Final report

Analysis of Amazon Stock, eBay Stock and Best Buy Stock

Instructor: Xinyu Song
TA: Jianchang Hu
Name:Qiuying Li
Date:14/12/2015

Overview:

Amazon and eBay are the two of the most famous online shopping companies, which have significant development in the past 10 years. Best Buy, as a well represented in store electronics corporation, normally has more sales during the holiday seasons. Assume I have one million dollars and I want to purchase a stock among these 3 stocks, I built time-series model, which is normal GARCH with distribution for each stock and made forecasting. Eventually I found out that Best Buy is the most risky one, so I choose eBay instead because it is the least risky one.

Introduction:

Amazon, EBay and the Best Buy are the 3 famous online shopping companies. In order to know more details about the development of these companies, studying and analyzing their stocks is one of the effective ways. So I collected the stock data from 2005-11-30 to 2015-11-30, and built proper models based on the data. In the rest of the paper, I would describe specifically how I develop models of Amazon stock, and then I would apply similar method to build models for EBay and Best Buy.

A. Analysis of the Amazon Stock

1. Amazon stock price

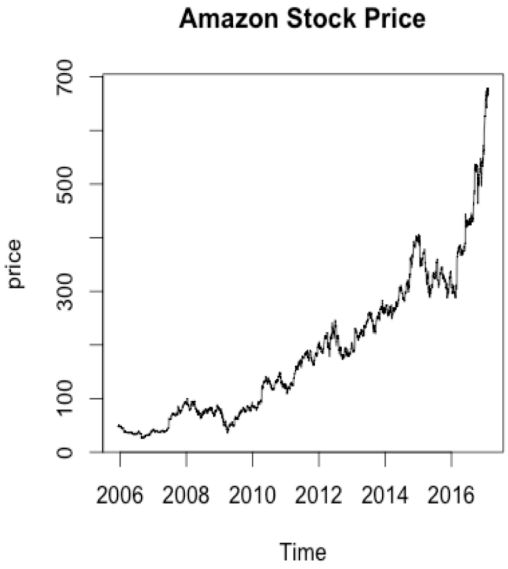


Figure 1

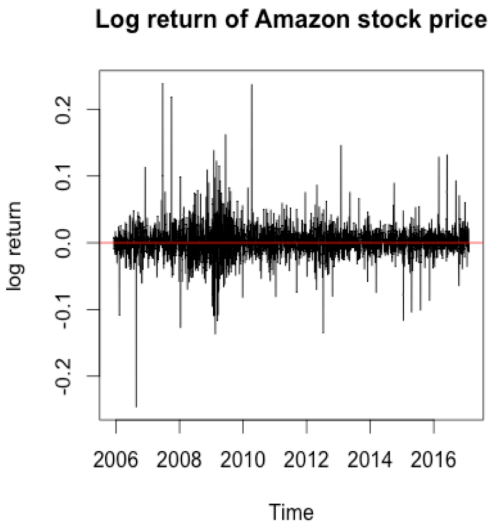


Figure 2

Figure 1 shows the general trend of the Amazon stock price. Basically, the stock price continues increasing, except the two period of time: 2009 and the first half year of 2015.

Figure 2 shows the log return of the Amazon stock price. The results of first difference of the log stock price seems stationary, but further unit root test was needed.

2. Test for Serial Correlation and Arch Effect of the time series.

Test for Constant Term	One Sample t-test data: log.rtn t = -1.9819, df = 2516, p-value = 0.0476 sample estimates: mean of x -0.001038268	P-value is slightly smaller than a=0.05, so we could still reject the null hypothesis, and conclude the mean is not equal to 0. In other word, we need a constant.
Serial Correlation	Box-Ljung test data: log.rtn X-squared = 39.8281, df = 20, p-value = 0.005251	P-value=0.05,which indicates serial correlation might exist in the data.
Arch Effect	Box-Ljung test data: res_rtn^2 X-squared = 84.4231, df = 20, p-value = 6.882e-10	P-value is close the 0, so we can reject the null hypothesis, and concludes there is an arch effect in the data.

