



Department of Statistics

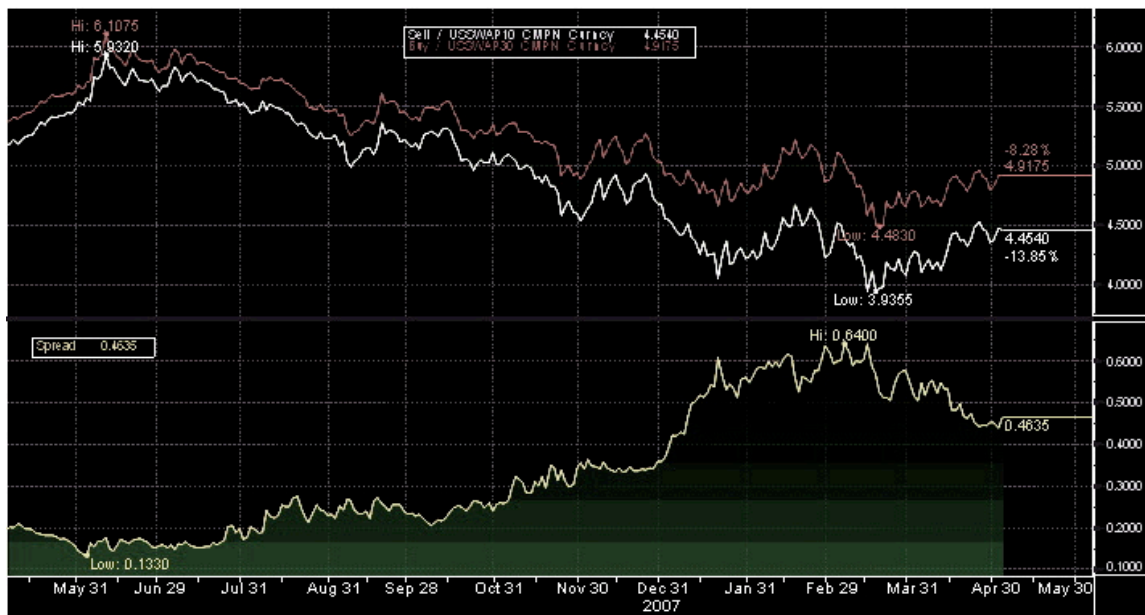
Advanced Data Analysis Final Project

May 5<sup>th</sup>, 2008

## ***Time Series Analysis of the US Swaps Markets***

Author:

Regina M. Flores



Source: Bloomberg



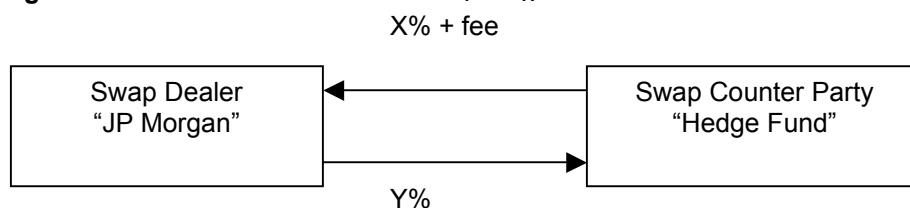
## ***Table of Contents***

<b>1. A Background in US Swaps Market.....</b>	<b>4</b>
<b>2. Objective.....</b>	<b>6</b>
<b>3. Data.....</b>	<b>6</b>
3.1 Description & Design.....	6
3.2 Evaluating Data Quality.....	7
<b>4. Methods.....</b>	<b>7</b>
4.1 Model Framework: Regression with Time Series Errors.....	7
4.2 Model Framework: Modeling Volatility.....	8
4.3 Time Series Regression & Model Selection.....	9
4.4 Model Fitting Procedure.....	10
<b>5. Results.....</b>	<b>17</b>
<b>6. Conclusions &amp; Discussion.....</b>	<b>27</b>
<b>7. Appendix.....</b>	<b>31</b>
7.1 Plots and Tables for 30yr Swap Rates.....	31
7.2 Plots and Tables for 7yr Swap Rates.....	31
7.3 Plots and Tables for 10yr Swap vs. 10yr Treasury.....	36
7.4 Plots and Tables for 7yr and 10yr Swap Spread.....	41
7.5 Plots and Tables for 30yr Swap vs. Fed Funds.....	45
7.6 R Procedures.....	55
<b>8. References.....</b>	<b>70</b>

## 1. A Background in US Swaps Market

Since the early 1990's the swaps market has experienced tremendous growth. By the year 2002 the outstanding notional on both interest and currency rate swaps exceeded \$100 trillion<sup>1</sup> USD. The swaps market attracts investors of all flavors since it allows investors broad access to the capital markets in order to express their macro-economic views. To this end, there are a variety of different types of swaps products that have been born to allow investors to express their views in ever more complex ways<sup>2</sup>. However, in the confines of this paper only the most basic type of interest rate swap product need be explained. An interest rate swap is defined as “a contractual agreement between two parties to exchange a series of payments for a stated period of time<sup>3</sup>” at an agreed upon rate. An interest rate swap, compared to a currency rate swap, is a series of payments made in the same currency. The most basic kind of interest rate swap is shown in Figure 1 below. A swaps dealer, say a large broker like JP Morgan, and a swaps counter party, say a hedge fund, enter into a contract whereby the swaps dealer will pay the counterparty Y% on a \$100 million notional for 5 years. In return the counterparty will pay X% + fee on the \$100 million for 5 years.

**Figure 1:** Plain vanilla interest rate swap diagram



The X% + fee the counterparty pays is a fixed rate, that is, it is locked and cannot change over the lifetime of the contract. The locked rate is determined by today's prevailing 5 year swap rate. The Y% the dealer pays is floating, that is, over the lifetime of the contract it will move in

<sup>1</sup> Corb, pg 372.

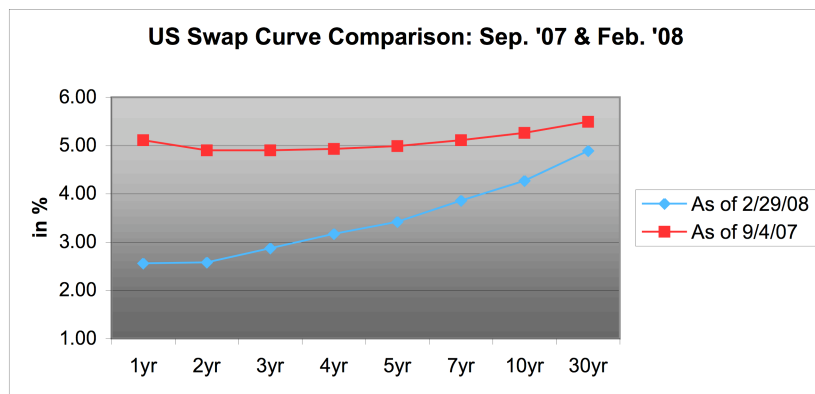
<sup>2</sup> Many complex swaps products now have embedded options, called swaptions, that allow investors to take on additional risk which of course could yield potentially greater return.

<sup>3</sup> Dattatreya, pg 1.

accordance with Libor<sup>4</sup>. The counterparty is betting that she knows more about the macro-economic variables that drive the swaps curve and that over the 5 years the X% + fee that she is paying will be less than the Y% she is receiving. Conversely, the dealer is betting that he will be paying less than he is receiving.

As mentioned above, the rate that an investor pays in a swaps transaction is called the swap rate and is based off of the swaps curve. The swaps curve is derived from the yield curve (the curve that represents the yield to maturity of on-the-run treasuries) and is in fact equal to the treasury yield plus a spread. The yield curve is a measure of the market's expectations of future interest rates given the current market conditions, thus, swap rates are very volatile and change day over day based on a variety of economic factors such as supply and demand and interest rate levels<sup>5</sup>. Figure 2 shows the swap curve as of September 2007 and February 2008 (after a series of rate cuts), which demonstrates how the swap curve changed under different interest rate environments.

Understanding what drives the swaps curve is crucial for players in the rates markets. Developing a model that exploited the term structure of interest rates to determine the relationship between interest rates with different maturities and potentially other economic factors would be a very powerful tool for investors and dealers alike.



**Figure 2:** Swap curve under different interest rate conditions. In September '07 interest rates were around 5% where as in February '08 rates were around 3%.

<sup>4</sup> Libor (London Inter-Bank Offer Rate) "is a daily reference rate based on the interest rates at which banks offer to lend unsecured funds to other banks in the London wholesale money market," (Wikipedia). It is the most common reference rate used in swap market, for the purposes of this paper no further information is necessary.

<sup>5</sup> Corb, pg 375.

## **2. Objective**

This paper attempts to determine appropriate models to predict long term swap rates. 'Long term' is defined to be greater than five years and 'short term' is defined to be less than or equal to five years. The term structure of interest rates will be used in order to determine a relationship between the front end and the back end of the swaps curve. That is, short term swap rates will be used as explanatory variables to determine long term rates<sup>6</sup>. An additional set of macro-economic explanatory variables will also be used as exogenous variables to predict long term rates. The exogenous variables are: the Federal Reserve effective rate, constant maturity US treasury rates, and US Baa<sup>7</sup> corporate bond yields. These will be employed to determine to what extent, if any, these help predict long term rates.

## **3. Data**

### **3.1. Description & Design**

The unit of analysis for all data examined in this paper is a daily rate represented in percentage terms. All data were retrieved from the United States Federal Reserve Website<sup>8</sup> compiled in the H.15 statistical data release during the period of 1/03/2006 – 2/29/2008. The below descriptions of the explanatory variables come from the Federal Reserve website.

The swaps data are provided by the International Swaps and Derivatives Association (ISDA®). Rates are mid-market par swap rates for a Fixed Rate Payer in return for receiving three month LIBOR, and are based on rates collected at 11:00 a.m. Eastern time by Garban Intercapital PLC and published on Reuters website as ISDAFIX®<sup>9</sup>. The data consist of 1, 2, 3, 4, 5, 7, 10 and 30 year rates.

The daily effective federal funds rate is a weighted average of rates on brokered trades. Weekly figures are averages of 7 calendar days ending on Wednesday of the current week;

---

<sup>6</sup> Tsay, 66.

<sup>7</sup> Baa bond yields may be more representative of economic effects since a Aaa company's performance may be less susceptible to fluctuations in economic conditions.

<sup>8</sup> <http://www.federalreserve.gov/releases/>

<sup>9</sup> Reuters Limited

monthly figures include each calendar day in the month. Rates are annualized using a 360-day year or bank interest.

Constant maturity treasury yields are taken from actively traded non-inflation-indexed issues adjusted to constant maturities. The 30-year Treasury constant maturity series was discontinued on February 18, 2002, and reintroduced on February 9, 2006. From February 18, 2002, to February 9, 2006, the U.S. Treasury published a factor for adjusting the daily nominal 20-year constant maturity in order to estimate a 30-year nominal rate. The historical adjustment factor can be found on the US Treasury website.

Moody's Baa rates through December 6, 2001, are averages of Baa utility and Baa industrial bond rates. As of December 7, 2001, these rates are averages of Baa industrial bonds only.

### **3.2. Evaluating Data Quality**

As the data comes from the Federal Reserve it is highly reliable. There are a few missing values in swaps data due to the markets being closed. An example is September 12, 2001 (the day after Sep.11). In all cases the gaps are only one day long and thus linear interpolation was used to fill-in missing rates. This approach seems to be an accurate method since there are not large movements in the rates between the gap days. The 30 year Treasury rate missing data (as outlined in the above section) was treated as per the US treasury website<sup>10</sup>.

## **4. Methods**

The methods described in this paper apply the Box-Jenkins approach using a time domain analysis. The next two sub-sections describe the mathematical framework of the models.

### **4.1. Model Framework: Regression with Time Series Errors**

A linear regression with ARMA(p,q) errors model was used. As per Brockwell and Davis, when using time series methods to analyze errors terms we can no longer assume that they are

---

<sup>10</sup> [www.treas.gov/offices/domestic-finance/debt-management/interest-rate/ltcompositeindex\\_historical.shtml](http://www.treas.gov/offices/domestic-finance/debt-management/interest-rate/ltcompositeindex_historical.shtml)

i.i.d (independent and identically distributed). A more common method is to assume that the errors are a mean zero second order stationary process. Thus, we can approximate these processes with an ARMA(p,q) model and we will assume the following model<sup>11</sup>:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \dots + \beta_k X_{tk} + W_t \quad (1)$$

where  $t=1,\dots,n$ ,  $k=1,\dots,n-1$  and  $W_t$  are values of a causal zero-mean ARMA(p,q) process satisfying:

$$\phi(B)W_t = \theta(B)Z_t \quad (2)$$

where  $\{Z_t\} \sim WN(0, \sigma^2)$ .

A generalization of the ARMA(p,d) model is the ARIMA process. The ARIMA(p,d,q) model is an alternative to the ARMA(p,q) model and reduces to the ARMA model when differenced finitely many times. Brockwell and Davis define the ARIMA(p,d,q) model as follows: If  $d$  is a nonnegative integer, then  $\{X_t\}$  is an ARIMA(p,d,q) process if  $Y_t := (1-B)^d X_t$  is a causal ARMA(p,q) process and  $\{X_t\}$  satisfies a difference equation of the form<sup>12</sup>

$$\phi^*(B)X_t = \phi(B)(1-B)^d X_t = \theta(B)Z_t \quad (3)$$

where  $\{Z_t\} \sim WN(0, \sigma^2)$ .  $\{X_t\}$  is only stationary when  $d=0$ , thus ARIMA models are best suited for data with trends. ARIMA models were used in the analysis of the data.

#### 4.2. Model Framework: Modeling Volatility

Since ARIMA processes assume constant variance they are not appropriate when modeling changing volatility. The GARCH process is often used in analyzing financial time series since it requires only a few parameters to properly describe the volatility structure. Tsay defines a GARCH(m,s) model as<sup>13</sup>:

$$a_t = \sigma_t \varepsilon_t \quad \text{and} \quad \sigma_t^2 = \alpha_0 + \sum_{i=1,m} \alpha_i a_{t-i}^2 + \sum_{j=1,s} \omega_j \sigma_{t-j}^2 \quad (4)$$

where  $\alpha_0 > 0$  and  $\alpha_i, \omega_j \geq 0$  for  $j = 1, 2, \dots$

This model is employed when the residuals of the fitted ARIMA model are not stationary.

<sup>11</sup> Brockwell and Davis, page 210-211

<sup>12</sup> Brockwell and Davis, pg 180.

<sup>13</sup> Tsay, pg 93.



#### 4.3. Time Series Regression & Model Selection

Model selection proved to be one of the most difficult aspects of the analysis mainly due to the fact there seem to exist different opinions on the appropriate procedure to use when regression errors are correlated. In Shumway and Stoffer's *Time Series Analysis and Its Applications*<sup>14</sup>, model selection is performed using standard linear regression techniques (assuming error terms are i.i.d normal variables with mean zero and variance  $\sigma^2$ ). Even though it is known that the error terms are correlated, this assumption is "relaxed" in order to perform model selection. The authors describe standard linear regression methods such as 'regsubsets' in R (maximizing  $R^2$ ) to evaluate the best models using time series data under these "relaxed" assumptions. Richard Davis<sup>15</sup>, professor of statistics at Columbia University and author of *Introduction to Time Series and Forecasting*, does not agree with the methods suggested by Shumway. In his co-authored paper, *Model Selection for Geostatistical Models*, Davis explains that "correlation is [often] ignored in the selection of explanatory variables [however] ignoring the autocorrelation structure in the data can influence model selection results<sup>16</sup>." Nobel laureate C.W.J Granger also agrees with Davis. In his paper *Spurious Regressions in Econometrics*, Granger describes how "nonsense regressions can arise<sup>17</sup>" when ignoring the time series properties of variables.

Based on the papers mentioned above and conversations with Davis, in order to determine which explanatory variables to include in the model a linear model was fit using all 8 explanatory variables (the 1,2,3,4 and 5 year swap rates, Baa bond yields, Fed Funds rate and Treasury rates) for each of the three response variables respectively: 7yr, 10yr and 30yr swap rates. Time series analysis was then employed to find the best ARIMA model that minimizes the AIC and maximizes the log likelihood (see section 4.4). The coefficients for  $\beta$  were then examined to check for significance. If  $\beta$  was not significant the explanatory variable was eliminated and the ARIMA model was refit. This was repeated until all coefficients were significant. Additional models

---

<sup>14</sup> Shumway, pg 54-71.

<sup>15</sup> Professor Davis was contacted and provided suggestions and personal opinions on this matter.

<sup>16</sup> Hoeting and Davis, pg 2.

<sup>17</sup> Granger, pg 117.

were also selected to test additional economic hypothesis. Refer to the Results section for further detail.

#### 4.4. Model Fitting Procedure

The following procedure was used for all of the models in this paper. The procedure<sup>18</sup> is a 9 step iterative method that is implemented as follows:

Step 1: Fit a linear regression model in the form  $Y = X\beta + W$ .

Step 2: Show that the residuals are correlated.

Step 3: Examine the ACF and PACF to determine if differencing procedure is needed.

Step 4: If differencing procedure is needed, apply it.

Step 5: Run autoregressive function minimizing AIC to find best ARMA(p,q) fit.

Step 6: Fit ARIMA(p,d,q) to  $\max L(\beta, \phi, \theta, \sigma^2)$ , verify that coefficients are significant<sup>19</sup>.

Step 7: Plot fitted ARIMA(p,d,q) residuals and verify variance is constant.

Step 8: If variance is not constant, fit GARCH(m,s) model to the residuals.

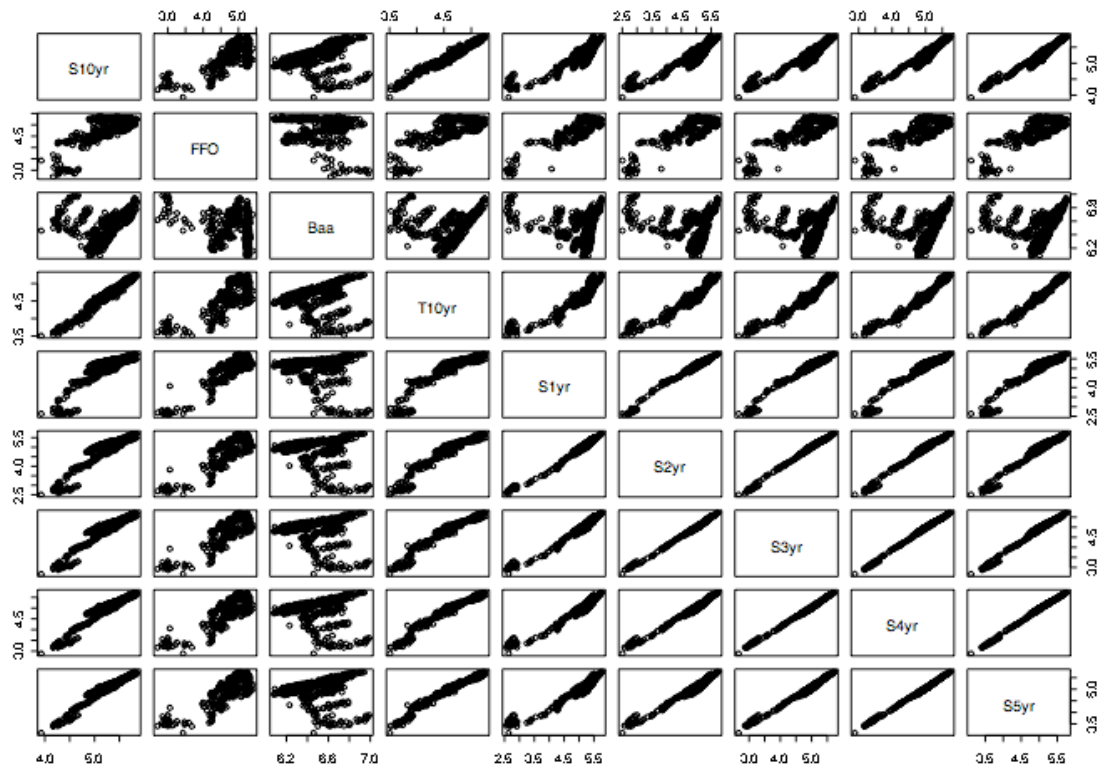
Step 9: Perform model diagnostics and validate assumptions.

In order to illustrate the above procedure, we will examine the linear fit of the 10year swap rates versus all eight variables as a detailed example. All other plots and tables for other models will be displayed in the Appendix section. Refer to the results section for the details on the performance of each model. Step 1: Figure 3(a) shows a matrix of scatter plots of all the eight variables plotted against each other to demonstrate the linear dependence of certain variables. Step 2: The residuals are then plotted to determine if they are in fact correlated. Figure 3(b) shows the plot of the residuals that clearly shows the correlation structure. Step 3: The ACF and PACF are plotted next in order to determine if there is significant trend and seasonality. Step 4: As shown from Figure 4(a) and (b), the residuals decay exponentially and thus need to be differenced. Figure 6 shows the ACF and the PACF of the differenced residuals at lag 1.

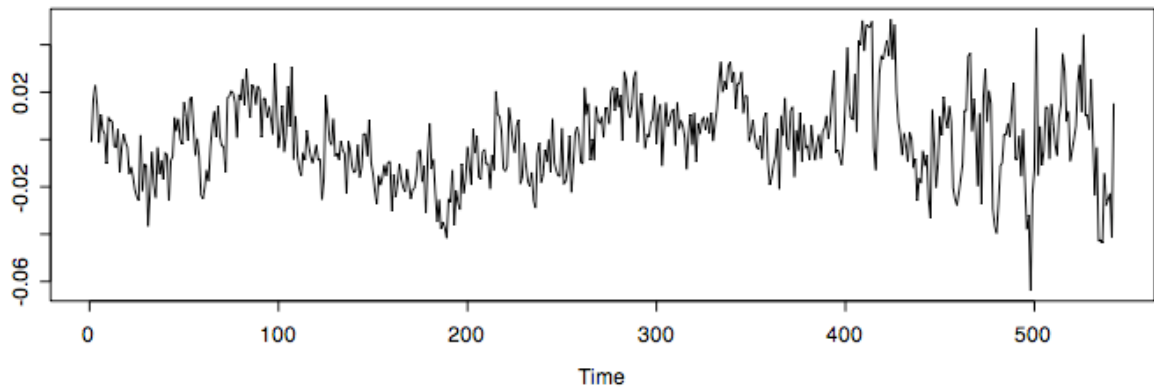
---

<sup>18</sup> Refer to Appendix XX: R Code Procedures

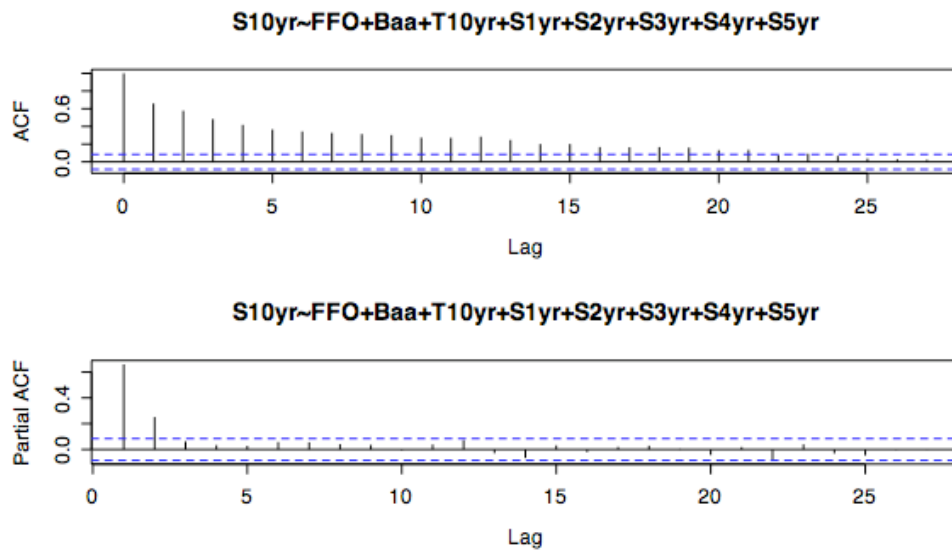
<sup>19</sup> Note that the 'arima0' function in R must be used since the 'arma' function does not return the estimates for  $\beta$ .



