APPLICATIONS OF ASYMMETRIC GARCH MODEL WITH VARIOUS CONDITIONAL DISTRIBUTIONS: THE EMPIRICAL CASE OF THE NASDAQ COMPUTER INDEX DAILY CLOSING RETURNS

RAN LI
SUMMER 2012

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Mathematics and Statistics
with interdisciplinary honors in Mathematics and Statistics

Reviewed and approved* by the following:

Jason R. Morton
Assistant Professor of Mathematics and Statistics
Thesis Supervisor

David R. Hunter
Professor of Statistics
Honors Adviser

Sergei Tabachnikov
Professor of Mathematics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.
Abstract

The purpose of this honors thesis is to find an appropriate GARCH (Generalized Autoregressive Conditional Heteroskedasticity) Model for the NASDAQ Computer Index Daily Closing Returns, given a ten-year time series of closing prices. On the one hand, Standard GARCH Models are not sufficient enough, if consider the leverage effects (the volatility responds to good news and bad news differently). Instead, asymmetric GARCH Models are better, and, in particular, Exponential GARCH (EGARCH) Model is the best. On the other hand, EGARCH Models with alternative conditional distributions perform better than that with the default Normal Conditional Distribution. In particular, the Generalized Hyperbolic Distribution is found to be good fit that generate large P-values against the null hypotheses in the various tests. In conclusion, among all of the models investigated, the EGARCH Model with the Skew Generalized Error Distribution is the best.
# Table of Contents

List of Figures ........................................................................................................... iii
List of Tables ............................................................................................................... iv
Acknowledgement ...................................................................................................... vi

1. Introduction ............................................................................................................. 1

2. Data Analysis ......................................................................................................... 3
   1) The Time Series ..................................................................................................... 3
   2) Summary Statistics ............................................................................................... 4
   3) Normal Test ........................................................................................................... 5
   4) Heteroskedasticity Test ......................................................................................... 7

3. Methodology Framework ....................................................................................... 9
   1) Standard GARCH ................................................................................................. 9
   2) EGARCH ............................................................................................................. 11
   3) GJR-GARCH ...................................................................................................... 12
   4) Conditional Distribution ..................................................................................... 13

4. Empirical Results ................................................................................................ 17
   1) Reviews of the Default Model ............................................................................. 17
   2) Variance Model ................................................................................................... 23
   3) Conditional Distribution ..................................................................................... 28

5. Conclusion ............................................................................................................. 31

6. Reference .............................................................................................................. 33

Appendix
List of Figures

Figure 1: Histogram of the Daily Closing Returns .............................................4
Figure 2: Normal Q-Q Plot of the Daily Closing Returns....................................5
Figure 3: Results of Jarque-Bera Normality Test................................................5
Figure 4: Time Series Plot of the Daily Closing Returns from Jan 3, 2000 to June 18, 2012........................................................................................................7
Figure 5: Results of Breusch - Pagan Test............................................................8
Figure 6: Normal Distributions ...........................................................................13
Figure 7: Skew-Normal Distributions..................................................................14
Figure 8: Student Distributions...........................................................................14
Figure 9: Generalized Error Distributions............................................................15
Figure 10: Normal Inverse Gaussian Distributions...............................................15
Figure 11: Johnson SU Distributions....................................................................16
Figure 12: R Statements of the Default Rugarch Model.......................................17
Figure 13(a): R Outcomes of the Default Rugarch Model....................................18
Figure 13(b): R Outcomes of the Default Rugarch Model....................................19
List of Tables

Table 1: Summary Statistics of the Daily Closing Returns........................................4

Table 2: P-Values of GARCH Parameters for ARMA (1, 1) – GARCH - Normal Distribution Models.................................................................23

Table 3: P-Values of Q-Statistics on Standardized Residuals for ARMA (1, 1) – GARCH - Normal Distribution Models.................................................................23

Table 4: P-Values of Q-Statistics on Standardized Squared Residuals for ARMA (1, 1) – GARCH - Normal Distribution Models.................................................................24

Table 5: P-Values of ARCH LM Tests for ARMA (1, 1) – GARCH - Normal Distribution Models.................................................................24

Table 6: P-Values of Q-Statistics on Standardized Residuals for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models.................................................................25

Table 7: P-Values of Q-Statistics on Standardized Squared Residuals for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models.................................................................25

Table 8: P-Values of ARCH LM Test for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models.................................................................26

Table 9: P-Values of Sign Bias Test for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models.................................................................26

Table 10: P-Values of Adjusted Pearson Goodness-of-Fit Test and GARCH Parameters for ARMA (1, 1) – GARCH (2, 1) – Normal Distribution Model.................................27

Table 11: P-Values of Sign Bias Test t for ARMA (1, 1) - EGARCH (2, 1) – Conditional Distribution Models.................................................................28

Table 12: P-Values of Adjusted Pearson Goodness-of-Fit Test t for ARMA (1, 1) - EGARCH (2, 1) - Conditional Distribution Models.................................................................28
Table 13: P-Values of Q-Statistics, ARCH LM Tests, and GARCH Parameters for ARMA (1, 1) - EGARCH (2, 1) - Conditional Distribution Models

Table 14: AICs, BICs, SICs and HQICs for ARMA (1, 1) - EGARCH (2, 1) - Conditional Distribution Models
Acknowledgement

First and foremost, I am heartily grateful to my thesis supervisor, Dr. Jason Morton, for his patient supervision and guidance throughout my honors research in Spring 2012.

Next, I would like to express my deep gratitude to my honors adviser Dr. David Hunter. Without his advice and academic support since my freshman year, I cannot have made such rapid progress and improvements. In addition, I would also like to record my gratitude to thank Dr. Megan Romer for her interesting time series course, which motivated me to independently study heteroscedasticity for my honors thesis.

Then, I sincerely appreciate the generous suggestions and motivations from Ruby Zhou, Catherine Ye, Gavin Feng and many other friends during my days in Penn State.

Last but not least, I thank Dr. Yun Zhou and my parents, Yue Li and Jian Ren, for their continuous encouragements and inspirations that have accompanied me during the years of my undergraduate study, throughout all the good days and bad days.
1. Introduction

The social networking company, Facebook, Inc. has been such a huge success. It held its initial public offering (IPO) on May 18, 2012, only 9 years after its predecessor, Facemash, was written. Depending on its functions, Facebook, Inc. is a component of the NASDAQ Computer Index since its IPO.

The NASDAQ Computer Index is a sector index, which came into being in Jan 3, 2000. Like the major index of the NASDAQ Composite, the NASDAQ Computer is also modeled from the aspects of its mean, its volatility and so on. I was motivated by Facebook, Inc.’s IPO to study the NASDAQ Computer Index in my honors thesis research. The main objects of this thesis are to describe the NASDAQ Computer Index’s volatility estimations, to model its leverage effects in asymmetric variance models and to select an appropriate conditional distribution in the GARCH Model.

