Summer 2015

Day of the week effect in returns and volatility of the S&P 500 sector indices

Juan Liu

DAY OF THE WEEK EFFECT IN RETURNS AND VOLATILITY

OF THE S&P 500 SECTOR INDICES

by

JUAN LIU

A THESIS

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN APPLIED MATHEMATICS

2015

Approved by

V.A. Samaranayake, Advisor
Gayla Olbricht
Gregory Gelles

ABSTRACT

Previous studies have shown that returns associated with the stock market or foreign exchange's futures show variations across the day of the week. On such study, that employs a modified GARCH model for estimation, shows that returns associated with the S&P 500 stock index is highest on Wednesday and lowest returns on Monday. The same study shows that volatility is highest on Fridays and lowest on Wednesdays. In this study we investigate if this day-of-the-week effect on returns and volatility is present in the different sectors that constitute the S&P 500 index. The data set used provides daily returns from February 2005 to February 2015 and is more recent than the data used for the original study on the S&P index. Results show mixed outcomes with some days showing higher returns or volatilities on certain days of the week depending on the sector.

Keywords: Day-of-the-week-effect, GARCH, Heteroscedasticity, S&P 500-index

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my advisor, Dr. V. A. Samaranayake, in the Department of Mathematics and Statistics at the Missouri University of Science and Technology, for his insight inspiration and guidance, for his patience, motivation, enthusiasm and immense knowledge. I first joined the university to do a doctoral degree in computer science, but was ambivalent about pursuing a Ph.D. in that field. During that difficult time when I was confused about my choices and future objectives, Dr. Samaranayake encouraged me make a selection based on my own aspirations and was willing to share his experience in academic research and career success. His guidance helped me in throughout this research and in writing this thesis. I could not have imagined having a better advisor and mentor for my master's study.

I also would like to thank the rest of my committee members, Dr. Gayla Olbricht and Dr. Gregory Gelles, for their encouragement and insightful comments.

Last but not least, I would like to thank my dear parents. I am very thankful for their support and continuous encouragement throughout my life.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

Statistical modeling of stock returns has had a long history. One of the early attempts at statistically modeling returns was by Kendall (1953). He analyzed twenty two time series consisting of asset prices observed on a weekly basis. He concluded that "The series looks like a 'wandering' one, almost as if once a week the Demon of Chance drew a random number from a symmetrical population of fixed dispersion and added it to the current price to determine next week's price," Kendall (1953, p. 13). Kendall was talking about price when he referred to a "wandering" series, but when he referred to a "random number" he was referring to returns. Returns of an asset are usually computed either as the ratio of the change in the price of a commodity to its previous price or as the difference of the natural logarithm of the current price and the previous price. Kendall's observations makes sense if the price of an asset or the log price at time t is the value of a random walk consisting of the sum of independent identically distributed *(i.i.d.)* random variables computed from the beginning of the series to time t. In such a case, the returns are independent identically distributed random variables. You would see a similar behavior even if the returns are uncorrelated random variables, which are normally referred to as white noise. This noise component can be considered as the noise component of the process and is sometimes referred to as the innovations.

Standard time series models, such as the Autoregressive Moving Average (ARMA) formulation, assume that the underlying error process is either strictly stationary or second order (weakly) stationary. That is that the mean and the variance (as well as the autocovariances) of the process remains constant over time. The above formulations also assume that the conditional variance of this error process is a constant. In other words it is assumed that the variance when conditioned on past values is homoscedastic. However, there are many empirical time series that display conditional heteroscedasticity. For example, financial time series such as stock returns and electricity prices show conditional variances that change over time. Stock returns in particular can show intermittent bursts of higher than normal volatility (variance) followed by calmer periods. One approach to modeling these types of time series is to use Autoregressive Conditional Heteroskedastic (ARCH) models proposed by Engle (1982) or the Generalized Conditional Heteroskedastic (GARCH) models suggested by Bollerslev (1986). These models allow conditional variance to change over time, with high volatility periods followed by periods of low volatility. The unconditional variance of ARCH and GARCH models, however, are constant over time.

Some authors, such as Cross (1973) contested the assumption that the mean returns would remain constant across the five days of the trading week. Others, such as Osborn and Smith (1989) as well as Harvey and Huang (1991), argued that the assumption of constant unconditional variance is violated by some empirical series. Of particular interest is a paper by Berument and Kiymaz (2001) who analyzed 6,409 daily observations from Standard and Poor's 500 (S&P 500) Index taken from January 3, 1973 through October 20, 1997. Their analysis showed volatility to be higher than normal on Fridays and lower than normal Wednesdays. In this study, daily returns from ten different sectors included in the S&P 500 Index are studied to determine if similar "day-of-the-week" effect exists in both means returns and their volatility in individual sectors and whether such patterns are consistent across sectors.

1.1. LITERATURE REVIEW

The day-of-the week effect can impact returns as well as volatility**.** The first study on the day-of-the-week effect on returns was carried out by Cross (1973). This study analyzed returns on the S&P 500 Index that covered the years 1953 through 1970. The findings indicate that the mean return on Friday is higher than that of Monday. French (1980) found a similar pattern on returns on the S&P 500 Index over the period 1953-1977. Gibbons and Hess analyzed 30 selected stocks from the Dow Jones Industrial Index and found negative returns for Mondays. Additional analysis was carried out by Keim and Stambaugh (1984), which found patterns similar to those found by the earlier studies.

Of particular interest to researchers was the Monday returns, which some suggested should be higher than for other days because of the gap that exist between Friday trading and Monday activities. For example French (1980) suggested that Monday returns should be higher than returns for other days. Other publications that investigated related issues are Gibbons and Hess (1981), Lakonishok and Levi (1982), and Rogalski (1984). In addition, Jaffe and Westerfield (1985) studied the day-of-the-week effect in stock markets in Australia, Canada, Japan, U.K. while and Solnik and Bousquet (1990) studied such effects for stocks traded in the Paris Bourse (a historic Paris stock exchange renamed Euronext Paris in 2000). The former study found the lowest returns for the Japanese and Australian stock markets to occur on Tuesdays. The latter study found negative returns on Tuesdays for the Paris market.

Investigations on the relationship between returns and on volatility were carried out by French, Schwert, and Stambaugh (1987) unusual stock market returns are negatively corrected with unexpected volatility changes. Campbell and Hentschel (1992) suggest that increase in volatility lowers stock prices. Others who studied the relationship between stock returns and volatility are: Baillie and DeGennaro (1990), Chan, Karolyi, and Stulz (1992), Glosten, Jagannathan and Runkle (1993), Corhay and Rad (1994), and Theodossiou and Lee (1995). These studies do not directly study the presence of a day-of-the-week effect on stock market volatility but looked at the relationship between stock price and volatility.

As mentioned earlier, Berument and Kiymaz (2001) found a day-of-the-week effect that increased volatility of Fridays and lowered it on Wednesdays. Other investigations also found such effects. For example, Harvey and Huang (1991) who studied interest rate and foreign exchange futures market found higher volatilities on Fridays while Ederington and Lee (1993) found such effects in the bond and stock markets. Choudhry (2000) studied data from seven Asian stock markets and found evidence of day-of-the-week effects on volatility, but these effects were not alike across the countries under study. Rodriguez (2012) who studied volatilities in the Latin American stock markets found Monday to have lower than normal volatility with Friday showing a higher than normal effects.

1.2. INTRODUCTION TO ARCH AND GARCH MODELS

The ARCH and GARCH models introduced by Engle (1982) and Bollerslev (1986) are defined in the following paragraphs. Following that, variations of these models that account for non-constant mean and unconditional volatility, such as that those due to day-of-the-week effect are given in Chapter 2.

Definition 1.2.1: A time series $\{\varepsilon_t : \ t=0, \ \pm 1, \ \pm 2, ...\}$ is said to be an Autoregressive Conditional Heteroskedastic Process of order q if it follows the formulation:

$$\varepsilon_t = \sigma_t e_t \text{, for } t=0, \ \pm 1, \ \pm 2, ..., \text{ where } e_t \sim i.i.d.\ N(0,1)$$

and
$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 \text{ for } t=0, \ \pm 1, \ \pm 2, ... \tag{1.2.1}$$

where $\sigma_t = \sqrt{\sigma_t^2}$, $\alpha_0 > 0$, $\alpha_j \geq 0$ for $j = 1, \ 2, ..., q$. The additional condition $\sum_{j=1}^{q} \alpha_j < 1$ is required for the time series $\{\varepsilon_t\}$ to be covariance stationary. Time series that satisfy the above condition are sometimes called by the acronym ARCH (q). The unconditional variance of a stationary ARCH (q) process can be easily derived, and is given in the following well-known theorem.

Theorem 1.2.1: Let the time series $\{\varepsilon_t : \ t=0, \ \pm 1, \ \pm 2, ...\}$ satisfy the conditions given under Definition 1.2.1. Then,

$$E\left[\varepsilon_t\right] = 0 \text{ for all } t=0, \ \pm 1, \ \pm 2, ... \text{ and}$$

$$E\left[\varepsilon_t^2\right] = \frac{1}{1-\sum_{j=1}^{q} \alpha_j} \text{ for all } t=0, \ \pm 1, \ \pm 2, ...$$

Proof: [This derivation is from Edirisinghe (2011).] First define $E\left[X_t \mid j<t\right]$ to denote the conditional expectation of any random variable X_t conditional on all its past values $X_{t-1}, \ X_{t-1}, \ X_{t-3}, \ ....$.

