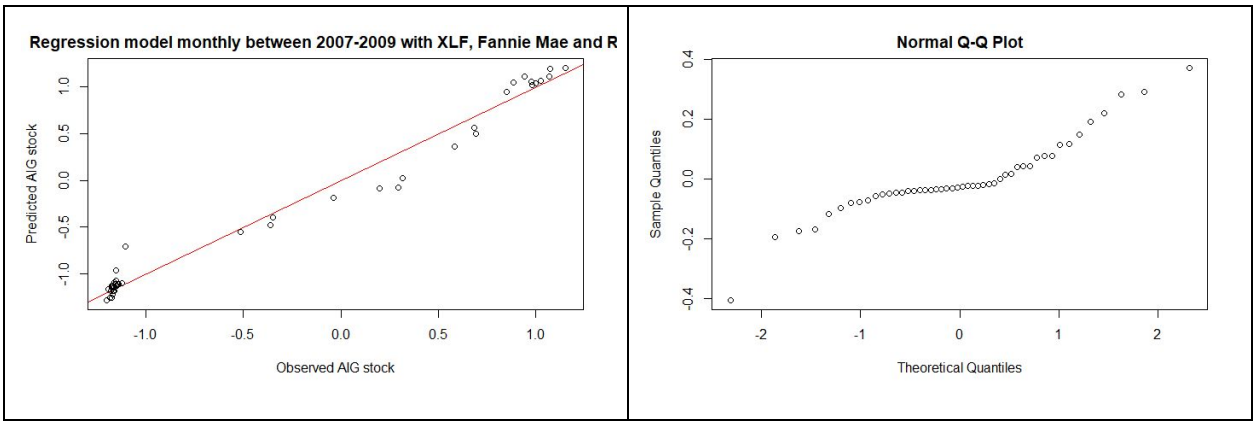


Correlations between the variables which include stock prices and interest rate for the period from 2007 to 2009

Linear regression using monthly stock data between AIG and XLF, Fannie Mae and Interest Rates from 2007 to 2009

```
Call:
lm(formula = aig ~ xlf + fnma + Rate, data =
stocks_monthly_between_2007_and_2009_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.40427 -0.04568 -0.02629  0.04244  0.37144

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.53828    0.14562  -3.697 0.000602 ***
xlf          0.14735    0.06653   2.215 0.032012 *
fnma         0.42410    0.11149   3.804 0.000436 ***
Rate         0.18816    0.05270   3.570 0.000877 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1337 on 44 degrees of freedom
Multiple R-squared:  0.9806,         Adjusted R-squared:  0.9793
```

F-statistic: 742.7 on 3 and 44 DF, p-value: < 0.00000000000000022



Inference
We can see that all predictors are significant at alpha = 0.05. This model explains around 97% of the variance in AIG stock. In general things appeared to be getting better, after AIG and other banks were bailed out to to avoid a financial depression.

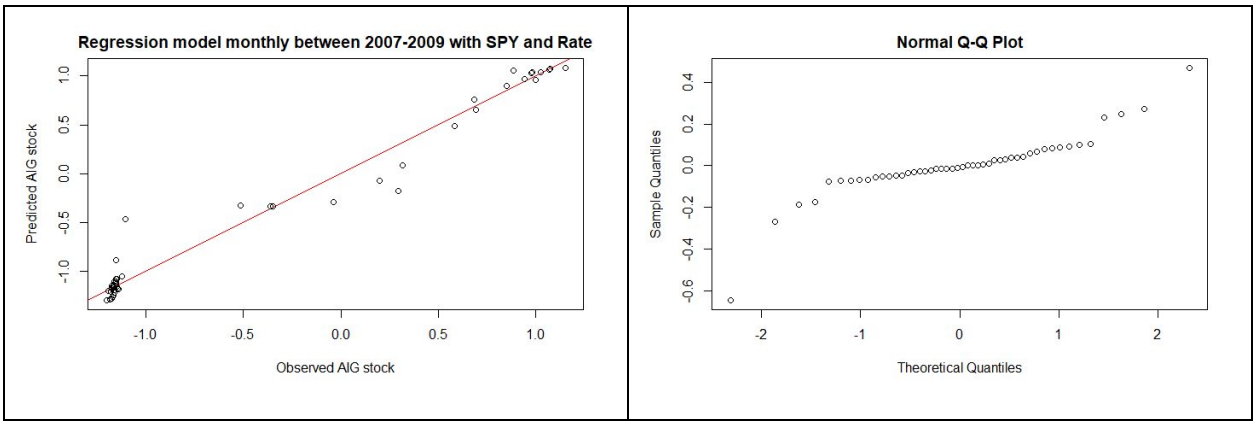
Linear regression using monthly stock data between AIG, SPY and Interest Rates from 2007 to 2009