On one hand, ARIMA (Autoregressive Integrated Moving Average) Models (Mills, 1990) fail to fit the time series because the assumption that the NASDAQ Computer should be stationary is frequently violated. GARCH (Generalized Autoregressive Conditional Heteroskedasticity) Models solved this problem for the heteroskedastic stock prices (Bollerslev, 1986).

On the other hand, Standard GARCH Model is not effective enough, because index prices may react to good news and bad news in different patterns. Then, asymmetric GARCH Models appear. There are EGARCH (Exponential GARCH) Models by Nelson and GJR-GARCH by Glosten, Jagannathan, and Runkle. When there are leverage effects in the NASDAQ Computer, asymmetric models either include absolute values (St. Pierre, 1998) or add indicator variables into the conditional variance equations (Glosten, Jagannathan, and Runcle, 1993).
Last but not least, some conditional distributions may improve the performance of the selected mean model and the selected variance model. Not only Normal Distributions but also Student Distributions and Generalized Error Distributions can be set as conditional distributions. What is more, the skew versions of unimodal and symmetric distributions mentioned above can also be considered (Zhang, 2009).

I appreciate the rugarch R package and the document, Introduction to the Rugarch Package (Version 1.0-8), contributed by Alexios Ghalanos. They have provided diverse R functions with effective arguments and detailed application instructions for my honors thesis analysis (Ghalanos, 2012).

The final results of this honor thesis indicate that the EGARCH Models with the Skew Student Distribution, the Generalized Error Distribution, the Skew Generalized Error Distribution, the Normal Inverse Gaussian Distribution and the Generalized Hyperbolic Distribution fit the NASDAQ Computer Daily Closing Returns. The one with the Skew Generalized Error Distribution is the best from the five for this ten-year return series.
2. Data Analysis

1) The Time Series

The stock prices of the NASDAQ Computer Index are gained from Yahoo Finance website, http://finance.yahoo.com/q/hp?s=^IXK+Historical+Prices (NASDAQ Computer). The daily closing prices have time range from Jan 3, 2000 to June 18, 2012.

Given the daily closing prices, $P_t$’s, the NASDAQ Computer closing returns can be computed as below,

$$r_t = \log(P_t) - \log(P_{t-1})$$

($r_t$) is the time series modeled throughout this honors thesis.
2) Summary Statistics

<table>
<thead>
<tr>
<th>Minimum</th>
<th>Lower Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>Upper Quartile</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.1444000</td>
<td>-0.0097540</td>
<td>-0.0001368</td>
<td>0.0006804</td>
<td>0.0093820</td>
<td>0.1661000</td>
<td>0.02142713</td>
</tr>
</tbody>
</table>

Table 1. Summary Statistics of the Daily Closing Returns

As expected, the daily closing return series, \( r_t \)'s, has mean 0.0006804, which is around 0. The median, -0.0001368, is around the mean, 0.0006804, so the sample distribution is very likely to be symmetric.

![Histogram of NASDAQ Computer Daily Closing Returns](image)

**Figure 1: Histogram of the Daily Closing Returns**

As expected, the histogram of the series is symmetric. Also, it is unimodal with skewness around 0. However, the sample distribution may be leptokurtic as displayed above.
3) Normal Test

![Normal Q-Q Plot](image)

Figure 2: Normal Q-Q Plot of the Daily Closing Returns

A Normal Q-Q (Quantile - Quantile) Plot above is presented. If the $r_i$’s from Jan 3, 2000 to June 18, 2012 is normally distributed, the points should converge to the straight line (Wilk, 1968). However, the points diverge from it, indicating the sample distribution may not be well approximated to normal distributions.

**Jarque-Bera Normality Test**

```r
data: composite_computer_returns[, 9]
JB = 3988.202, p-value < 2.2e-16
alternative hypothesis: greater
```

Figure 3: Results of Jarque-Bera Normality Test
In addition to the Q-Q Plot Normality Test, jarque.test function in the moments R package can also be utilized for normality test. If the data come from a normal distribution, the Jarque-Bera statistic asymptotically follows a Chi-Squared Distribution in degrees of freedom 2 (Jarque and Bera, 1980). The null hypothesis is that both of the skewness and the excess kurtosis are zero. Since Jarque-Bera = 3988.202 and P-values=2.2e-16, the null hypothesis is rejected. Either the skewness or the excess kurtosis is greater than 0. By the definition of normal distributions, the time series is not normally distributed.

Since the conclusion from the Normal Q-Q Plot and the Jarque Test are consistent, it is pretty sure that the time series is not normally distributed.
4) Heteroskedasticity Test

![Time Series Plot of NASDAQ Computer Daily Closing Returns](image)

Figure 4: Time Series Plot of the Daily Closing Returns from Jan 3, 2000 to June 18, 2012

The NASDAQ Daily Closing Return Plot has time variable in the x axis and the daily closing return variable in the y axis. The empirical observation of the plot shows that the time series, \( r_t \), is not stationary, because the sample distribution is not consistent when shifted in time.
To support the observation of heteroskedasticity, the bptest function in the lmtest R package can work. The Breusch–Pagan Test is a chi-squared test and it tests for conditional homoscedasticity (Breusch and Pagan, 1979). The P-value for the Breusch-Pagan Test, 2.2e-16, is way smaller than 0.05 against $H_0$, so $(r_t)$ is heteroskedasticity.

ARIMA Models only work for homoscedasticity (Mills, 1990), so it is of course useless here. In conclusion, all of the properties tested above require GARCH Models to fit the non-stationary time series, $(r_t)$. 

**Figure 5: Results of Breusch-Pagan Test**

```
studentized Breusch-Pagan test

data:  y ~ x
BP = 141.6255, df = 1, p-value < 2.2e-16
```


3. Methodology Framework

1) Standard GARCH

The NASDAQ Computer Daily Closing Return series, \((r_t)\), is non-normally distributed and heteroskedastic as stated above. Hence, GARCH Models could be good solutions.

GARCH Models should have input of no trend and constant mean 0, but the original return series, \((r_t)\), does not have to satisfy these requirements. Therefore, firstly, \((r_t)\) is fitted into a mean process to generate a series of error terms. The series of error terms does not have to be stationary but it does not have a trend and its mean should be 0. Then, GARCH Models are the second step for these residuals of mean process, in order to generate stationary residuals with constant volatility (Bollerslev, 1986).

For example, the default mean process in the rugarch R package is ARIMA(1,0,1). Then, given a time series, it may generate non-stationary error terms, \(\epsilon_t\), of ARIMA(1,0,1).

Express \(\epsilon_t\) as

\[ \epsilon_t = \sigma_t z_t \quad \text{where} \]

\(z_t\)'s follow independent and identically distributions (i.i.d.). We call it conditional distribution, which will be discussed in details in 3.4. \(\sigma_t^2\)'s are the conditional variance, which is being discussed in 3.1, 3.2 and 3.3.