Note that, $E_{\mathsf{F}_t}(\varepsilon_t^2) = E(e_t^2 \sigma_t^2)$

$$= E(E(e_t^2\sigma_t^2 \mid j < t))$$

$$= E(E(e_t^2(\alpha_0 + \sum_{j=1}^{q}\alpha_j\varepsilon_{t-j}^2) \mid j < t))$$

$$= E(\alpha_0 + \sum_{j=1}^{q}\alpha_j\varepsilon_{t-j}^2)E(e_t^2 \mid j < t).$$

Now, $E(e_t^2 \mid j < t) = E(e_t^2) = 1$ because e_t are independent identically distributed random variables with variance equal to one.

This implies, $E_{\mathrm{F}_t}(\varepsilon_t^2) = \alpha_0 + \sum_{j=1}^{q}\alpha_j E(\varepsilon_{t-j}^2).$

Also, $E(\sigma_t^2) = \alpha_0 + \sum_{j=1}^{q}\alpha_j E(\varepsilon_{t-j}^2).$

Which implies $E(\varepsilon_t^2) = E(\sigma_t^2)$ and hence, $E(\varepsilon_{t-1}^2) = E(\sigma_{t-1}^2)$.

Now, $E_{\mathrm{F}_t}(\varepsilon_t^2) = \alpha_0 + \sum_{j=1}^{q}E(\varepsilon_{t-1}^2).$

Since $\{\varepsilon_t\}$ is covariance stationary, $E_{\mathrm{F}_t}(\varepsilon_t^2) = E_{\mathrm{F}_t}(\varepsilon_{t-1}^2).$

Now, $E_{\mathrm{F}_t}(\varepsilon_t^2) = \alpha_0 + \sum_{j=1}^{q}\alpha_j E(\varepsilon_t^2).$

Therefore, $E_{\mathrm{F}_t}(\varepsilon_t^2) = \dfrac{\alpha_0}{(1 - \sum_{j=1}^{q}\alpha_j)}.$

The definition of the GARCH process introduced by Bollerslev (1986) is as follows:

Definition 1.2.2: A time series $\{\varepsilon_t :\ t = 0,\ \pm 1,\ \pm 2, ...\}$ is said to be a Generalized Autoregressive Conditional Heteroskedastic Process or orders p and q if it follows the formulation:

$$\varepsilon_t = \sigma_t e_t \text{, for } t = 0, \ \pm 1, \ \pm 2, \ \ldots\text{, where } e_t \sim i.i.d.\ N(0,1)$$

and $$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \qquad \text{for} \qquad t = 0, \ \pm 1, \ \pm 2, \ \ldots \quad .$$

(1.2.2)

where $\sigma_t = \sqrt{\sigma_t^2}$, $\alpha_0 > 0$, $\alpha_j \geq 0$ for $j = 1, 2, \ldots, q$, and $\beta_i \geq 0$ for $i = 1, 2, \ldots, p$. The additional condition $\sum_{j=1}^{q} \alpha_j + \sum_{i=1}^{p} \beta_i < 1$ is required for the time series $\{\varepsilon_t\}$ to be covariance stationary. Time series that satisfy the above condition are sometimes called by the acronym GARCH (q, p). The unconditional variance of a stationary GARCH (q, p) process can be derived similar to that of the ARCH processes, and is given in the following well-known theorem.

Theorem 1.2.2: Let the time series $\{\varepsilon_t : \ t = 0, \ \pm 1, \ \pm 2, \ \ldots\}$ satisfy the conditions given under Definition 1.2.2. Then,

$$E[\varepsilon_t] = 0 \text{ for all } t = 0, \ \pm 1, \ \pm 2, \ \ldots \text{ and}$$

$$E[\varepsilon_t^2] = \frac{1}{1 - \left[\sum_{j=1}^{q} \alpha_j + \sum_{i=1}^{p} \beta_i \right]} \text{ for all } t = 0, \ \pm 1, \ \pm 2, \ \ldots .$$

Proof: The proof is similar to that of Theorem 1.2.1 and hence is not given here.

Even though many empirical time series that are modeled as a GARCH process satisfies the covariance stationarity condition $\sum_{j=1}^{q} \alpha_j + \sum_{i=1}^{p} \beta_i < 1$, in some cases this may not be the case. Lindner (2009) states that “For real data one often estimates parameters α_j and β_i such that $\sum_{j=1}^{q} \alpha_j + \sum_{i=1}^{p} \beta_i < 1$ is close to one, when assuming noise variance 1.”

He implies that modeling such processes as an Integrated GARCH (IGARCH) process where $\sum_{j=1}^{q}\alpha_j + \sum_{i=1}^{p}\beta_i = 1$. Such processes may still have strictly stationary solutions according to Lindner (2009).

2. MODELS WITH TIME VARING MEAN AND UNCONDITIONAL VOLATILITY

As you can see from the results of Theorems 1.2.1 and 1.2.2, the unconditional variance of ARCH and GARCH processes are both independent of the time index *t* and hence are constant over time. Therefore, the day-of-the week effects observed in empirical time series by many researchers cannot be modeled using the standard ARCH and GARCH models. This is because if for example Fridays consistently have higher volatility than other days, then the average volatility observed for Fridays across many years must be higher than similar averages obtained for other days. This means that the unconditional volatility for Fridays must be higher than that for other days.

In addition, the mean of a GARCH process is zero. Also, unconditionally, the process is uncorrelated. Both these properties can be a drawback when using a GARCH process to model a time series with a non-zero mean and whose terms are correlated. This can be easily overcome by fitting a non-zero mean ARMA process to the time series, but under the assumption that the error terms are GARCH. The mean of the ARMA process can be allowed to vary, say according to the day of the week by introducing dummy variables.

2.1. AN AUTOREGRESSIVE- GARCH MODEL

One way to introduce a non-constant unconditional variance is to use the formulation adopted by Choudhry (2000) as well as by Berument and Kiymaz (2001). In this formulation, the constant term α_0 found in the ARCH and GARCH models (1.2.1) and (1.2.2) respectively, is replaced by terms specific to each day. Thus, the modified GARCH model is as follows:

$$\varepsilon_t = \sigma_t e_t \text{, for } t = 0,\ \pm 1,\ \pm 2,\ \ldots\text{, where } e_t \sim i.i.d.\ N(0,1)$$

and
$$\sigma_t^2 = \sum_{k=1}^{5} d_k + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \text{ for } t = 0,\ \pm 1,\ \pm 2,\ \ldots. \quad (1.2.2)$$

where $\sigma_t = \sqrt{\sigma_t^2}$, $\alpha_0 > 0$, $\alpha_j \geq 0$ for $j = 1,\ 2,\ \ldots,\ q$, and $\beta_i \geq 0$ for $i = 1,\ 2,\ \ldots,\ p$, with the d_k representing a dummy variable for the k^{th} trading day of the week, k=1, 2, 3, 4, 5. The additional condition $\sum_{j=1}^{q} \alpha_j + \sum_{i=1}^{p} \beta_i < 1$ is required for the time series $\{\varepsilon_t\}$ to be covariance stationary.

Given the closing value X_t of a stock on day t, it is common to compute the return, R_t, for day t by $R_t = \ln(X_t / X_{t-1})$. The above ARCH and GARCH processes are zero mean processes because it can be shown easily that $E[\varepsilon_t] = 0$ for all values of t and this may be too restrictive to model the returns of a given stock. Researchers such as Berument and Kiymaz (2001) as well as Rodriguez (2012) extended this model to an Autoregressive Model (AR) with a mean that varies with the day-of-the-week, with errors that are a GARCH process given by (1.2.2). Their formulation for R_t, the return observed on day t, is given by

$$R_t = \sum_{k=1}^{5} \gamma_k d_k + \sum_{l=1}^{m} \phi_l R_{t-l} + \varepsilon_t \quad (2.1.1)$$

with
$$\varepsilon_t = \sigma_t e_t \text{, for } t = 0,\ \pm 1,\ \pm 2,\ \ldots\text{, where } e_t \sim i.i.d.\ N(0,1)$$

and $$\sigma_t^2 = \alpha_0 + \sum_{k=1}^{4}\delta_k d_k + \sum_{j=1}^{q}\alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p}\beta_i \sigma_{t-i}^2$$ for

$t = 0,\ \pm 1,\ \pm 2, \ldots.$

Observe that $\sigma_t = \sqrt{\sigma_t^2}$, $\alpha_0 > 0$, $\alpha_j \geq 0$ for $j = 1,\ 2, \ldots, q$, and $\beta_i \geq 0$ for $i = 1, 2, \ldots, p$, with the d_k representing a dummy variable for the k^{th} trading day of the week, k=1, 2, 3, 4, 5. The additional condition $\sum_{j=1}^{q}\alpha_j + \sum_{i=1}^{p}\beta_i < 1$ is required for the time series $\{\varepsilon_t\}$ to be covariance stationary. Only four of the five d_k terms are included in the intercept term of the GARCH portion of Equation (2.1.1) because including all five dummy variables together with the constant term α_0 will result in collinearity. All five dummy variables were, however, fitted in the regression portion of Equation (2.1.1) which has no intercept.

The above formulation will be used in this study to model the returns computed from the S&P 500 sector indices. One advantage of the above formulation is that it allows the modeling of returns as an autoregressive process and also account for the conditional heteroskedasticity of the error process. It also accounts for any day-of-the-week effect on the returns as well as on volatility. Another advantage is that this model can be fitted to data using existing software such as the Statistical Analysis System (SAS©).