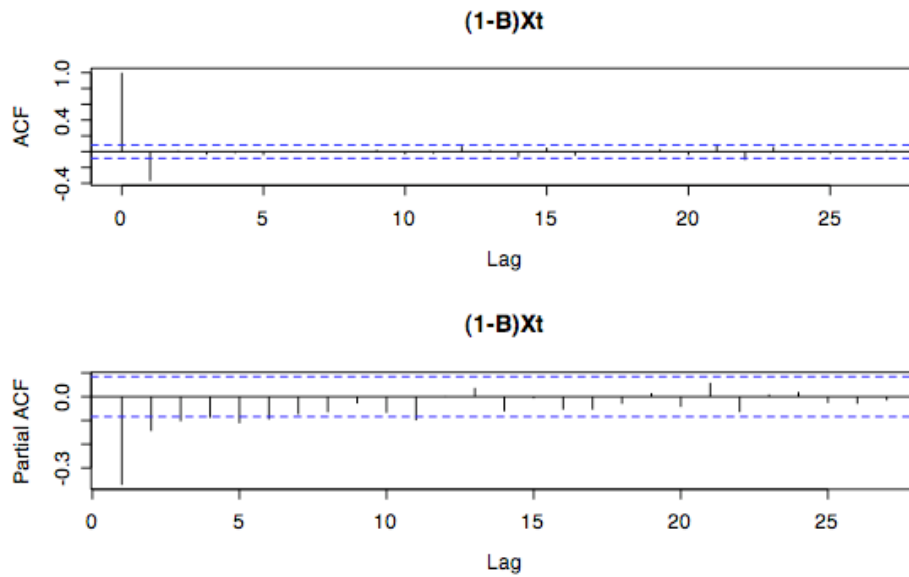
**Residuals - All variables**



**Figure 3 (a) & (b):** Relationship between all eight variables. When linear regression is fit residuals are clearly correlated.



**Figure 4 (a) & (b):** ACF and PACF of the residuals.



**Figure 5 (a) & (b):** ACF and PACF of the differenced residuals.

**Table 1:** AIC values for different ARIMA models. The smallest AIC is selected.

AR	MA	AIC
0	0	-3036.217
0	1	-3142.816
0	2	-3146.372
0	3	-3155.161
0	4	-3157.885
1	0	-3117.318
1	1	-3153.411
1	2	-3164.694
1	3	-3162.989
1	4	-3163.4
2	0	-3127.485
2	1	-3164.254
2	2	-3163.087
2	3	-3161.153
2	4	-3162.978
3	0	-3131.746
3	1	-3162.952
3	2	-3163.178
3	3	-3159.165
3	4	-3157.232
4	0	-3134.364
4	1	-3161.07
4	2	-3159.04
4	3	-3157.039
4	4	-3155.129

Step 5: An autoregressive function is run in order to find the best ARMA fit by minimizing the AIC.

In this case an ARIMA(1,1,2) is found to have the smallest AIC. However, when the coefficients of the model are investigated, 'NAN' errors are found. Looking at the next smallest AIC, an ARIMA(2,1,1) model is selected. Step 6: Fitting the model, the coefficients are examined. If  $|\hat{\beta}_k|^{20} > 2*SE$ , then the explanatory variable is eliminated from the mode. Based on this model fitting procedure, three variables were removed and the following model was fit:

$$10yrSwap = \beta_0 + \beta_1 \text{Baa bond yields} + \beta_2 \text{2yrSwap} + \beta_3 \text{3yrSwap} + \beta_4 \text{4yrSwap} + \beta_5 \text{5yrSwap} + \varepsilon$$

Table 2 shows the estimates of both the linear regression coefficients and the correlated error coefficients modeled as an ARIMA(2,1,1).

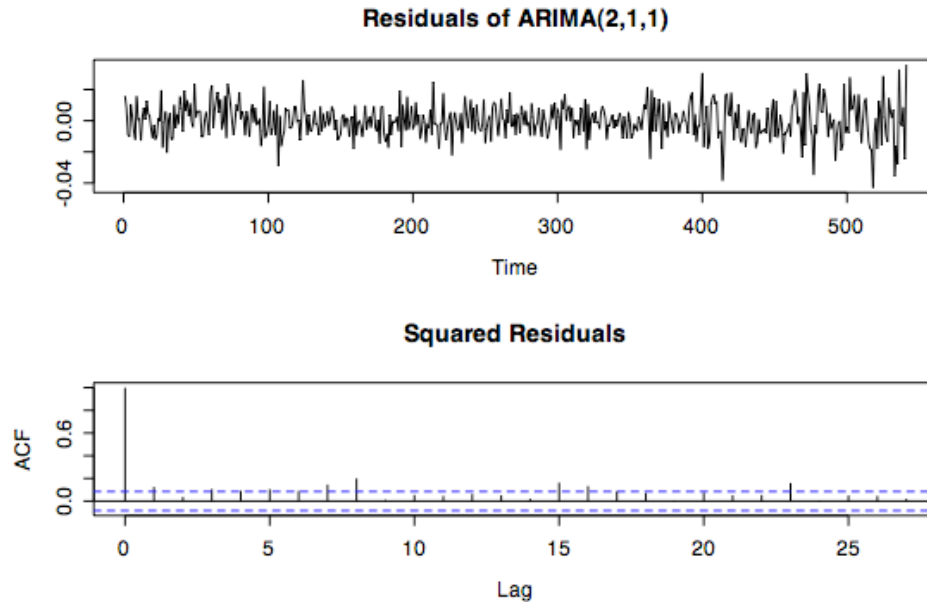
---

<sup>20</sup> This is the estimated value of  $\beta$ , what we refer to as "beta hat".

**Table 2:** Estimates of the ARIMA(2,1,1) and the regression coefficients.

ARIMA(2,1,1) & Regression Estimates:									
	$\phi_1$	$\phi_2$	$\theta_1$	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$
	0.5657	0.1333	-0.8552	1.00E-04	0.1255	-0.3465	-0.3024	0.3209	1.1129
SE	0.0459	0.0523	0.2253	2.00E-04	0.0134	0.0497	0.0839	0.0882	0.0652
$\sigma^2$	0.0001072								
$L(\beta, \phi, \theta, \sigma^2)$	1704.91								
AIC	-3389.82								

Step 7: The residuals of the ARIMA fit are then examined to verify that they are stationary. The ACF of the squared residuals is also examined. If the error term is GARCH then the square of the residuals should decay exponentially like a non-Gaussian AR(1) process<sup>21</sup>. As seen in Figure 6, there is significant autocorrelation from the squared residuals and a GARCH model is needed.



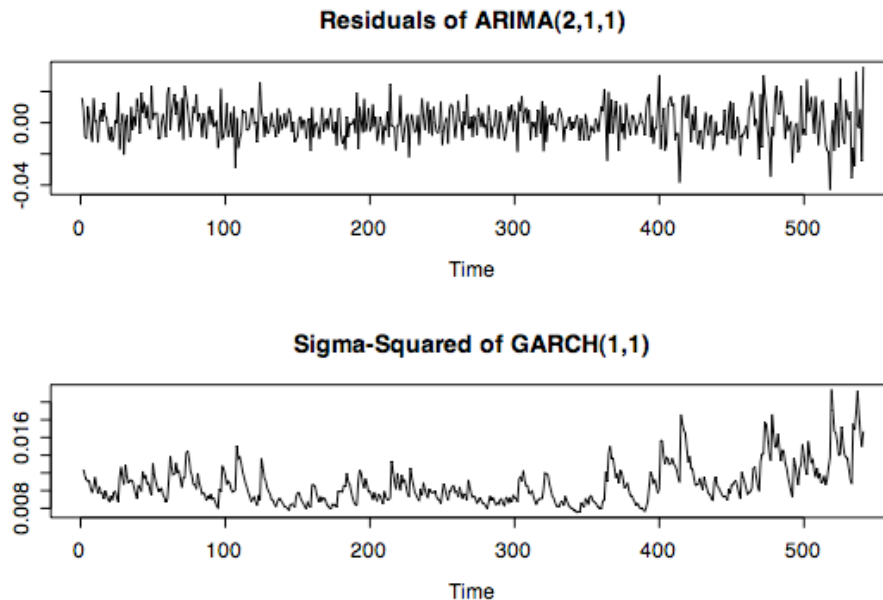
**Figure 6 (a) & (b):** Testing the changing variance in the residuals. ACF of the squared residuals from the ARIMA(2,1,1) fit.

<sup>21</sup> Shumway, pg.283.

Step 8: A GARCH(1,1) model is fit to the ARIMA residuals. Table 3 shows the estimates of the GARCH coefficients. Figure 7 shows the estimated volatility produced by the GARCH(1,1) fit to the ARIMA residuals.

**Table 3:** Model estimates of the GARCH(1,1) fit.

<b>GARCH(1,1) Estimates:</b>				
	<b>Estimate</b>	<b>Std. Error</b>	<b>t value</b>	<b>Pr(&gt; t )</b>
$\alpha_0$	1.28E-05	4.96E-06	2.569	0.010205
$\alpha_1$	1.67E-01	4.76E-02	3.498	0.000468
$\omega_j$	7.38E-01	6.01E-02	12.288	<2e-16



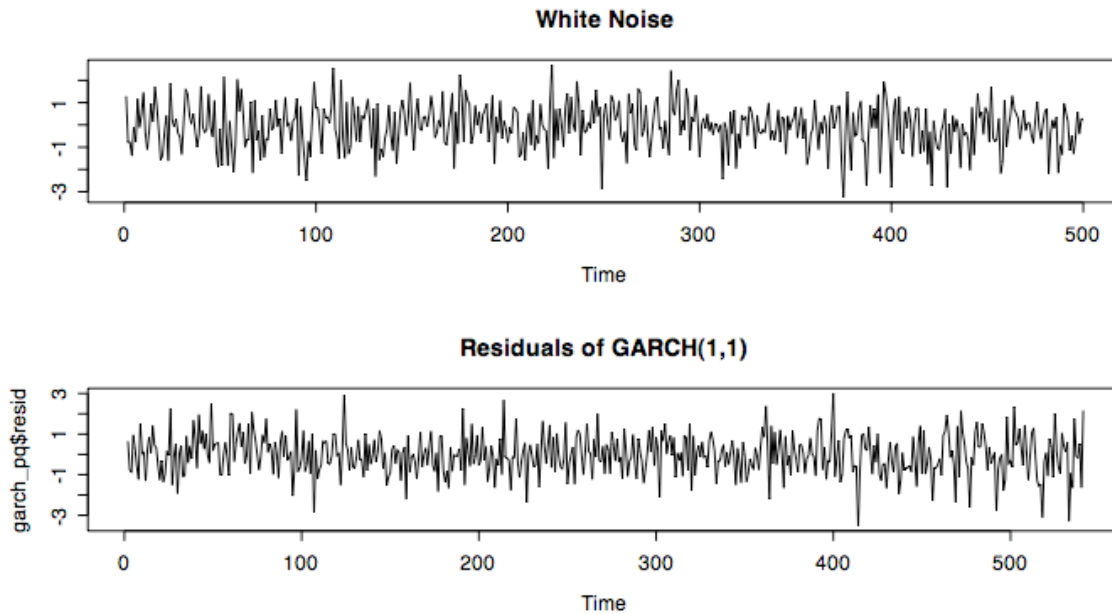
**Figure 7(a) & (b):** Estimated volatility shown with the residuals of the ARIMA fit to the error term.

Step 9: Diagnostic test are run to validate the assumptions of the GARCH model. As shown in Table 4, the Box-Ljung test, under the null that the series has a unit root, gives a p-value of 0.2403 and thus we fail to reject the null hypothesis and conclude it is a stationary process. Thus, the GARCH(1,1) model removes all the correlation structure from the error term. Looking further at Figure 8, the time series of the residuals of the GARCH(1,1) closely resembles the time series of white noise. The Jarque Barra test, under the null that the residuals are normal, gives a p-value

of 0.0959 so we fail to reject the null hypothesis. Looking at the histogram and the Q-Q plot of the residuals, shown in Figure 9, they appear to be normal with slight deviations at the tails. For further discussion about the fitted model refer to the Results section.

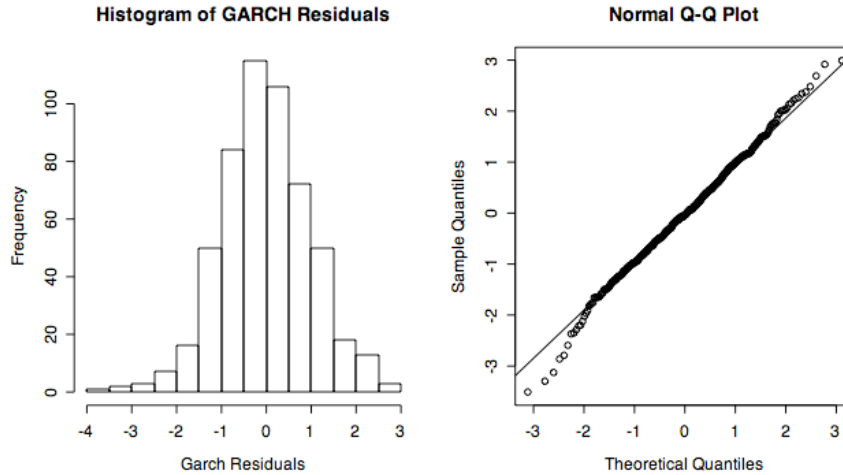
**Table 4:** Diagnostic test for the GARCH model.

Diagnostic Tests				
Jarque Bera Test				
Residuals				
X-squared	4.6889	df = 2,	p-value =	0.0959
Box-Ljung Test				
Squared.Residuals				
X-squared	1.3141	df = 1,	p-value =	0.2516



**Figure 8:** Tests for stationary residuals from the GARCH(1,1) model.





**Figure 9:** Tests for normality for the GARCH model.

## 5. Results

In the previous section the basic method for fitting the models was outlined. In this section, based on those procedures, the results for all the models are presented. A total of 6 models were examined in order to experiment with various ways to predict long term swap rates. Table 5 shows all of the models used. As discussed in the previous section a model was fit against the 7, 10 and 30yr swap rate response variables respectively using all the 8 explanatory variables. In addition three models were fit in order to explore macro-economic hypotheses.

As per Table 5, models (1), (2) and (3) refer to the 7, 10 and 30yr swap rates against all 8 explanatory variables. Refer to Appendix 7.1 and 7.2 for the pairs plots that show the relationships between them for the 7 and 30yr swap rates (recall the 10yr plots were shown in section 4.4). ARIMA-GARCH models were used to fit the 7, 10 and 30yr swap rate models. All the ARIMA models follow the following form:

$$\Delta Y_t = \sum \phi_j \Delta Y_{t-j} + \sum \beta_i \Delta X_{it} \quad (5)$$

Table 6 and 7 show the ARIMA and GARCH fits respectively to each of the long-term rates. In all three of these models, the regression residuals showed strong decaying autocorrelations functions and thus differencing was required. Appendix 7.1 and 7.2 show the plots of the ACF and PACF before and after differencing for the 30yr and 7yr swap rates respectively. As would be

**Table 5:** Six models used. S7yr, S10yr and S30yr mean the 7, 10 and 30yr swap rate. Baa means Baa bond yields. FFO means Fed Funds rate. T7yr, T10yr and T20yr means constant maturity 7, 10 and 30yr treasury rates. S10s7s and S5s2s mean the swap spread between the 10 and 7yr swap rate and the 5 and 2yr swap rate.

Model	Response Variable	Explanatory Variables
1	S7yr	Baa, T7yr, S2yr, S3yr, S4yr, S5yr
2	S10yr	Baa, S2yr, S3yr, S4yr, S5yr
3	S30yr	FFO, T30yr, S2yr, S3yr, S4yr, S5yr
4	S10yr	T10yr
5	S10s7s	Baa, S5s2s, S1yr, S3yr
6	S30yr	FFO

expected due to the term structure of interest rates, there is a strong linear relationship between the swap rates for different maturities. Based on the model selection process some variables were eliminated in order to find the best-fit model. In all three models the 1yr swap rate drops out of the model, but the 2, 3, 4 and 5yr rates are significant for all three models. This seems to agree with economic theory since the 1yr swap rate is a reflection of current market conditions. As the maturity of the swap rates increase, the rates tend to be a reflection of expectations of future market conditions<sup>22</sup>.

Because of the term structure of interest rates, one might assume that the variables that predict the 7yr swap rates would be the same for the 10 and 30yr rates. However that does not seem to be the case. For example, there does not appear to be a clear pattern as to how the exogenous variables, Baa bond yields and Fed Funds, behave. The Baa bond yields were a significant explanatory variable in the 7 and 10yr, but not in the 30yr. Conversely, Fed Funds rates were significant in the 30yr, but not the 7 and 10yr. However, the estimated  $\beta$  coefficient for the Fed Funds variable in the 30yr model is very weak. Furthermore, the corresponding year Treasury rates also do not exhibit any clear patterns. They are significant for predicting the 7 and 30yr swap rates, but not the 10yr swap rates. In all three models the 5yr swap rate has an increasing  $\beta$  coefficient. That is to say, that it is larger in the 30yr model than in the 7yr. This makes sense due to the term structure of interest rates.

<sup>22</sup> Lauschke, pg. 3.

**Table 6:** Comparing the ARIMA fit for the 7, 10 and 30yr swap rates.

7yr Swap: ARIMA(1,1,2) & Regression Estimates:										
	$\phi_1$	$\theta_1$	$\theta_2$	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$
	0.3905	-0.8484	0.1302	0.00E+00	0.1011	-0.0332	-0.2259	-0.1338	0.39	0.8681
S.E.	0.0466	0.0387	0.0394	1.00E-04	0.0121	0.011	0.0318	0.0596	0.0607	0.045
$\sigma^2$	0.0000445									
L( $\beta, \phi, \theta, \sigma^2$ )	1942.68									
AIC	-3863.35									
10yr Swap: ARIMA(2,1,1) & Regression Estimates:										
	$\phi_1$	$\phi_2$	$\theta_1$	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	
	0.5657	0.1333	-0.8552	1.00E-04	0.1255	-0.3465	-0.3024	0.3209	1.1129	
S.E.	0.0459	0.0523	0.2253	2.00E-04	0.0134	0.0497	0.0839	0.0882	0.0652	
$\sigma^2$	0.0001072									
L( $\beta, \phi, \theta, \sigma^2$ )	1704.91									
AIC	-3389.82									
30yr Swap: ARIMA(2,1,1) & Regression Estimates										
	$\phi_1$	$\phi_2$	$\theta_1$	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$
	0.62	0.1484	-0.8702	1.00E-04	0.0158	0.1895	-0.6341	-0.3894	0.2745	1.3453
S.E.	0.0195	0.0212	0.025	4.00E-04	0.0068	0.0202	0.0748	0.1249	0.1344	0.0983
$\sigma^2$	0.0002569									
L( $\beta, \phi, \theta, \sigma^2$ )	1468.47									
AIC	-2914.95									

**Table 7:** Comparing the GARCH fit fo the 7, 10 and 30yr swap rates.

<b>7yr Swap: GARCH(1,1) Estimates:</b>				
	Estimate	Std. Error	t value	Pr(> t )
$\alpha_0$	1.61E-05	6.78E-06	2.374	0.01759
$\alpha_1$	1.56E-01	6.56E-02	2.373	0.01766
$\omega_1$	4.81E-01	1.82E-01	2.64	0.00828
<b>10yr Swap: GARCH(1,1) Estimates:</b>				
	Estimate	Std. Error	t value	Pr(> t )
$\alpha_0$	1.28E-05	4.96E-06	2.569	0.010205
$\alpha_1$	1.67E-01	4.76E-02	3.498	0.000468
$\omega_1$	7.38E-01	6.01E-02	12.288	<2e-16
<b>30yr Swap: GARCH(1,1) Estimates:</b>				
	Estimate	Std. Error	t value	Pr(> t )
$\alpha_0$	3.27E-06	2.39E-06	1.369	0.170899
$\alpha_1$	5.72E-02	1.53E-02	3.743	0.000182
$\omega_1$	9.32E-01	2.03E-02	45.838	<2e-16

From Table 6 it appears that the 7yr model performed the best since it has the smallest AIC and  $\sigma^2$  with the largest log-likelihood. For each model the AR coefficients, which relates the past and current values of the response, are positive and increasing in magnitude as the response maturities increase.

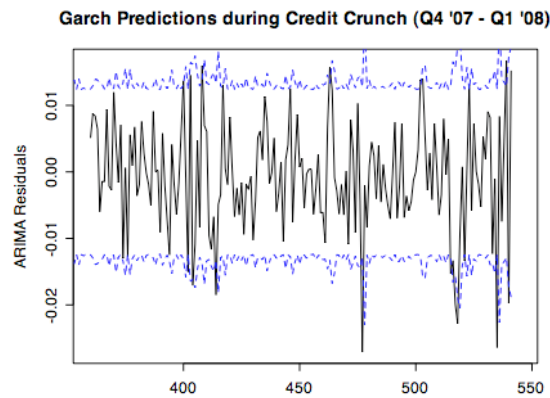
For all three long term swap rates, a GARCH(1,1) was selected as the best fit to model the volatility. The p-values given in Table 7 show that all the higher order estimated coefficients are significant, thus validating the order of the model. The Box-Ljung test was performed and all three models were found to have white noise residuals under the GARCH(1,1) model. The p-values for the 7 and 30yr are shown in Appendix 7.1 and 7.2. The 10yr model was the only one that passed the Jarque Bera test for normality as a condition of the Gaussian likelihood. However, as seen in the histogram and Q-Q plots, the residuals appear to be normal except for extreme values at the tails. The failure of the test may be a limitation in R (the software use to analyze all the data). R does not allow for long tailed distributions to be fit to the data<sup>23</sup>. This issue seems to propagate throughout all the results of this paper.