In Standard GARCH Models, \(\sigma_t^2\)'s are modeled as

\[ \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \]

\(\epsilon_t^2\)'s are called ARCH terms while \(\sigma_t^2\)'s are called GARCH terms. Hence, \(p\) is the order of GARCH terms, while \(q\) is the order of ARCH terms (Bollerslev, 1986). In other words, current volatility depends on the previous volatility, \(\sigma_{t-i}^2\)'s, as well as the previous divergence in
the mean process, $\varepsilon_{t,i}^2$. These models can make sense in many situations for different $p$ and $q$ values.

However, since $\varepsilon_t$ is involved in the conditional variance equation as $\varepsilon_t^2$, the previous $\varepsilon_{t,i}$ and the previous $-\varepsilon_{t,i}$ of the same absolute value affect the current volatility in a similar pattern. Therefore, Standard GARCH Models fail to respond to situations, where there are significant differences between good news $\varepsilon_t$ and bad news $-\varepsilon_t$ of the same magnitude.
2) **EGARCH**

EGARCH Models as asymmetric models can deal with this problem.

It is the same that

\[ \epsilon_t = \sigma_t z_t. \]

\( \epsilon_t \)'s are the error terms with mean 0 from mean process. \( z_t \)'s follow the independent and identical conditional distributions. \( \sigma_t^2 \)'s are the conditional variances, like the \( \sigma_t^2 \)'s in the Standard GARCH Models.

The difference is that, for specific \( (p, q) \), the conditional variance, \( \sigma_t^2 \)'s, are modeled as

\[
\log(\sigma_t^2) = \alpha_0 + \sum \alpha_i g(z_{t-k}) + \sum \beta_i \log(\sigma_{t-i}^2)
\]

The logarithm of current conditional variance, \( \log(\sigma_t^2) \), is affected by the logarithm of the previous conditional variances, \( \log(\sigma_{t-i}^2) \)'s, and \( g(z_{t-k}) \)'s, where \( g(z_{t-k}) = \theta z_{t-k} + \lambda (|z_{t-k} - E(z_{t-k})|) \), and \( z_t = \epsilon_t / \sigma_t \).

Based on the definition of \( g(z_t) \),

If \( \epsilon_t \) is positive, then \( g(z_t) = \theta z_t + \lambda (z_t - E(z_t)) = \theta \epsilon_t / \sigma_t + \lambda \left( \epsilon_t / \sigma_t - E(\epsilon_t / \sigma_t) \right) \);

If \( \epsilon_t \) is negative, then \( g(Z_t) = \theta z_t + \lambda (-z_t - E(-z_t)) = \theta \epsilon_t / \sigma_t - \lambda \left( \epsilon_t / \sigma_t - E(\epsilon_t / \sigma_t) \right) \).

The formulation of \( g(Z_t) \) allows the signs as well as the magnitudes of the previous \( \epsilon_t \)'s to affects the current conditional variance, \( \sigma_t^2 \) (St. Pierre, 1998). This inovation is a creative solution of leverage effects.
3) GJR-GARCH

GJR-GARCH by Glosten, Jagannathan and Runkle is another solution for asymmetric effects. Likewise,

$$\epsilon_t = \sigma_t Z_t$$

$\epsilon_t$’s are error terms generated by the mean process. The mean of the $\epsilon_t$’s is 0. $Z_t$’s follow the independent and identical conditional distributions, while $\sigma_t^2$’s are the conditional variances.

The constructions of the conditional variances are very much similar to that of Standard GARCH Models, but with additional sum, $\sum \varphi_i \sigma_{t-i}^2 I_t$.

$$\sigma_t^2 = \alpha_0 + \sum \alpha_i \sigma_{t-i}^2 + \sum \beta_i \epsilon_{t-i}^2 + \sum \varphi_i \epsilon_{t-i}^2 I_t$$

$I_t$ is an indicator function where $I_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$, and $I_{t-1} = 1$ if $\epsilon_{t-1} < 0$ (Glosten, Jagannathan, and Rucnkle, 1993). Therefore,

If $\epsilon_{t-i}$ is not negative, then the coefficient of $\epsilon_{t-i}^2$ is $\beta_i$;

If $\epsilon_{t-i}$ is negative, then the coefficient of $\epsilon_{t-i}^2$ is $(\beta_i + \varphi_i)$.

The current conditional variance, $\sigma_t^2$, is not only affected by the magnitudes of the previous $\epsilon_{t-1}$’s but also by their signs. Though this innovation of GJR-GARCH is different than that of EGARCH, they have the same goals for asymmetric effects.
4) Conditional Distribution

Coming back to $\zeta_t = \epsilon_t / \sigma_t$ in Standard GARCH Models, EGARCH Models or GJR-GARCH Models, $\zeta_t$'s are said to follow the independent and identically conditional distributions. They include but are not limited to Normal Distributions, Skew-Normal Distributions, Student Distributions, Skew Student Distributions, Generalized Error Distributions, Skew Generalized Error Distributions, Normal Inverse Gaussian Distributions, Generalized Hyperbolic Distributions, and Johnson's SU Distributions. The rugarch R package provides arguments of alternative conditional distributions as mentioned above. Though they are not suitable everywhere, considering some alternative conditional distributions sometimes can be very helpful.

Below are listed some sample alternative distribution plots:

![Figure 6: Normal Distributions](image-url)
Figure 7: Skew-Normal Distributions

Figure 8: Student Distributions
Figure 9: Generalized Error Distributions

Figure 10: Normal Inverse Gaussian Distributions
Figure 11: Johnson SU Distributions
4. Empirical Results

1) Reviews of the Default Model

A regular GARCH Model in the rugarch R package has three parts. They are the variance model, the mean model, and the distribution model. As it is said in the names, the variance model looks at the construction of the current conditional variance, $\sigma_t^2$. It can be a Standard GARCH, an EGARCH, a GJR-GARCH or any other alternative GARCH Model. The mean model determines the process to remove any non-zero mean. ARMA(p,q) or more generalized ARIMA(p,r,q) are always good ideas for time series. The distribution model controls the conditional distribution. In some specific cases, distributions other than the Normal Distribution may be more helpful.