Conclusion: Due to the both serial correlation and Arch Effect in the Amazon log return, so we need a Garch model to deal with the Arch Effect, and the Garch model has to contain ARMA model, which could deal with serial correlation.

3. Generate ARIMA(p,d,q) Model

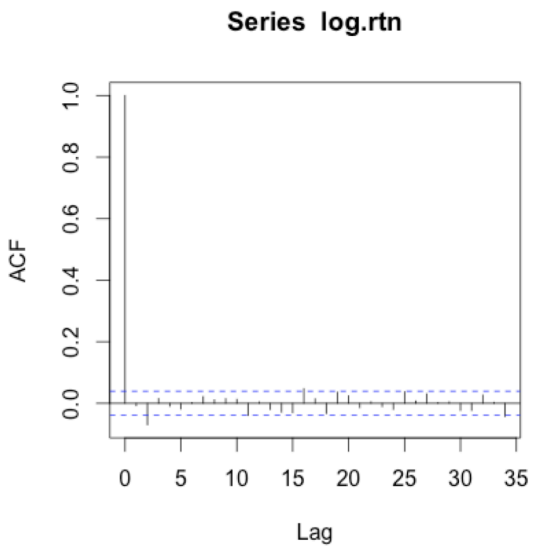


Figure 3

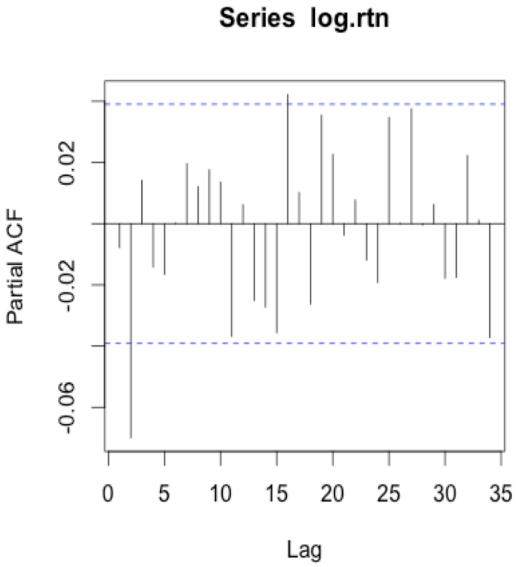


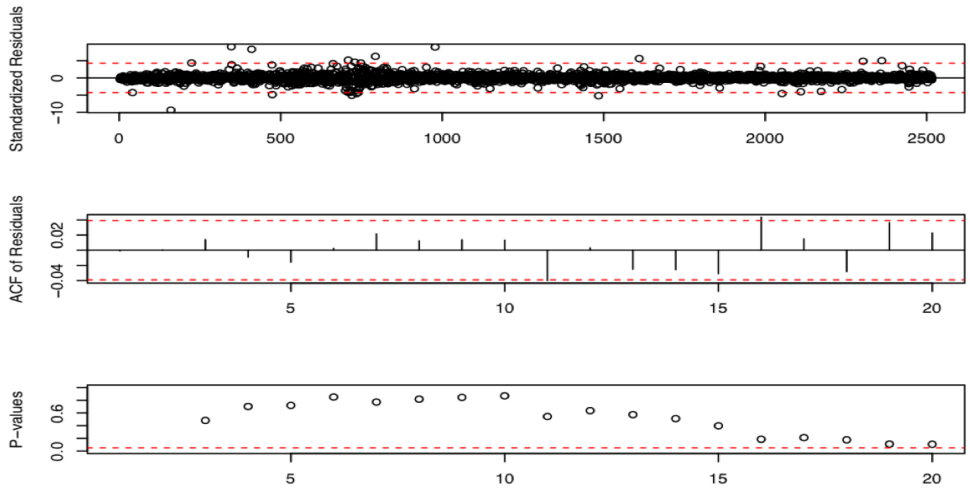
Figure 4

Based on the figure 3 and figure 4, the Amazon log return has a cuts off at log 2 in the ACF, and the PACF decays, so MA(2) could be a good candidate model.

```
fitb=arima(log.rtn, order=c(0,0,2) );fitb###AIC=-11182
```

```
##
## Call:
## arima(x = log.rtn, order = c(0, 0, 2))
##
## Coefficients:
##           ma1      ma2  intercept
##       -0.0065  -0.0713      1e-03
## s.e.   0.0199   0.0201      5e-04
##
## sigma^2 estimated as 0.000687:  log likelihood = 5594.39,  aic = -11182.77
```