Call:
lm(formula = aig ~ spy + Rate, data = stocks_monthly_between_2007_and_2009_df)

Residuals:
```
     Min       1Q   Median       3Q      Max
-0.64593 -0.05012 -0.00829  0.04662  0.47080
```

Coefficients:
```
            Estimate Std. Error t value            Pr(>|t|)
(Intercept) -1.18472    0.03783 -31.315 <0.0000000000000002 ***
spy          0.10493    0.04236   2.477              0.0171 *
Rate         0.40561    0.01815  22.349 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 0.1545 on 45 degrees of freedom
Multiple R-squared: 0.9735, Adjusted R-squared: 0.9724
F-statistic: 828.1 on 2 and 45 DF, p-value: < 0.00000000000000022

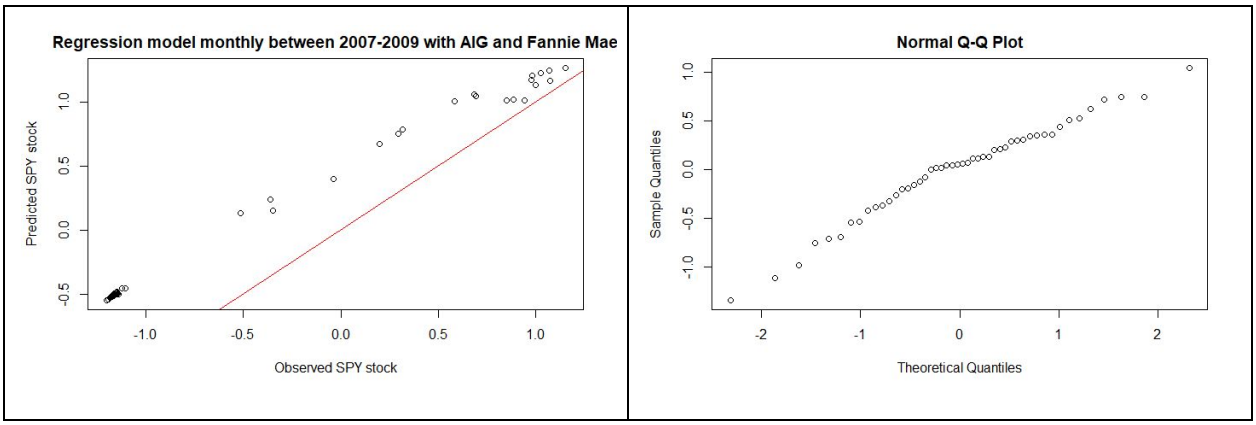


Inference

We can see that all predictors are significant at alpha = 0.05. This model explains around 97% of the variance in AIG stock. In general things appeared to be getting better after AIG and other banks were bailed out to avoid a financial depression.

Linear regression using monthly stock data between SPY, AIG and Fannie Mae from 2007 to 2009

```
Call:
lm(formula = spy ~ aig + fnma, data = stocks_monthly_between_2007_and_2009_df)

Residuals:
     Min       1Q   Median       3Q      Max
-1.33826 -0.27391  0.05893  0.31393  1.04790