The default model of the rugarch package has its three components are the ARMA(1,1), the Standard GARCH(1,1), and the Normal Conditional Distribution, respectively. The combination of the statements below calls the fitting of this model.

```r
> spec=ugarchspec()
> data=composite_computer_returns[,9]
> fit=ugarchfit(spec=spec, data=data)
> show(fit)
```

Figure 12: R Statements of the Default Rugarch Model

The key of these statements is ugarchfit. It provides the estimated parameters, their respective P-values, and a variety of tests from different aspects of the NASDAQ Computer series, $(r_t)$. After the time series is fitted, the R analysis results come out as below (Ghalanos, 2012):
**Conditional Variance Dynamics**

**GARCH Model:** sGARCH(1,1)

**Mean Model:** ARFIMA(1,0,1)

**Distribution:** norm

**Optimal Parameters**

| Parameter | Estimate | Std. Error | t value | Pr(>|t|) |
|-----------|----------|------------|---------|----------|
| mu        | 0.000730 | 0.000232   | 3.1424  | 0.001675 |
| e1        | 0.377105 | 0.237039   | 1.5909  | 0.111633 |
| n1        | -0.416752| 0.232769   | -1.7630 | 0.076126 |
| omega     | 0.000002 | 0.000001   | 3.4814  | 0.000499 |
| alpha1    | 0.075266 | 0.008853   | 8.5017  | 0.000000 |
| beta1     | 0.920130 | 0.008893   | 101.4682| 0.000000 |

**Robust Standard Errors:**

| Parameter | Estimate | Std. Error | t value | Pr(>|t|) |
|-----------|----------|------------|---------|----------|
| mu        | 0.000730 | 0.000252   | 2.9020  | 0.003706 |
| e1        | 0.377105 | 0.155200   | 2.4298  | 0.015107 |
| n1        | -0.416752| 0.154987   | -2.6889 | 0.007168 |
| omega     | 0.000002 | 0.000001   | 3.2435  | 0.001180 |
| alpha1    | 0.075266 | 0.012049   | 6.2450  | 0.000000 |
| beta1     | 0.920130 | 0.011052   | 83.2537 | 0.000000 |

**LogLikelihood:** 8121.195

**Information Criteria**

Akaike      -5.3495
Bayes       -5.3376
Shibata     -5.3495
Hannan-Quinn -5.3452

**Q-Statistics on Standardized Residuals**

<table>
<thead>
<tr>
<th>statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag10</td>
<td>7.935</td>
</tr>
<tr>
<td>Lag15</td>
<td>17.635</td>
</tr>
<tr>
<td>Lag20</td>
<td>20.994</td>
</tr>
</tbody>
</table>

**HO : No serial correlation**

**Q-Statistics on Standardized Squared Residuals**

<table>
<thead>
<tr>
<th>statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag10</td>
<td>22.24</td>
</tr>
<tr>
<td>Lag15</td>
<td>23.25</td>
</tr>
<tr>
<td>Lag20</td>
<td>27.38</td>
</tr>
</tbody>
</table>

Figure 13(a): R Outcomes of the Default Rugarch Model
Figure 13(b): R Outcomes of the Default Rugarch Model

Though the default model is not working as explained below, the outcomes provide much useful information.

As it says in the name, the Robust Standard Errors provide robust estimations for the parameters in the ARMA(1,1) mean model and the GARCH(1,1) variance model. The estimations’ standard errors and the P-values are also provided. The robust estimations are
another reference besides the Optimal Parameters to understand the parameter numbers (Huber, 1981). Though the P-values in the Optimal Parameters are greater than 0.05, those in the Robust Standard Errors are smaller than 0.05. Therefore, the only AR and the only MA parameters are kept to guarantee that the input of the variance model does not have serial correlation.

The Information Criteria of the default model displays the Akaike (AIC), Bayesian (BIC), Hannan-Quinn (HQIC) and Shibata (SIC) statistics to enable model selection (Ghalanos, 2012). According to Akaike (1974), “AIC is a relative measure of the information lost when a given model is used to describe reality”. Since it is a relative measure, AIC only works while comparing models (Akaike, 1974). Similar to AIC, BIC fits models based on the likelihood functions, but it adds a penalty term concerning the parameter number in the model candidates (Schwarz, 1978). HQIC (Hannan and Quin, 1979) and SIC (Schwarz, 1978) statistics performs the same function. For the Information Criteria statistics, smaller values are looking forwards to because they indicate relatively less lost information and, hence, better models.

In a GARCH model, besides the residual series of the mean process, ($\epsilon_t$), there is another residual series of the variance process, namely the stationary error series. The Q-Statistics on the Standardized Residuals tests the efficiency of the AR and MA parameters in the mean process, while the Q-Statistics on the Standardized Square Residuals tests the efficiency of the ARCH and GARCH parameters in the variance process. High Q-Statistics P-values means that there is no serial correlation or no heteroscedasticity at the 10th, the 15th or the 20th lags of the variance process’ error series, and, hence, both processes are successful. Similarly, the ARCH LM Test looks at the residuals of the variance process against their own lags, too (Lee, 1991). The null hypothesis is $H_0$: No ARCH effects, so high P-values are looking forwards to (Ghalanos, 2012). For the NASDAQ Computer, ($r_t$), the Q-Statistics on the Standardized Squared Residuals and the
ARCH LM Tests have P-values less than 0.05. It means that the default variance model is not working well.

The Sign Bias Test proposed by Engle and Ng (1993) tells the significance of leverage effects in the residuals of the variance model. Because of the null hypothesis $H_0$: there is no leverage effects remained, high P-values over 0.05 are wanted. The daily closing returns’ statistics of the Sign Bias Test is bad, because some P-values are much less than 0.05.

The Adjusted Pearson Goodness-of-Fit Test tells if the conditional distribution is chosen appropriately (Palm, 1996). As seen in the outcomes where $1.682e-05 << 0.05$, $2.609e-05 << 0.05$, $3.565e-03 << 0.05$, and $9.262e-03 << 0.05$, Normal Conditional Distribution is not suitable in the case of the $(r_t)$ and the ARMA(1,1) – GARCH(1,1). After the default variance model and the default mean model are adjusted, other statistics should improve. If the P-values of the Adjusted Pearson Goodness-of-Fit Tests are so small still, the later chose mean model and the later chosen variance model should be match with some other distribution model.

To sum up, given the statistics from the R outcomes, the default combination of the ARMA(1,1), the Standard GARCH(1,1) and the Normal Conditional Distribution is not working. Firstly, to remove the heteroskedasticity patterns of the variance model residuals, some other combinations of ARCH and GARCH parameter numbers should be tried. Secondly, if then the problems of the Q-Statistics and the ARCH LM Tests are fixed but the Sign Bias Tests, the Standard GARCH model should be replaced by an asymmetric GARCH Model with some similar ARCH and GARCH parameter numbers. Thirdly, if then both of the heteroskedasticity and the leverage effects are removed from the variance modal residuals but the Adjusted Pearson Goodness-of-Fit Tests problems, some other conditional distribution model could substitute the default distribution model. Finally, only if all of the Q-Statistics Test, the ARCH LM Tests, the
Sign Bias Tests and the Adjusted Pearson Goodness-of-Fit Tests have P-values greater than 0.05, the combination of the mean model, the variance model and the distribution model is qualified.