Figure 5




```
Box.test(fitb$resid, lag = 20, type = "Ljung", fitdf = 2)###P-value=0.1
```

```
##
##  Box-Ljung test
##
## data:  fitb$resid
## X-squared = 25.7084, df = 18, p-value = 0.1066
```

```
Box.test(fitb$resid^2, lag = 20, type = "Ljung", fitdf = 2)#####ARCH EFFECT
```

```
##
##  Box-Ljung test
##
## data:  fitb$resid^2
## X-squared = 80.9291, df = 18, p-value = 5.89e-10
```

MA(2) model conclusion:

1. The model for MA(2)
2. Most of the standard residuals are around the 0 with little variation, despite some points are closed to -10, which can be categorized as outlier. So the standard residual plot show the assumption of the independence.
3. In the ACF of the residuals, almost every lags are in the range(-0.04, 0.02).
4. Most of the P-values are greater than a=0.05, so we can not reject the null hypothesis of the serial correlation, so there is no more serial correlation in the data anymore;

5. The Box. Test of residual gives a values of the p-value of 0.1006, which is greater than the a=0.05, so we can not reject the null hypothesis of the no serial correlation, which indicates that there is no more serial correlation in the data anymore;
6. The Box. Test of squared residual gives a values of the p-value closed to 0, which is greater than the a=0.05, so we can reject the null hypothesis of the no Arch effect, so a GARCH model is needed if we want to deal with the arch effect.\

4. GARCH MODEL

In order to deal with the ARCH effect in the Amazon log return series, I tried standard GARCH model first, however, there were significant leverage effect remained in the model, so I tried both T-GRACH model with the intention to get rid of the leverage effect, and maintain the assumption of the normality. However, the eta11 term is not significant, so normal GARCH with t-distribution.

```
*----------------------------------*
*          GARCH Model Fit         *
*----------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,2)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000673    0.000357  -1.88719 0.059134
ma1    -0.003472    0.017849  -0.19449 0.845789
ma2    -0.039366    0.017714  -2.22235 0.026260
omega   0.000005    0.000003   1.84699 0.064749
alpha1  0.022124    0.003742   5.91178 0.000000
beta1   0.969187    0.002189 442.80075 0.000000
shape   3.671147    0.259397  14.15261 0.000000
```

$r_t = u_t + a_t$;

$u_t = -0.00673 - 0.003472a_{t-1} - 0.039366\theta_2 a_{t-2}$

$a_t = \sigma_t \varepsilon_t$,

$\varepsilon_t \sim iid\ t_{3.671147}$

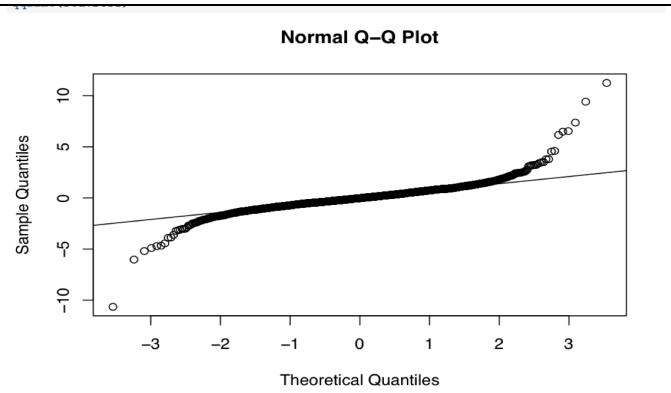
$\sigma t^2 = 0.00005 + 0.091612at-1^2 + 0.865247\sigma t-1^2$

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.05068  0.8219
Lag[2*(p+q)+(p+q)-1][5]   0.36031  1.0000
Lag[4*(p+q)+(p+q)-1][9]   1.38849  0.9982
d.o.f=2
H0 : No serial correlation
```

All the p-values are greater than a=0.05, so there are no more serial correlation.