GARCH predictions were made to model the volatility of the error terms. Predictions were examined during volatile market conditions to test model robustness. The recent credit crunch period, during Q4 of 2007 through Q1 of 2008, was selected. Figure 10 shows the ARIMA residuals with the one-step-ahead predictions of the corresponding volatility. For the most part, in all three cases the predictions fall within the 95% confidence bounds and thus seem to model the volatility fairly well. When there are large one day movements, however, the model leaves something to be desired. Further discussions in the limitations to the GARCH model will be addressed in the Conclusions section.

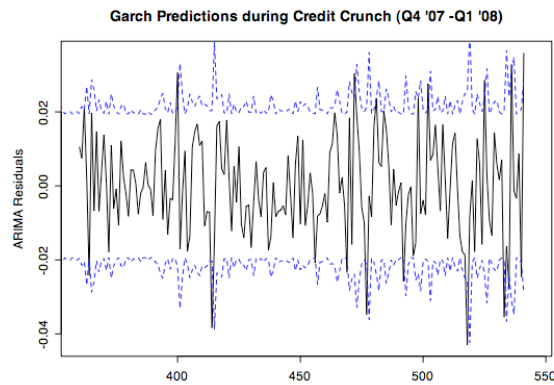
---

<sup>23</sup> Shumway, pg.287.

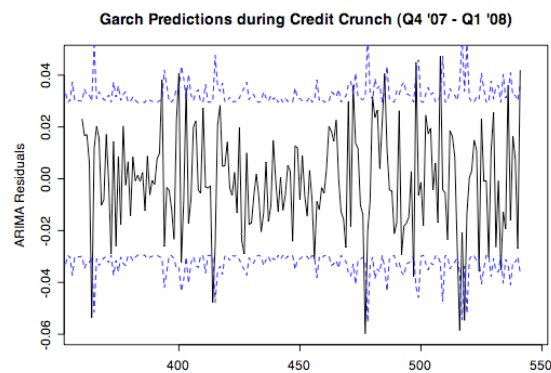
**Figure 10:** GARCH predictions. The  $\pm 2\sigma$  estimates are shown as dashed lines.



**(a):** GARCH predictions for the errors from 7yr swap rates regression.

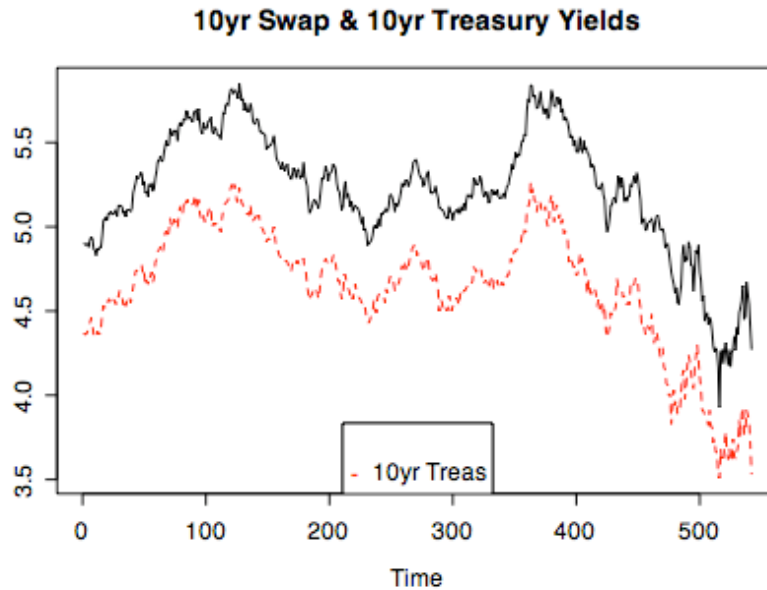


**(b):** GARCH predictions for the errors from 10yr swap rates regression.



**(c):** GARCH predictions for the errors from 30yr swap rates regression.

Three additional models were examined to verify various economic hypotheses. Figure 11 shows a strong relationship with treasury yields and swap rates demonstrated here by the 10yr curve. The goal of model (4) was to determine the relationship between these two rates and the potential for the treasury rates to predict swap rates. The 10yr curve was selected arbitrarily.



**Figure 11:** 10yr Swap rates plotted with 10yr Treasury rates.

Appendix 7.3 shows the plots of the ACF and PACF. This model also required differencing at lag 1. Applying the model fitting procedure yielded an ARIMA(0,1,1) fit with GARCH(0,1) errors as shown in Tables 8 and 9.

**Table 8:** Error fit to 10yr Swap vs 10yr Treasury regression.

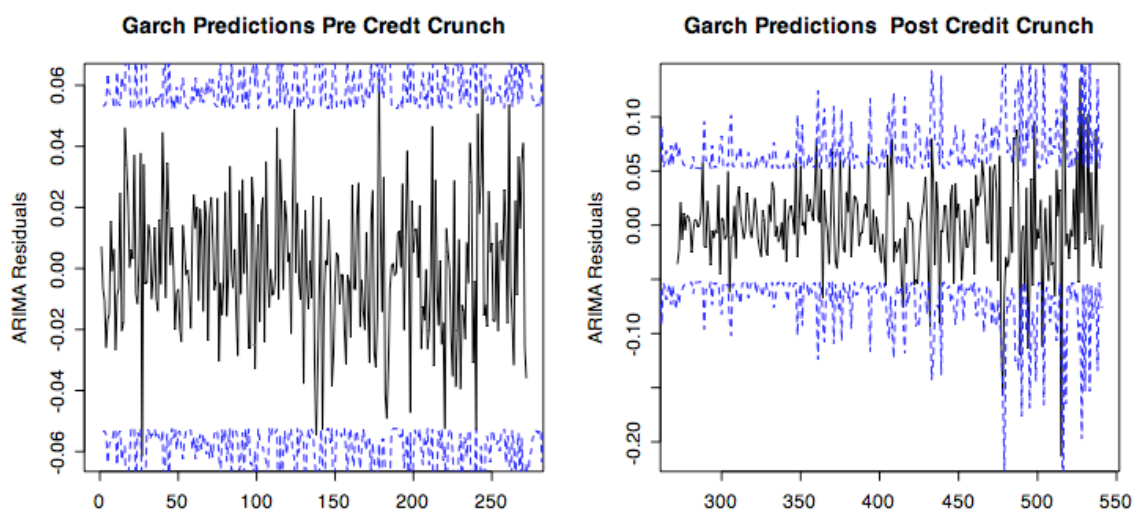
<b>ARIMA(0,1,1) &amp; Regression Estimates:</b>			
	$\theta_1$	$\beta_0$	$\beta_1$
	-0.5545	2.00E-04	0.8473
S.E.	0.0405	7.00E-04	0.0281
$\sigma^2$	0.001138		
$L(\beta, \phi, \theta, \sigma^2)$	1065.76		
AIC	-2123.52		

**Table 9:** GARCH(0,1)

<b>GARCH(0,1) Estimates:</b>				
	<b>Estimate</b>	<b>Std. Error</b>	<b>t value</b>	<b>Pr(&gt; t )</b>
$\alpha_0$	6.88E-04	4.81E-05	14.289	<2e-16
$\omega_1$	4.94E-01	6.72E-02	7.352	1.96E-13

A GARCH(1,1) model was originally fit, but it failed the Box-Ljung test. The GARCH(0,1) had a p-value of 0.3048 and thus passes the test for stationary residuals. Although the null hypothesis for the Jarque Bera test for normality was rejected, looking at the histogram and Q-Q plots shown in Appendix 7.3 the residuals appear to be normal except at the tails.

GARCH predictions were made during both volatile and normal market conditions as shown in Figure 12. The model does not seem to hold up very well during the credit crisis in the latter part of 2008 when swap rates were experiencing extreme volatility. Prior to the credit crisis, the model does seem to hold up well, as values fall between the 95% confidence bounds, which is in agreement with models (1), (2) and (3).

**Figure 12 (a) & (b):** Pre and Post credit crunch with  $\pm 2\sigma$  estimates shown as dashed lines.

Model 5 examines the spread, or the difference, between the 7yr and the 10yr swap rates. This model was selected because investors will often take views in the market by trading the difference between two swap rates. As an example, it is often said that the spread between the 2yr and 3yr swap rates is the markets expectation of where 1yr swap rate should be 2 years in the future<sup>24</sup>. The explanatory variables were not selected using the procedure outlined in section 4, rather they were selected based on hypothesis of how the spread should behave. Table 10 shows the estimates from the ARIMA(0,1,2) model that produced the best fit.

**Table 10:** Coefficient estimates for the 10yr and 7yr swap spread.

<b>ARIMA(0,1,2) &amp; Regression Estimates:</b>							
	$\theta_1$	$\theta_2$	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$
	-0.4341	-0.1091	0.00E+00	-0.0637	0.2156	-0.0568	0.1577
S.E.	0.0442	0.0453	2.00E-04	0.0091	0.0248	0.0187	0.0157
$\sigma^2$	5.20E-05						
$L(\beta, \phi, \theta, \sigma^2)$	1900.51						
AIC	-3785.02						

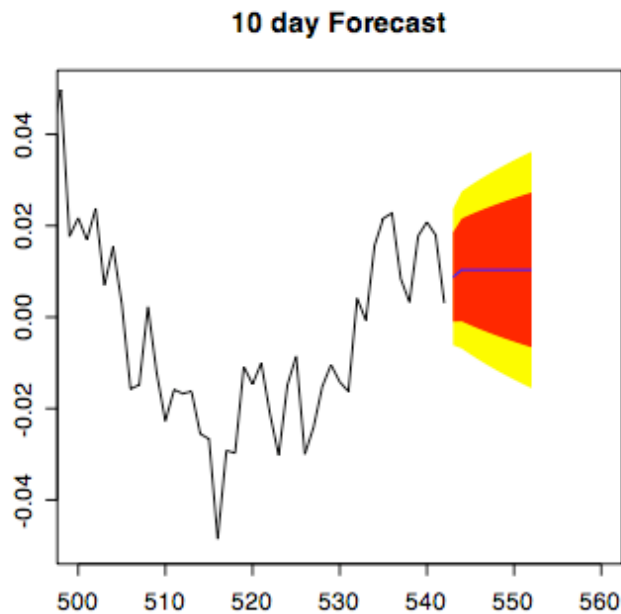
Appendix 7.4 shows the ACF and PACF plots for this model. The ACF plot seems to exhibit some seasonality behavior. However, differencing at lag 1 removes this pattern. For this model a GARCH fit was not needed since the ACF of squared residuals showed no significant autocorrelation. The diagnostic plots are shown in Appendix 7.4. All the p-values fall above 0.5 for the Box-Ljung test. Looking at the Augmented Dickey-Fuller and Phillips-Perron test, the null hypothesis that there is a unit root is rejected, thus the ARIMA model has stationary residuals. This is further revealed when comparing the residuals to white noise as seen in Appendix 7.4. The residuals appear to be normal as per the histogram and Q-Q plots even though the null hypothesis is rejected for the Shapiro-Wilk test.

ARIMA forecasting was performed as show in Figure 13. Appendix 7.4 shows the 10 day forecasted values with their respective 95% confidence intervals. They forecast does seem to

<sup>24</sup> Lauschke, pg. 3



capture general regime of the vales however further forecasting procedures may need to be investigated to determine directional movements.

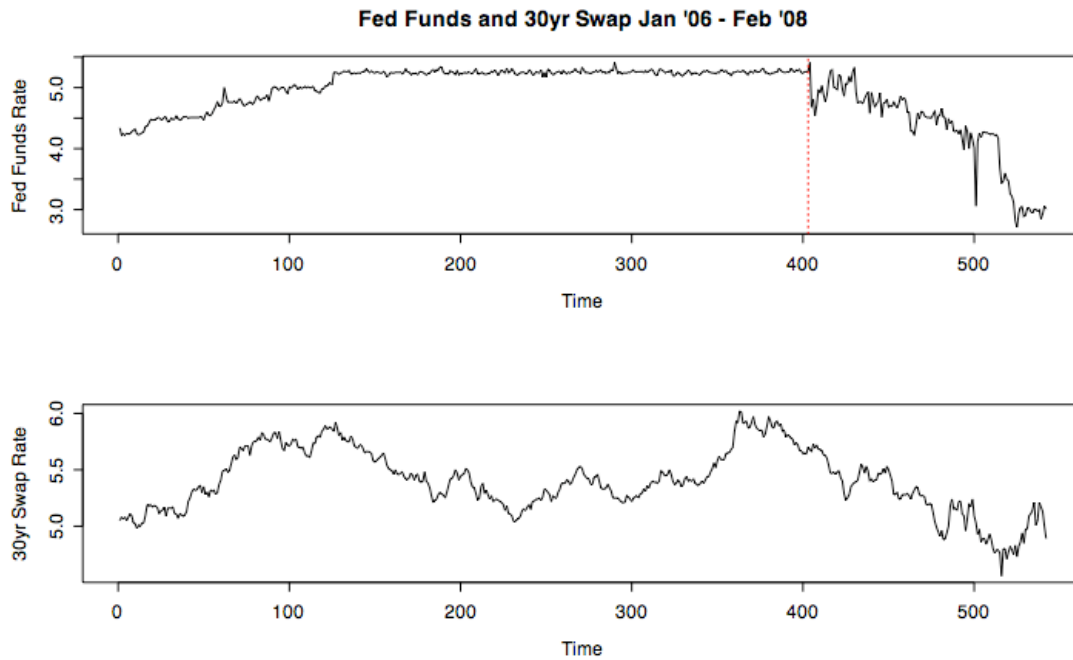


**Figure 13:** ARIMA 10 day forecasting for the 10yr and 7yr swap spread.

Model 6 explores the relationship between swap rates and Fed Funds rates. The 30yr swap curve was selected arbitrarily. Referring to Figure 14, notice how Fed Funds rates remained relatively constant prior to the second half of '07 (indicated by red dashed line). During the '07-'08 credit crisis Fed Funds rates experienced extreme volatility. Looking at the back end of the curve and the back end of the 30yr swaps curve, there seem to be similar movements. Thus model 6 was selected in order to explore if Fed Funds are an appropriate predictor of long term rates in extreme market conditions.

As shown in Appendix 7.5, there does not appear to be a strong linear relationship between Fed Funds rates and 30yr swap rates. (Other swap rate maturities were also explored with the same result.) Looking at the ACF and PACF of the differenced residuals reveals no significant autocorrelations, thus the residuals were not differenced. In fact fitting an ARIMA model to the differenced residuals yielded no significant coefficients. An ARIMA(1,0,0) with

GARCH(1,1) errors was fit. Table 11 shows the estimated coefficients. Although the null for the Jarque Bera test for normality was rejected, referring to the histogram and the Q-Q plot of the GARCH residuals appear to be normal.

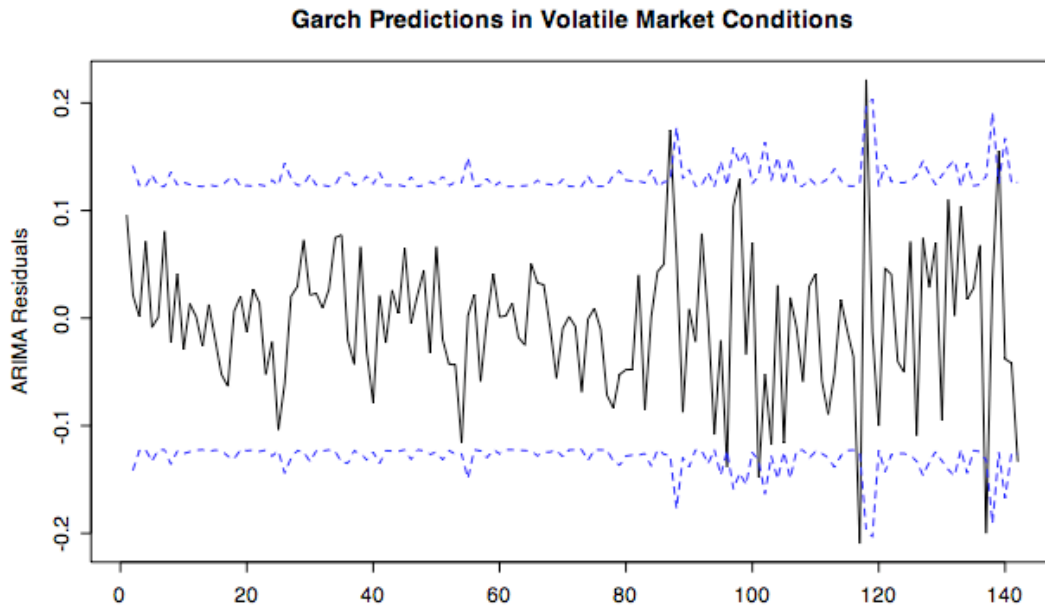


**Figure 14:** Relationship between Fed Funds rates and Swap rates during times of market volatility.

**Table 11 (a) & (b):** ARIMA and GARCH estimates.

<b>ARIMA(1,0,0) &amp; Regression Estimates:</b>			
	$\phi_1$	$\beta_0$	$\beta_1$
	0.9723	5.1675	0.0118
S.E.	0.0188	0.2038	0.0288
$\sigma^2$	0.004367		
$L(\beta, \phi, \theta, \sigma^2)$	182.85		
AIC	-357.7		

<b>GARCH(1,1) Estimates:</b>				
	<b>Estimate</b>	<b>Std. Error</b>	<b>t value</b>	<b>Pr(&gt; t )</b>
$\alpha_0$	0.0004234	0.0003881	1.091	0.2753
$\alpha_1$	0.2642829	0.1361484	1.941	0.0522
$\omega_1$	0.6938609	0.1372865	5.054	4.32E-07



**Figure 15:** GARCH predictions during credit crisis,  $\pm 2\sigma$  estimates shown as dashed lines.

GARCH predictions were made during the period between Q4 '07 and Q1 '08. As seen in Figure 15, the model seems to fit best in less volatile market conditions. In the later part of the curve when the Federal Reserve was aggressively cutting rates, the model does not hold up as well.

## 6. Conclusions & Discussion

Previous research has been done to model the swaps curves themselves, but in general there do not seem to be many papers that look at the linear relationships between long dated rates and other factors. In researching this paper, only one published report was found that attempted to use linear regression to model the swaps curve. The report, *Parsimonious Models for the USD Swap Curve*, examined the term structure of interest rates and used 3mo Libor as an exogenous variable to predict the 2 and 10yr swap rate. Since it was published by DBS Bank and intended for portfolio managers and traders, there are very few statistical results given. In fact, it is unclear if that paper takes correlated errors into account since there is no mention of it nor coefficient estimates provided. Thus, the results cannot be verified. Here, in this paper however, a

strict statistical method was employed in order to explain and predict the behavior of long term swap rates. Models were selected based on minimizing the Akaike Information Criterion and maximizing the Gaussian likelihood. In almost all cases ARIMA-GARCH models were fit since most residuals exhibited time-varying and persistent volatility. Diagnostic tests were made and assumptions were verified.

Two exogenous variables were explored in this paper for modeling 7,10 and 30yr swap rates: Baa bond yields and Fed Funds rates. Both of these variables are strongly correlated to the movement of interest rates. In fact, the Fed Funds rate is considered the benchmark of interest rates. Based on the findings in this paper, the relationship between these exogenous variables and all three long term rates is not the same. One would expect, based on the term structure of interest rates, that these variables would interact with the response variables in the same manner. In the models used in this paper, Baa bond yields were statistically significant in predicting 7 and 10yr swap rates, but not in the 30yr. It is important to recognize, however, that the average maturity of the Baa bonds from the H.15 Federal Reserve statistical release is not given. Thus, it is plausible that the Baa bond yields in this data set account for bonds with an average maturity between 7 and 10 years. Similarly, the Fed Funds rate was statistically significant only in predicting 30yr swap rates. However, in the H.15 data release the Fed Funds rates were determined based on the weighted average of rates on brokered trades. It is possible that there is a bias in the sample of brokered trades towards longer term maturities. In order to fully identify why these relationships did not hold on all three of the long dated swap rates, a broader data set may be required.

In examining models (1), (2) and (3), it is clear that there is a linear relationship between long term and short term swap rates. However, in all three models the 1yr swap rate was not significant. The most likely reason for this is due to the fact that 1yr rates are more closely tied to today's current market conditions while longer term rates are tied to market expectations of future economic environments. Both the 2 and 3yr swap rates had negative  $\beta$  terms in the models while the 4 and 5yr  $\beta$  terms were positive. Based on the estimated AIC, log-likelihood and the variance values, the 7yr model seemed to be the best fit. From model (4), a strong linear relationship

between the 10yr treasury and 10yr swap rate, with a  $\beta$  of 0.8, was found. This was expected since the swaps curve is derived from the treasury yield curve. The  $\beta$  coefficients tend to be smaller in model (5), thus it is not clear if the explanatory variables used were appropriate. Further investigation into the drivers of swap spreads would be particularly interesting since investors are attracted to trading both swap rates and the spreads between them. Contrary to expectations the findings from model (6) show that there is not a strong relationship between Fed Funds and swap rates (recall other swap maturities were investigated). Since Fed Funds are a driver of almost every rate in the financial markets, one would expect a strong linear relationship. In this model the Fed Funds Effective rate was used. Further tests should be performed looking at other variations of Fed Funds rates to verify this result.