If more than one model achieve these qualifications, their AIC’s, BIC’s, SIC’s and HQIC’s should be compared. The one with the smallest AIC, the smallest BIC, the smallest SIC and the smallest HQIC at the same time or the one with the most number of the smallest information criterion should win the model selection.
2) Variance Model

Of the NASDAQ Computer Index Daily Closing Return series, \((r_t)\), ARMA Model’s P-values are greater than 0.05 in the Optimal Parameters results. However, it does not hurt to still keep the only AR term and the only MA term. Turning to the Alpha and the Beta, neither of the P-Values is greater than 0.05. This fact shows that at least one parameter of each kind is necessary but there may not be enough. To explore the effective and efficient parameter numbers of the variance model, I try Standard GARCH(1,2), Standard GARCH(2,1), and Standard GARCH(3,1) with the default mean model, ARMA(1,1), and the default distribution model, Normal Distribution.

<table>
<thead>
<tr>
<th>(ARCH, GARCH)</th>
<th>P–Value for Alpha 1</th>
<th>P–Value for Alpha 2</th>
<th>P–Value for Alpha 3</th>
<th>P–Value for Beta 1</th>
<th>P–Value for Beta 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td>0.000000</td>
<td>NA</td>
<td>NA</td>
<td>0.000000</td>
<td>NA</td>
</tr>
<tr>
<td>(1,2)</td>
<td>0.000000</td>
<td>NA</td>
<td>NA</td>
<td>0.000000</td>
<td>0.999910</td>
</tr>
<tr>
<td>(2,1)</td>
<td>0.912176</td>
<td>0.000000</td>
<td>NA</td>
<td>0.000000</td>
<td>NA</td>
</tr>
<tr>
<td>(3,1)</td>
<td>0.954801</td>
<td>0.000001</td>
<td>0.999754</td>
<td>0.000000</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 2: P-Values of GARCH Parameters for ARMA (1, 1) – GARCH - Normal Distribution Models

<table>
<thead>
<tr>
<th>(ARCH, GARCH)</th>
<th>Q-Statistics on Standardized Residuals Lag10</th>
<th>Q-Statistics on Standardized Residuals Lag15</th>
<th>Q-Statistics on Standardized Residuals Lag20</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td>0.4399</td>
<td>0.1719</td>
<td>0.2797</td>
</tr>
<tr>
<td>(1,2)</td>
<td>0.4354</td>
<td>0.1752</td>
<td>0.2851</td>
</tr>
<tr>
<td>(2,1)</td>
<td>0.4620</td>
<td>0.2031</td>
<td>0.3143</td>
</tr>
<tr>
<td>(3,1)</td>
<td>0.4689</td>
<td>0.1924</td>
<td>0.2953</td>
</tr>
</tbody>
</table>

Table 3: P-Values of Q-Statistics on Standardized Residuals for ARMA (1, 1) – GARCH - Normal Distribution Models
(ARCH, GARCH) & Q-Statistics on Standardized Squared Residuals & Q-Statistics on Standardized Squared Residuals & Q-Statistics on Standardized Squared Residuals \\
Lag10 & Lag15 & Lag20 \\
(1,1) & 0.004483 & 0.038728 & 0.072110 \\
(1,2) & 0.004373 & 0.037914 & 0.070383 \\
(2,1) & 0.6606 & 0.8491 & 0.8334 \\
(3,1) & 0.6677 & 0.8509 & 0.8307 \\

Table 4: P-Values of Q-Statistics on Standardized Squared Residuals for ARMA (1, 1) – GARCH - Normal Distribution Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td>0.003583</td>
<td>0.024476</td>
<td>0.014126</td>
</tr>
<tr>
<td>(1,2)</td>
<td>0.00352</td>
<td>0.02418</td>
<td>0.01386</td>
</tr>
<tr>
<td>(2,1)</td>
<td><strong>0.8334</strong></td>
<td><strong>0.9922</strong></td>
<td><strong>0.8277</strong></td>
</tr>
<tr>
<td>(3,1)</td>
<td>0.8208</td>
<td>0.9943</td>
<td>0.8299</td>
</tr>
</tbody>
</table>

Table 5: P-Values of ARCH LM Tests for ARMA (1, 1) – GARCH – Normal Distribution Models

From the outcomes, the Standard GARCH(1,2) is not helping because its Beta2 P-value, 0.999910, is greater than 0.05 but there are still some P-Values less than 0.05 of Q-Statistics on Standardized Squared Residuals and of ARCH LM Tests.

The Standard GARCH(2,1) looks better than the Standard GARCH(1,1). Since the P-values of Alpha2 is less than 0.05, another ARCH parameter is necessary. What is more important, all of the P-values in the Q-Statistics on Standardized Residuals, the Q-Statistics on Standardized Squared Residuals and the ARCH LM Tests are greater 0.05. Therefore, two ARCH and one GARCH parameters efficiently solve the serial correlation and the heteroskedasticity problems of the variance model residuals.
The Standard GARCH(3,1)’s Alpha3 P-value is greater than 0.05, so it is not necessary to add another Alpha parameter to the Standard GARCH(2,1). Therefore, the P-values presented above in the tables reveal a preferable variance model of the Standard GARCH(2,1).

The next problem is the leverage effect in the variance model residuals. The P-values in the Sign Bias Test should all be greater than 0.05. It is good that, by the Standard GARCH(2,1), the Sign Bias P-value is 0.1672, and the Negative Sign Bias P-value is 0.4919, but it is not perfect the P-values of the Positive Sign Bias and of the Joint Effect are way less than 0.05 (see Figure 13(b)). What it means is that the symmetric Standard GARCH(2,1) should be replaced by some asymmetric variance model to eliminate the leverage effect in the variance model residuals. For example, EGARCH Model and GJR-GARCH Model as introduced above are can be effective.

<table>
<thead>
<tr>
<th>GARCH Model</th>
<th>Q-Statistics on Standardized Residuals Lag10</th>
<th>Q-Statistics on Standardized Residuals Lag15</th>
<th>Q-Statistics on Standardized Residuals Lag20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard GARCH(2,1)</td>
<td>0.4620</td>
<td>0.2031</td>
<td>0.3143</td>
</tr>
<tr>
<td>GJR – GARCH(2,1)</td>
<td>0.4129</td>
<td>0.1338</td>
<td>0.2099</td>
</tr>
<tr>
<td>EGARCH(2,1)</td>
<td>0.34362</td>
<td>0.07657</td>
<td>0.11333</td>
</tr>
</tbody>
</table>

Table 6: P-Values of Q-Statistics on Standardized Residuals for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models

<table>
<thead>
<tr>
<th>GARCH Model</th>
<th>Q-Statistics on Standardized Squared Residuals Lag10</th>
<th>Q-Statistics on Standardized Squared Residuals Lag15</th>
<th>Q-Statistics on Standardized Squared Residuals Lag20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard GARCH(2,1)</td>
<td>0.6606</td>
<td>0.8491</td>
<td>0.8334</td>
</tr>
<tr>
<td>GJR-GARCH(2,1)</td>
<td>0.2770</td>
<td>0.6377</td>
<td>0.7162</td>
</tr>
<tr>
<td>EGARCH(2,1)</td>
<td>0.3903</td>
<td>0.7697</td>
<td>0.6561</td>
</tr>
</tbody>
</table>