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.37842    0.09798   3.862 0.000357 ***
aig          1.10755    0.42138   2.628 0.011694 *
fnma        -0.36142    0.46773  -0.773 0.443731
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5102 on 45 degrees of freedom
Multiple R-squared:  0.6841,         Adjusted R-squared:  0.6701
F-statistic: 48.73 on 2 and 45 DF,  p-value: 0.000000000005483
```

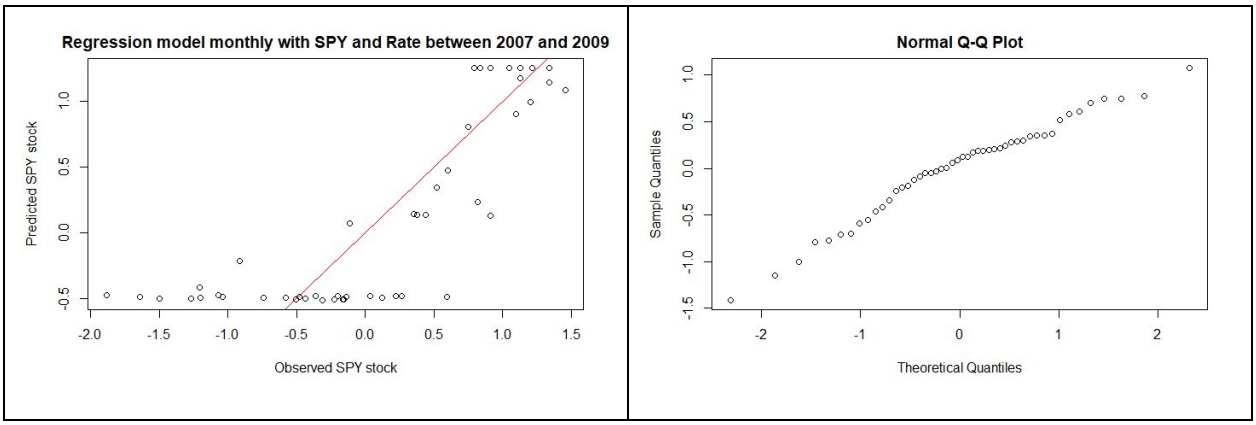


Inference

We can see that all predictors are significant at alpha = 0.05. This model explains around 67% of the variance in SPY stock. Interestingly enough Fannie Mae still has a negative influence on SPY, which is contrary to expectation.

Linear regression using monthly stock data between SPY, and Interest Rates from 2007 to 2009

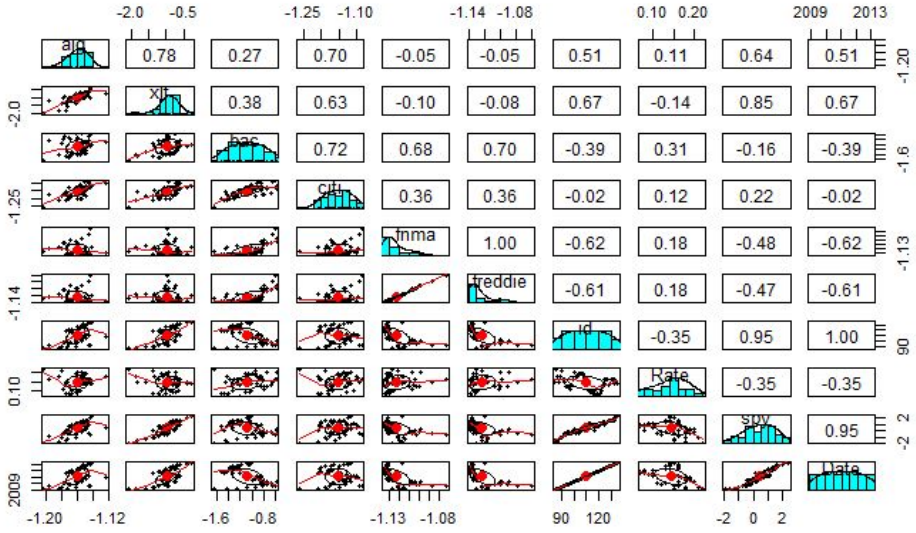
```
Call:
lm(formula = spy ~ Rate, data = stocks_monthly_between_2007_and_2009_df)