```
Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.418  0.2337
Lag[2*(p+q)+(p+q)-1][5]     1.782  0.6708
Lag[4*(p+q)+(p+q)-1][9]     2.147  0.8868
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.3024 0.500 2.000  0.5824
ARCH Lag[5]    0.4434 1.440 1.667  0.9002
ARCH Lag[7]    0.7667 2.315 1.543  0.9483
```

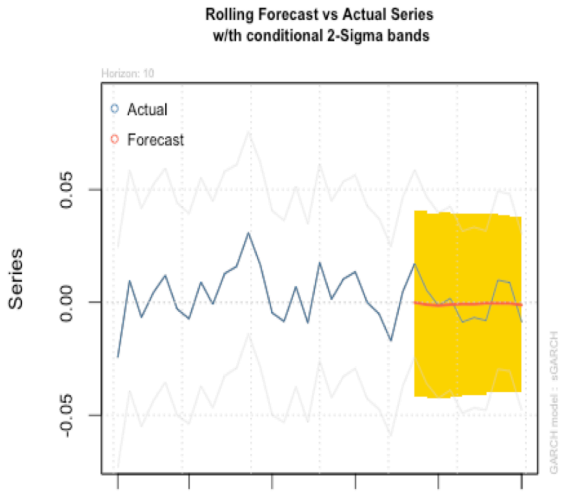
All the p-values are greater than the a=0.05, so there are no more ARCH effect in the s-GARCH model.

Sign Bias Test ---------- t-value prob sig Sign Bias 0.006988 0.994425 Negative Sign Bias 2.854807 0.004342 *** Positive Sign Bias 1.293124 0.196087 Joint Effect 10.729847 0.013280 **	**There are still leverage effect in the standardized residuals of the s-GARCH model**
Adjusted Pearson Goodness-of-Fit Test: ---------- group statistic p-value(g-1) 1 20 40.47 0.0028417 2 30 60.15 0.0005914 3 40 65.39 0.0050986 4 50 72.81 0.0152461	**Even though I use the t-distubution, there is only one p-value greater than a=0.05, which reject the assumption of the normality.**
	**From the QQ plot, there are heavy tails on the both side, so the data is not satisfied with the assumption of the normality.**

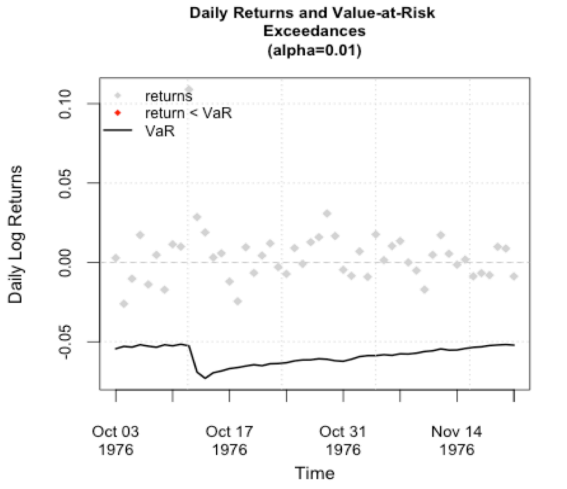
In conclusion, TCARCH model preforms well in getting rid of the serial correlation and ARCH effect, but the shortcomings are there are still slightly leverage effect in the TGARCH model, and the model is failed to meet the assumption of the normality.

5. Forecast

```
*-----------------------------------*
*       GARCH Model Forecast        *
*-----------------------------------*
Model: sGARCH
Horizon: 10
Roll Steps: 10
Out of Sample: 10