There may be several limitations in the models presented in this paper. First, it has been suggested that *Regime-Switching* models are better in explaining the behavior of financial time-series data. Regime-switching models take into account that the time series processes behave differently in different economic environments. This is most apparent in the Fed Funds time series shown in Figure 14. Fed Funds rates were at relatively constant levels prior to 2007, however during the recent credit crisis the Fed underwent a series of rate cuts. Thus, there exist two regimes in the time series. To avoid this issue in this paper only the last two years of the series were used. It is possible that a regime-switching model might more accurately account for these drastic changes. Second, the models in this paper do not address the possibility of cross-correlated time series. It is possible that some or all of the variables are correlated with past values of each other and thus multivariate time series approaches may be more effective in predicting long dated swap rates. Third, the property of co-integration was not addressed in this paper. In Figure 11 the movement of the 10yr swap and treasury rates are shown. Surprisingly, since this appears to be a classical example of co-integrated series, the residuals of their regression was not stationary. This may be the case that linear regression is not the best way to capture the relationship between these two series and that cointegration techniques would be more effective. Forth, as seen in the results, the GARCH models did not seem to model extremely volatile markets well. This may be due to some of the limitations of the GARCH model

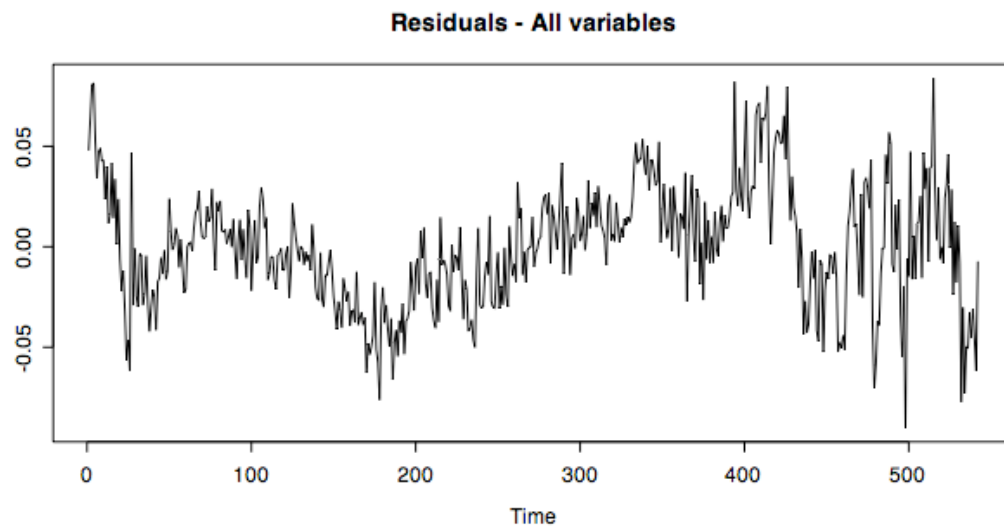
that Shumway addresses, in particular, that the model tends to respond slowly to large isolated movements. Other models in the ARCH family may prove to be better at dealing with these properties. Fifth, further research should be made to determine if there are other exogenous variables that are better predictors of long dated swap rates. As mentioned above, 3mo Libor has been used. Additional macro-economic variables such as Non-Farm Payroll, Consumer Confidence Index, Jobless Claims, Consumer Price Index or Retail Sales Index could be used as all of these dramatically effect market expectations, which in turn tend to move the market.

## 7. Appendix

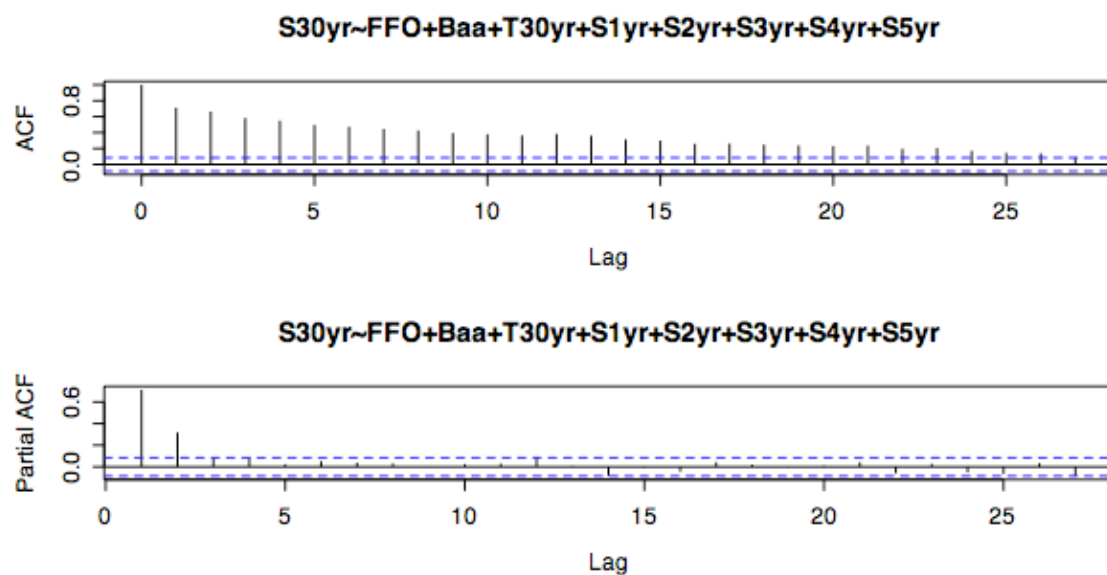
### 7.1. Plots and Tables for 30yr Swap Rates



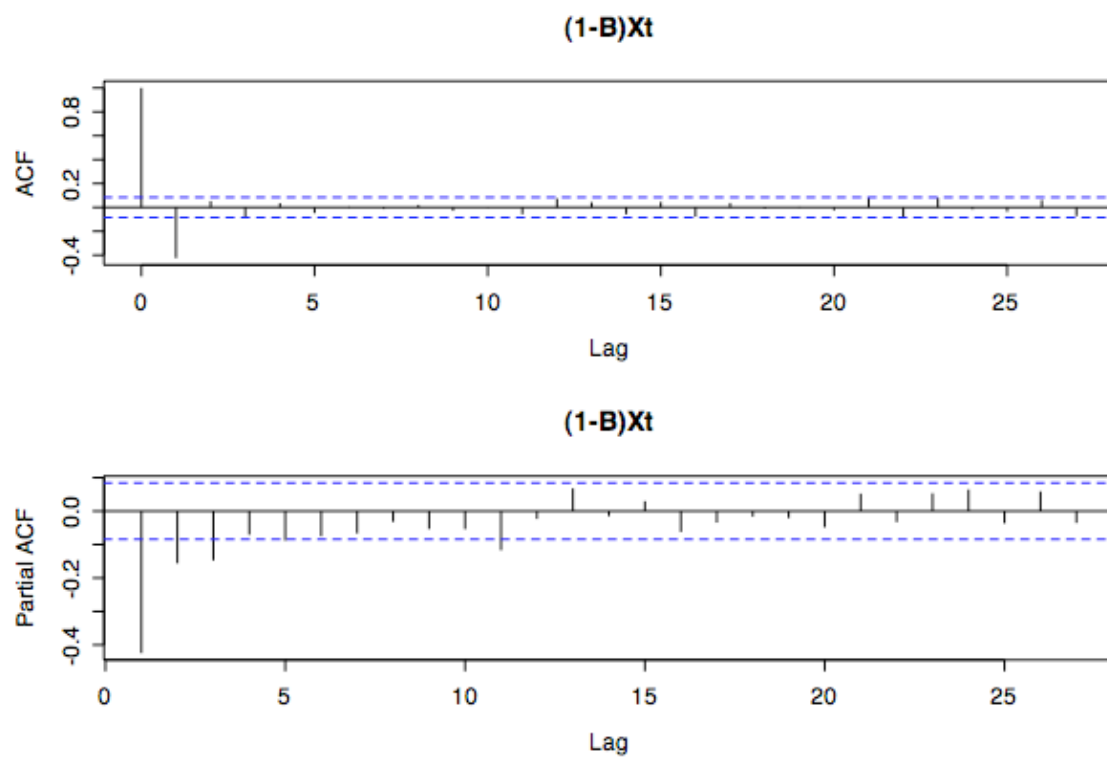
**Figure 16:** Pairs plot for the 30yr swap rate and all eight explanatory variables.



**Figure 17:** Residuals from linear regression using all 8 variables shows strong correlation.



**Figure 18:** ACF and PACF from fit of all eight variables.

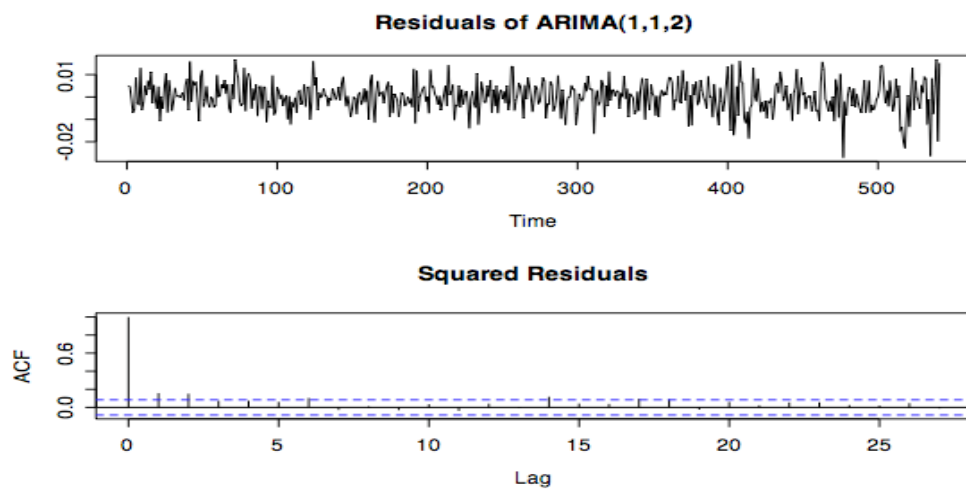


**Figure 19:** ACF and PACF after differencing.

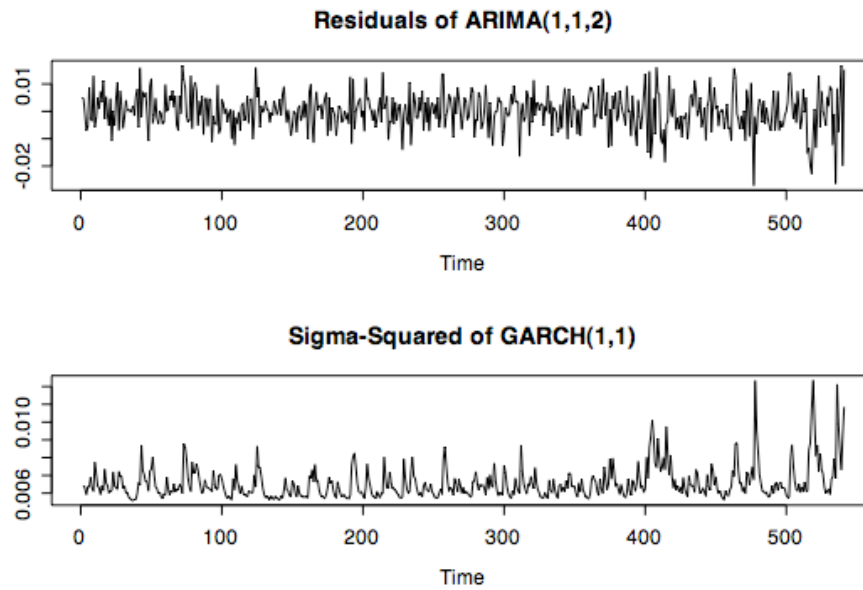


**Table 12:** AIC values from the autoregressive function.

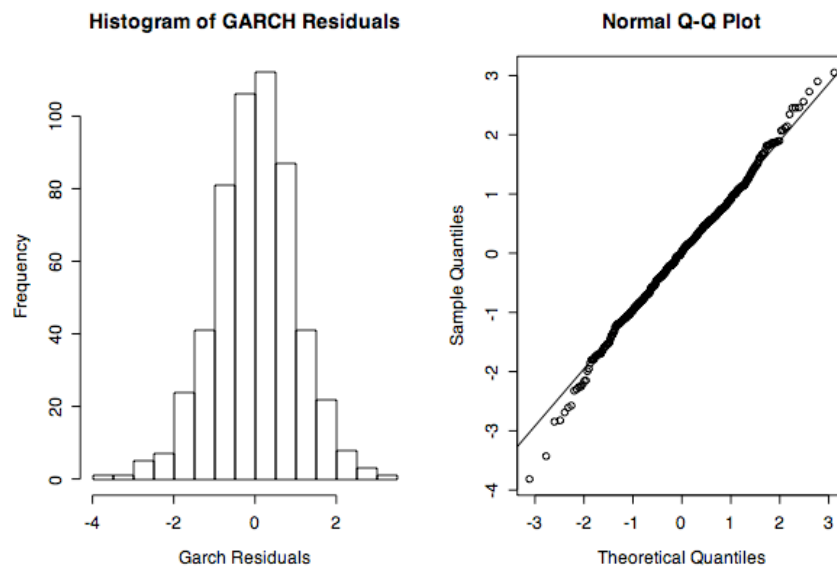
AR	MA	AIC
0	0	-2575.690
0	1	-2714.482
0	2	-2713.942
0	3	-2719.509
0	4	-2718.037
1	0	-2682.471
1	1	-2715.054
1	2	-2721.052
1	3	-2719.458
1	4	-2717.623
2	0	-2694.339
2	1	-2721.833
2	2	-2719.836
2	3	-2718.464
2	4	-2716.493
3	0	-2704.477
3	1	-2719.835
3	2	-2718.231
3	3	-2716.485
3	4	-2714.499
4	0	-2705.297
4	1	-2718.246
4	2	-2716.535
4	3	-2714.557
4	4	-2720.864



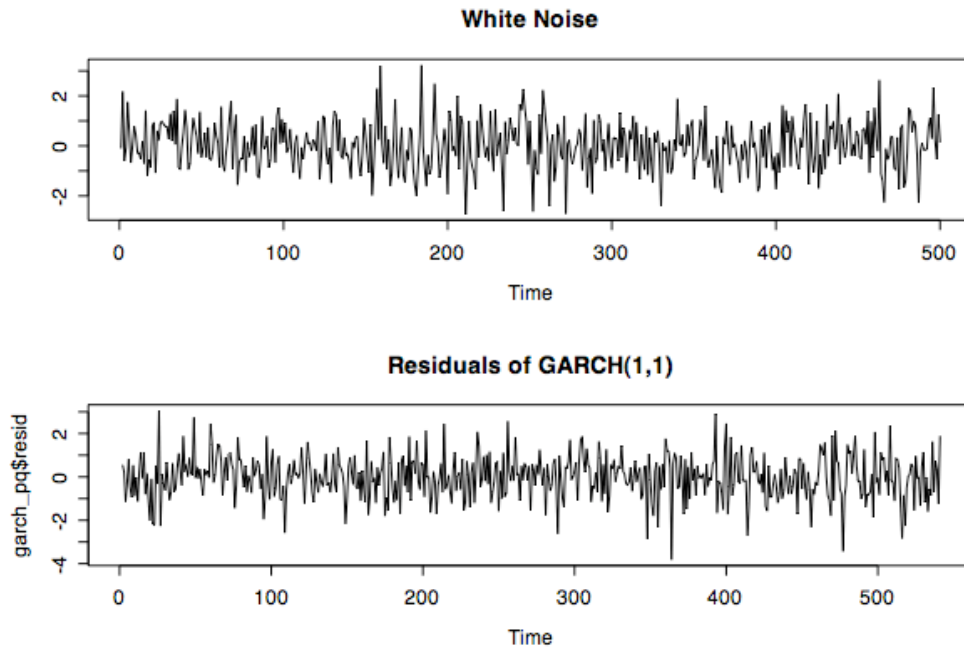
**Figure 20:** Looking for changing variance in the residuals of the ARIMA fit.



**Figure 21:** Matching variance to residuals.



**Figure 22:** Checking for normality under Gaussian likelihood.

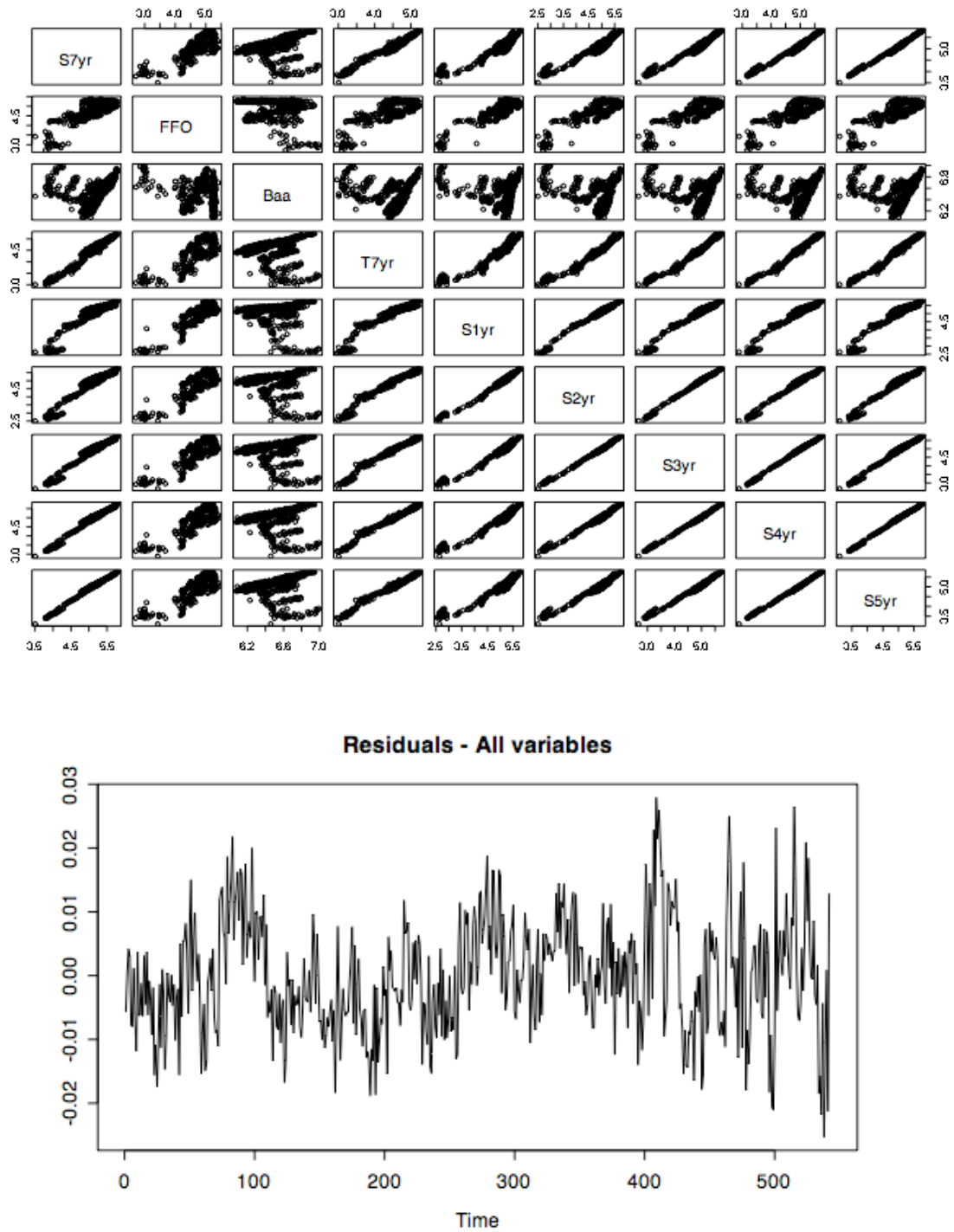


**Figure 23:** Verifying stationary residuals.

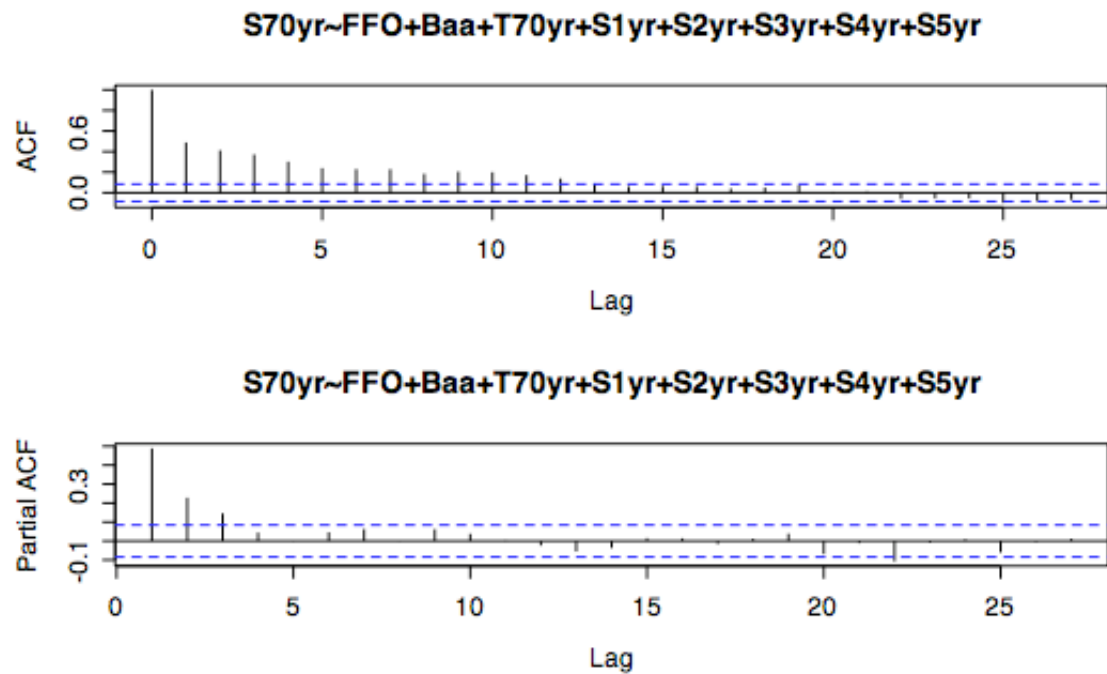
**Table 13:** Diagnostics tests for GARCH model.

Diagnostic Tests			
Jarque Bera Test			
Residuals			
X-squared	9.6953	df = 2, p-value =	0.007847
Box-Ljung Test			
Squared.Residuals			
X-squared	0.0869	df = 1, p-value =	0.7682

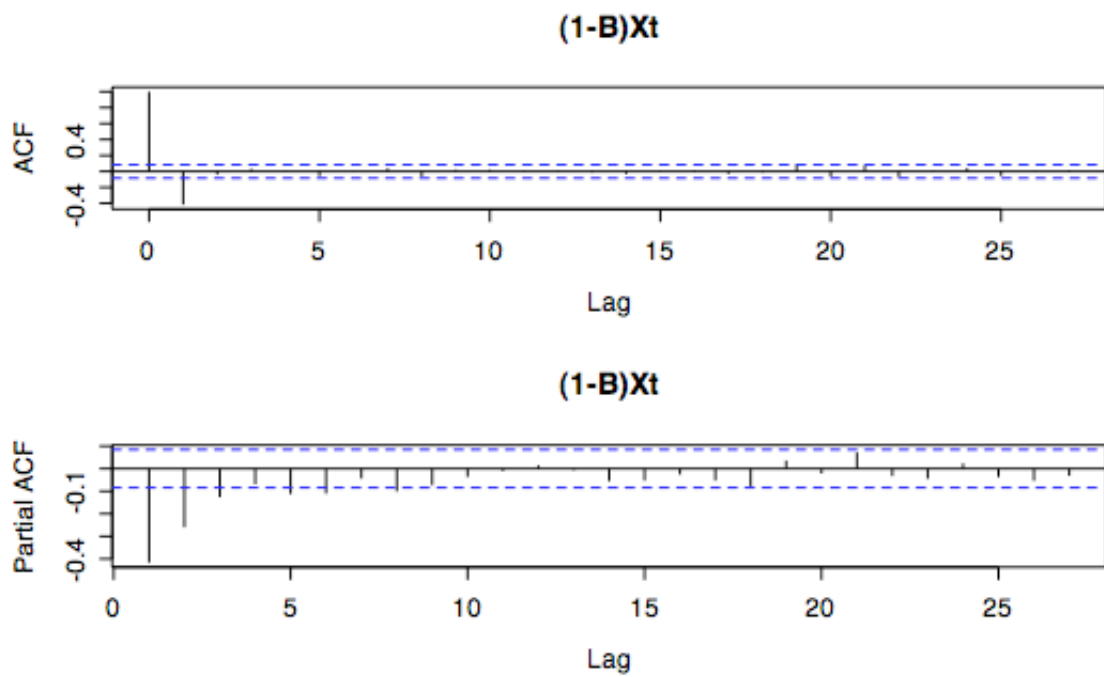
## 7.2. Plots and Tables for 7yr Swap Rates



**Figure 24 (a) & (b):** Pairs plot showing relationship between all variables and their correlated errors after regression.