Residuals:
    Min      1Q  Median      3Q     Max
-1.4116 -0.2692  0.1049  0.3127  1.0793

Coefficients:
            Estimate Std. Error t value         Pr(>|t|)
(Intercept) -0.54964    0.10379  -5.296 0.00000323789917 ***
Rate         0.34306    0.03784   9.066 0.00000000000836 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5379 on 46 degrees of freedom
Multiple R-squared:  0.6412,         Adjusted R-squared:  0.6334
F-statistic:  82.2 on 1 and 46 DF,  p-value: 0.000000000008358
```



Inference

We can see that interest rates is significant at alpha = 0.05. This model explains around 63% of the variance in SPY's stock price. It makes sense that interest rates had a positive influence on stocks at this point. They were relatively low to enable the economy to recover from the financial crisis.

Linear Regression Models for the period from 2009 to 2013.

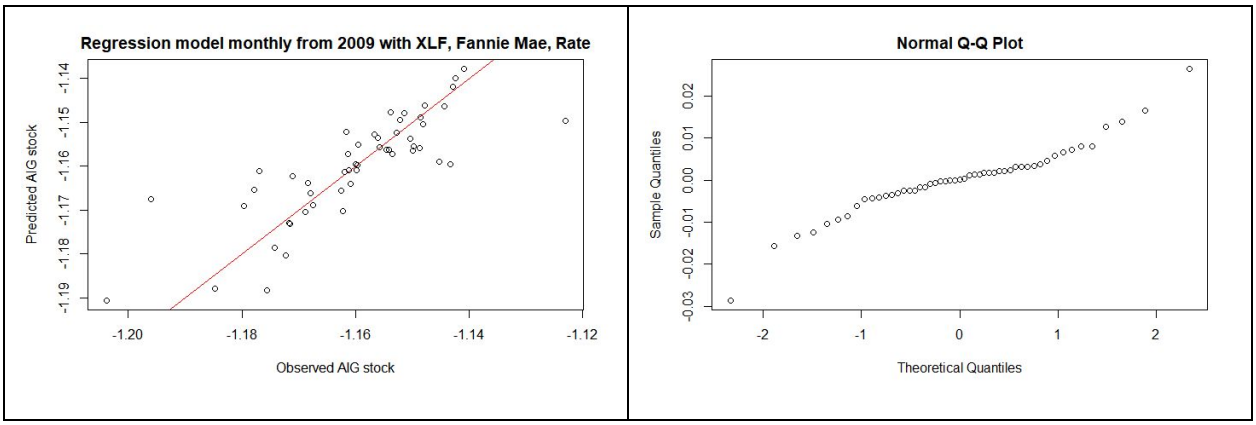


Correlations between the variables which include stock prices and interest rate for the period from 2009 to 2013.

We can see that there is significant positive correlation between AIG and XLF and a negative correlation between SPY and Fannie Mae. This is a worrying sign. Even after all the steps taken to ensure that the financial crisis does not repeat, it appears that Fannie Mae still has a significant negative effect on SPY. Something to keep in mind going forward.

Linear regression using monthly stock data between AIG and XLF, Fannie Mae and Interest Rates from 2009 to 2013.