0-roll forecast [T0=1976-11-11 18:00:00]:
         Series   Sigma
T+1  -8.252e-05 0.02072
T+2  -8.998e-04 0.02075
T+3  -6.960e-04 0.02078
T+4  -6.960e-04 0.02082
T+5  -6.960e-04 0.02085
T+6  -6.960e-04 0.02088
T+7  -6.960e-04 0.02091
T+8  -6.960e-04 0.02094
T+9  -6.960e-04 0.02097
T+10 -6.960e-04 0.02100
```



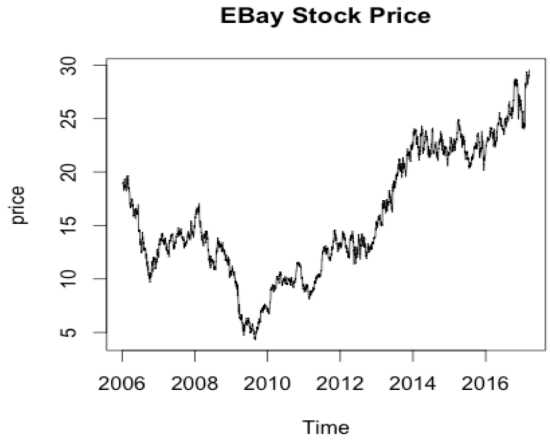

```
     Mu  Sigma Skew  Shape Shape(GIG) Realized
-0.0019 0.0198    0 3.7049          0   0.0028
-0.0008 0.0196    0 3.7049          0  -0.0261
-0.0008 0.0198    0 3.7049          0  -0.0102
 0.0004 0.0197    0 3.7049          0   0.0173
-0.0005 0.0196    0 3.7049          0  -0.0139
-0.0014 0.0196    0 3.7049          0   0.0047