2.2. OTHER ALTERNATIVE MODELS

While there are advantages to using the above model, there are other models proposed by researchers. Edirisinghe (2011) in his doctoral dissertation proposed several models. In brief his models assumed an underlying process $\{\varepsilon_t\}$ given by the GARCH model (1.2.2) but assumed that the observed process $\{R_t\}$ is given by $Y_t = \gamma(t)\varepsilon_t$ where

$\gamma(t)$ is a deterministic function that changes over time. Edirisinghe (2011) showed that the unconditional variance of this process equals

$$E\left[R_t^2\right]=\gamma(t)E\left[\varepsilon_t^2\right]=\gamma(t)\left\{\frac{1}{1-\left[\sum_{j=1}^{q}\alpha_j+\sum_{i=1}^{p}\beta_i\right]}\right\} \tag{2.2.1}$$

So if $\gamma(t)$ takes different values based on the day-of-the-week t fall into, then we have a process whose unconditional variance changes with time. A model similar to (2.2.1) above was independently proposed by Amado and Terasvirta (2013).

Another very similar approach to modeling day-of-the-week effects was implemented by Thilakaratne and Samaranayake (2013). In this formulation, the process $\{R_t\}$ is given by

$$R_t=\theta_{S(t)}\varepsilon_t \text{ for } t = 1, 2, \ldots, n \tag{2.2.2}$$

where $\theta_{s(t)}$ are constants and $s(t)$ takes values 1, 2, 3, 4, and 5 depending on which day-of-the-week t falls on. The authors assumed without loss of generality that $\theta_5=1$ and estimated the values of θ_i, I = 1, 2, 3, 4 by dividing the average of the returns for Monday through Thursday by the average of the Friday returns. Then they modeled the resulting scaled returns as a regular GARCH process. Their main aim was to obtain prediction intervals for returns and volatilities and the intervals they obtained using GARCH modeling were re-scaled by multiplying them by the estimated quantities $\hat{\theta}_i$, I = 1, 2, 3, 4. They did not, however, come up with a procedure to test the null hypothesis that $\theta_1=\theta_2=\theta_3=\theta_4=1$.

3. STANDARD AND POOR'S 500 STOCK INDEX AND THE DATA

The Standard and Poor's 500 (S&P 500) is based on the weighted stock prices of 500 large companies. The criteria for selection as one of the 500 companies include: (1) must be a U.S. Company, (1) have an unadjusted market capitalization of at or above $5.3 billion, (2) the ratio of annual dollar value traded to float adjusted market capitalization should be 1.0 or greater and trade a minimum of 250,000 shares in each of the six months leading up to the evaluations date, (3) at least 50% of outstanding shares must be available for trading, (4) have positive as-reported earnings over the most recent quarter, (5) initial public offerings should be seasons for 6 to 12 months before being considered for addition to the index, (6) consist of highly tradable common stocks, with active and deep markets (quoted from S&P Dow Jones Indices: Index Methodology (2015)). Companies listed in the S&P 500 can be deleted if they no longer meet the above criteria but violations of a temporary nature may not result in deletion. The method of calculating the index and the mathematical details of determining the weights assigned to each company in the index is quite complicated and will not be discussed here. These details can be found at the website www.spdji.com.

3.1. DESCRIPTION OF SECTORS OF S&P 500 INDEX

The S&P 500 index consist of companies that can be broadly categorized into ten sectors: (1) Consumer Discretionary, (2) Consumer Staples, (3) Energy, (4) Financials, (5) Health Care, (6) Industrials, (7) Materials, (8) Technology, (9) Telecommunications Services, and (10) Utilities. Based on this standard, the above sectors consist of the industries given in Table 3.1 given below.

Table 3.1. List of Industries Belonging to S&P 500 Sectors

Sector	**Industry**
Consumer Discretionary	Auto Components, Automobiles, Household Durables, Leisure Equipment & Products, Textiles Apparel & Luxury Goods, Hotels, Restaurants & Leisure, Diversified Consumer Services, Media, Distributors, Internet and catalog Retail, Multiline Retail, Specialty Retail
Consumer Staples	Food staples and Retailing, Beverages, Food Products, Tobacco, Household products, Personal Products
Energy	Energy Equipment & Services, Oil, Gas, & Consumable Fuels
Financials	Commercial banks, Thrift & Mortgage Finance, Diversified financial services, Consumer Finance, Capital markets, Insurance, Real Estate (discontinued effective 04/30/2006), Real Estate Investment Trusts, Real Estate management & Development
Healthcare	Healthcare Providers & Services, Healthcare Equipment & Supplies, Healthcare Technology, Biotechnology, Pharmaceuticals, Life Sciences Tools & services
Industrials	Aerospace & Defense, Building Products, Construction & engineering, Electrical Equipment, Industrial Conglomerates, Machinery, Trading companies & Distributors, commercial services & Supplies, Professional Services, Air Freight & Logistics, Airlines, Marine, Road & Rail, Transportation Infrastructure
Information Technology	Internet Software & Services, IT Services, Software, Communications Equipment, Computers & Peripherals, Electronic Equipment & Components, Office Electronics, Semiconductor Equipment and Products (discontinued effective 04/30/2003), Semiconductors & Semiconductor Equipment
Materials	Chemicals, Construction Materials, Containers & Packaging, Metals & Mining, Paper & Forest Products
Telecommunications Services	Diversified Telecommunication Services, Wireless Telecommunication Services
Utilities	Electric Utilities, Gas Utilities, Multi-Utilities, Water Utilities, Independent Power Producers & Energy Traders

It is important to note that sometimes financial analysts consider Consumer Staples and Discretionary Sectors as one. Also some combine Materials and Industrial sectors. The ten-sector classification given above is defined based on the Global Industry Classification Standard (GICS®) which was jointly developed by Standard and Poor's and MSCI Barra in 1999 (S&P Indices (2008)).

3.2. DESCRIPTION OF DATA

The price data for each sector was obtained from the website http://us.spindices.com/indices/equity/sp-500. Sector breakdowns can be obtained from the same site. The index data provides prices computed using total returns, which include dividends and based on total net returns, which does not count dividends. The analysis conducted in this research used total net returns series. The website also provides data from other indices such as S&P 100, S&P Small Caps 600, S&P 900, S&P 1000 and S&P Composite 1500.

The graphs of the returns for each sector over a ten-year period from February 15, 2005 through February 12, 2015 are given in Figures 3.1 through 3.10. Note that the horizontal axis is labeled starting at one through 2,517 to reflect the 2,517 returns computed from 2,518 prices. Note that the return for the first day in the price series, namely February 14, 2005, could not be computed because the price of the index for the previous day was not available in the data set. Since the 2008/2009 financial crisis affected all stocks in some way or another, the behavior of the returns during that time may be of interest. September 2, 2008 corresponds to data point 894 (t=894). October 1, 2008 corresponds to t=915 and the corresponding t value for December 31, 2008 is 978.