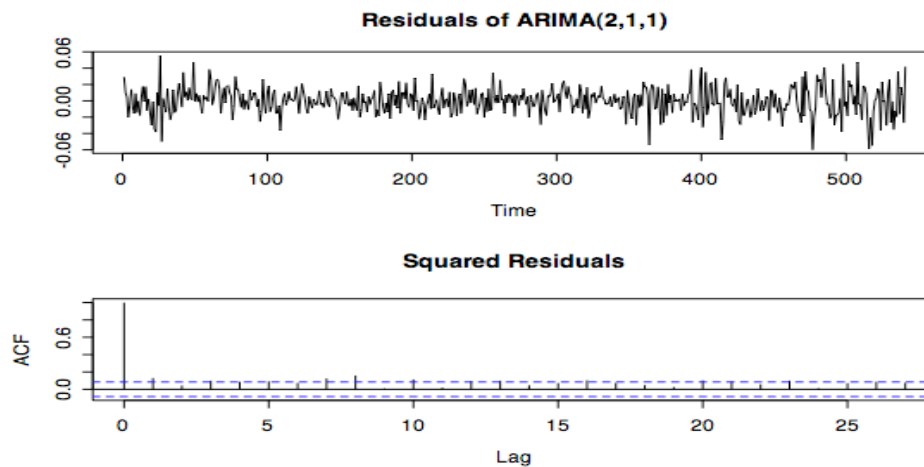
**Figure 25:** ACF and PACF of residuals.



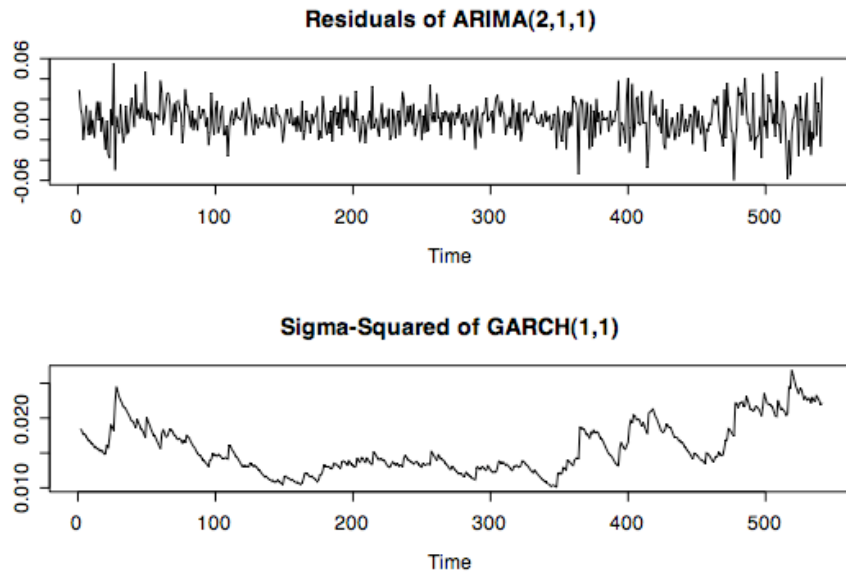
**Figure 26:** ACF and PACF of differenced residuals at lag 1.

**Table 14:** AIC results for all models run.

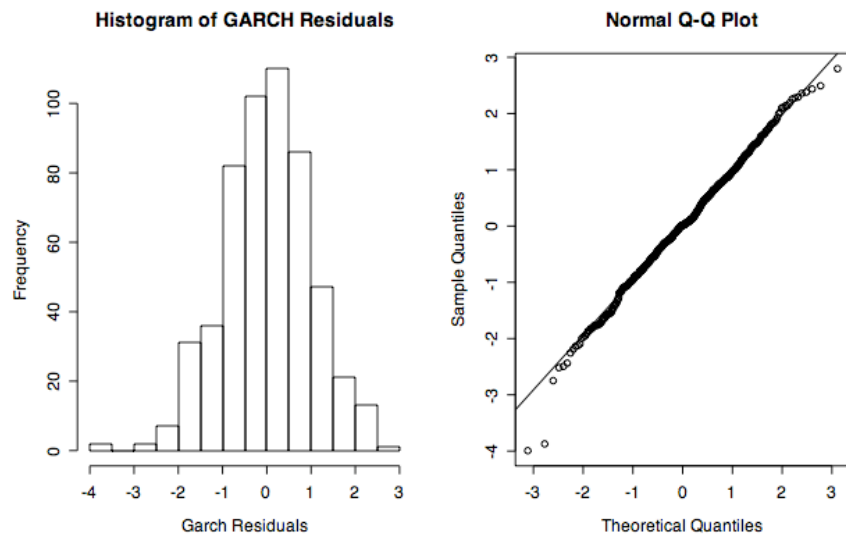
AR	MA	AIC
0	0	-3555.252
0	1	-3725.219
0	2	-3728.288
0	3	-3728.252
0	4	-3731.342
1	0	-3659.501
1	1	-3729.629
1	2	-3745.003
1	3	-3743.153
1	4	-3741.153
2	0	-3696.077
2	1	-3732.036
2	2	-3743.153
2	3	-3741.422
2	4	-3739.306
3	0	-3703.592
3	1	-3741.66
3	2	-3741.165
3	3	-3739.292
3	4	-3738.818
4	0	-3705.075
4	1	-3740.648
4	2	-3738.672
4	3	-3736.134
4	4	-3744.964



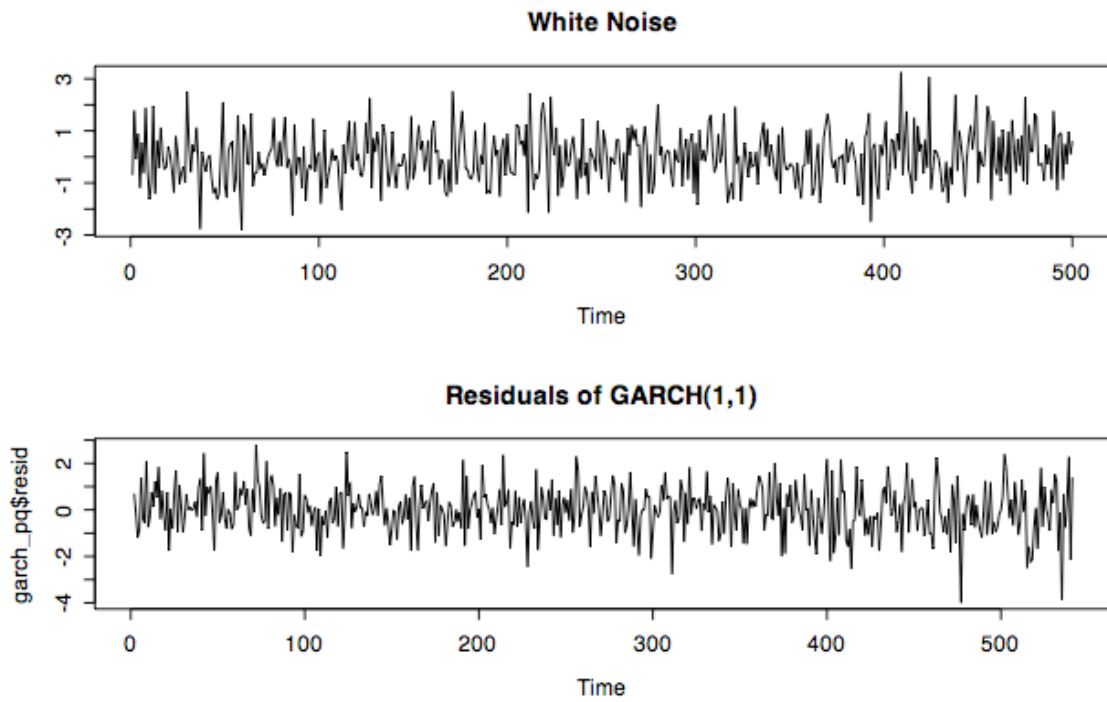
**Figure 27:** Examining ARIMA residuals for changing variance.



**Figure 28:** Looking at estimated volatility from the GARCH model.



**Figure 29:** Checking for normality of the residuals under the Gaussian likelihood.



**Figure 30:** Comparing GARCH residuals to white noise.

**Table 15:** Diagnostics tests for the GARCH model.

Diagnostic Tests			
Jarque Bera Test			
Residuals			
X-squared	7.9647,	df = 2,	p-value = 0.01864
Box-Ljung Test			
Squared.Residuals			
X-squared	0.8859,	df = 1,	p-value = 0.3466



### 7.3. Plots and Tables for 10yr Swap vs. 10yr Treasury

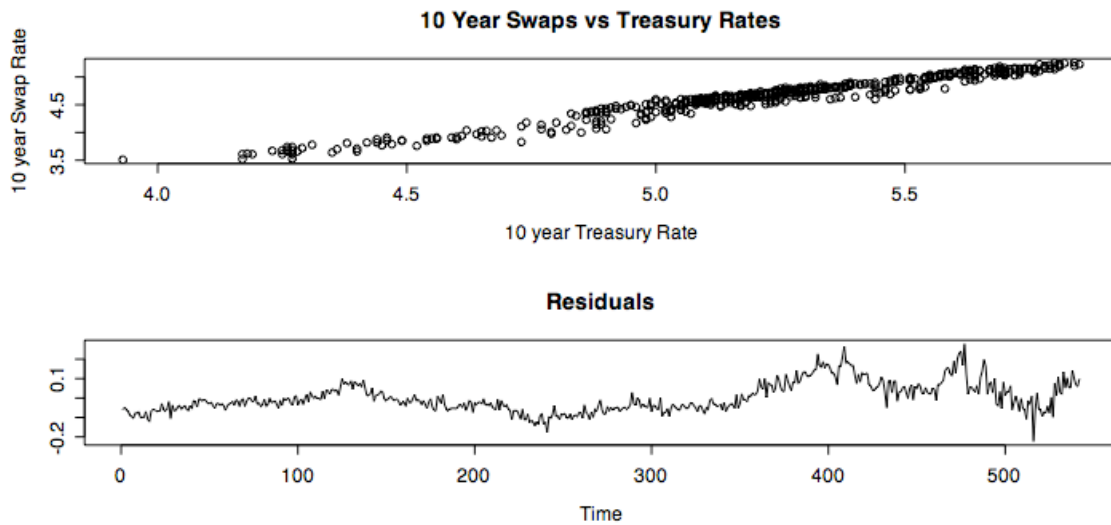


Figure 31 (a) & (b): Linear fit of 10yr swaps and treasuries and their correlated errors.

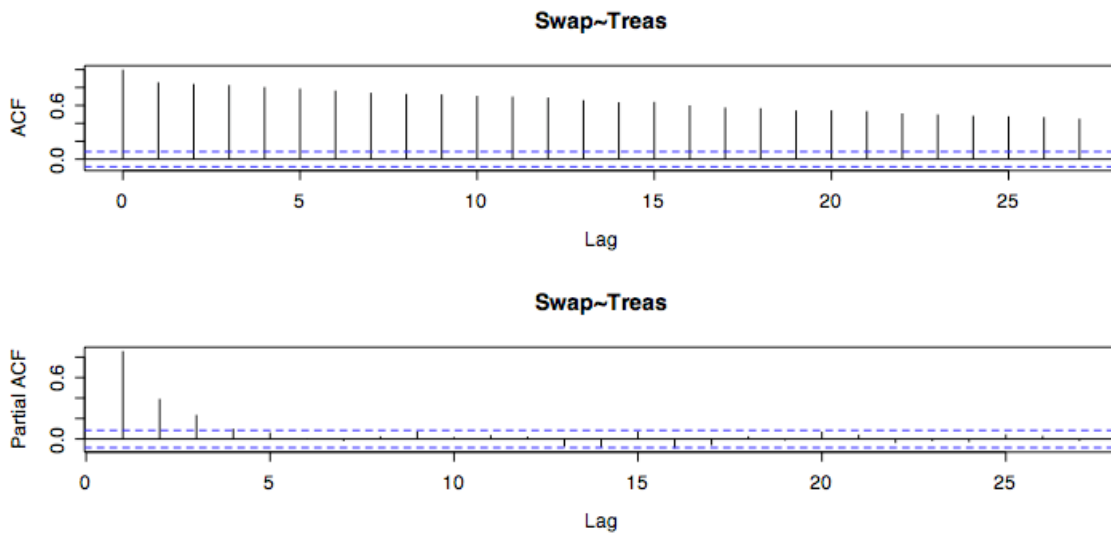
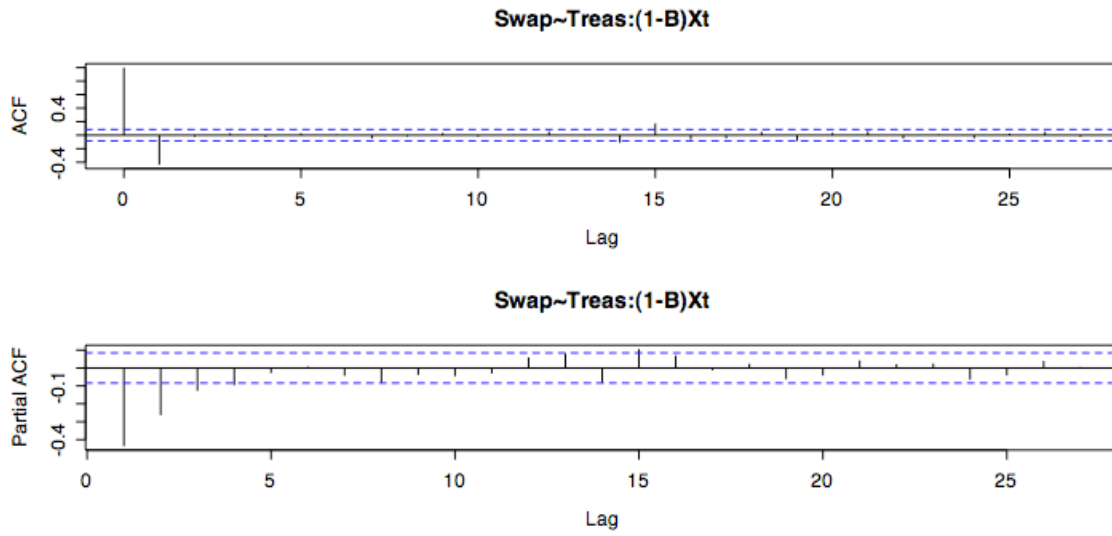


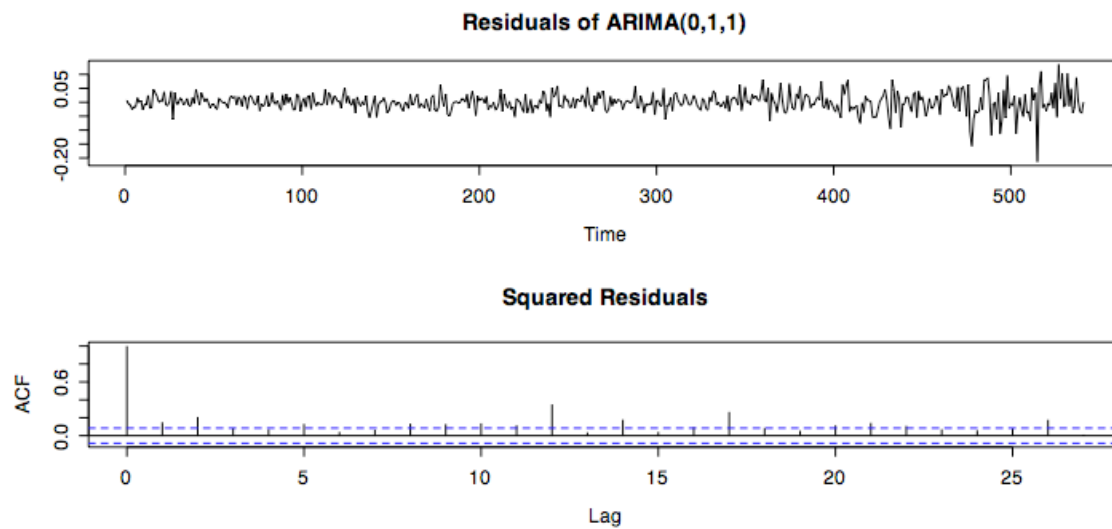
Figure 32: ACF and PACF of the residuals show strong need for differencing.



**Figure 33:** ACF and PACF of the differenced residuals.

**Table 16:** AIC values for all the fitted models.

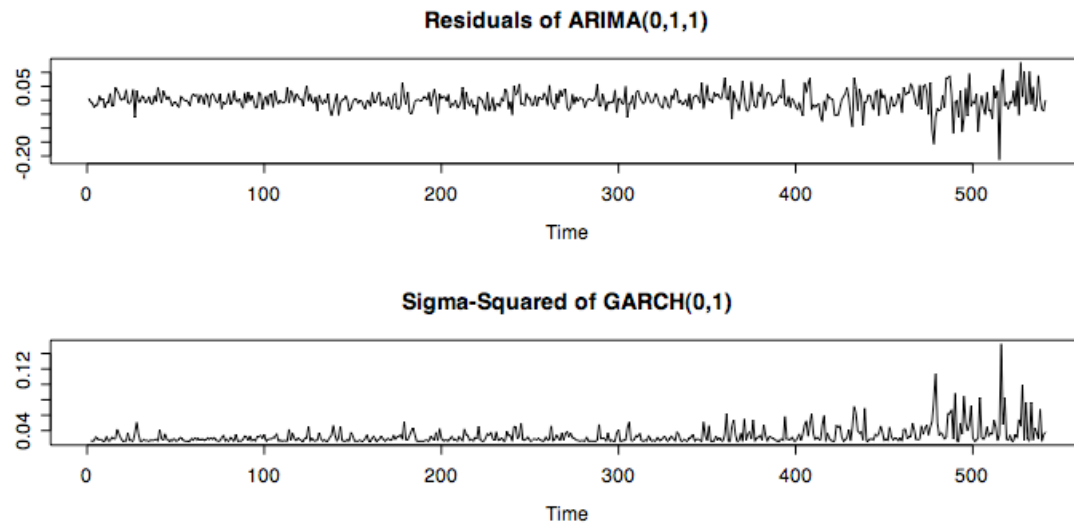
AR	MA	AIC
0	0	-1948.249
0	1	-2114.282
0	2	-2112.303
0	3	-2110.368
0	4	-2108.843
1	0	-2060.771
1	1	-2112.302
1	2	-2110.334
1	3	-2108.702
1	4	-2106.805
2	0	-2097.683
2	1	-2110.371
2	2	-2108.609
2	3	-2106.658
2	4	-2105.170
3	0	-2104.449
3	1	-2108.955
3	2	-2106.874
3	3	-2105.113
3	4	-2103.070
4	0	-2107.345
4	1	-2107.020
4	2	-2104.952
4	3	-2103.161
4	4	-2101.083



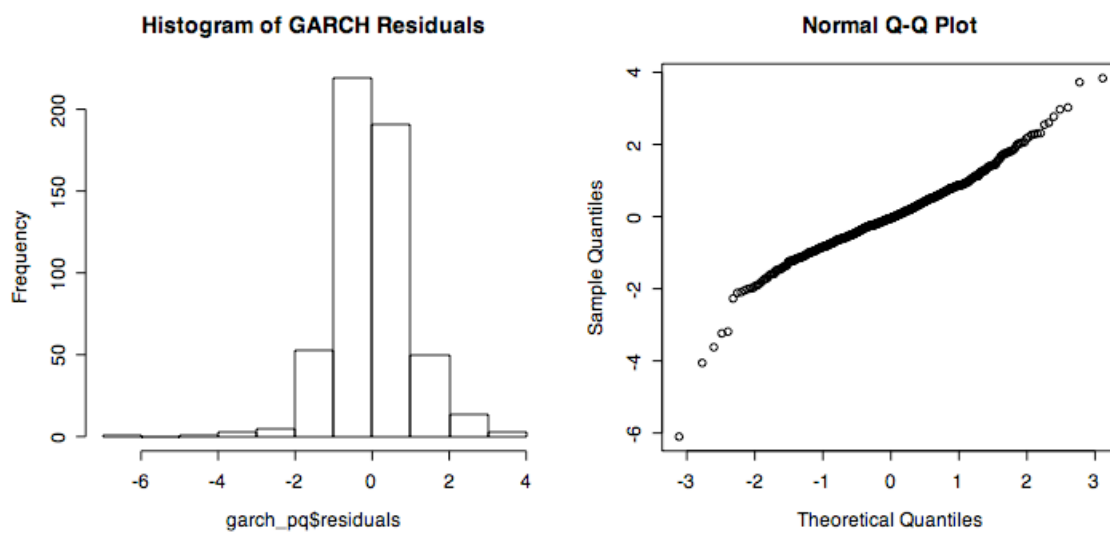
**Figure 34:** Checking ARIMA residuals for changing volatility.