.......
    Mu  Sigma Skew  Shape Shape(GIG) Realized
-6e-04 0.0199    0 3.6965          0  -0.0087
-7e-04 0.0197    0 3.6965          0  -0.0067
-3e-04 0.0196    0 3.6965          0  -0.0080
-3e-04 0.0194    0 3.6965          0   0.0099
-4e-04 0.0193    0 3.6965          0   0.0088
-1e-03 0.0192    0 3.6965          0  -0.0088
```



Summary of the Forecast: For the rolling forecast, based on the plot, we can see that confidence interval of the predicted value are within the range of the real values, so the it is a good prediction;

For the Daily Return and the VaR, we can see all the returns are above the upper VaR, which provides the upper bound of the lost. So the prediction is accurate.

B. Analysis of eBay Stock:



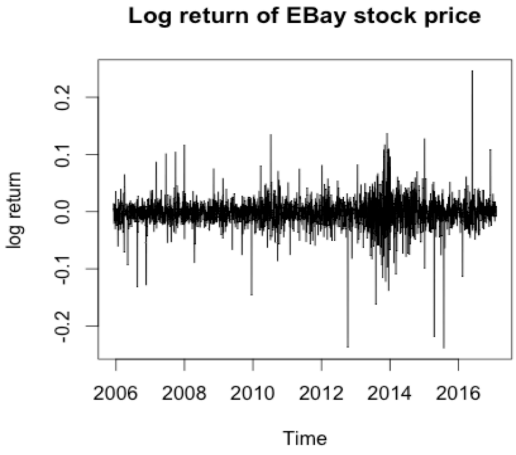


Figure 6 **Figure 7**

From the figure5, we can see the general time series trend for EBay stock can be divided into 2 parts. From 2005-2009, the price of stock is not stable. However, from 2009-2015, the price of the stock continues increasing.

Figure 6 is the log return of the eBay stock price, and the pattern seems stationary.

Final Model and Forecast

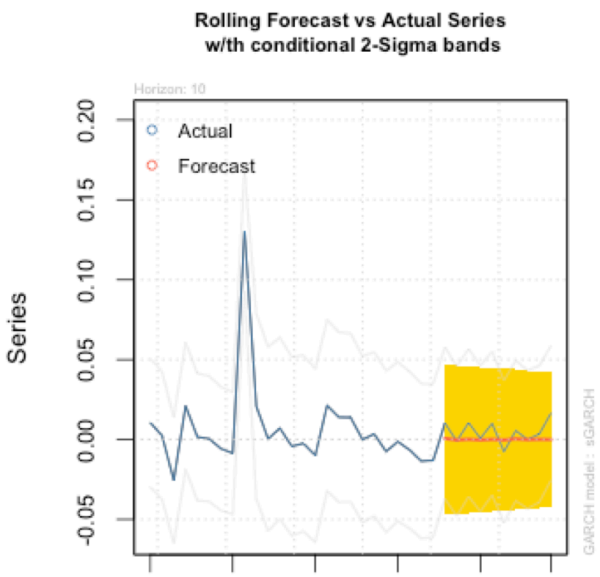
```
Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000121    0.000326   0.37111 0.710558
ma1    -0.035541    0.020146  -1.76420 0.077698
omega   0.000002    0.000001   1.65046 0.098850
alpha1  0.028666    0.002937   9.75886 0.000000
beta1   0.966867    0.002554 378.49911 0.000000
shape   4.821685    0.495578   9.72942 0.000000
```

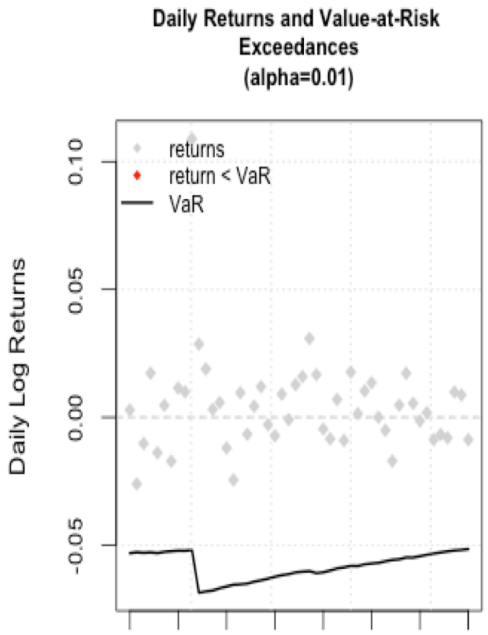
$r_t = u_t + a_t$;

$u_t = 0.000121 - 0.035541 a_{t-1}$

$a_t = \sigma_t \varepsilon_t$, $\varepsilon_t \sim_{iid} t_{4.821685}$

$\sigma_t^2 = 0.00002 + 0.028666 a_{t-1}^2 + 0.96687 \sigma_{t-1}^2$





From the rolling Forecast plot, we can see that the forecast value is within the actual value, and the confidence of the forecast value is within the range of the real values. Moreover, the returns are all above the VaR, so it is good sign for forecast.

Model conclusion for the EBay:

Among all the models I tried, MA (2) had the best fit for the mean function, and the s-GARCH model had the best fit for the residual. The serial correlation and ARCH effect were been removed successfully. However, eta11 term is not significant in the TGARCH model. Considering that it is real data, it is reasonable that the model would not be exactly perfect. So s-GARCH with t-distribution is a good model.