```
Call:
lm(formula = aig ~ xlf + fnma + Rate, data = stocks_monthly_from_2009_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.0285583 -0.0032909  0.0000714  0.0031607  0.0264736

Coefficients:
            Estimate Std. Error t value            Pr(>|t|)
(Intercept) -1.157357   0.103800 -11.150 0.00000000000000851 ***
xlf          0.033742   0.003553   9.496 0.00000000000165543 ***
fnma        -0.016083   0.091346  -0.176              0.8610
Rate         0.083400   0.031590   2.640              0.0112 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.008634 on 47 degrees of freedom
Multiple R-squared:  0.6634,         Adjusted R-squared:  0.6419
F-statistic: 30.88 on 3 and 47 DF,  p-value: 0.0000000000353
```



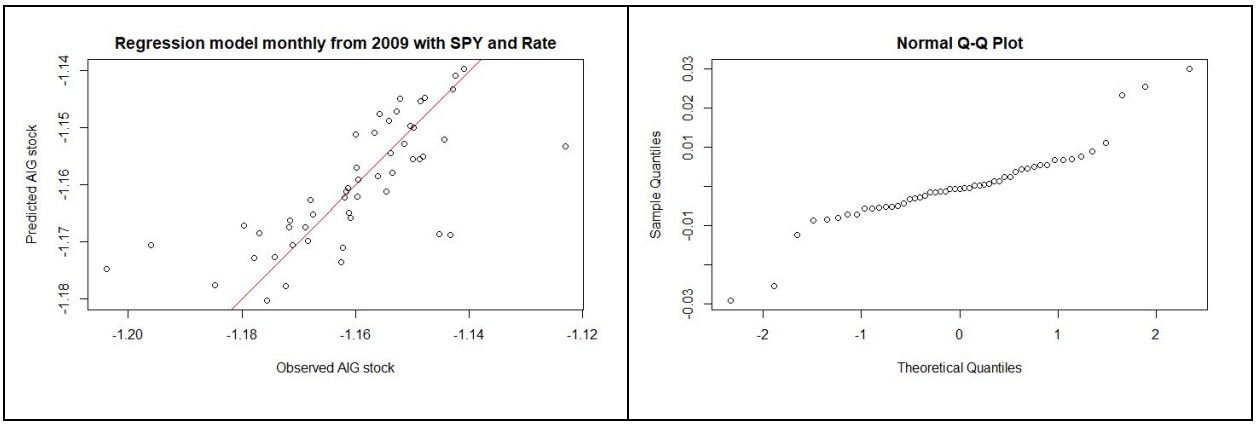
Inference

We can see that Fannie Mae is not significant in this model. It appears that AIG is now on the good road to recovery. This may have resulted from reduced exposure to the mortgage industry. https://money.cnn.com/2018/01/22/investing/aig-acquisitions/index.html

64% of the variance in AIG’s stock price is explained by this model.

Linear regression using monthly stock data between AIG, SPY and Interest Rates, from 2009 to 2013.

```
Call:
lm(formula = aig ~ spy + Rate, data = stocks_monthly_from_2009_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.0291584 -0.0051638 -0.0005807  0.0044669  0.0300961