**Table 17:** Diagnostic test for the GARCH model.

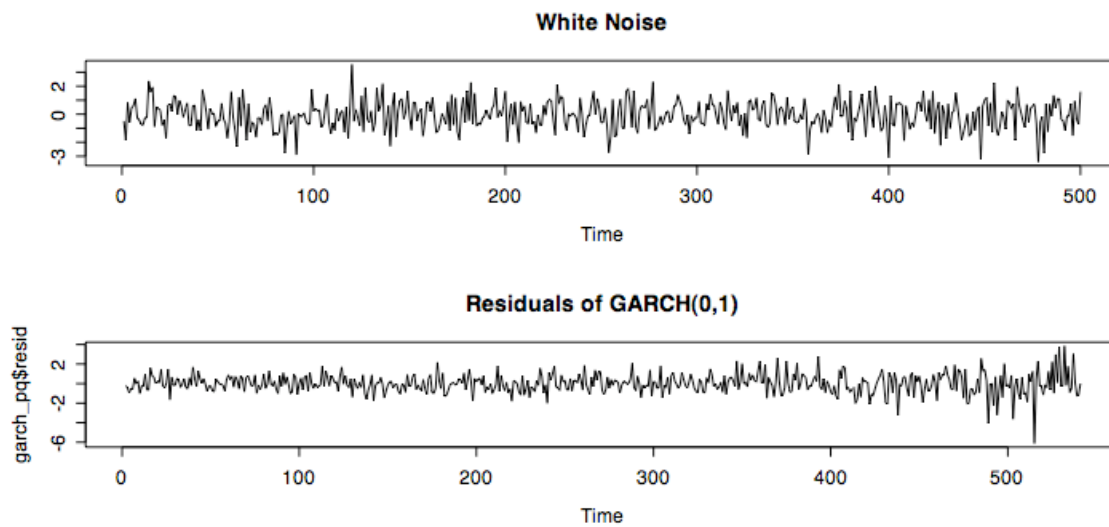
Diagnostic Tests			
Jarque Bera Test			
Residuals			
X-squared	312.7643,	df = 2,	p-value = < 2.2e-16
Box-Ljung Test			
Squared.Residuals			
X-squared	1.0532,	df = 1,	p-value = 0.3048



**Figure 35:** Modeling the estimated variance of the GARCH model.

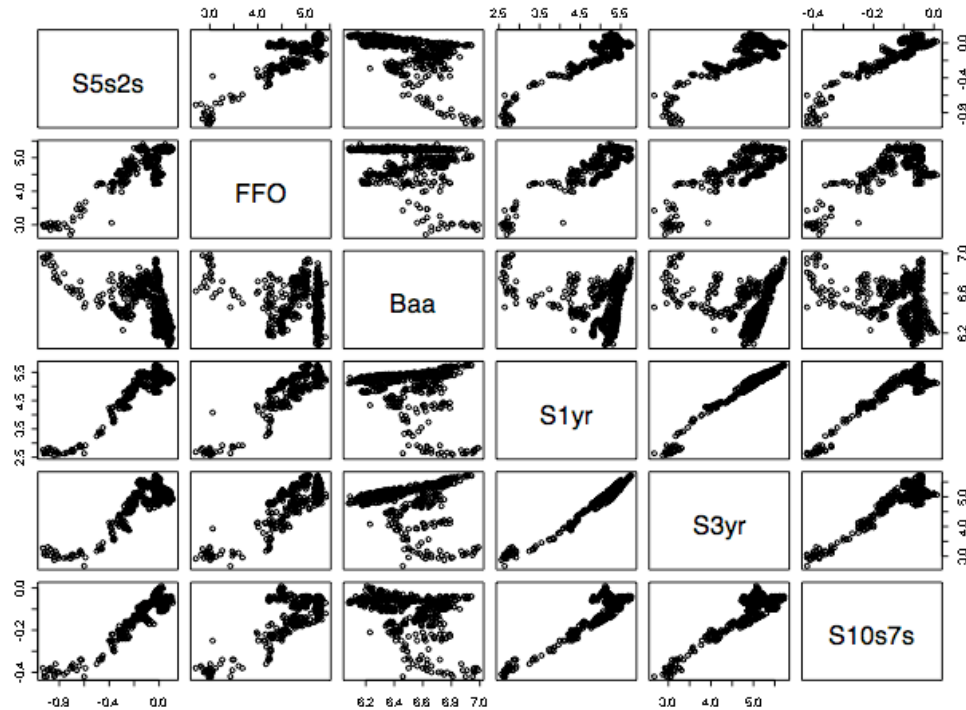


**Figure 35:** Checking for normality of the GARCH residuals.

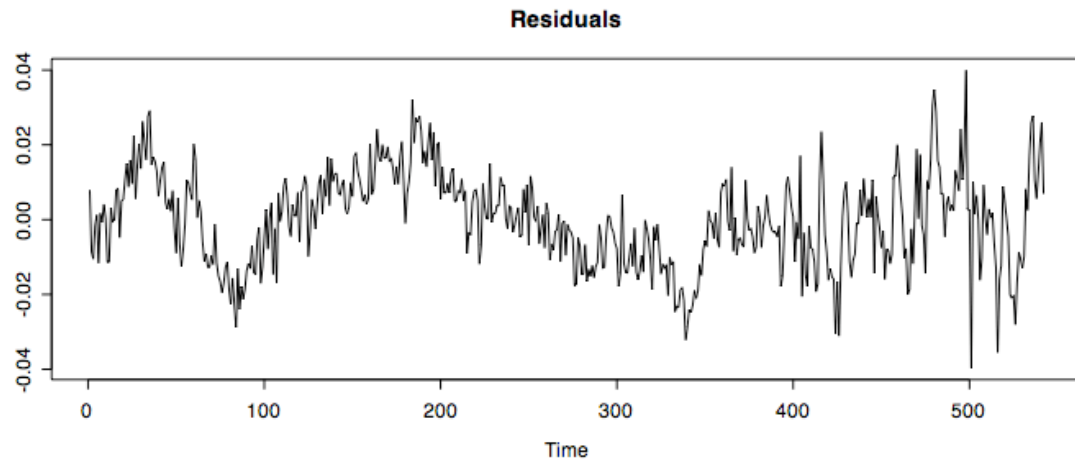


**Figure 36:** Verifying GARCH residuals follow white noise.

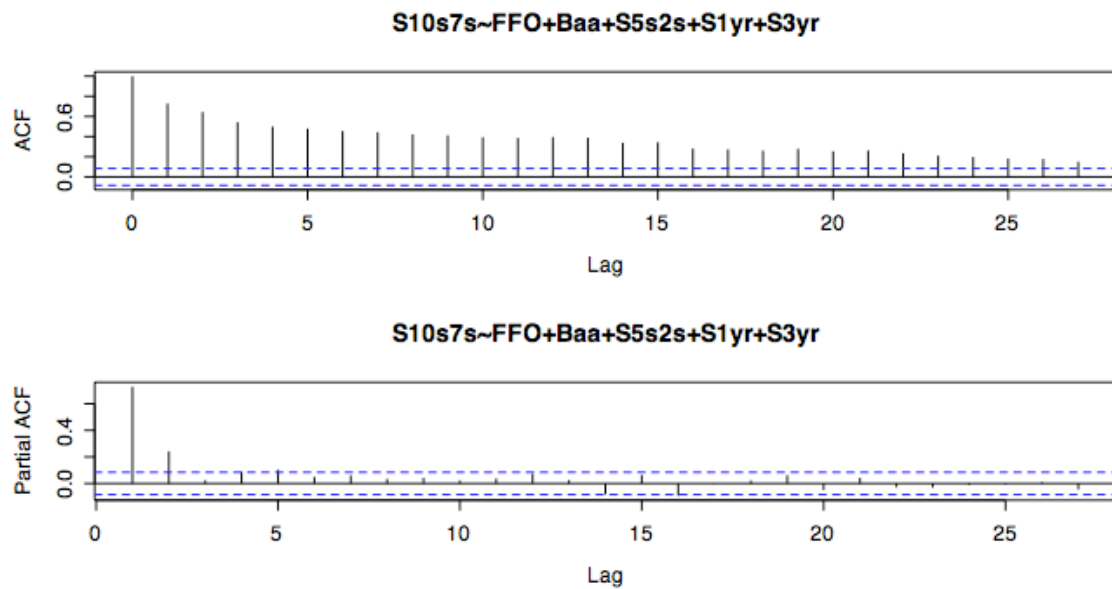
#### 7.4. Plots and Tables for 7yr and 10yr Swap Spread



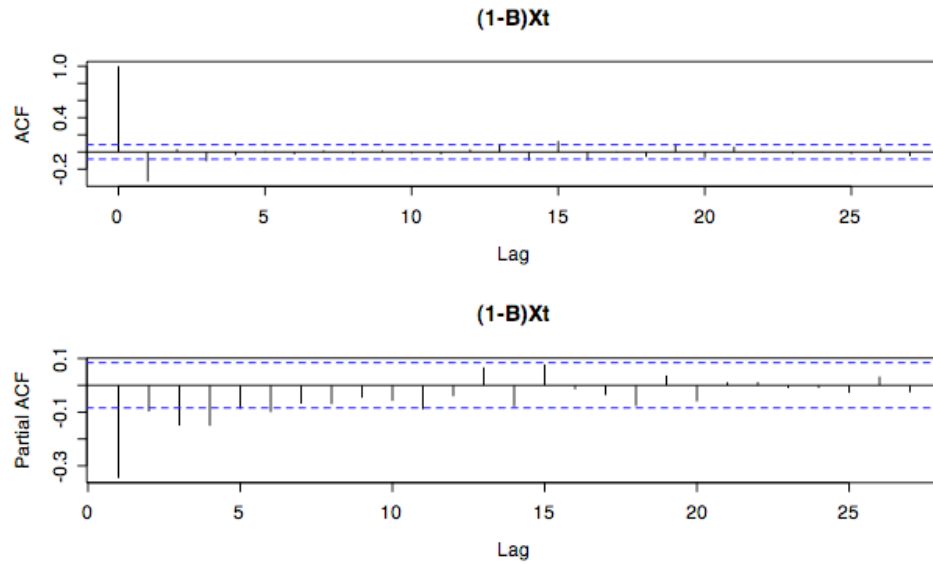
**Figure 37:** Pairs plot of the 10yr and 7yr swap spreads against explanatory variables.



**Figure 38:** Residuals of the linear fit are correlated.



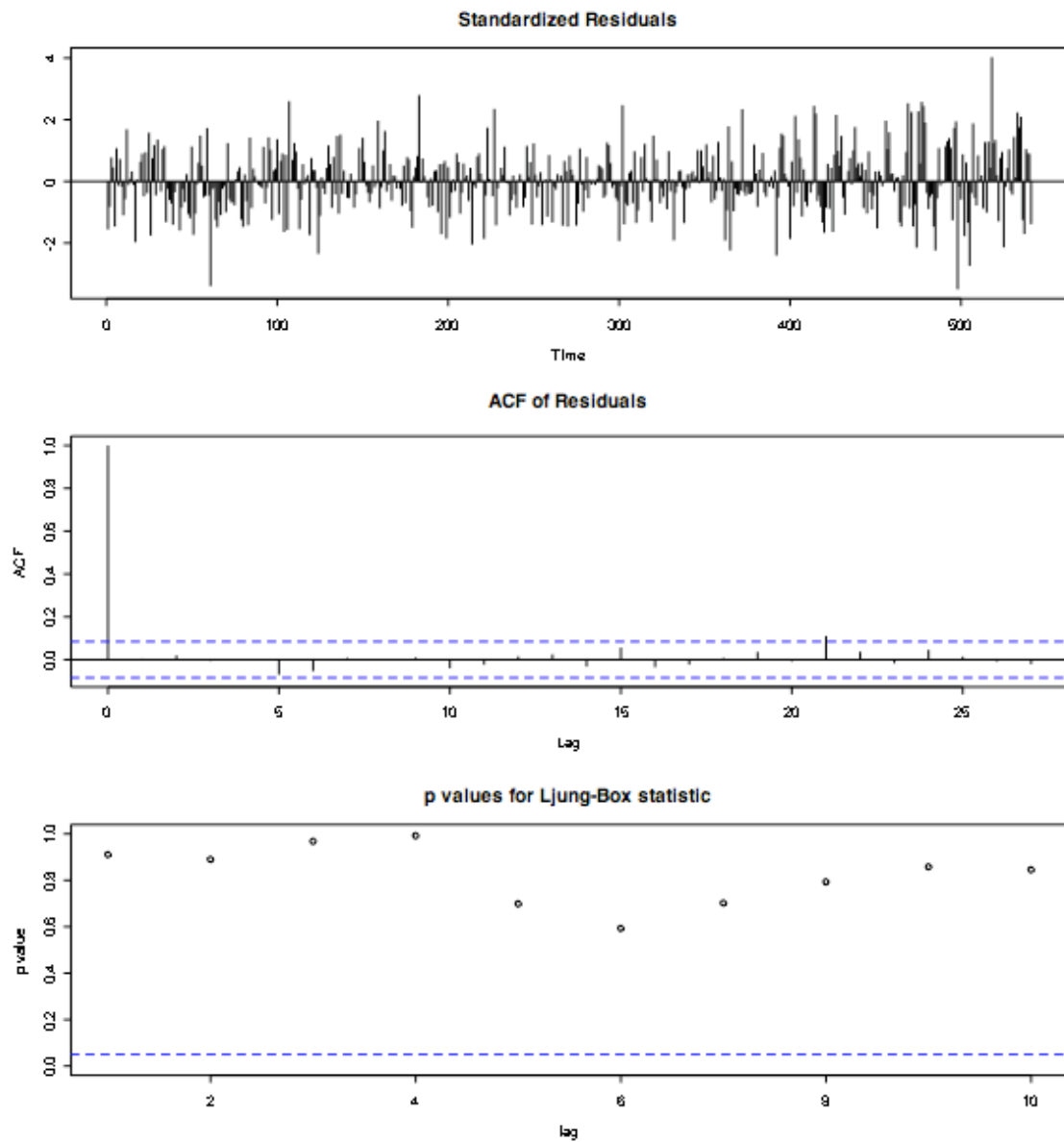
**Figure 39:** ACF and PACF of the residuals.



**Figure 40:** ACF and PACF of the differenced residuals at lag 1.

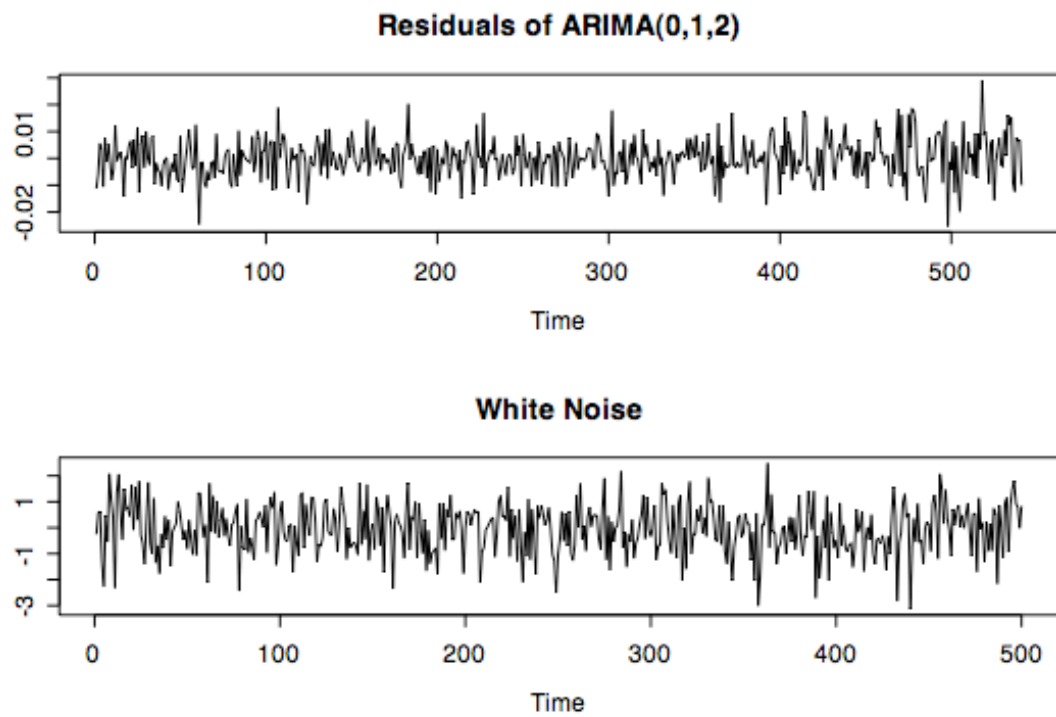
**Table 18:** AIC values for all ARIMA models fit.

AR	MA	AIC
0	0	-3535.235
0	1	-3622.21
0	2	-3629.67
0	3	-3647.498
0	4	-3647.12
1	0	-3603.054
1	1	-3641.144
1	2	-3644.264
1	3	-3647.324
1	4	-3645.34
2	0	-3606.484
2	1	-3646.538
2	2	-3646.203
2	3	-3645.336
2	4	-3643.327
3	0	-3616.662
3	1	-3647.06
3	2	-3645.131
3	3	-3643.403
3	4	-3641.45
4	0	-3627.055
4	1	-3645.184
4	2	-3643.426
4	3	-3645.351
4	4	-3639.611

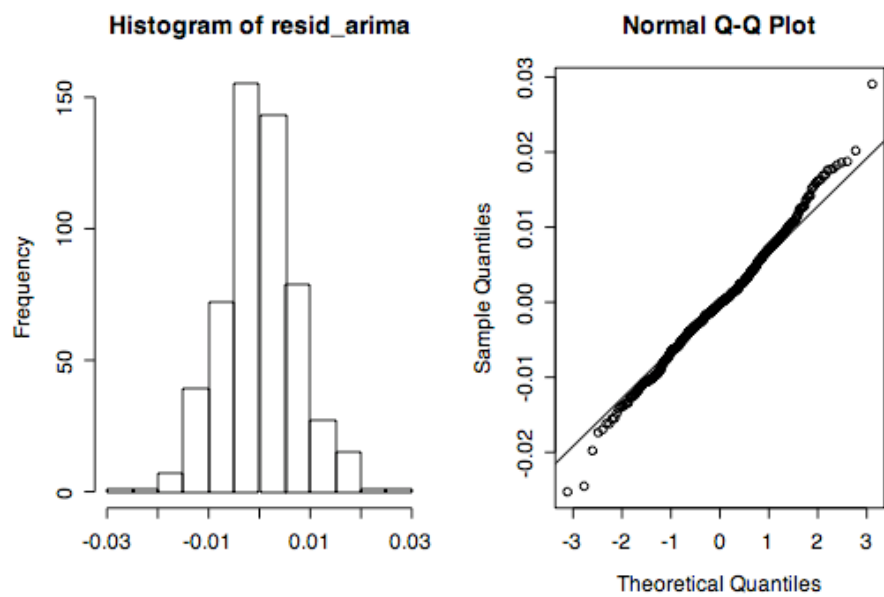


**Figure 41:** Diagnostic plots for the ARIMA residuals.





**Figure 42:** Verifying that the ARIMA residuals compare to White Noise.



**Figure 43:** Verifying the normality of the residuals under the Gaussian likelihood.

**Table 18:** Diagnostic tests for normal and stationary residuals.

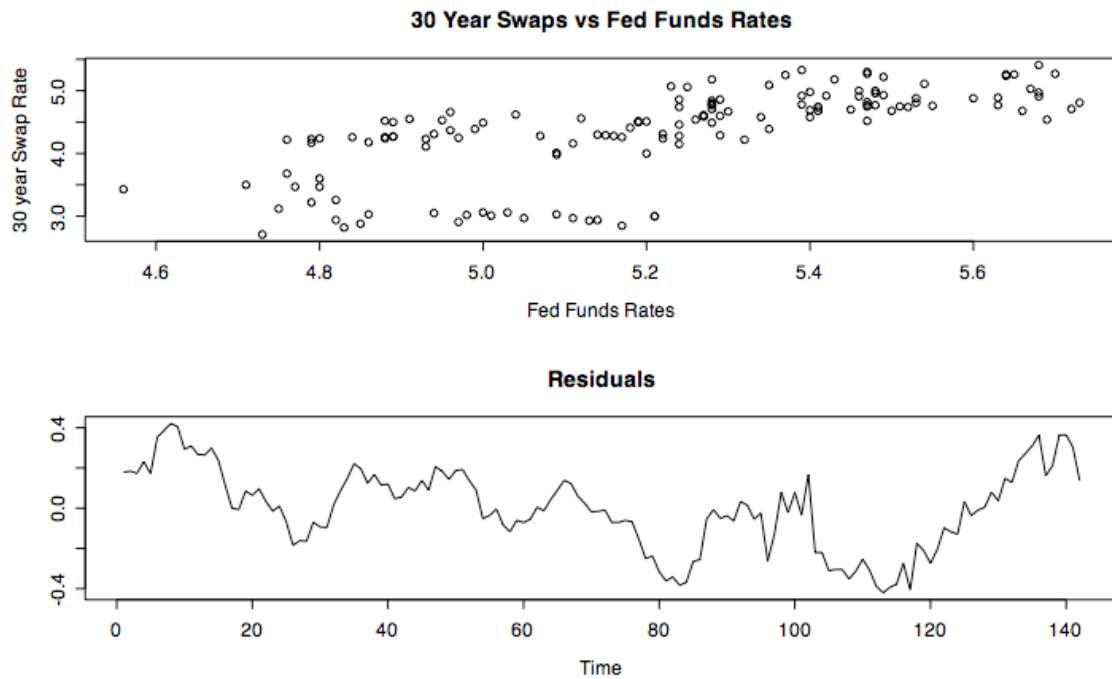
Diagnostics: Shapiro-Wilk Normality Test			
W	=	0.9935,	p-value = 0.01906

Unit Root Diagnostics:									
Augmented	Dickey-Fuller	Test							
Dickey-Fuller	=	-8.3425,	Lag	order	=	8,	p-value	=	0
Phillips-Perron	Unit	Root	Test						
Dickey-Fuller	Z(alpha)	=	-523.4287,	Truncation lag	parameter	=	6,	p-value	= 0.01

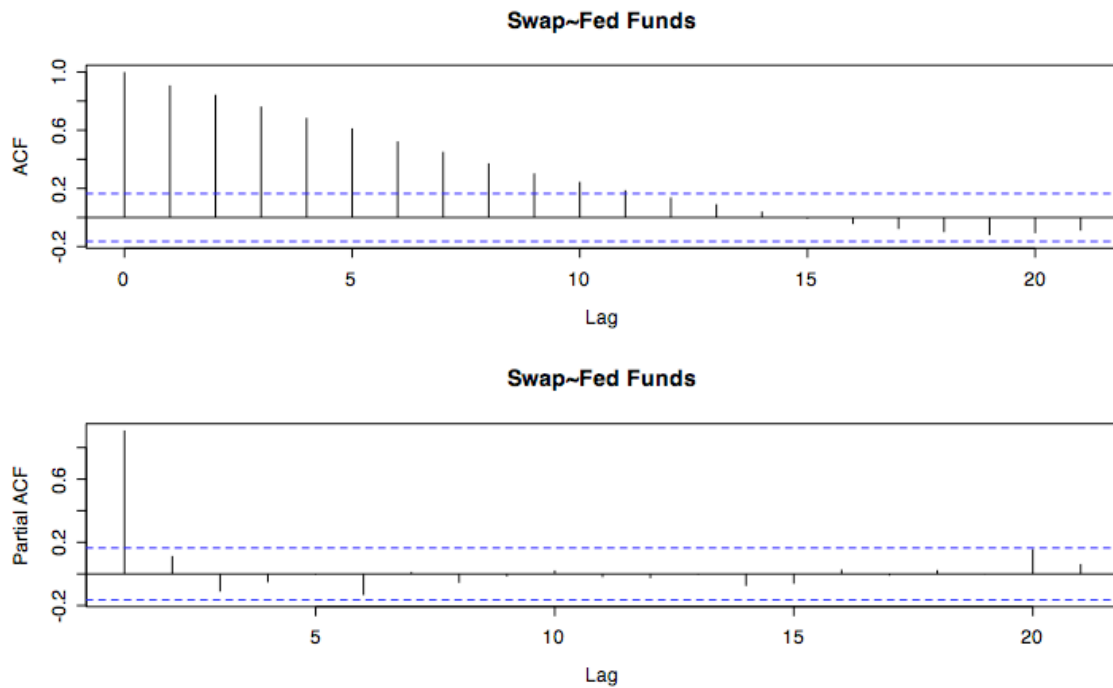
**Table 19:** 95% Confidence Intervals for the forecasted ARIMA values.