Table 7: P-Values of Q-Statistics on Standardized Squared Residuals for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models
<table>
<thead>
<tr>
<th>GARCH Model</th>
<th>ARCH LM Tests Lag 2</th>
<th>ARCH LM Tests Lag 5</th>
<th>ARCH LM Tests Lag 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard GARCH(2,1)</td>
<td>0.8334</td>
<td>0.9922</td>
<td>0.8277</td>
</tr>
<tr>
<td>GJR-GARCH(2,1)</td>
<td>0.02091</td>
<td>0.13527</td>
<td>0.40104</td>
</tr>
<tr>
<td>EGARCH(2,1)</td>
<td>0.1498</td>
<td>0.5019</td>
<td>0.6560</td>
</tr>
</tbody>
</table>

Table 8: P-Values of ARCH LM Test for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models

The table above indicates (2,1) are effective parameter numbers for Standard GARCH and EGARCH, but not perfect for GJR-GARCH. There is one P-value of the GJR-GARCH(2,1) smaller than 0.05.

<table>
<thead>
<tr>
<th>GARCH Model</th>
<th>Sign Bias</th>
<th>Negative Sign Bias</th>
<th>Positive Sign Bias</th>
<th>Joint Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard GARCH(2,1)</td>
<td>0.2035067</td>
<td>0.1807567</td>
<td>0.0595225</td>
<td>0.0001472</td>
</tr>
<tr>
<td>GJR – GARCH(2,1)</td>
<td>0.175850</td>
<td>0.361276</td>
<td>0.017305</td>
<td>0.003024</td>
</tr>
<tr>
<td>EGARCH(2,1)</td>
<td>0.8094</td>
<td>0.9998</td>
<td>0.9573</td>
<td>0.9872</td>
</tr>
</tbody>
</table>

Table 9: P-Values of Sign Bias Test for ARMA (1, 1) - GARCH (2, 1) - Normal Distribution Models

At the same time, the GJR-GARCH(2,1) is not sufficient to remove all heteroscedasticity, because, like the Standard GARCH(2,1), the P-values of the Positive Sign Bias and the Joint Effect are both less than 0.05.

However, for the EGARCH(2,1) Model, the Positive Sign Bias Test and the Joint Effect Test perform as well as the EGARCH(2,1) Model’s Sign Bias Test and its Negative Sign Bias Test. Therefore, the EGARCH(2,1) Model performs way better than the GJR-GARCH(2,1) Model when it removes all asymmetric effects in the variance model residuals.
However, Adjusted Pearson Goodness-of-Fit Tests are still problems for ARMA(1,1) - EGARCH(2,1) – Normal Distribution Model. Otherwise, the P-values should all be greater than 0.05. However, the variance model determined so far has accomplished its responsibilities. The remaining problems are the conditional distribution model’s responsibilities.

<table>
<thead>
<tr>
<th>GARCH Model</th>
<th>Adjusted Pearson Goodness-of-Fit Test (20 bins)</th>
<th>Adjusted Pearson Goodness-of-Fit Test (30 bins)</th>
<th>Adjusted Pearson Goodness-of-Fit Test (40 bins)</th>
<th>Adjusted Pearson Goodness-of-Fit Test (50 bins)</th>
<th>Alpha’s and Beta’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard GARCH(2,1)</td>
<td>0.0000330</td>
<td>0.0003331</td>
<td>0.0023728</td>
<td>0.0026857</td>
<td>Satisfy</td>
</tr>
<tr>
<td>GJR-GARCH(2,1)</td>
<td>0.001262</td>
<td>0.007108</td>
<td>0.004547</td>
<td>0.009325</td>
<td>Satisfy</td>
</tr>
<tr>
<td>EGARCH(2,1)</td>
<td><strong>0.003310</strong></td>
<td><strong>0.008169</strong></td>
<td><strong>0.014868</strong></td>
<td><strong>0.001863</strong></td>
<td>Satisfy</td>
</tr>
</tbody>
</table>

Table 10: P-Values of Adjusted Pearson Goodness-of-Fit Test and GARCH Parameters for ARMA (1, 1) – GARCH (2, 1) - Normal Distribution Models
3) Conditional Distribution

To select the distribution model for the NASDAQ Computer Index Daily Closing Returns, \( r_t \)’s, all possible conditional distributions are listed below in a table. They are the Normal Distribution, the Skew Normal Distribution, the Student Distribution, the Skew Student Distribution, the Generalized Error Distribution, the Skew Generalized Error Distribution, the Normal Inverse Gaussian Distribution, the Generalized Hyperbolic Distribution, and the Johnson's SU Distribution.

<table>
<thead>
<tr>
<th>Conditional Distribution</th>
<th>Sign Bias</th>
<th>Negative Sign Bias</th>
<th>Positive Sign Bias</th>
<th>Joint Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.8094</td>
<td>0.9998</td>
<td>0.9573</td>
<td>0.9872</td>
</tr>
<tr>
<td>Skew Normal</td>
<td>0.8666</td>
<td>0.8571</td>
<td>0.9623</td>
<td>0.9800</td>
</tr>
<tr>
<td>Student</td>
<td>0.8523</td>
<td>0.7584</td>
<td>0.9042</td>
<td>0.9386</td>
</tr>
<tr>
<td>Skew Student</td>
<td>0.7995</td>
<td>0.7458</td>
<td>0.9081</td>
<td>0.9100</td>
</tr>
<tr>
<td>Generalized Error</td>
<td>0.7908</td>
<td>0.8890</td>
<td>0.8673</td>
<td>0.9467</td>
</tr>
<tr>
<td>Skew Generalized Error</td>
<td>0.8013</td>
<td>0.8161</td>
<td>0.8531</td>
<td>0.9234</td>
</tr>
<tr>
<td>Normal Inverse Gaussian</td>
<td>0.7928</td>
<td>0.7133</td>
<td>0.8990</td>
<td>0.8873</td>
</tr>
<tr>
<td>Generalized Hyperbolic</td>
<td>0.7922</td>
<td>0.7448</td>
<td>0.8656</td>
<td>0.8918</td>
</tr>
<tr>
<td>Johnson SU</td>
<td>0.167228</td>
<td>0.206953</td>
<td>0.011718</td>
<td>0.002088</td>
</tr>
</tbody>
</table>

Table 11: P-Values of Sign Bias Test \( t \) for ARMA (1, 1) - EGARCH (2, 1) – Conditional Distribution Models

Almost all of the possible conditional distributions, except Johnson SU, satisfy the Sign Bias Tests for ARMA(1,1) – EGARCH(2,1). Among the ones that satisfy, some may even also satisfy the Adjusted Pearson Goodness-of-Fit Tests and, hence, be better than the Normal Conditional Distribution, which does not satisfy.
Table 12: P-Values of Adjusted Pearson Goodness-of-Fit Test t for ARMA (1, 1) - EGARCH (2, 1) - Conditional Distribution Models