Coefficients:
            Estimate Std. Error  t value            Pr(>|t|)
(Intercept) -1.184723   0.005953 -199.001 < 0.0000000000000002 ***
spy          0.010916   0.001494    7.306        0.0000000025 ***
Rate         0.139160   0.038397    3.624            0.000699 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01007 on 48 degrees of freedom
Multiple R-squared:  0.5322,         Adjusted R-squared:  0.5127
F-statistic: 27.31 on 2 and 48 DF,  p-value: 0.00000001205
```



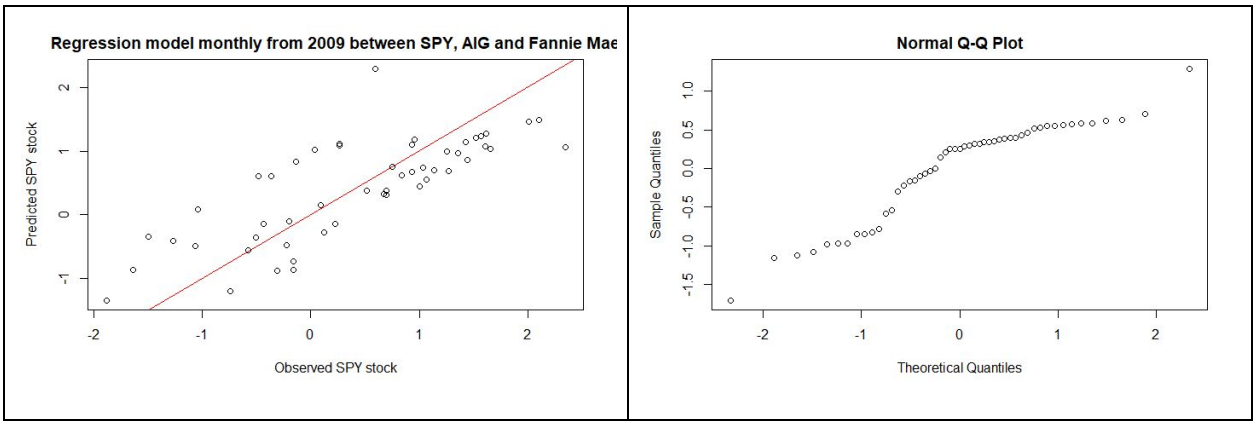
Inference

We can see that all predictors are significant at alpha = 0.05. Interest rates have a positive effect on AIG's stock price, which is all good news for AIG. This model appears to reinforce the findings of the earlier model. This may have resulted from reduced exposure to the mortgage industry.

https://money.cnn.com/2018/01/22/investing/aig-acquisitions/index.html

Linear regression using monthly stock data between SPY, AIG and Fannie Mae from 2009 to 2013.

```
Call:
lm(formula = spy ~ aig + fnma, data = stocks_monthly_from_2009_df)

Residuals:
    Min      1Q  Median      3Q     Max
-1.7036 -0.4224  0.2555  0.4457  1.2906

Coefficients:
            Estimate Std. Error t value     Pr(>|t|)
(Intercept)   12.895     10.930   1.180        0.244
aig           43.235      6.414   6.740 0.0000000184 ***
fnma         -33.542      6.786  -4.943 0.0000097848 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6535 on 48 degrees of freedom
Multiple R-squared:  0.6052,         Adjusted R-squared:  0.5887
F-statistic: 36.79 on 2 and 48 DF,  p-value: 0.0000000002058
```



Inference

We can see that all predictors in this model are significant at alpha = 0.05. On a good note AIG has a positive influence on S&P 500. However on a worrying note, Fannie Mae continues to have a negative influence on the S&P 500. This is something to keep an eye on. It appears that the financial sector is continuing to be weighed down by Fannie Mae. If interest rates were to rise again, this may point to further headwinds down the road.

Linear regression using monthly stock data between SPY and Interest Rates from 2009 to 2013.

```
Running linear regression using monthly stock data between SPY and Rate from 2009
Call:
lm(formula = spy ~ Rate, data = stocks_monthly_from_2009_df)

Residuals:
    Min      1Q  Median      3Q     Max
-1.86024 -0.79899  0.00159  0.72419  1.89105