c. Analysis of Best Buy Stock

Figure 7 **Figure 8**

According to the figure 7, the general trend of the Best Buy stock can be divided to in two parts. From the 2005—2013, the price of stock keeps decreasing. However, the price of stock keeps increasing from 2013-2015.

From figure8, we can see the log return of the Best Buy stock price seems stationary.

The final Model for Best Buy:

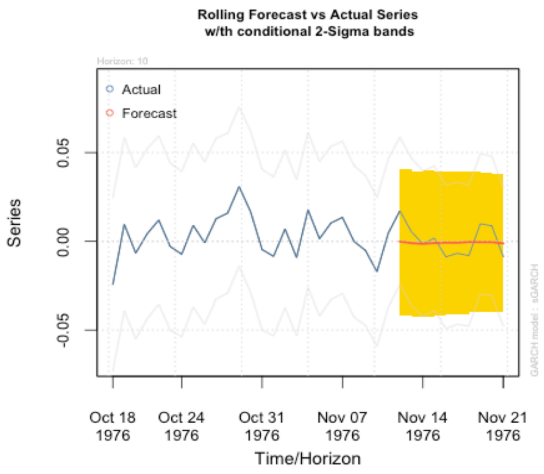
```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model      : sGARCH(1,1)
Mean Model       : ARFIMA(0,0,2)
Distribution     : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000252    0.000420  -0.5989 0.549241
ma1    -0.024442    0.020016  -1.2211 0.222029
ma2     0.021683    0.019182   1.1304 0.258312
omega   0.000019    0.000010   1.8614 0.062691
alpha1  0.047747    0.018533   2.5763 0.009985
beta1   0.927117    0.030180  30.7199 0.000000
shape   3.819087    0.320791  11.9052 0.000000
```

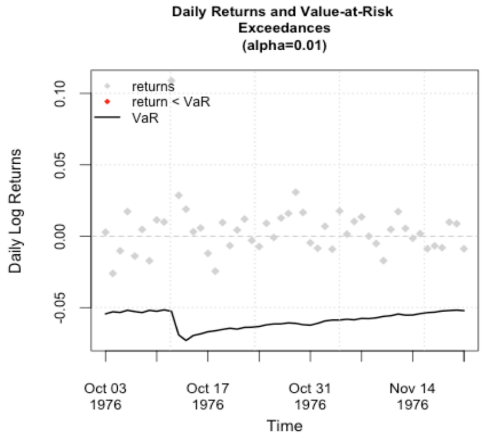
$r_t = u_t + a_t$;

$u_t = -0.000252 - 0.024442a_{t-1} + 0.021683a_{t-2}$

$a_t = \sigma_t \varepsilon_t$, $\varepsilon_t \sim$ iid $t_{3.819087}$

$\sigma t^2 = 0.000019 + 0.047747at-1^2 + 0.927117\sigma t-1^2$



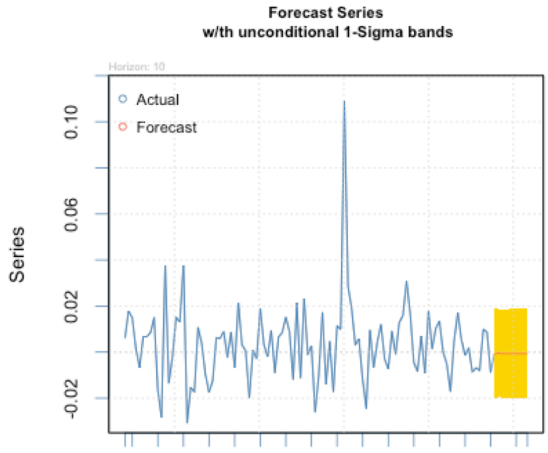
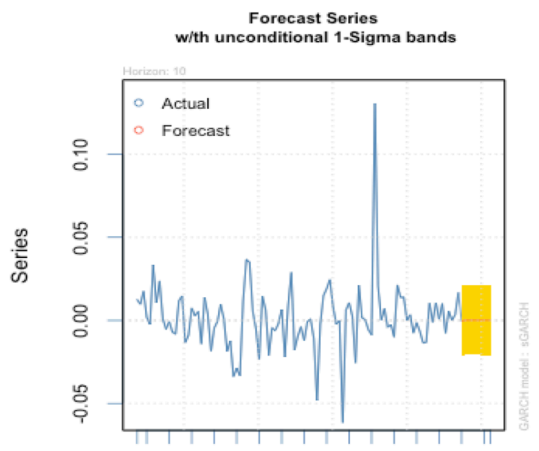
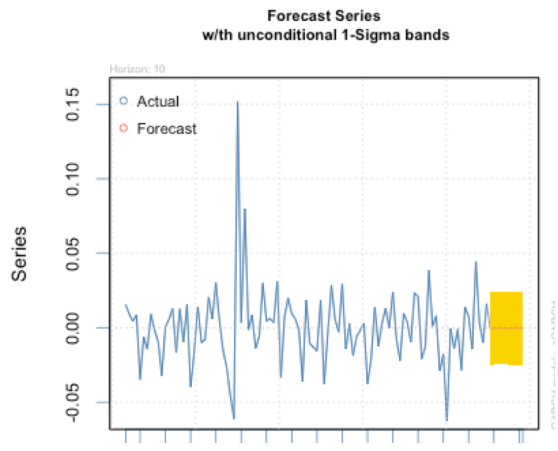


From the Rolling Forecast plot, the confidence interval is within the range of the real values, so the prediction is acceptable. In addition, all the returns are above the VaR.

Model summary for Best Buy

Among all the models I tried, MA (2) had the best fit for the mean function, and the s-GARCH model had the best fit for the residual. The serial correlation and ARCH effect were been removed successfully. However, there is still slight leverage effect in the s-GARCH model. When I tried TGARCH,eta11 is not significant. Considering that it is real data, it is reasonable that the model would not be exactly perfect. In addition, another advantage of the TGACH model for the Best Buy log return series is that the model meets the need of the normality assumption.

D. Trading Strategies:

	Forecast Plot	Sigma	VaR &ES
Amazon		Sigma 0.02072 0.02075 0.02078 0.02082 0.02085 0.02088 0.02091 0.02094 0.02097 0.02100	For one step ahead: $VaR_{0.95}$ =0.0341 $ES_{0.95}$ = 0.0484
EBay		Sigma 0.02372 0.02372 0.02372 0.02371 0.02371 0.02371 0.02371 0.02370 0.02370 0.02370	$VaR_{0.95}$ = 0.0318 $ES_{0.95}$ = 0.0398