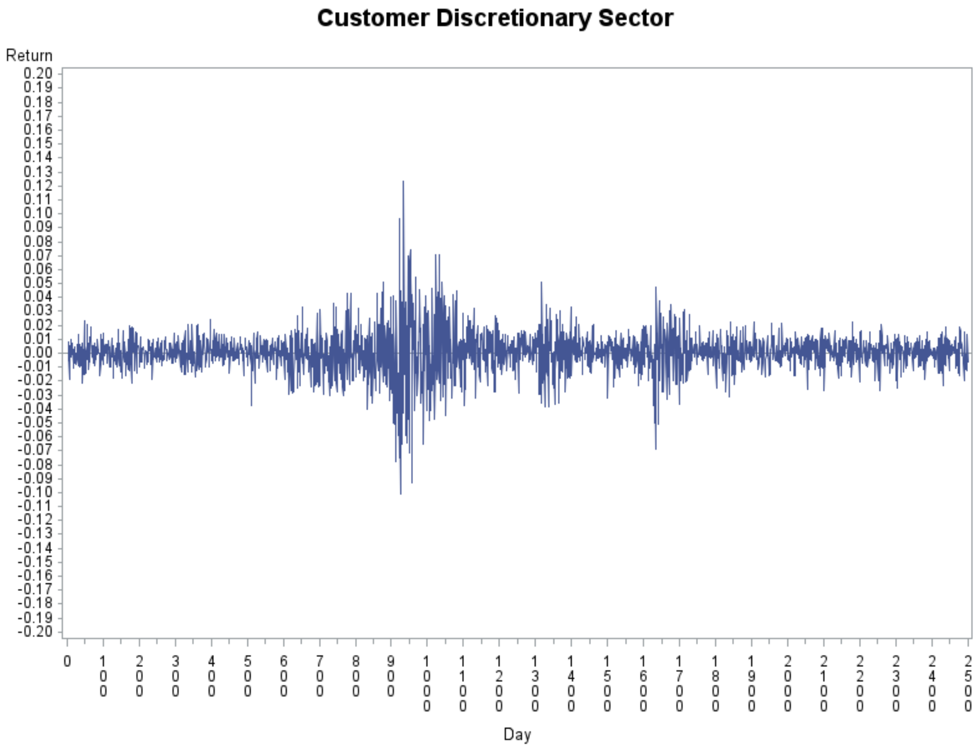

Figure 3.1. The Plot of Return by Time for Customer Discretionary Sector

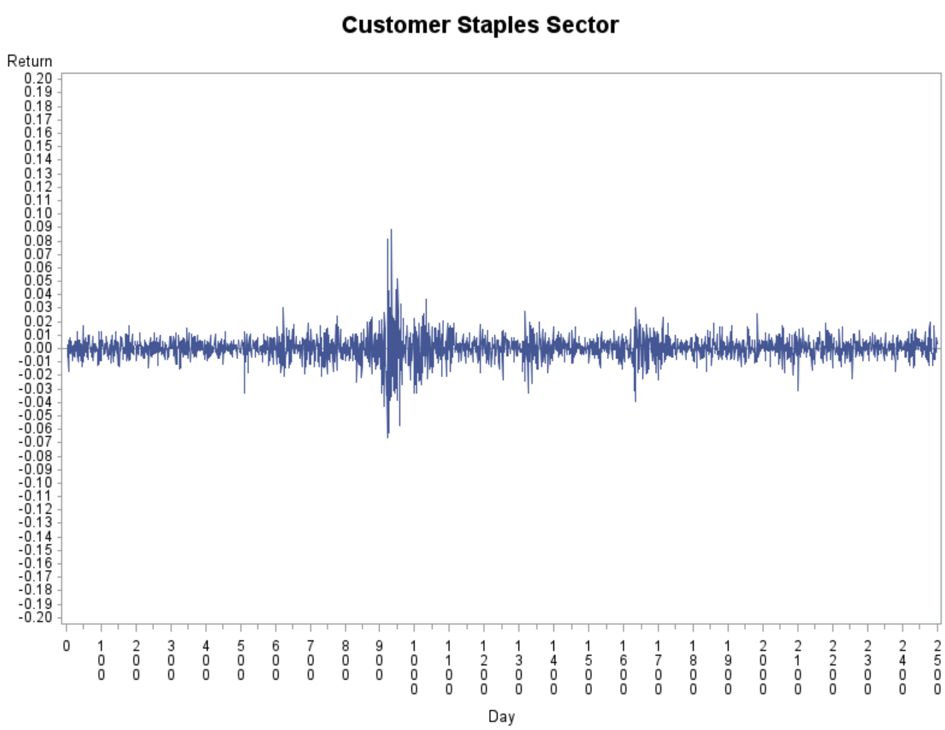

Figure 3.2. The Plot of Return by Time for Customer Staples Sector

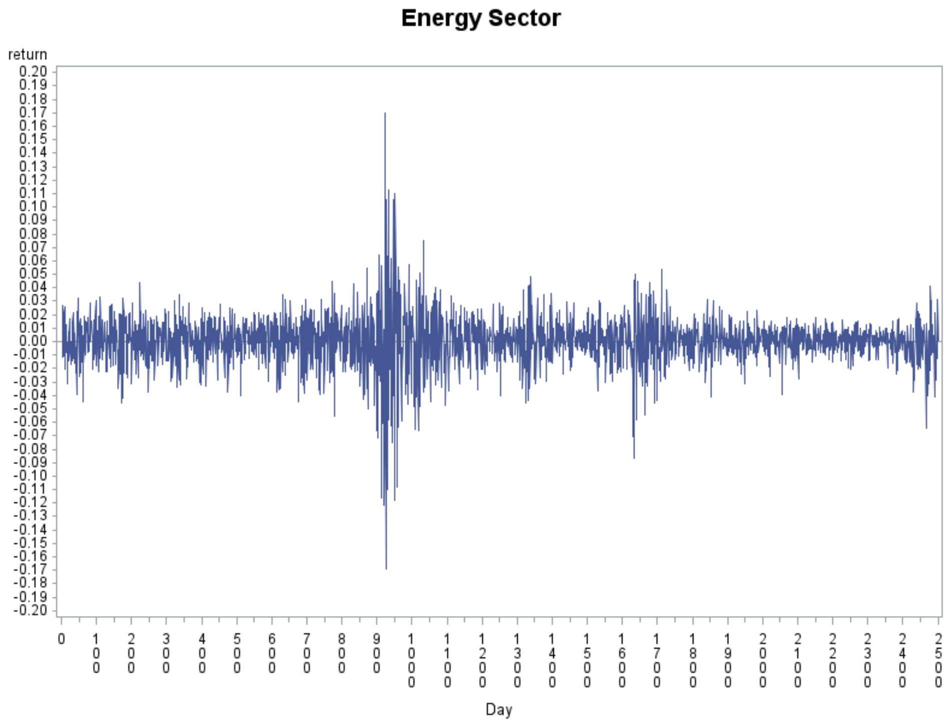

Figure 3.3. The Plot of Return by Time for Energy Sector

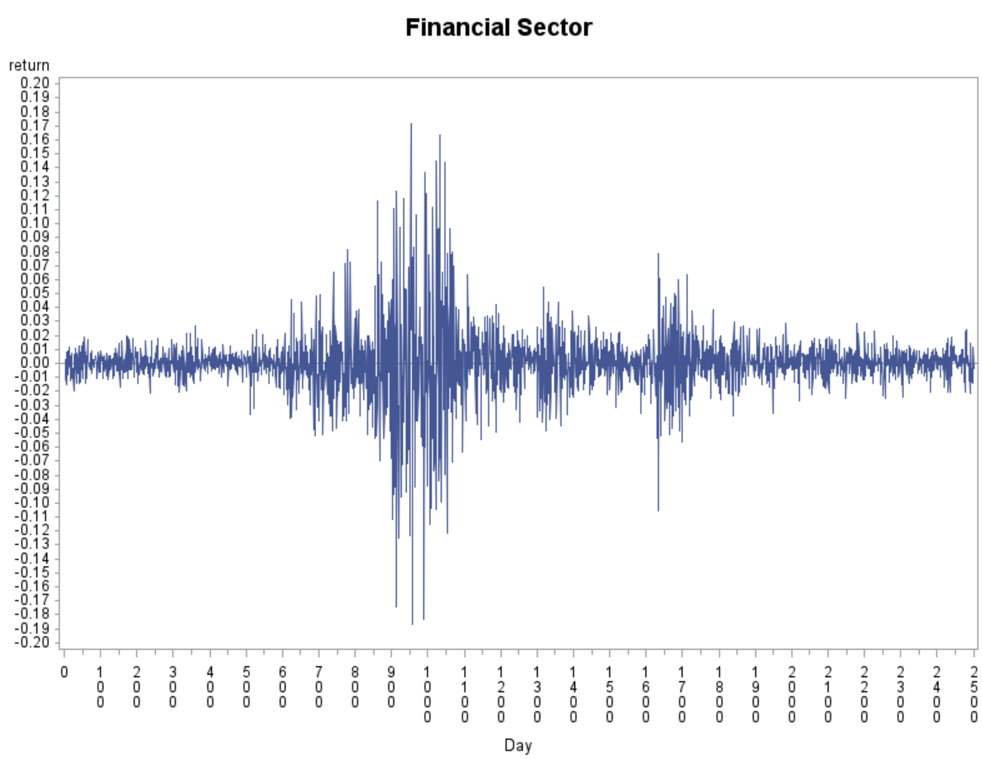

Figure 3.4. The Plot of Return by Time for Financial Sector

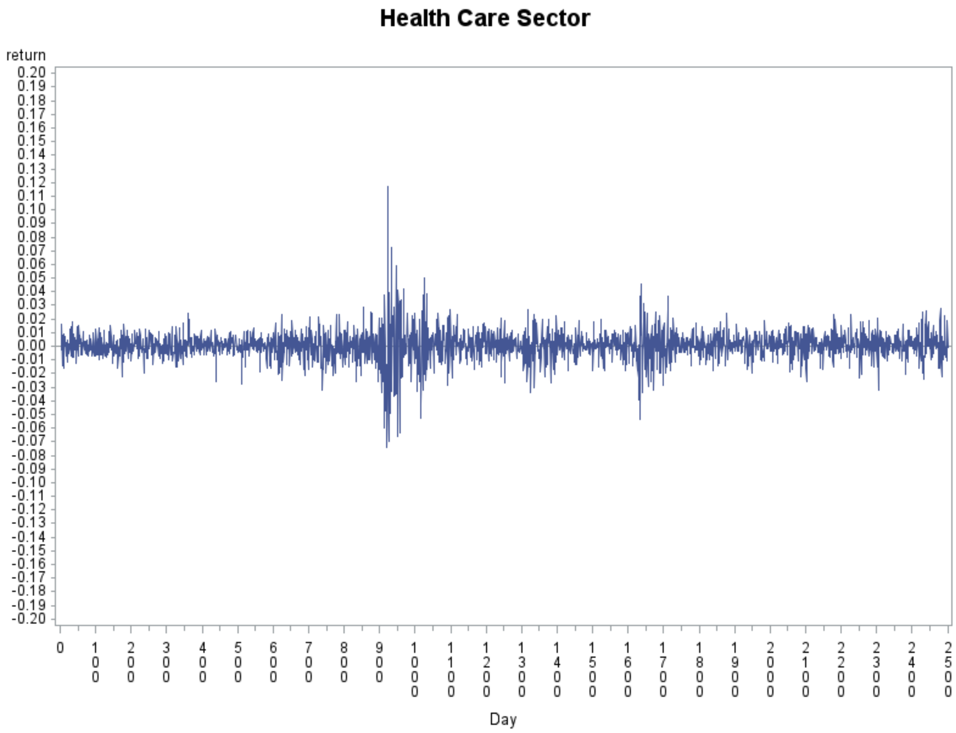

Figure 3.5. The Plot of Return by Time for Health Care Sector

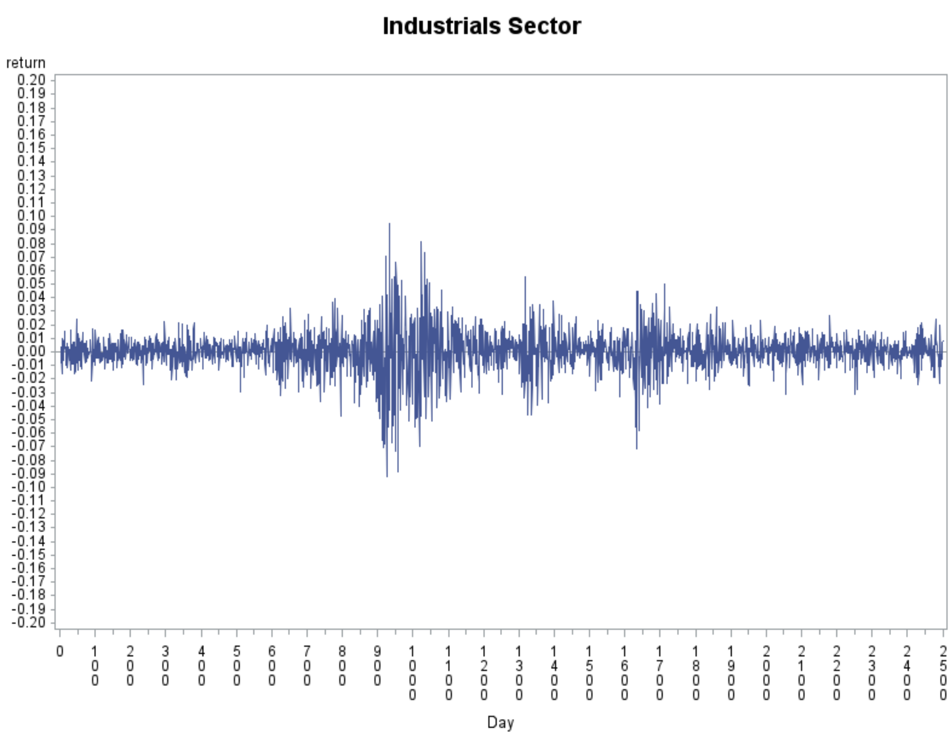

Figure 3.6. The Plot of Return by Time for Industrials Sector

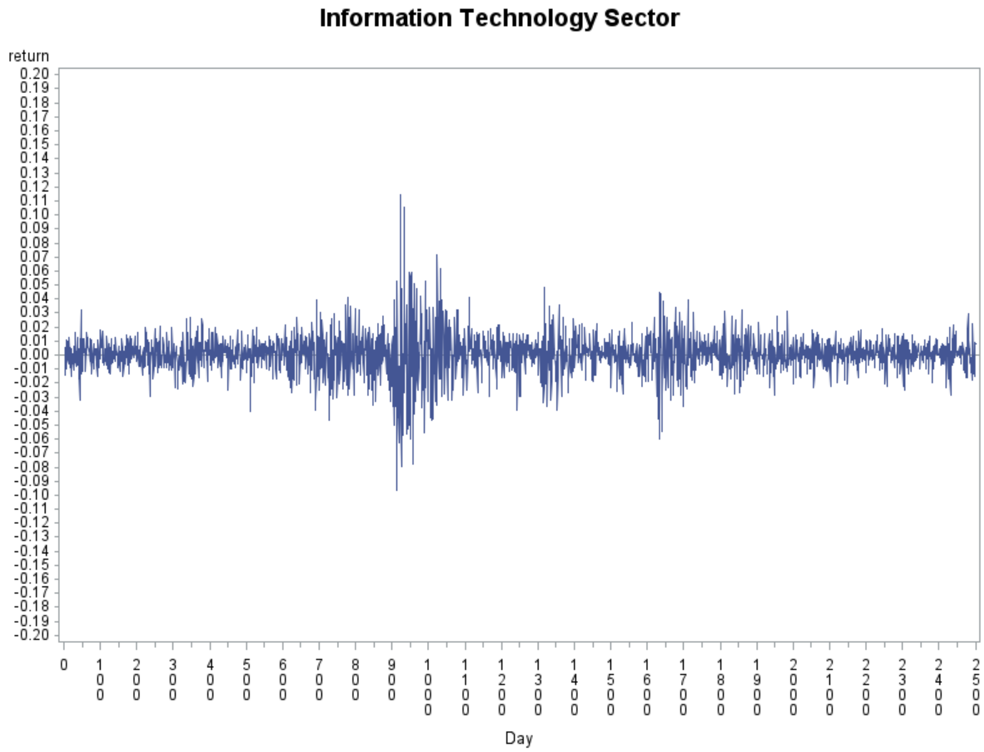

Figure 3.7. The Plot of Return by Time for Information Technology Sector

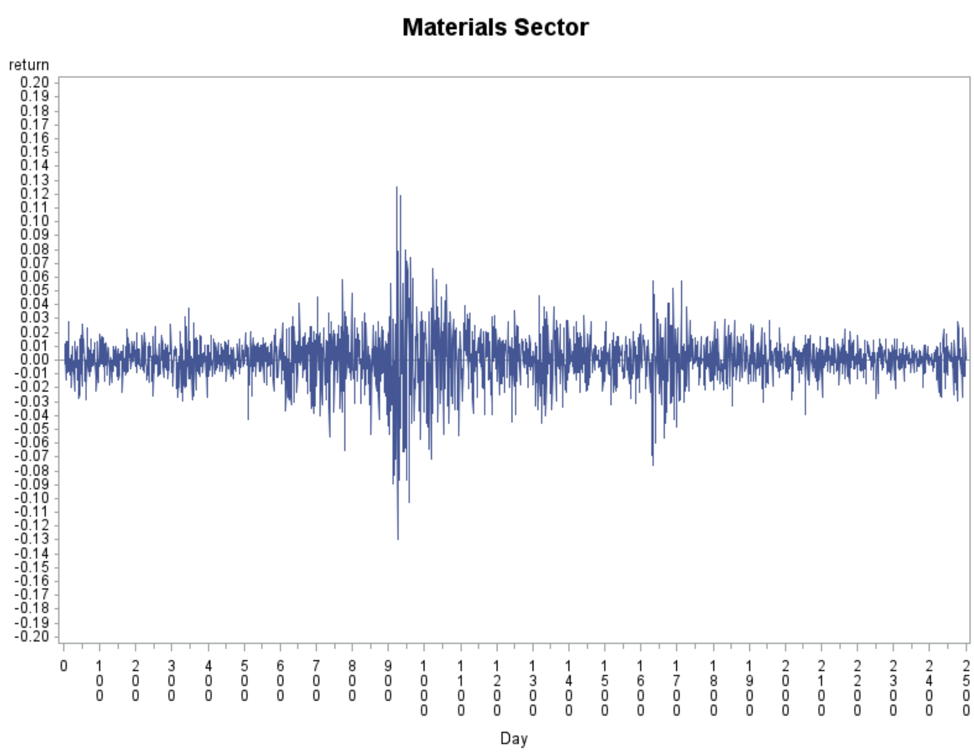

Figure 3.8. The Plot of Return by Time for Materials Sector

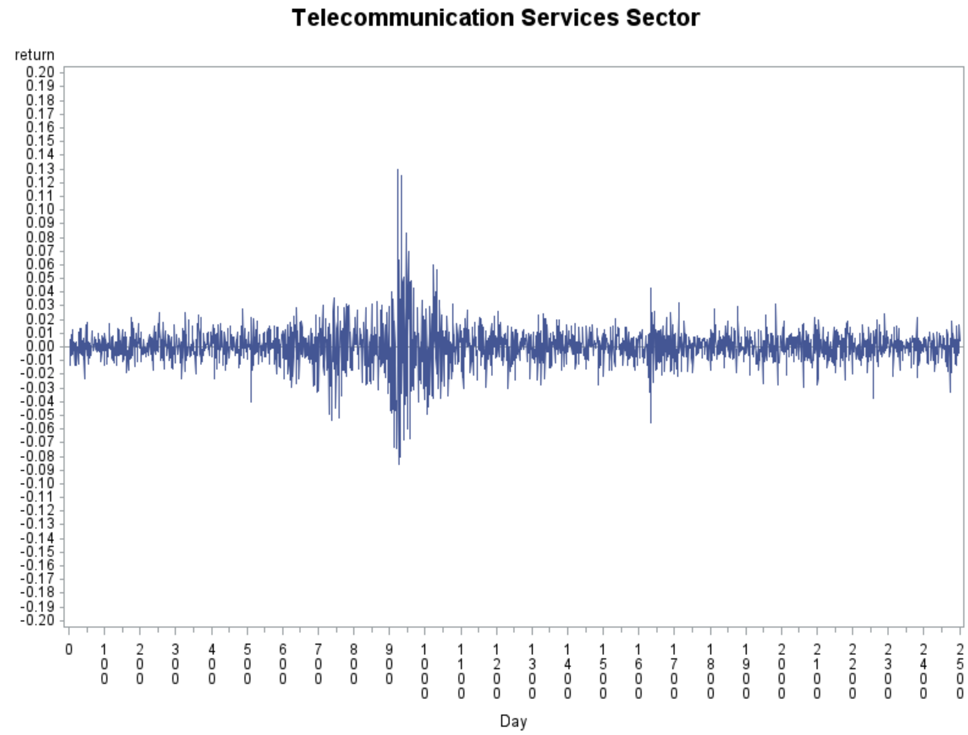

Figure 3.9. The Plot of Return by Time for Telecommunication Services Sector

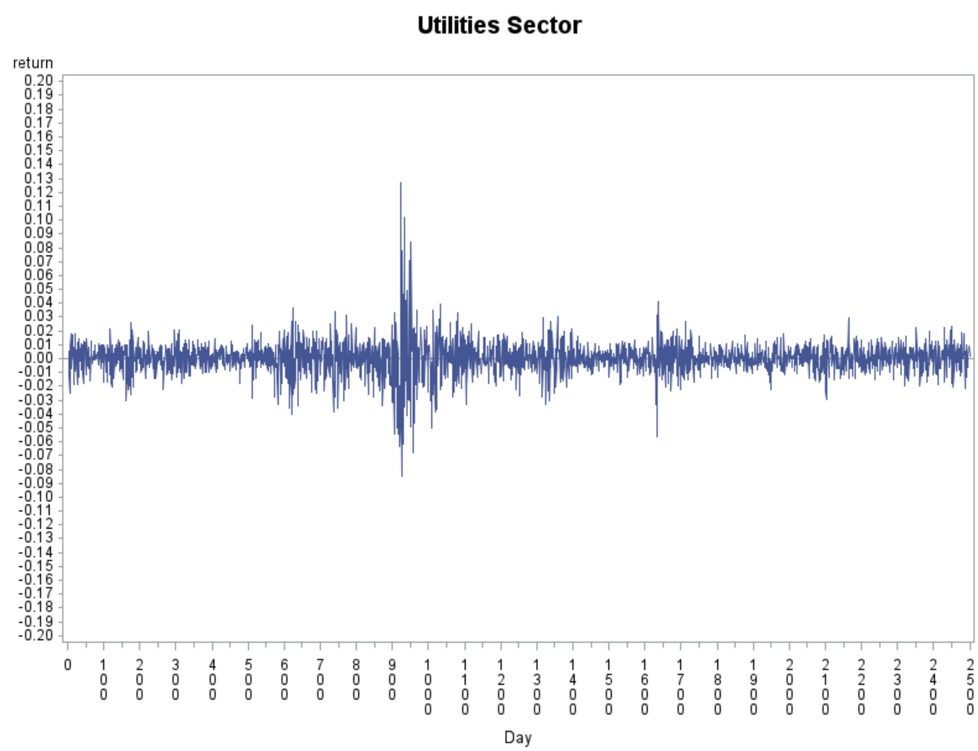

Figure 3.10. The Plot of Return by Time for Utilities Sector

The observation 1,000 corresponds to February 3, 2009. As expected, this period shows very high volatility across all sectors. For the Consumer Discretionary spending Sector, there are two other periods of smaller but yet prominent period of volatility centered around observation number 1,300 (April 14, 2010) and observation number 1,625 (July 27, 2011). Volatility levels seem to return to pre-2008 levels after observation number 1,750 (January 25, 2012). A similar pattern is observed for other sectors as well.

4. STATISTICAL MODELING OF S&P 500 SECTOR DATA AND RESULTS

The data obtained from the S&P website consisted of the date and the ending price for each sector index for that day based on total as well as net returns. The data set for each sector was first pre-processed to include the day of the week using an algorithm that used the calendar date to determine the day. The returns, R_t for day *t* was computed using the formula $R_t = \ln(P_t) - \ln(P_{t-1})$, where P_t is the price for day *t*.

4.1. THE MODELING PROCEDURE

The volatility was modeled using the Autoregressive-GARCH formulation given in Equation Set (2.1.1). The AUTOREG Procedure available in SAS (Version 9.4) was employed to carry out the model fitting. The conditions $\sum_{j=1}^{q} \alpha_j + \sum_{i=1}^{p} \beta_i < 1$, which is sufficient to ensure the covariance stationary assumption was imposed and the assumption $e_t \sim i.i.d.\, N(0,1)$ was made for the underlying innovations e_t that drive the GARCH process. In addition, the orders of the GARCH process was assumed to be *p*=1 and *q*=1 as is commonly done. Inspection of the Akaike information criterion (AIC) and the corrected Akaike information criterion (AICC) showed that assuming the e_t to be independently distributed as *t* random variables gave a better fit except for one sector. Note that the AUTOREG procedure in SAS automatically determines the degrees of freedom associated with the *t*-distribution.