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.7270     0.5129   3.367  0.00149 **
Rate         -9.0850     3.4339  -2.646  0.01093 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9629 on 49 degrees of freedom
Multiple R-squared:  0.125,  Adjusted R-squared:  0.1071
F-statistic:     7 on 1 and 49 DF,  p-value: 0.01093
```

Inference

While the interest rate is significant in this model, we appear to be missing some predictors. Adjusted R square is relatively low at 10%. We decided to drop this model for this period.

Conclusion

This project was an attempt to understand the reasons behind the financial crisis of 2007-2008 and whether it could have been predicted earlier. This could have softened the blow the economy suffered. Millions of homes could have been prevented from foreclosure, etc.

We used linear regression models between the stock prices of the major players of the financial crisis like AIG, XLF, Fannie Mae, Interest Rates, SPY, etc. Based on the results of our models, it appears that there were a number of indicators that there was trouble in the financial industry starting around late 2006. Perhaps this crisis could have been averted, by taking preventive steps earlier?

On a positive note, it appears that AIG has reformed its business practices and has reduced its exposure to the mortgage insurance industry. However the continued negative influence of Fannie Mae on SPY (S&P 500) is a worrying sign. Steps need to be taken to ensure that this crisis does not repeat.

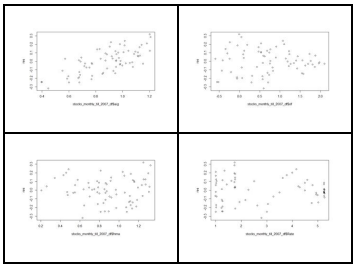
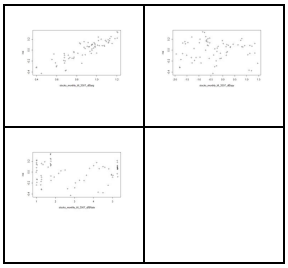
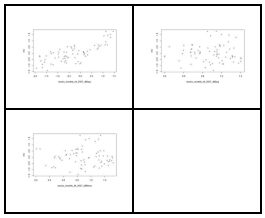
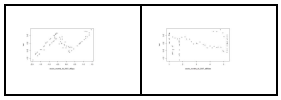
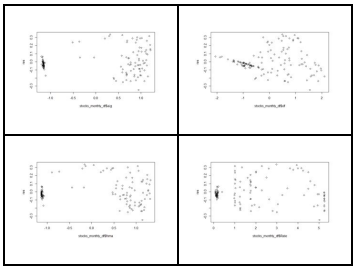
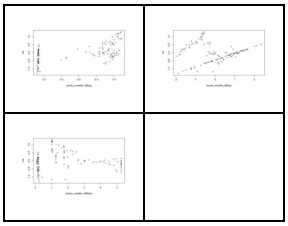
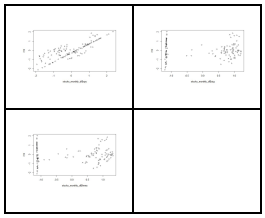
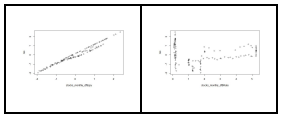
References

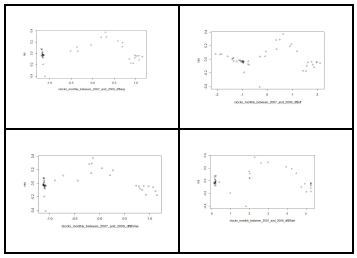
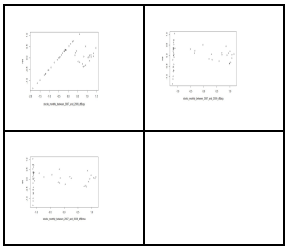
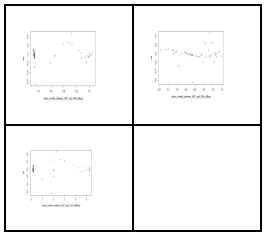
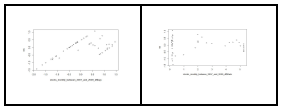
https://en.wikipedia.org/wiki/Subprime_mortgage_crisis
https://en.wikipedia.org/wiki/Bear_Stearns
https://fred.stlouisfed.org/series/TDSP
https://www.thebalance.com/mortgage-crisis-overview-315684
https://www.intheblack.com/articles/2015/07/07/6-economists-who-predicted-the-global-financial-crisis-and-why-we-should-listen-to-them-from-now-on
http://r.789695.n4.nabble.com/Plotting-Time-Series-Data-by-Month-td4162534.html

Data sources

https://fred.stlouisfed.org
https://finance.yahoo.com/

Appendix
Residual Plots:

	AIG~XLF+FNMA+RATE	AIG~SPY+RATE	SPY~AIG+FNMA	SPY~RATE
2002-2007				
2002-2013				

2007-2009				
2009-2013