Best Buy		Sigma 0.02563 0.02568 0.02573 0.02578 0.02583 0.02588 0.02592 0.02596 0.02601 0.02605	**VaR**=0.0386 $ES_{0.95}$ = 0.0484

Trading strategies and discussion:

From the model comparison above, we can see the stock of the Best Buy has the largest sigma, which indicates it is a relatively risky stock. While other two stocks, amazon and the EBay has smaller sigma, if a customer wants to have a relatively stable profit in the future, amazon and EBay would be the better choice, if the customer is willing to take risk, Best Buy is a good option considering that the stock price of the best Buy continues increasing since 2009.

In the terms of the VaR, value of the risk, for the amazon stock, if the I have 1 million dollar, there is 3.41% chance that I will lose more than 1 million dollars in the next day; If I buy the stock of the EBay, there is 3.18% chance that I will lose more than 1 million dollars. Moreove, if I choose Best Buy, there is 3.86% chance that I will lose more than 1 million dollar in the next day. In a short, choose EBay, I will have smallest risk of value.

More importantly, when we buy a stock, we should also bring the company background, and future supportive policies into consideration. For Amazon, its web services is providing a nice cushion of profits to allow Amazon to pursue a low-cost strategy with e-commerce, which is an incredible advantage.

For EBay, due to the separation of the PayPal business, eBay is able to focus exclusively on its core business of developing and operating an online marketplace. As of the most recent quarter, eBay had 157 million active buyers, a number that's grown at a 12% annualized rate over the past three years. Moreover, for Best Buy, as the holiday shopping season approaches, it

worth looking at how Best Buy managed its rocky but dogged recovery.

In conclusion, both amazon and eBay are well reprehensive E-commerce stock, due to the fast development of the online shopping; amazon and eBay have more room to develop. On the other hand, as the traditional electronic device seller, Best Buy also has its advantages especially during the holiday season. As the model I built for the 3 stocks, the most risky stock is the Best Buy, while Amazon and eBay are less risky and are worth trading in the future. I choose eBay eventually because it has the smallest VaR and ES.