Fitting the full model created estimability problems because the model was overparameterized. Therefore, a step-by-step approach was employed to do the modeling. First Model (2.1.1) was fitted without the GARCH component. That is, the error terms were assumed to be conditionally homoscedastic. Then the insignificant terms in the model $R_t = \sum_{k=1}^{5} \gamma_k d_k + \sum_{l=1}^{m} \phi_l R_{t-l} + \varepsilon_t$ were eliminated using significance level 0.05 as the cut-off criteria. This elimination was done one term at a time, with the most insignificant term (that with the highest *p*-value) considered for eliminated first. When two terms had p-values close to one another, each of the terms were eliminated in two separate runs and the AIC values for each model were compared. The elimination that reduces the AIC by the most amount was then selected.

Once the model was reduced in this manner, the GARCH portion $\sigma_t^2 = \alpha_0 + \sum_{k=1}^{4} \delta_k d_k + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$ of the model was added to the remaining Autoregressive (AR) part. The terms $\sum_{k=1}^{4} \delta_k d_k$ were introduced into the model using the HETERO command available in SAS. Then the dummy variables d_k that were not significant at 0.05 level were eliminated. Fitting of these dummy variables sometimes caused identification problems. Therefore, these terms were fitted one at a time. First the significant term that reduced the AIC by the most amount was fitted. Then another term was considered for inclusion using the significance level and AIC value as criteria.

4.2. DATA ANALYSIS AND RESULTS

Results from the above described modeling process are listed in Tables 4.2 through 4.11. Three important conclusions can be made from the results. One fact the results revealed is that the sum of the ARCH and GARCH terms (i.e. $\alpha_j + \beta_i$) is very close to one. Therefore, as Lindner (2009) suggested, fitting an IGARCH model may be more appropriate. Results from the IGARCH models are given in Tables 4.12 through Table 4.21. Results of the stationary GARCH analysis are reported in Table 4.1.

Table 4.1. Days of the Week with Significant Differences in Returns and Volatility

Sector		**Day of the Week**				
		Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Customer Discretionary	Percentage Change in Return			12.0482%	10.2609%	
	Percentage Change in Volatility		20.2908%			
Customer Staples	Percentage Change in Return		10.6560%	11.1863%	13.0259%	
	Percentage Change in Volatility					
Energy	Percentage Change in Return		9.7971%			11.2659%
	Percentage Change in Volatility					