<table>
<thead>
<tr>
<th>Conditional Distribution</th>
<th>Q-Statistics on Standardized Residuals</th>
<th>Q-Statistics on Standardized Squared Residuals</th>
<th>ARCH LM Tests</th>
<th>Alpha’s and Beta’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skew Student</td>
<td>0.2256</td>
<td>0.1907</td>
<td>0.4192</td>
<td>0.2683</td>
</tr>
<tr>
<td>Generalized Error</td>
<td>0.07304</td>
<td>0.30819</td>
<td>0.28965</td>
<td>0.21056</td>
</tr>
<tr>
<td>Skew Generalized Error</td>
<td>0.3344</td>
<td>0.3523</td>
<td>0.5695</td>
<td>0.2097</td>
</tr>
<tr>
<td>Normal Inverse Gaussian</td>
<td>0.2756</td>
<td>0.2761</td>
<td>0.6108</td>
<td>0.5756</td>
</tr>
<tr>
<td>Generalized Hyperbolic</td>
<td>0.1978</td>
<td>0.2605</td>
<td>0.5586</td>
<td>0.1912</td>
</tr>
</tbody>
</table>

There are five conditional distributions that satisfy both so they are kept for further evaluations. They are the Skew Student Distribution, the Generalized Error Distribution, the Skew Generalized Error Distribution, the Normal Inverse Gaussian Distribution and the Generalized Hyperbolic Distribution. To find out the conditional distributions, whose combinations with ARMA(1,1) – GARCH(1,1) are also good concerning the ARCH and GARCH parameters, the Q-Statistics and the ARCH LM Tests, the summary of the statistics are listed as below in the table.

Table 13: P-Values of Q-Statistics, ARCH LM Tests, and GARCH Parameters for ARMA (1, 1) - EGARCH (2, 1) - Conditional Distribution Models
Given the information in the Table 11, the Table 12 and the Table 13, the Skew Student Distribution, the Generalized Error Distribution, the Skew Generalized Error Distribution, the Normal Inverse Gaussian Distribution and the Generalized Hyperbolic Distribution are nice conditional distributions for the ARMA(1,1) – EGARCH(1,1) and the NASDAQ Computer Index Daily Closing Returns, \(r_t\)'s. In order to choose the best one, Information Criteria statistics provides good tools to compare the effectiveness and the efficiency.

<table>
<thead>
<tr>
<th>Conditional distribution</th>
<th>Aikake</th>
<th>Bayes</th>
<th>Shibata</th>
<th>Hannan-Quinn</th>
<th>Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skew Student</td>
<td>-5.3906</td>
<td>-5.3688</td>
<td>-5.3906</td>
<td>-5.3827</td>
<td>3</td>
</tr>
<tr>
<td>Generalized Error</td>
<td>-5.3877</td>
<td>-5.3679</td>
<td>-5.3877</td>
<td>-5.3806</td>
<td>5</td>
</tr>
<tr>
<td><strong>Skew Generalized Error</strong></td>
<td><strong>-5.3924</strong></td>
<td><strong>-5.3706</strong></td>
<td><strong>-5.3924</strong></td>
<td><strong>-5.3845</strong></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td>Normal Inverse Gaussian</td>
<td>-5.3906</td>
<td>-5.3688</td>
<td>-5.3907</td>
<td>-5.3828</td>
<td>2</td>
</tr>
<tr>
<td>Generalized Hyperbolic</td>
<td>-5.3904</td>
<td>-5.3666</td>
<td>-5.3904</td>
<td>-5.3818</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 14: AICs, BICs, SICs, HQICs and Ranks for ARMA (1, 1) - EGARCH (2, 1) - Conditional Distribution Models

According to the Information Criteria in the Table 14, since the ARMA (1, 1) - EGARCH (2, 1) - Skew Generalized Error Distribution Model has the smallest AIC, BIC, SIC and HQIC at the same time, it is the winner of the model selection. This is all due to the fact that it has good Optimal Parameters, Robust Standard Errors, Q-Statistics, ARCH LM Tests, Sign Bias Tests, Adjusted Pearson Goodness-of-Fit Tests, and Information Criterion at the same time.
5. Conclusion

The NASDAQ Computer Index over the daily closing returns, \((r_t)\), in the past twelve and a half years, is studied over its volatility.

As preliminary results, the default mean process ARMA (1,1) generates a residual time series with serial correlation removed.

Given the mean model residuals of constant mean 0, Alpha and Beta numbers of the GARCH Model are determined to be 2 and 1, in order to remove the heteroskedasticity patterns in the variance model residuals.

It is empirically acknowledged that the NASDAQ Composite Index has leverage effects and it is always better fitted by asymmetric GARCH Models than by symmetric variance models. This fact leads to the applications of EGARCH and GJR-GARCH Models in the \((r_t)\), because \((r_t)\) has many similarities with the major NASDAQ Composite Index. Also, the Sign Bias Tests determines the EGARCH(2,1) is better one than the GJR-GARCH(2,1).

The Adjusted Pearson Goodness-of-Fit Tests are the last step and they chooses good fitting alternative distribution. Though a different conditional distribution cannot make big difference to the whole package of the mean model, the GARCH variance model, and the distribution model, some good ones, like the Skew Generalized Error Conditional Distribution, can at least improve the whole GARCH Model at the end by increasing the P-Values for the Adjusted Pearson Goodness-of-Fit Tests.

Thanks to the rugarch R package, an effective and efficient model, the EGARCH(2,1) Variance Model with the ARMA(1,1) Mean Process and the Skew Generalized Error Conditional Distribution, is selected to model the \((r_t)\). It removes the heteroskedasticity and asymmetric problems from the variance model residuals, and generates a new series of
homoscedastic and symmetric variance model residuals. This application gives the NASDAQ Computer daily closing returns, $r_t$’s, a better estimation than what ARIMA Models or regular GARCH Models do. To sum up, it improves the NASDAQ Computer Sector Index’s future return predictions.
6. Reference


Appendix

Academic Vita
Ran Li

CONTACT INFORMATION
Department of Statistics, Harvard University, Cambridge, MA, 02138
Email: ranli02@fas.harvard.edu
Phone: 8147530321

EDUCATION
- AM (2014), Statistics, Harvard University
- BA (2012), Statistics (Honors), Pennsylvania State University
- BA (2012), Mathematics (Honors), Pennsylvania State University

Thesis: Applications of Asymmetric GARCH Models with Various Conditional Distributions: The Empirical Case of the NASDAQ Computer Index Daily Closing Returns
Adviser: Dr. Jason Morton