	Point Forecast	Lower 95% CI	Upper 95% CI
543	0.008776922	-0.006153207	0.02370705
544	0.010330282	-0.006868144	0.02752871
545	0.010330282	-0.008178876	0.02883944
546	0.010330282	-0.009402736	0.0300633
547	0.010330282	-0.010555002	0.03121557
548	0.010330282	-0.011646937	0.0323075
549	0.010330282	-0.01268713	0.03334769
550	0.010330282	-0.013682305	0.03434287
551	0.010330282	-0.014637846	0.03529841
552	0.010330282	-0.015558142	0.0362187

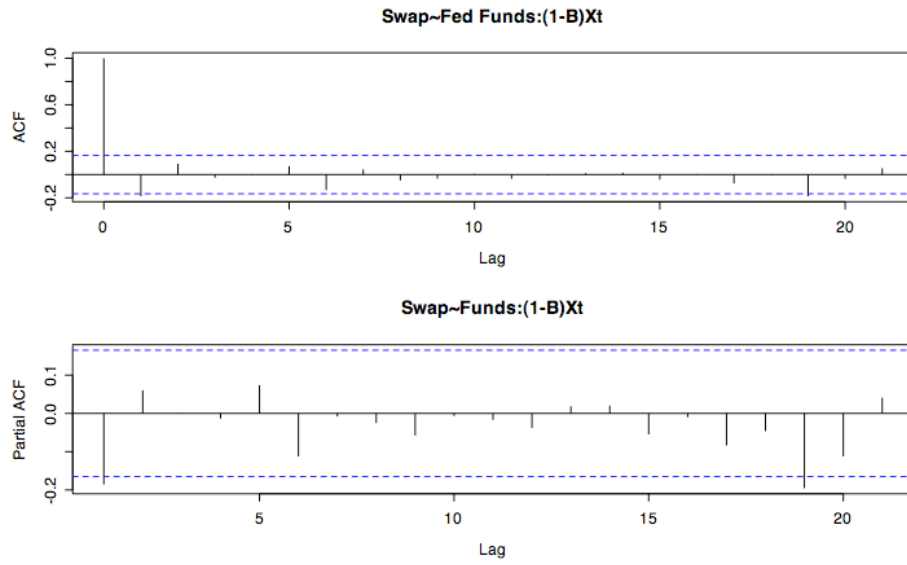
## 7.5. Plots and Tables for 30yr Swap vs Fed Funds



**Figure 44:** Linear fit and the correlated errors.



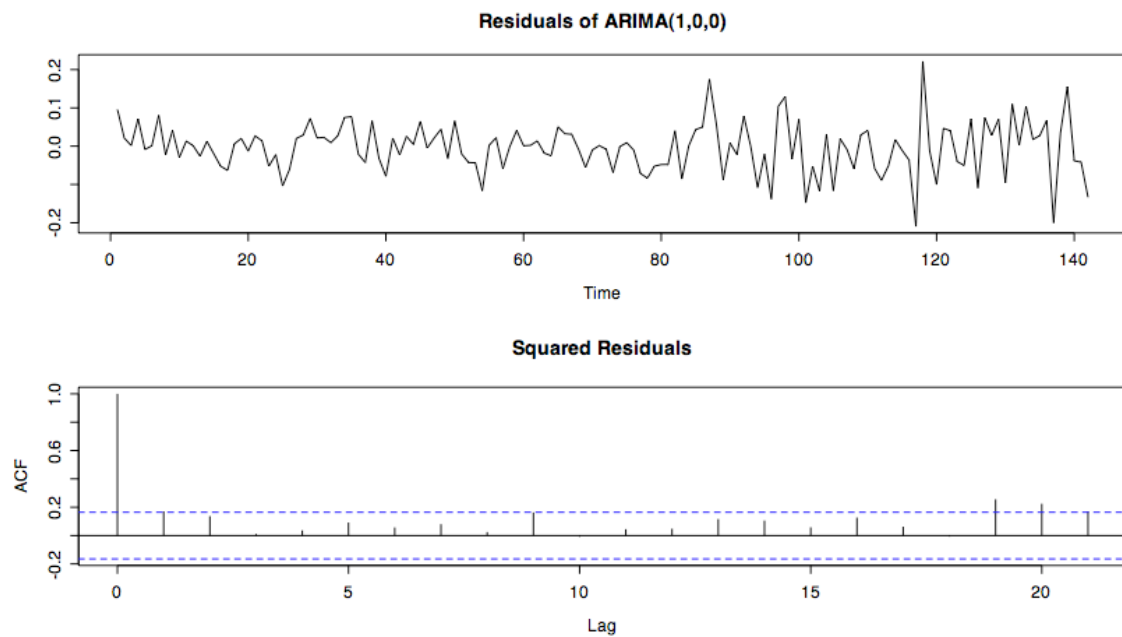
**Figure 45:** ACF and PACF of the residuals.



**Figure 46:** ACF and PACF of the differenced residuals.

**Table 20:** AIC values for the fitted models.

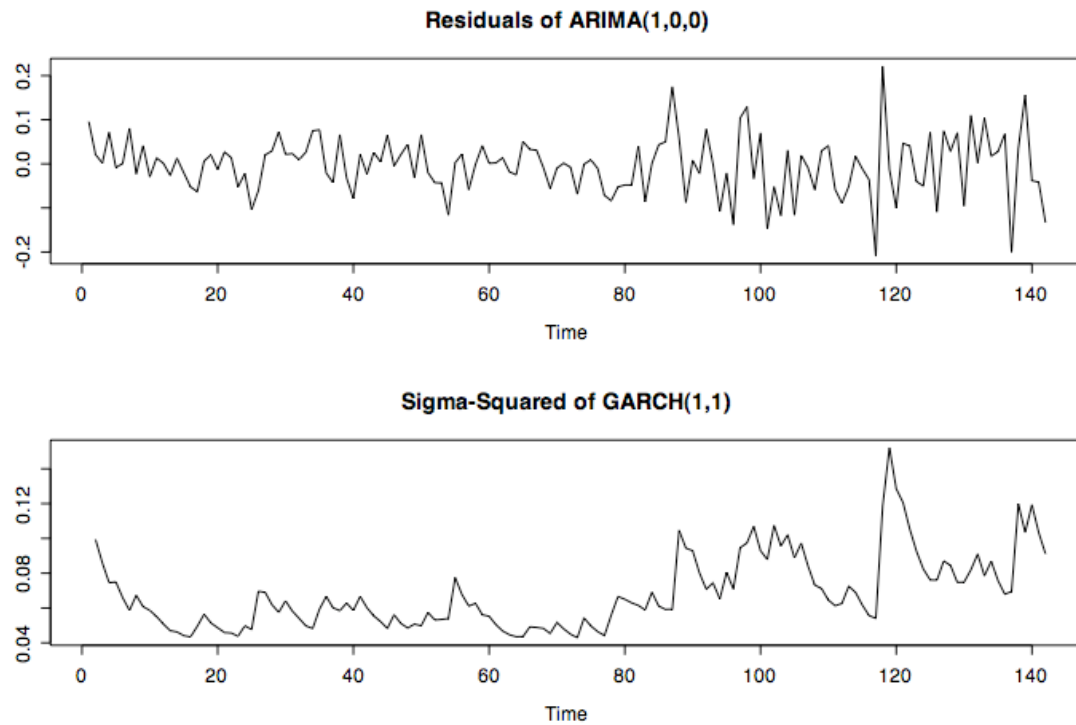
AR	MA	AIC
0	0	-49.40515
0	1	-159.1335
0	2	-199.8775
0	3	-122.9277
0	4	-197.3933
1	0	-298.7101
1	1	-299.2182
1	2	-299.2001
1	3	-283.797
1	4	-286.5323
2	0	-299.866
2	1	-298.597
2	2	-297.5855
2	3	-295.9629
2	4	-285.8597
3	0	-299.0922
3	1	-296.4855
3	2	-295.5393
3	3	-294.0479
3	4	-290.4923
4	0	-297.597
4	1	-296.8858
4	2	-292.9194
4	3	-294.8968
4	4	-297.4241



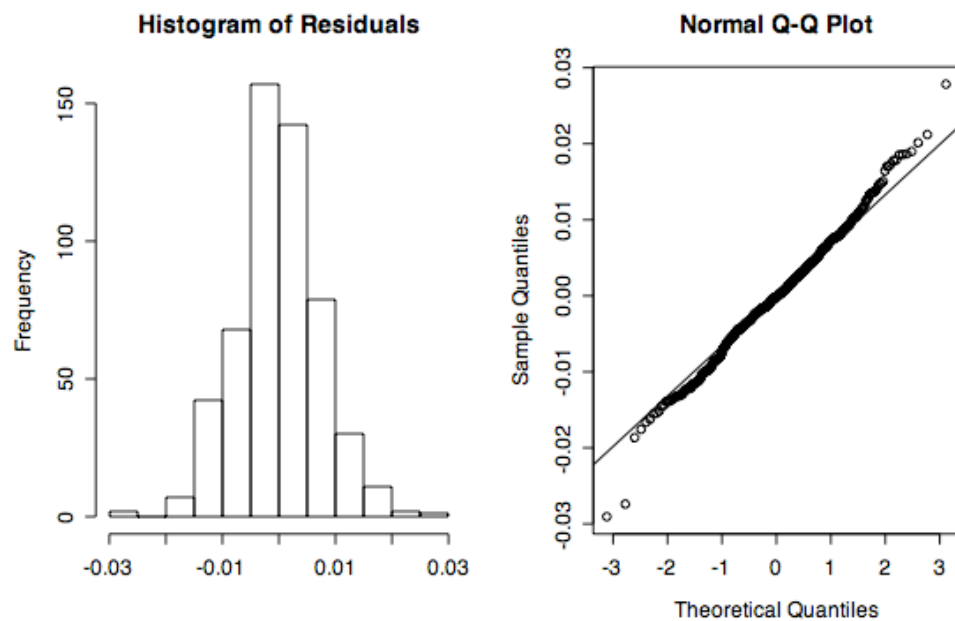
**Figure 47:** Verifying the varying volatility of the residuals

**Table 21:** Diagnostic tests for the GARCH model.

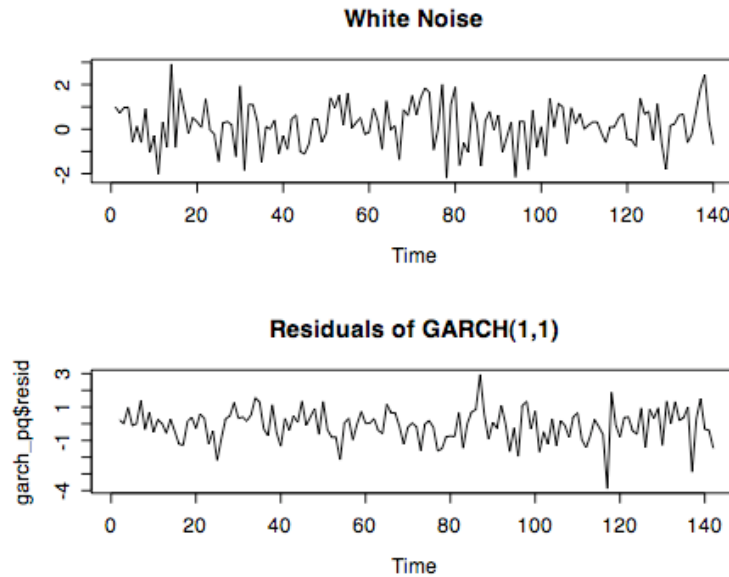
Diagnostic Tests				
Jarque Bera Test				
Residuals				
X-squared	16.0529,	df = 2,	p-value =	0.0003267
Box-Ljung Test				
Squared.Residuals				
X-squared	0.0388,	df = 1,	p-value =	0.844



**Figure 48:** Modeling the estimated variance of the GARCH model.



**Figure 49:** Verifying the normality of the residuals.



**Figure 50:** Showing the residuals resemble white noise.

## 7.6. R Procedures

```
#####
#####TIME SERIES ANALYSIS OF THE SWAPS MARKET#####
```

#Note: Data cleaning was performed in Excel - linear interpolation of missing days due to holidays and Sep.11 was performed. 30 yr treasury data cleaning performed here as noted below.

```
#Required Libaraies:
```

```
library(leaps)
library(zoo)
library(quadprog)
library(tseries)
library(MASS)
library(mgcv)
library(fCalendar)
library(fEcofin)
library(fBasics)
library(fSeries)
library(nlme)
library(stats4)
library(forecast)
```

```
#####Clean data for 30yr#####
```

```
clean=read.table("/Users/reginaflores/Documents/School/Spring_08/AdvancedDataAnalysis/the
sis/data/Factor30yr.txt",header=T)
attach(clean)
```

```
clean30yr=Composite+Factor
as.matrix(clean30yr)
```

```
#####
alldata=read.table("/Users/reginaflores/Documents/School/Spring_08/AdvancedDataAnalysis/t
hesis/data/allData.txt",header=T)
attach(alldata)

subset<-alldata[1104:1645,]
attach(subset)

#Fit all 8 variables
```

### 7.6.1. 10yr Swap vs All 8 Variables

```
#####
#####ERROR FITTING PROCEDURE#####
#####

#STEP 1 - PLOT LINEAR RELATIONSHIP
#plot(Tres10yr, SwapTres10yr,xlab="10 year Treasury Rate",ylab="10 year Swap
Rate",main="10 Year Swaps vs Treasury Rates")
#abline(lm(SwapTres10yr ~ Tres10yr),col=2)
pairs(cbind(S10yr,FF0,Baa,T10yr,S1yr,S2yr,S3yr,S4yr,S5yr))

#STEP 2 - LINEAR REGRESSION/CHECK CORRELATED ERRORS
linreg<-lm(S10yr~FF0+Baa+T10yr+S1yr+S2yr+S3yr+S4yr+S5yr)
linreg_res=resid(linreg)
ts.plot(linreg_res, main="Residuals", ylab="")

#STEP 3 - PLOT ACF/PACF
par(mfrow=c(2,1))
acf(linreg_res, main="S10yr~FF0+Baa+T10yr+S1yr+S2yr+S3yr+S4yr+S5yr")
pacf(linreg_res, main="S10yr~FF0+Baa+T10yr+S1yr+S2yr+S3yr+S4yr+S5yr")

#STEP 4 - DIFF BY LAG 1/CHECK ACF/PACF
res_dif<-diff(linreg_res,1)
par(mfrow=c(2,1))
acf(res_dif, main="(1-B)Xt")
pacf(res_dif, main="(1-B)Xt")

#STEP 5 - RUN AUTOREGRESSIVE FUNCTION MINIMIZING AIC

for (i in 0:4){
  for(j in 0:4){

    AIC=arima(res_dif,c(i,0,j),method="ML")$aic
    print(c(i,j,AIC))
  }
}

AR=1
MA=2

#STEP 6 - FIT MODEL OF RESIDUALS
#check which model gives most stable results
arima_pdq<-arima0(diff(S10yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(Baa),diff(T10yr),diff(S1yr),diff(S2yr),diff(S3yr),diff(S4yr),di
ff(S5yr)),include.mean=TRUE, method="ML")
arima_pdq

#STEP 7 - VALIDATE MODEL
```



```

#may need to adjust order if there are NAN results - refer back to AIC output and select
second lowest AIC
AR=2
MA=1
arima_pdq<-arima0(diff(S10yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(Baa),diff(T10yr),diff(S1yr),diff(S2yr),diff(S3yr),diff(S4yr),di
ff(S5yr)),include.mean=TRUE, method="ML")
arima_pdq

#check that all beta estimates are significant - eliminate those that are not based on
2*SE test
arima_pdq<-arima0(diff(S10yr),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean=TRUE,
method="ML")
arima_pdq

#STEP 8 - RESIDUALS OF ARIMA MODEL - CHECK ACF
resid_arima<-arima0(diff(S10yr),order=c(AR,0,MA),
xreg=cbind(diff(T10yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean=TRUE,
method="ML")$resid
par(mfrow=c(2,1))
acf(resid_arima)
pacf(resid_arima)

#STEP 9 - CHECK CONSTANT VARIANCE
par(mfrow=c(2,1))
ts.plot(resid_arima,main="Residuals of ARIMA(2,0,1)",ylab="")
acf(resid_arima^2,main="Squared Residuals")

#STEP 10 - FIT GARCH MODEL IF VARIANCE NOT CONSTANT
garch_pq<-garch(resid_arima,order=c(1,1))
garch_pq
summary(garch_pq)

#STEP 11 - PLOT VARIANCE
sigma.t<-garch(resid_arima,order=c(1,1))$fitted.values[,2]
par(mfrow=c(2,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,1)", ylab="")
ts.plot(sigma.t,main="Sigma-Squared of GARCH(1,1)",ylab="")

#####DIAGNOSTICS#####
#Checking diagnostics of the ARIMA model
tsdiag(arima_pdq)

#Checking normality - maximizing the Gaussian likelihood
# if the residuals are normal then the L(phi, theta, beta) is a good fit
par(mfrow=c(1,2))
hist(resid_arima)
qqnorm(resid_arima)
qqline(resid_arima)
shapiro.test(resid_arima)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(3,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,3)", ylab="")
ts.plot(rnorm(2000),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

```

```

adf.test(resid_arma)
pp.test(resid_arma)

prod=predict(fit,n.ahead=20)
ts.plot(linreg_res[1600:1645],xlim=c(0,70),ylim=c(0,.5))
abline(v=50,lty="dotted")
lines(prod$pred,col="red",type="o")
u=prod$pred+2*prod$se
d=prod$pred-2*prod$se
lines(u,col="blue",lty="dashed")
lines(d,col="blue",lty="dashed")

###For GARCH models:
###GARCH(0,1) MODEL SATISFIES THE BOX LJUNG TEST

par(mfrow=c(1,2))

hist(garch_pq$residuals,main="Histogram of GARCH Residuals",xlab="Garch Residuals")
qqnorm(garch_pq$residuals)
qqline(garch_pq$residuals)
summary(garch_pq)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(2,1))
ts.plot(rnorm(500),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

#GARCH PREDICTIONS - looking at sumer of 2008
#(PG 287 SHUMWAY)
u=predict(garch(resid_arma,order=c(0,1)))
plot(360:545,resid_arma[360:545],type="l",main="Garch Predictions during Credit Crunch
(Q4 '07 - Q1 '08)",xlab="",ylab="ARIMA Residuals")
lines(u[,1],col="blue",lty="dashed")
lines(u[,2],col="blue",lty="dashed")

```

### 7.6.2. 30yr Swap vs All 8 Variables

```

#####
#####ERROR FITTING PROCEDURE#####
#####

#STEP 1 - PLOT LINEAR RELATIONSHIP
#plot(Tres10yr, SwapTres10yr,xlab="10 year Treasury Rate",ylab="10 year Swap
Rate",main="10 Year Swaps vs Treasury Rates")
#abline(lm(SwapTres10yr ~ Tres10yr),col=2)
pairs(cbind(S30yr,FF0,Baa,T30yr,S1yr,S2yr,S3yr,S4yr,S5yr))

#STEP 2 - LINEAR REGRESSION/CHECK CORRELATED ERRORS
linreg<-lm(S30yr~FF0+Baa+T30yr+S1yr+S2yr+S3yr+S4yr+S5yr)
linreg_res=resid(linreg)
ts.plot(linreg_res, main="Residuals - All variables", ylab="")

#STEP 3 - PLOT ACF/PACF
par(mfrow=c(2,1))
acf(linreg_res, main="S30yr~FF0+Baa+T30yr+S1yr+S2yr+S3yr+S4yr+S5yr")

```

```

pacf(linreg_res, main="S30yr~FF0+Baa+T30yr+S1yr+S2yr+S3yr+S4yr+S5yr")

#STEP 4 - DIFF BY LAG 1/CHECK ACF/PACF
res_dif<-diff(linreg_res,1)
par(mfrow=c(2,1))
acf(res_dif, main="(1-B)Xt")
pacf(res_dif, main="(1-B)Xt")

#STEP 5 - RUN AUTOREGRESSIVE FUNCTION MINIMIZING AIC

for (i in 0:4){
  for(j in 0:4){

    AIC=arima(res_dif,c(i,0,j),method="ML")$aic
    print(c(i,j,AIC))
  }
}

AR=2
MA=1

#STEP 6 - FIT MODEL OF RESIDUALS
#check which model gives most stable results
arima_pdq<-arima0(diff(S30yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(Baa),diff(T30yr),diff(S1yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean=TRUE, method="ML")
arima_pdq

#STEP 7 - VALIDATE MODEL
#may need to adjust order if there are NAN results - refer back to AIC output and select second lowest AIC
AR=2
MA=1
arima_pdq<-arima0(diff(S30yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(Baa),diff(T30yr),diff(S1yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean=TRUE, method="ML")
arima_pdq

#MODEL SELECTION
#check that all beta estimates are significant - eliminate those that are not based on 2*SE test
arima_pdq<-arima0(diff(S30yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(T30yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean=TRUE, method="ML")
arima_pdq

#STEP 8 - RESIDUALS OF ARIMA MODEL - CHECK ACF
resid_arima<-arima0(diff(S30yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(T30yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean=TRUE, method="ML")$resid
par(mfrow=c(2,1))
acf(resid_arima)
pacf(resid_arima)

#STEP 9 - CHECK CONSTANT VARIANCE
par(mfrow=c(2,1))
ts.plot(resid_arima,main="Residuals of ARIMA(2,0,1)",ylab="")
acf(resid_arima^2,main="Squared Residuals")

```

```

#STEP 10 - FIT GARCH MODEL IF VARIANCE NOT CONSTANT
garch_pq<-garch(resid_arima,order=c(1,1))
garch_pq
summary(garch_pq)

#STEP 11 - PLOT VARIANCE
sigma.t<--garch(resid_arima,order=c(1,1))$fitted.values[,2]
par(mfrow=c(2,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,1)", ylab="")
ts.plot(sigma.t,main="Sigma-Squared of GARCH(1,1)",ylab="")

#####DIAGNOSTICS#####
#Checking daiagnostics of the ARIMA model
tsdiag(arima_pdq)

#Checking normality - maximizing the Gaussian liklihood
# if the residuals are normal then the L(phi, theta, beta) is a good fit
par(mfrow=c(1,2))
hist(resid_arima)
qqnorm(resid_arima)
qqline(resid_arima)
shapiro.test(resid_arima)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(3,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,3)", ylab="")
ts.plot(rnorm(2000),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

adf.test(resid_arima)
pp.test(resid_arima)

prod=predict(fit,n.ahead=20)
ts.plot(linreg_res[1600:1645],xlim=c(0,70),ylim=c(0,.5))
abline(v=50,lty="dotted")
lines(prod$pred,col="red",type="o")
u=prod$pred+2*prod$se
d=prod$pred-2*prod$se
lines(u,col="blue",lty="dashed")
lines(d,col="blue",lty="dashed")

###For GARCH models:
###GARCH(0,1) MODEL SATISFIES THE BOX LJUNG TEST

par(mfrow=c(1,2))

hist(garch_pq$residuals,main="Histogram of GARCH Residuals",xlab="Garch Residuals")
qqnorm(garch_pq$residuals)
qqline(garch_pq$residuals)
summary(garch_pq)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(2,1))
ts.plot(rnorm(500),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

#GARCH PREDICTIONS - looking at sumer of 2008