Financial	Percentage Change in Return		6.5369%	9.2996%		
	Percentage Change in Volatility		5.4219%			
Health Care	Percentage Change in Return		14.1262%	11.7858%	13.3321%	
	Percentage Change in Volatility		16.1218%			
Industrials	Percentage Change in Return			10.0121%	9.8537%	9.6108%
	Percentage Change in Volatility		14.4586%			
Information Technology	Percentage Change in Return		11.1687%	16.9764%	9.2860%	
	Percentage Change in Volatility					
Materials	Percentage Change in Return			12.6571%		13.3500%
	Percentage Change in Volatility		6.6704%			
Telecommunication Services	Percentage Change in Return				10.0708%	
	Percentage Change in Volatility		29.3133%			
Utilities	Percentage Change in Return		12.9303%			12.0616%
	Percentage Change in Volatility					

Note: Percent change in returns computed as the ratio of change in return to mean total return multiplies by 100; percent change in volatility is computes as 100 times the coefficient of the respective dummy variable divided by the unconditional volatility.

Table 4.2. Analysis Results for Customer Discretionary Sector – Stationary Model

<table>
<tr><th colspan="6">Customer Discretionary Sector</th></tr>
<tr><th colspan="6">Stationary GARCH Estimates</th></tr>
<tr><td colspan="2">SSE</td><td colspan="2">0.51392693</td><td>Observations</td><td>2516</td></tr>
<tr><td colspan="2">MSE</td><td colspan="2">0.0002043</td><td>Uncond Var</td><td>0.00003731</td></tr>
<tr><td colspan="2">Log Likelihood</td><td colspan="2">7802.12058</td><td>Total R-Square</td><td>0.0006</td></tr>
<tr><td colspan="2">SBC</td><td colspan="2">-15549.428</td><td>AIC</td><td>-15590.241</td></tr>
<tr><td colspan="2">MAE</td><td colspan="2">0.00947821</td><td>AICC</td><td>-15590.197</td></tr>
<tr><td colspan="2">MAPE</td><td colspan="2">114.898256</td><td>HQC</td><td>-15575.429</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td>Normality Test</td><td>157.0012</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td>Pr > ChiSq</td><td><.0001</td></tr>
<tr><th colspan="6">Parameter Estimates</th></tr>
<tr><th>Variable</th><th>DF</th><th>Estimate</th><th>Standard Error</th><th>t Value</th><th>Approx Pr > |t|</th></tr>
<tr><td>DW</td><td>1</td><td>0.001146</td><td>0.000447</td><td>2.56</td><td>0.0104</td></tr>
<tr><td>DH</td><td>1</td><td>0.000976</td><td>0.000395</td><td>2.47</td><td>0.0135</td></tr>
<tr><td>AR10</td><td>1</td><td>-0.0452</td><td>0.0198</td><td>-2.29</td><td>0.0220</td></tr>
<tr><td>ARCH0</td><td>1</td><td>5.2577E-7</td><td>3.4644E-7</td><td>1.52</td><td>0.1291</td></tr>
<tr><td>ARCH1</td><td>1</td><td>0.0864</td><td>0.008548</td><td>10.11</td><td><.0001</td></tr>
<tr><td>GARCH1</td><td>1</td><td>0.8995</td><td>0.009454</td><td>95.15</td><td><.0001</td></tr>
<tr><td>HET DT</td><td>1</td><td>7.5705E-6</td><td>1.9486E-6</td><td>3.89</td><td>0.0001</td></tr>
</table>