RELEVANT COURSES

RESEARCH EXPERIENCE
Department of Civil and Environmental Engineer, PSU University Park, PA
Extremes Values Analysis for an Extreme Environment (REU) Advisor: Dr. Michael Gooseff
Statistical Analyst Summer 2012
- Read into R and clean around 200 time series, mostly containing 400,000 observations or more, from the MCMLTER and USDA soils databases of NSF McMurdo Dry Valleys Long Term Ecological Research
- Produced standard statistics as well as histograms for majority of the time series
- Used Generalized Extreme Values (GEV) Distribution and Generalized Pareto (GP) Distribution in the extRemes R package to analyze the time series
• Assisted Dr. Michael Gooseff to characterize and classify the means and extremes of McMurdo Dry Valleys, Antarctica polar desert ecosystem and to provide a baseline against which to compare future events.
• Prepared the abstract for the American Geophysical Union’s 45th annual Fall Meeting (AGU), San Francisco, where 20,000 Earth and space scientists, educators, students and other leaders gather to present groundbreaking research and connect with colleagues

Department of Psychology, PSU University Park, PA
Undergraduate Research Assistant Fall 2010 - Spring 2011
• Utilized knowledge of Asian languages (Mandarin and Cantonese) to record conversations in videos where Asian-American children were playing blocks and talking with their mothers
• Acquainted myself with two coding language systems that categorized sentences spoken into specific groups

TEACHING EXPERIENCE
Department of Statistics, PSU University Park, PA
Lab Intern (Undergraduate Teaching Assistant) Fall 2010 - Spring 2011
• Worked with instructors and Graduate Teaching Assistants to hold lab activities for courses Introduction to Statistics and courses introduction to Biostatistics
• Developed skills in Minitab and provided helpful instructions to students from diverse academic backgrounds and statistics understanding levels

North Tutoring Center, State College High School University Park, PA
Private Mathematics Tutor/Private Physics Tutor Fall 2010 - Spring 2011
• Tutored students in different senior high grades and had witnessed rapid improvements
• Helped a senior high student with a Dyscalculia learning disability to navigate AP Calculus, and brought his D failing grade up to a mid C (sometimes middle B)

Information Technology Services, PSU University Park, PA
Technology Learning Assistant Spring 2011-Present
• Taught Penn State faculty members how to create and manage web pages, grade books, and so on

Department of Mathematics, PSU University Park, PA
Grader Spring 2011
• Graded and managed scores of homework assignments and quizzes for course Ordinary Differential Equations and Partial Differential Equations
• Worked an average of 12 hours/week while maintaining a full course load of 25 credits
PROFESSIONAL INVOLVEMENT
The Future Actuary, Sponsored by the Society of Actuary (SOA)  
Schaumburg, IL  
Student Representative  
Summer 2011 - Present
- Interviewed Professor Ron Gebhardtsbauer in the Department of Risk Management, PSU
- Based on the interview, wrote an article, “Advice from a Role Model” for The Future Actuary 2012 Summer Issue
- Attended the 2011 Editorial Board Meeting and offered ideas for the incoming publications

Penn State–Peking University Summer Program in Dynamical Systems  
Beijing, China  
PSU Student President  
Summer 2012
- Attended the lectures about Floquet Theory, Lyapunov-Poincare Theorem, Euler-Lagrange Equations, Hamilton's Equations, Poincare's Geometric Theorem, Lorenz Attractor and so on
- Organized extracurricular activities to the Forbidden City, the Great Wall of China, the Summer Palace, and so on

Church and Dwight Co., Inc.  
Princeton, NJ  
Externship  
May 2011
- Analyzed and compared the growth trends and the sales history of Kaboom and Lysol
- Presented the analysis result and provided an opinion on which sub-category was the best opportunity to grow the Kaboom brand

Statistics and Applied Mathematics Sciences Institute (SAMSI)  
Research Triangle Park, NC  
Participant of Undergraduate Workshops  
February 2011/May 2011
- Attended tutorials on mathematical and statistical methods that analyze functional data and epidemiological data
- Worked in team to accomplish an epidemiological scenario project in MATLAB and R

Pre Mathematics Advanced Study Semester (PMASS), PSU  
University Park, PA  
Participant  
Spring 2011
- Participated the immersive honors program, which consisted of Real Analysis honors course, Modern Geometry honors course, weekly seminars and biweekly colloquiums
- Performed the roles of team leader for the two courses’ final projects
- Collaborated with my classmate to record one Symbolic Dynamics colloquium lecture

The Shenzhen Development Bank, Guangzhou Branch  
Guangzhou, China  
Intern of Retail Department  
Summer 2010
- Sold financial products to diverse clients
- Gained competence in financial investments and management information system

HONORS AND AWARDS
- Penn State Research Experience for Undergraduates Stipend  
Summer 2012
• Penn State Schreyer Honors College Travel Grants  Summer 2012
• Penn State Eberly College of Science Travel Grants  Summer 2012
• Penn State Department of Mathematics Travel Grants  Summer 2012
• Penn State Department of Mathematics Internal Scholarship  Summer 2012
• Penn State College of Science Alumni Scholarship  Fall 2011
• The Future Actuary Editorial Board Meeting Travel Grants  Summer 2011
• SAMSI Undergraduate Workshop Travel Grants  May 2011
• SAMSI Undergraduate Workshop Travel Grants  Feb 2011
• Penn State Pre Mathematics Advanced Study Semester Fellowship  Spring 2011
• The National Society of Collegiate Scholar  Spring 2010
• Penn State President Freshmen Award  Spring 2009
• Dean’s list (6/6)  Fall 2009 – Spring 2012

LEADERSHIP
Welcome Week, PSU  University Park, PA
Welcome Week Captain  August 2010
• Collaborated with Welcome Week Leaders and the other Welcome Week Captains to assist over 7000 freshmen moving in campus

Fresh START Day of Service, PSU  University Park, PA
Team Leader  September 2010
• Collaborated with the 2010 Fresh START Executive Committee to organize over 700 freshmen volunteers in the Fresh START Day of Service
• Led the volunteers to clean THON mats and to get involved in the THON Dance Marathon

American Red Cross, Centre Community Chapter  State College, PA
Instructor  Spring 2010 - Fall 2010
• Got trained of First Aid/CPR/AED and instructor skills
• Taught and co-taught First Aid/CPR/AED courses to introduce the basic emergency measures to the community

Student Red Cross Club, PSU  University Park, PA
On-Site Coordinator  Spring 2010
• Organized campus blood drives to ensure comfortable and pleasant blood donating experience

SKILLS
Language
• Fluent in spoken and written English
- Fluent in spoken and written Chinese (Mandarin and Cantonese)

**Software**
- Intermediate: SAS, R, Minitab, LaTeX, Interact
- Basic: MATLAB, Visual Studio, BASIC, Word, Excel, PowerPoint, Access

**Validation by Educational Experience**
- VEE Economics
- VEE Applied Statistics

**SOA Exams**
- Probability Exam
- Financial Mathematics Exam

**Credential**
- SAS Certified Base Programmer for SAS 9 Credential