```

```

#(PG 287 SHUMWAY)
u=predict(garch(resid_arma,order=c(0,1)))
plot(360:545,resid_arma[360:545],type="l",main="Garch Predictions during Credit Crunch
(Q4 '07 - Q1 '08)",xlab="",ylab="ARIMA Residuals")
lines(u[,1],col="blue",lty="dashed")
lines(u[,2],col="blue",lty="dashed")

```

### 7.6.3. 7yr Swap vs All 8 Variables

```

#####
#####ERROR FITTING PROCEDURE#####
#####

#STEP 1 - PLOT LINEAR RELATIONSHIP
#plot(Tres10yr, SwapTres10yr,xlab="10 year Treasury Rate",ylab="10 year Swap
Rate",main="10 Year Swaps vs Treasury Rates")
#abline(lm(SwapTres10yr ~ Tres10yr),col=2)
pairs(cbind(S7yr,FF0,Baa,T7yr,S1yr,S2yr,S3yr,S4yr,S5yr))

#STEP 2 - LINEAR REGRESSION/CHECK CORRELATED ERRORS
linreg<-lm(S7yr~FF0+Baa+T7yr+S1yr+S2yr+S3yr+S4yr+S5yr)
linreg_res=resid(linreg)
ts.plot(linreg_res, main="Residuals - All variables", ylab="")

#STEP 3 - PLOT ACF/PACF
par(mfrow=c(2,1))
acf(linreg_res, main="S70yr~FF0+Baa+T70yr+S1yr+S2yr+S3yr+S4yr+S5yr")
pacf(linreg_res, main="S70yr~FF0+Baa+T70yr+S1yr+S2yr+S3yr+S4yr+S5yr")

#STEP 4 - DIFF BY LAG 1/CHECK ACF/PACF
res_dif<-diff(linreg_res,1)
par(mfrow=c(2,1))
acf(res_dif, main="(1-B)Xt")
pacf(res_dif, main="(1-B)Xt")

#STEP 5 - RUN AUTOREGRESSIVE FUNCTION MINIMIZING AIC
for (i in 0:4){
  for(j in 0:4){

    AIC=arima(res_dif,c(i,0,j),method="ML")$aic
    print(c(i,j,AIC))
  }
}

AR=1
MA=2

#STEP 6 - FIT MODEL OF RESIDUALS
#check which model gives most stable results
arima_pdq<-arima0(diff(S7yr),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(Baa),diff(T7yr),diff(S1yr),diff(S2yr),diff(S3yr),diff(S4yr),dif
f(S5yr)),include.mean=TRUE, method="ML")
arima_pdq

#STEP 7 - VALIDATE MODEL

```

```

#may need to adjust order if there are NAN results - refer back to AIC output and select
second lowest AIC
#AR=2
#MA=1
#arima_pdq<-arima0(diff(S7yr),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(T7yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean
=TRUE, method="ML")
#arima_pdq

#MODEL SELECTION
#check that all beta estimates are significant - eliminate those that are not based on
2*SE test
arima_pdq<-arima0(diff(S7yr),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(T7yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean
=TRUE, method="ML")
arima_pdq

#STEP 8 - RESIDUALS OF ARIMA MODEL - CHECK ACF
resid_arima<-arima0(diff(S7yr),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(T7yr),diff(S2yr),diff(S3yr),diff(S4yr),diff(S5yr)),include.mean
=TRUE, method="ML")$resid
par(mfrow=c(2,1))
acf(resid_arima)
pacf(resid_arima)

#STEP 9 - CHECK CONSTANT VARIANCE
par(mfrow=c(2,1))
ts.plot(resid_arima,main="Residuals of ARIMA(1,0,2)",ylab="")
acf(resid_arima^2,main="Squared Residuals")

#STEP 10 - FIT GARCH MODEL IF VARIANCE NOT CONSTANT
garch_pq<-garch(resid_arima,order=c(1,1))
garch_pq
summary(garch_pq)

#STEP 11 - PLOT VARIANCE
sigma.t<-garch(resid_arima,order=c(1,1))$fitted.values[,2]
par(mfrow=c(2,1))
ts.plot(resid_arima, main="Residuals of ARIMA(1,0,2)", ylab="")
ts.plot(sigma.t,main="Sigma-Squared of GARCH(1,1)",ylab="")

#####DIAGNOSTICS#####
#Checking diagnostics of the ARIMA model
tsdiag(arima_pdq)

#Checking normality - maximizing the Gaussian likelihood
# if the residuals are normal then the L(phi, theta, beta) is a good fit
par(mfrow=c(1,2))
hist(resid_arima)
qqnorm(resid_arima)
qqline(resid_arima)
shapiro.test(resid_arima)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(3,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,3)", ylab="")
ts.plot(rnorm(2000),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

```

```

adf.test(resid_arma)
pp.test(resid_arma)

prod=predict(fit,n.ahead=20)
ts.plot(linreg_res[1600:1645],xlim=c(0,70),ylim=c(0,.5))
abline(v=50,lty="dotted")
lines(prod$pred,col="red",type="o")
u=prod$pred+2*prod$se
d=prod$pred-2*prod$se
lines(u,col="blue",lty="dashed")
lines(d,col="blue",lty="dashed")

###For GARCH models:
###GARCH(0,1) MODEL SATISFIES THE BOX LJUNG TEST

par(mfrow=c(1,2))

hist(garch_pq$residuals,main="Histogram of GARCH Residuals",xlab="Garch Residuals")
qqnorm(garch_pq$residuals)
qqline(garch_pq$residuals)
summary(garch_pq)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(2,1))
ts.plot(rnorm(500),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

#GARCH PREDICTIONS - looking at sumer of 2008
#(PG 287 SHUMWAY)
u=predict(garch(resid_arma,order=c(0,1)))
plot(360:545,resid_arma[360:545],type="l",main="Garch Predictions during Credit Crunch
(Q4 '07 - Q1 '08)",xlab="",ylab="ARIMA Residuals")
lines(u[,1],col="blue",lty="dashed")
lines(u[,2],col="blue",lty="dashed")

```

#### 7.6.4. 10yr Swap vs 10yr Treasury

```

data=read.table("/Users/reginaflores/Documents/School/Spring_08/AdvancedDataAnalysis/thesis/data/10yrswaptreas.txt",header=T)
#attach(data)

subset<-data[1104:1645,]
attach(subset)

#####
#####ERROR FITTING PROCEDURE#####
#####

#STEP 1 - PLOT LINEAR RELATIONSHIP
par(mfrow=c(2,1))
plot(SwapTres10yr, Tres10yr,xlab="10 year Treasury Rate",ylab="10 year Swap
Rate",main="10 Year Swaps vs Treasury Rates")

```

```

#abline(lm(S30yr ~ T30yr),col=2)

#STEP 2 - LINEAR REGRESSION/CHECK CORRELATED ERRORS
linreg=lm(SwapTres10yr ~ Tres10yr)
linreg_res=resid(linreg)
ts.plot(linreg_res, main="Residuals", ylab="")

#STEP 3 - PLOT ACF/PACF
par(mfrow=c(2,1))
acf(linreg_res, main="Swap~Treas")
pacf(linreg_res, main="Swap~Treas")

#STEP 4 - DIFF BY LAG 1/CHECK ACF/PACF
res_dif<-diff(linreg_res,1)
par(mfrow=c(2,1))
acf(res_dif, main="Swap~Treas:(1-B)Xt")
pacf(res_dif, main="Swap~Treas:(1-B)Xt")

#STEP 5 - RUN AUTOREGRESSIVE FUNCTION MINIMIZING AIC

for (i in 0:4){
  for(j in 0:4){

    AIC=arima0(res_dif,c(i,0,j),method="ML")$aic
    print(c(i,j,AIC))
  }
}

AR=0
MA=1

#STEP 6 - FIT MODEL OF RESIDUALS
#check model gives most stable results
arima_pdq<-arima0(diff(SwapTres10yr),order=c(AR,0,MA),
xreg=diff(Tres10yr),include.mean=TRUE, method="ML")
arima_pdq

#STEP 7 - VALIDATE MODEL
#check that all estimates are significant - may need to adjust order
#AR
#MA
#arima_pdq<-arima0(diff(SwapTres10yr),order=c(AR,0,MA),
xreg=diff(Tres10yr),include.mean=TRUE, method="ML")
#arima_pdq

#STEP 8 - RESIDUALS OF ARIMA MODEL - CHECK ACF
resid_arima<-arima0(diff(SwapTres10yr),order=c(AR,0,MA),
xreg=diff(Tres10yr),include.mean=TRUE, method="ML")$resid
acf(resid_arima)

#STEP 9 - CHECK CONSTANT VARIANCE
par(mfrow=c(2,1))
ts.plot(resid_arima,main="Residuals of ARIMA(0,1,1)",ylab="")
acf(resid_arima^2,main="Squared Residuals")

#STEP 10 - FIT GARCH MODEL IF VARIANCE NOT CONSTANT
garch_pq<-garch(resid_arima,order=c(0,1))
garch_pq
summary(garch_pq)

```



```

#STEP 11 - PLOT VARIANCE
sigma.t<-garch(resid_arima,order=c(0,1))$fitted.values[,2]
par(mfrow=c(2,1))
ts.plot(resid_arima, main="Residuals of ARIMA(0,1,1)", ylab="")
ts.plot(sigma.t,main="Sigma-Squared of GARCH(0,1)",ylab="")

#####DIAGNOSTICS#####
#Checking daiagnostics of the ARIMA model
tsdiag(arima_pdq)

#Checking normality - maximizing the Gaussian liklihood
# if the residuals are normal then the L(phi, theta, beta) is a good fit
par(mfrow=c(1,2))
hist(resid_arima)
qqnorm(resid_arima)
qqline(resid_arima)
shapiro.test(resid_arima)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(3,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,3)", ylab="")
ts.plot(rnorm(2000),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

adf.test(resid_arima)
pp.test(resid_arima)

prod=predict(fit,n.ahead=20)
ts.plot(linreg_res[1600:1645],xlim=c(0,70),ylim=c(0,.5))
abline(v=50,lty="dotted")
lines(prod$pred,col="red",type="o")
u=prod$pred+2*prod$se
d=prod$pred-2*prod$se
lines(u,col="blue",lty="dashed")
lines(d,col="blue",lty="dashed")

###For GARCH models:
###GARCH(0,1) MODEL SATISFIES THE BOX LJUNG TEST
garch_pq<-garch(resid_arima,order=c(0,1))

par(mfrow=c(1,2))

hist(garch_pq$residuals,main="Histogram of GARCH Residuals")
qqnorm(garch_pq$residuals)
summary(garch_pq)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(2,1))
ts.plot(rnorm(500),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(0,1)")

#GARCH PREDICTIONS (PG 287 SHUMWAY)
u=predict(garch(resid_arima,order=c(0,1)))
#post credit crunch
plot(272:541,resid_arima[272:541],type="l",main="Garch Predictions ",xlab="",ylab="ARIMA
Residuals")

```

```

lines(u[,1],col="blue",lty="dashed")
lines(u[,2],col="blue",lty="dashed")

#pre credit crunch
plot(1:272,resid_arma[1:272],type="l",main="Garch Predictions ",xlab="",ylab="ARIMA
Residuals")
lines(u[,1],col="blue",lty="dashed")
lines(u[,2],col="blue",lty="dashed")

```

### 7.6.5. 7yr and 10yr Swap Spread

```

subset<-alldata[1104:1645,]
attach(subset)

#Define new variable:
#swap spread between 2yr and 5yr, etc.
S5s2s<-S2yr-S5yr
S10s7s<-S7yr-S10yr
#####
#####ERROR FITTING PROCEDURE#####
#####

#STEP 1 - PLOT LINEAR RELATIONSHIP
pairs(cbind(S5s2s,FF0,Baa,S1yr,S3yr,S10s7s))

#STEP 2 - LINEAR REGRESSION/CHECK CORRELATED ERRORS
linreg<-lm(S10s7s~FF0+Baa+S5s2s+S1yr+S3yr)
linreg_res=resid(linreg)
ts.plot(linreg_res, main="Residuals", ylab="")

#STEP 3 - PLOT ACF/PACF
par(mfrow=c(2,1))
acf(linreg_res, main="S10s7s~FF0+Baa+S5s2s+S1yr+S3yr")
pacf(linreg_res, main="S10s7s~FF0+Baa+S5s2s+S1yr+S3yr")

#STEP 4 - DIFF BY LAG 1/CHECK ACF/PACF
res_dif<-diff(linreg_res,1)
par(mfrow=c(2,1))
acf(res_dif, main="(1-B)Xt")
pacf(res_dif, main="(1-B)Xt")

#STEP 5 - RUN AUTOREGRESSIVE FUNCTION MINIMIZING AIC
for (i in 0:4){
  for(j in 0:4){
    AIC=arima(res_dif,c(i,0,j),method="ML")$aic
    print(c(i,j,AIC))
  }
}

AR=0
MA=3

#STEP 6 - FIT MODEL OF RESIDUALS
#check which model gives most stable results
arma_pdq<-arima0(diff(S10s7s),order=c(AR,0,MA),
xreg=cbind(diff(FF0),diff(Baa),diff(S5s2s),diff(S1yr),diff(S3yr)),include.mean=TRUE,
method="ML")

```

```

arima_pdq

#STEP 7 - VALIDATE MODEL
#may need to adjust order if there are NAN results - refer back to AIC output and select
second lowest AIC
AR=0
MA=2
arima_pdq<-arima0(diff(S10s7s),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(S5s2s),diff(S1yr),diff(S3yr)),include.mean=TRUE, method="ML")
arima_pdq

#check that all beta estimates are significant - eliminate those that are not based on
2*SE test
arima_pdq<-arima0(diff(S10s7s),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(S5s2s),diff(S1yr),diff(S3yr)),include.mean=TRUE, method="ML")
arima_pdq

#STEP 8 - RESIDUALS OF ARIMA MODEL - CHECK ACF
resid_arima<-arima_pdq<-arima0(diff(S10s7s),order=c(AR,0,MA),
xreg=cbind(diff(Baa),diff(S5s2s),diff(S1yr),diff(S3yr)),include.mean=TRUE,
method="ML")$resid
par(mfrow=c(2,1))
acf(resid_arima)
pacf(resid_arima)

#STEP 9 - CHECK CONSTANT VARIANCE
par(mfrow=c(2,1))
ts.plot(resid_arima,main="Residuals of ARIMA(0,1,2)",ylab="")
acf(resid_arima^2,main="Squared Residuals")

#####DIAGNOSTICS#####
#Checking diagnostics of the ARIMA model
tsdiag(arima_pdq)

#Checking normality - maximizing the Gaussian likelihood
# if the residuals are normal then the L(phi, theta, beta) is a good fit
par(mfrow=c(1,2))
hist(resid_arima,xlab="")
qqnorm(resid_arima)
qqline(resid_arima)
shapiro.test(resid_arima)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(2,1))
ts.plot(resid_arima, main="Residuals of ARIMA(2,0,3)", ylab="")
ts.plot(rnorm(500),ylab="",main="White Noise")

adf.test(resid_arima)
pp.test(resid_arima)

fit<-lm(S10s7s~ Baa+ S5s2s+ S1yr+ S3yr)$resid
fit2<-arima(fit,order=c(0,1,2))
plot(forecast(fit2,h=10),xlim=c(500,560),ylim=c(-.05,.05),main="10 day Forecast")

```

### 7.6.6. 30yr Swap vs. Fed Funds

```
Subset<-data[1104:1645,]
attach(subset)

#####

par(mfrow=c(2,1))
ts.plot(FF0,ylab="Fed Funds Rate",main="Fed Funds and 30yr Swap Jan '06 - Feb '08")
abline(v=403,lty="dotted",col="red")
ts.plot(S30yr,ylab="30yr Swap Rate")

#CREDIT CRUNCH VOLATILITY
FF0new<-FF0[400:541]
S30yrnew<-S30yr[400:541]

#####
#####ERROR FITTING PROCEDURE#####
#####

#STEP 1 - PLOT LINEAR RELATIONSHIP
par(mfrow=c(2,1))
plot(S30yrnew, FF0new,xlab="Fed Funds Rates",ylab="30 year Swap Rate",main="30 Year Swaps
vs Fed Funds Rates")
#abline(lm(S30yr ~ T30yr),col=2)

#STEP 2 - LINEAR REGRESSION/CHECK CORRELATED ERRORS
linreg=lm(S30yrnew ~ FF0new)
linreg_res=resid(linreg)
ts.plot(linreg_res, main="Residuals", ylab="")

#STEP 3 - PLOT ACF/PACF
par(mfrow=c(2,1))
acf(linreg_res, main="Swap~Fed Funds")
pacf(linreg_res, main="Swap~Fed Funds")

#STEP 4 - DIFF BY LAG 1/CHECK ACF/PACF
res_dif<-diff(linreg_res,1)
par(mfrow=c(2,1))
acf(res_dif, main="Swap~Fed Funds:(1-B)Xt")
pacf(res_dif, main="Swap~Funds:(1-B)Xt")

#STEP 5 - RUN AUTOREGRESSIVE FUNCTION MINIMIZING AIC
for (i in 0:4){
  for(j in 0:4){
    AIC=arima0(res_dif,c(i,0,j),method="ML")$aic
    print(c(i,j,AIC))
  }
}

AR=2
MA=0

#STEP 6 - FIT MODEL OF RESIDUALS
#check model gives most stable results
arima_pdq<-arima0(S30yrnew,order=c(AR,0,MA), xreg=FF0new,include.mean=TRUE, method="ML")
```

```

arima_pdq

#STEP 7 - VALIDATE MODEL
#check that all estimates are significant - may need to adjust order
AR=1
MA=0
arima_pdq<-arima0(S30yrnew,order=c(AR,0,MA), xreg=FF0new,include.mean=TRUE, method="ML")
arima_pdq

#STEP 8 - RESIDUALS OF ARIMA MODEL - CHECK ACF
resid_arima<-arima0(S30yrnew,order=c(1,0,0), xreg=FF0new,include.mean=TRUE,
method="ML")$resid
acf(resid_arima)

#STEP 9 - CHECK CONSTANT VARIANCE
par(mfrow=c(2,1))
ts.plot(resid_arima,main="Residuals of ARIMA(1,0,0)",ylab="")
acf(resid_arima^2,main="Squared Residuals")

#STEP 10 - FIT GARCH MODEL IF VARIANCE NOT CONSTANT
garch_pq<-garch(resid_arima,order=c(1,1))
garch_pq
summary(garch_pq)

#STEP 11 - PLOT VARIANCE
sigma.t<--garch(resid_arima,order=c(1,1))$fitted.values[,2]
par(mfrow=c(2,1))
ts.plot(resid_arima, main="Residuals of ARIMA(1,0,0)", ylab="")
ts.plot(sigma.t,main="Sigma-Squared of GARCH(1,1)",ylab="")

#####DIAGNOSTICS#####
###For GARCH models:
###GARCH(1,1) MODEL SATISFIES THE BOX LJUNG TEST

par(mfrow=c(1,2))
hist(garch_pq$residuals,main="Histogram of GARCH Residuals")
qqnorm(garch_pq$residuals)
summary(garch_pq)

#Checking that the residuals of fitted model are ~ WN(0,sigma^2)
par(mfrow=c(2,1))
ts.plot(rnorm(140),ylab="",main="White Noise")
ts.plot(garch_pq$resid,main="Residuals of GARCH(1,1)")

#GARCH PREDICTIONS (PG 287 SHUMWAY)
u=predict(garch(resid_arima,order=c(0,1)))
#credit crunch
plot(resid_arima,type="l",main="Garch Predictions in Volatile Market
Conditions",xlab="",ylab="ARIMA Residuals")
lines(u[,1],col="blue",lty="dashed")
lines(u[,2],col="blue",lty="dashed")

```

## 8. References

- Aas, Kjersti and Xenia K. Dimakos. Statistical Modeling of Financial Time Series: An Introduction. Norwegian Computing Center Applied Research and Development: Note SMABA/08/04, March 2004.
- Bernadell, Carlos et. al. Yield Curve Prediction for the Strategic Investor. European Central Bank Working Paper Series: No. 472, April 2005.
- Bishop, Paul and Don Dixon. Foreign Exchange Handbook: Managing Risk and Opportunity in Global Currency Markets. New York: McGraw Hill, 1992.
- Brockwell, Peter J. and Richard A. Davis. Introduction to Time Series and Forecasting. New York: Springer, 2002.
- Corb, Howard. Columbia Business School Readings: Fixed Income Derivatives. Course Number: B8399-017. Spring, 2008.
- Crawford, Gordon W. and Micheal C. Fratantoni. Assessing the Forecasting Performance of Regime-Switching, ARIMA and GARCH Models of House Prices. Real Estate Economics: Vol. 31, pg 223-243.
- Dattatreya, Ravi E., Raj E. S. Venkatesh, and Vijaya E. Venkatesh. Interest Rate and Currency Swaps: The Markets, Products and Applications. Chicago: Probus Publishing Company, 1994.
- Estrella, Arturo and May R. Trubin. The Yield Curve as a Leading Indicator: Some Practical Issues. Federal Reserve Bank of New York Current Issues: Vol. 12, No. 5, July 2006.
- Granger, C.W.J and P. Newbold. Spurious Regressions in Econometrics. Journal of Econometrics: Volume 2, Issue 2, July 1974, Pages 111-120.
- Hoeting, Jennifer A. and Richard A. Davis, et. al. Model Selection for Geostatistical Models. Ecological Society of America Online Journal: Vol. 16, Sep. 2004.
- Lauschke, Jens and Lee Boon Keng. Parsimonious Models for the USD Swap Curve. DBS Group Research: DBS Bank Ltd, Interest Rate Strategy, Aug. 11, 2006.

- Lekkos, Ilias, et. al. *Forecasting Interest Rate Spreads Using Domestic and International Risk Factors: Evidence from Linear and Non-Linear Models*. Journal of Forecasting: Vol.26, pg. 601-619.
- Mills, Terence C. *The Econometric Modelling of Financial Time Series*. New York: Cambridge University Press, 1999.
- Schwager, Jack D. *Technical Analysis*. Toronto: John Wiley & Sons, 1996.
- Shumway, Robert H. and David S. Stoffer. *Time Series and Its Applications*. New York: Springer, 2006.