Table 4.3. Analysis Results for Customer Staples Sector – Stationary Model

Customer Staples Sector			
Stationary GARCH Estimates			
SSE	0.19502292	**Observations**	2516
MSE	0.0000775	**Uncond Var**	0.0000645
Log Likelihood	8876.90851	**Total R-Square**	0.0159
SBC	-17675.513	**AIC**	-17733.817
MAE	0.00596693	**AICC**	-17733.729
MAPE	121.396521	**HQC**	-17712.657
		Normality Test	408.4783
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.000643	0.000258	2.49	0.0129	
DW	1	0.000675	0.000263	2.57	0.0103	
DH	1	0.000786	0.000252	3.11	0.0019	
AR1	1	0.0731	0.0213	3.43	0.0006	
AR4	1	0.0493	0.0204	2.41	0.0158	
AR5	1	0.0575	0.0204	2.83	0.0047	
ARCH0	1	1.5887E-6	4.132E-7	3.84	0.0001	
ARCH1	1	0.1038	0.0148	7.02	<.0001	
GARCH1	1	0.8716	0.0180	48.41	<.0001	
TDFI	1	0.1212	0.0179	6.77	<.0001	Inverse of t DF

Table 4.4. Analysis Results for Energy Sector – Stationary Model

Energy Sector			
Stationary GARCH Estimates			
SSE	0.83431705	**Observations**	2516
MSE	0.0003316	**Uncond Var**	0.00034906
Log Likelihood	7098.88293	**Total R-Square**	0.0078
SBC	-14142.953	**AIC**	-14183.766
MAE	0.01235697	**AICC**	-14183.721
MAPE	117.826541	**HQC**	-14168.954
		Normality Test	220.9218
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.001214	0.000531	2.29	0.0223	
DF	1	0.001396	0.000562	2.48	0.0131	
AR1	1	0.0466	0.0212	2.20	0.0279	
ARCH0	1	1.9756E-6	6.952E-7	2.84	0.0045	
ARCH1	1	0.0762	0.0103	7.43	<.0001	
GARCH1	1	0.9181	0.0107	85.95	<.0001	
TDFI	1	0.0939	0.0164	5.72	<.0001	Inverse of t DF

Table 4.5. Analysis Results for Financial Sector – Stationary Model

Financial Sector			
Stationary GARCH Estimates			
SSE	1.28722734	**Observations**	2516
MSE	0.0005116	**Uncond Var**	0.00010999
Log Likelihood	7308.36294	**Total R-Square**	0.0217
SBC	-14530.591	**AIC**	-14594.726
MAE	0.01308942	**AICC**	-14594.62
MAPE	118.00104	**HQC**	-14571.45
		Normality Test	287.7646
		Pr > ChiSq	<.0001

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DT	1	0.000866	0.000402	2.16	0.0310	
DW	1	0.001232	0.000419	2.94	0.0033	
AR1	1	0.0828	0.0209	3.95	<.0001	
AR4	1	0.0422	0.0200	2.11	0.0347	
AR5	1	0.0515	0.0199	2.60	0.0094	
AR6	1	0.0502	0.0196	2.57	0.0102	
ARCH0	1	1.154E-8	$<10^{-20}$	-	<.0001	
ARCH1	1	0.0935	0.0108	8.67	<.0001	
GARCH1	1	0.9064	0.0108	84.10	<.0001	
TDFI	1	0.1564	0.0191	8.20	<.0001	Inverse of t DF
HET DT	1	5.9636E-6	1.7904E-6	3.33	0.0009	

Table 4.6. Analysis Results for Health Care Sector – Stationary Model

Health Care Sector			
Stationary GARCH Estimates			
SSE	0.28305086	**Observations**	2516
MSE	0.0001125	**Uncond Var**	0.00004369
Log Likelihood	8399.47628	**Total R-Square**	0.0122
SBC	-16720.648	**AIC**	-16778.953
MAE	0.00714023	**AICC**	-16778.865
MAPE	129.471419	**HQC**	-16757.792
		Normality Test	176.1272
		Pr > ChiSq	<.0001

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DT	1	0.001014	0.000334	3.04	0.0024	
DW	1	0.000846	0.000339	2.50	0.0125	
DH	1	0.000957	0.000336	2.85	0.0044	
AR1	1	0.0518	0.0217	2.39	0.0169	
AR2	1	0.0444	0.0206	2.16	0.0306	
ARCH0	1	1.1249E-6	7.9629E-7	1.41	0.1578	
ARCH1	1	0.1061	0.0148	7.19	<.0001	
GARCH1	1	0.8682	0.0174	49.88	<.0001	
TDFI	1	0.1343	0.0227	5.90	<.0001	Inverse of t DF
HET DT	1	7.0436E-6	3.5266E-6	2.00	0.0458	

Table 4.7. Analysis Results for Industrial Sector – Stationary Model

Industrials Sector			
Stationary GARCH Estimates			
SSE	0.50423657	**Observations**	2516
MSE	0.0002004	**Uncond Var**	0.0000444
Log Likelihood	7801.08261	**Total R-Square**	0.0005
SBC	-15531.691	**AIC**	-15584.165
MAE	0.00943457	**AICC**	-15584.093
MAPE	119.170277	**HQC**	-15565.121
		Normality Test	214.9902
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DW	1	0.000948	0.000385	2.46	0.0139	
DH	1	0.000933	0.000379	2.46	0.0138	
DF	1	0.000910	0.000391	2.33	0.0200	
AR5	1	0.0408	0.0199	2.05	0.0403	
ARCH0	1	3.6576E-7	3.9593E-7	0.92	0.3556	
ARCH1	1	0.0890	0.0124	7.16	<.0001	
GARCH1	1	0.9028	0.0125	72.26	<.0001	
TDFI	1	0.1261	0.0213	5.93	<.0001	Inverse of t DF
HET DT	1	6.4196E-6	2.3642E-6	2.72	0.0066	

Table 4.8. Analysis Results for Information Technology Sector – Stationary Model

Information Technology Sector			
Stationary GARCH Estimates			
SSE	0.47743077	**Observations**	2516
MSE	0.0001898	**Uncond Var**	0.0001544
Log Likelihood	7750.88706	**Total R-Square**	0.0038
SBC	-15439.131	**AIC**	-15485.774
MAE	0.00935689	**AICC**	-15485.717
MAPE	127.739573	**HQC**	-15468.846
		Normality Test	201.6371
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.001050	0.000432	2.43	0.0152	
DW	1	0.001596	0.000426	3.75	0.0002	
DH	1	0.000873	0.000419	2.09	0.0370	
AR10	1	-0.0420	0.0192	-2.18	0.0290	
ARCH0	1	2.3011E-6	6.7565E-7	3.41	0.0007	
ARCH1	1	0.0821	0.0119	6.88	<.0001	
GARCH1	1	0.9030	0.0136	66.22	<.0001	
TDFI	1	0.1261	0.0208	6.05	<.0001	Inverse of t DF

Table 4.9. Analysis Results for Materials Sector – Stationary Model

Materials Sector			
Stationary GARCH Estimates			
SSE	0.70997991	**Observations**	2516
MSE	0.0002822	**Uncond Var**	0.00010155
Log Likelihood	7323.56026	**Total R-Square**	.
SBC	-14584.477	**AIC**	-14631.121
MAE	0.01135964	**AICC**	-14631.063
MAPE	119.106467	**HQC**	-14614.192
		Normality Test	217.4533
		Pr > ChiSq	<.0001

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DW	1	0.001443	0.000486	2.97	0.0030	
DF	1	0.001522	0.000512	2.97	0.0029	
AR4	1	0.0416	0.0205	2.03	0.0425	
ARCH0	1	3.8518E-7	6.6913E-7	0.58	0.5649	
ARCH1	1	0.0870	0.0117	7.43	<.0001	
GARCH1	1	0.9092	0.0116	78.71	<.0001	
TDFI	1	0.1110	0.0198	5.60	<.0001	Inverse of t DF
HET DT	1	6.7738E-6	2.8187E-6	2.40	0.0163	

Table 4.10. Analysis Results for Telecommunication Services Sector – Stationary Model

Telecommunication Services Sector			
Stationary GARCH Estimates			
SSE	0.42565402	**Observations**	2516
MSE	0.0001692	**Uncond Var**	0.00003329
Log Likelihood	7965.34595	**Total R-Square**	0.0003
SBC	-15883.709	**AIC**	-15918.692
MAE	0.0086601	**AICC**	-15918.658
MAPE	103.483944	**HQC**	-15905.996
		Normality Test	188.0245
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DH	1	0.000873	0.000393	2.22	0.0262	
ARCH0	1	6.8118E-7	6.3332E-7	1.08	0.2821	
ARCH1	1	0.0835	0.0129	6.45	<.0001	
GARCH1	1	0.8960	0.0153	58.67	<.0001	
TDFI	1	0.1290	0.0205	6.30	<.0001	Inverse of t DF
HET DT	1	9.7584E-6	3.7621E-6	2.59	0.0095	

Table 4.11. Analysis Results for Utilities Sector – Stationary Model

Utilities Sector			
Stationary GARCH Estimates			
SSE	0.3532221	**Observations**	2516
MSE	0.0001404	**Uncond Var**	0.00013081
Log Likelihood	8181.23727	**Total R-Square**	.
SBC	-16315.492	**AIC**	-16350.475
MAE	0.00802491	**AICC**	-16350.441
MAPE	106.858475	**HQC**	-16337.778
		Normality Test	112.9219
		Pr > ChiSq	<.0001

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.001042	0.000371	2.81	0.0050	
DF	1	0.000972	0.000387	2.51	0.0121	
ARCH0	1	1.7273E-6	4.8329E-7	3.57	0.0004	
ARCH1	1	0.0971	0.0123	7.91	<.0001	
GARCH1	1	0.8897	0.0136	65.36	<.0001	
TDFI	1	0.0940	0.0185	5.07	<.0001	Inverse of t DF

Table 4.12. Analysis Results for Customer Discretionary Sector – IGARCH Model

Customer Discretionary Sector					
Integrated GARCH Estimates					
SSE			0.51395338	**Observations**	2516
MSE			0.0002043	**Uncond Var**	.
Log Likelihood			7794.96916	**Total R-Square**	0.0006
SBC			-15550.786	**AIC**	-15579.938
MAE			0.00947885	**AICC**	-15579.914
MAPE			114.539363	**HQC**	-15569.358
				Normality Test	200.8201
				Pr > ChiSq	<.0001
Parameter Estimates					
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**
DW	1	0.001189	0.000449	2.65	0.0081
DH	1	0.000934	0.000387	2.41	0.0159
AR10	1	-0.0434	0.0209	-2.08	0.0380
ARCH1	1	0.0996	0.008792	11.33	<.0001
GARCH1	1	0.9004	0.008792	102.41	<.0001
HET DT	1	6.6709E-6	1.1225E-6	5.94	<.0001

Table 4.13. Analysis Results for Customer Staples Sector – IGARCH Model

Customer Staples Sector			
Integrated GARCH Estimates			
SSE	0.19517721	**Observations**	2516
MSE	0.0000776	**Uncond Var**	.
Log Likelihood	8792.48935	**Total R-Square**	0.0151
SBC	-17522.335	**AIC**	-17568.979
MAE	0.00597165	**AICC**	-17568.921
MAPE	119.501787	**HQC**	-17552.05
		Normality Test	1345.1505
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
dt	1	0.000112	0.000156	0.72	0.4711	
dw	1	0.000863	0.000227	3.80	0.0001	
dh	1	0.000570	0.000203	2.81	0.0050	
AR1	1	0.0834	0.0189	4.40	<.0001	
AR4	1	0.0304	0.0176	1.73	0.0843	
AR5	1	0.0462	0.0180	2.57	0.0102	
ARCH1	1	0.0743	0.004403	16.88	<.0001	
GARCH1	1	0.9257	0.004403	210.23	<.0001	
TDFI	1	1.0537E-8	0	Infty	<.0001	Inverse of t DF

Table 4.14. Analysis Results for Energy Sector – IGARCH Model

Energy Sector			
Integrated GARCH Estimates			
SSE	0.83439667	**Observations**	2516
MSE	0.0003316	**Uncond Var**	.
Log Likelihood	7063.06457	**Total R-Square**	0.0077
SBC	-14086.977	**AIC**	-14116.129
MAE	0.01235844	**AICC**	-14116.105
MAPE	113.777301	**HQC**	-14105.549
		Normality Test	231.7022
		Pr > ChiSq	<.0001

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DT	1	0.000881	0.000458	1.93	0.0541	
DF	1	0.001125	0.000478	2.35	0.0187	
AR1	1	0.0437	0.0194	2.25	0.0245	
ARCH1	1	0.0621	0.004568	13.59	<.0001	
GARCH1	1	0.9379	0.004568	205.31	<.0001	
TDFI	1	1.0537E-8	0	Infty	<.0001	Inverse of t DF

Table 4.15. Analysis Results for Financial Sector – IGARCH Model

Financial Sector			
Integrated GARCH Estimates			
SSE	1.28728689	**Observations**	2516
MSE	0.0005116	**Uncond Var**	.
Log Likelihood	7305.11277	**Total R-Square**	0.0217
SBC	-14539.752	**AIC**	-14592.226
MAE	0.01308959	**AICC**	-14592.154
MAPE	118.167531	**HQC**	-14573.181
		Normality Test	290.4599
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.000891	0.000402	2.22	0.0265	
DW	1	0.001254	0.000420	2.99	0.0028	
AR1	1	0.0829	0.0210	3.95	<.0001	
AR4	1	0.0427	0.0200	2.14	0.0324	
AR5	1	0.0513	0.0198	2.59	0.0097	
AR6	1	0.0506	0.0195	2.59	0.0095	
ARCH1	1	0.0984	0.0112	8.78	<.0001	
GARCH1	1	0.9016	0.0112	80.44	<.0001	
TDFI	1	0.1614	0.0191	8.44	<.0001	Inverse of t DF
HET DT	1	6.5299E-6	1.9214E-6	3.40	0.0007	

Table 4.16. Analysis Results for Health Care Sector – IGARCH Model

Health Care Sector			
Integrated GARCH Estimates			
SSE	0.28311325	**Observations**	2516
MSE	0.0001125	**Uncond Var**	.
Log Likelihood	8395.79984	**Total R-Square**	0.0119
SBC	-16728.956	**AIC**	-16775.6
MAE	0.00714018	**AICC**	-16775.542
MAPE	129.507921	**HQC**	-16758.671
		Normality Test	208.7595
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.001012	0.000333	3.04	0.0024	
DW	1	0.000858	0.000335	2.56	0.0104	
DH	1	0.000968	0.000330	2.94	0.0033	
AR1	1	0.0504	0.0223	2.26	0.0239	
AR2	1	0.0437	0.0212	2.06	0.0392	
ARCH1	1	0.1235	0.0153	8.09	<.0001	
GARCH1	1	0.8765	0.0153	57.39	<.0001	
TDFI	1	0.1563	0.0208	7.52	<.0001	Inverse of t DF
HET DT	1	7.6491E-6	1.9992E-6	3.83	0.0001	

Table 4.17. Analysis Results for Industrials Sector – IGARCH Model

Industrials Sector			
Integrated GARCH Estimates			
SSE	0.50388383	**Observations**	2516
MSE	0.0002003	**Uncond Var**	.
Log Likelihood	7765.39479	**Total R-Square**	0.0012
SBC	-15475.977	**AIC**	-15516.79
MAE	0.00943725	**AICC**	-15516.745
MAPE	115.769342	**HQC**	-15501.977
		Normality Test	217.5449
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DW	1	0.000844	0.000409	2.06	0.0394	
DH	1	0.000669	0.000376	1.78	0.0749	
DF	1	0.000737	0.000402	1.83	0.0670	
AR5	1	0.0432	0.0217	1.99	0.0468	
ARCH1	1	0.0975	0.008929	10.92	<.0001	
GARCH1	1	0.9025	0.008929	101.07	<.0001	
TDFI	1	1.0537E-8	$<10^{-20}$	-	<.0001	Inverse of t DF
HET DT	1	6.3957E-6	1.1909E-6	5.37	<.0001	

Table 4.18. Analysis Results for Information Technology Sector – IGARCH Model

Information Technology Sector			
Integrated GARCH Estimates			
SSE	0.47698863	**Observations**	2516
MSE	0.0001896	**Uncond Var**	.
Log Likelihood	7695.25071	**Total R-Square**	0.0047
SBC	-15343.519	**AIC**	-15378.501
MAE	0.00936933	**AICC**	-15378.468
MAPE	119.511637	**HQC**	-15365.805
		Normality Test	254.0921
		Pr > ChiSq	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
DT	1	0.000677	0.000355	1.90	0.0569	
DW	1	0.001283	0.000383	3.35	0.0008	
DH	1	0.000298	0.000352	0.85	0.3977	
AR10	1	-0.0529	0.0186	-2.84	0.0045	
ARCH1	1	0.0630	0.004261	14.78	<.0001	
GARCH1	1	0.9370	0.004261	219.90	<.0001	
TDFI	1	1.0537E-8	$<10^{-20}$	-	<.0001	Inverse of t DF

Table 4.19. Analysis Results for Materials Sector – IGARCH Model

Materials Sector						
Integrated GARCH Estimates						
SSE		0.70998577		**Observations**		2516
MSE		0.0002822		**Uncond Var**		.
Log Likelihood		7322.50514		**Total R-Square**		.
SBC		-14598.028		**AIC**		-14633.01
MAE		0.01135983		**AICC**		-14632.977
MAPE		119.149861		**HQC**		-14620.314
				Normality Test		229.3117
				Pr > ChiSq		<.0001
Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DW	1	0.001447	0.000486	2.98	0.0029	
DF	1	0.001507	0.000509	2.96	0.0031	
AR4	1	0.0420	0.0207	2.03	0.0427	
ARCH1	1	0.0910	0.0113	8.05	<.0001	
GARCH1	1	0.9090	0.0113	80.42	<.0001	
TDFI	1	0.1164	0.0184	6.33	<.0001	Inverse of t DF
HET DT	1	7.3048E-6	2.1591E-6	3.38	0.0007	

Table 4.20. Analysis Results for Telecommunication Services Sector – IGARCH Model

Telecommunication Services Sector			
Integrated GARCH Estimates			
SSE	0.42566389	**Observations**	2516
MSE	0.0001692	**Uncond Var**	.
Log Likelihood	7960.31677	**Total R-Square**	0.0003
SBC	-15889.312	**AIC**	-15912.634
MAE	0.00865999	**AICC**	-15912.618
MAPE	103.650218	**HQC**	-15904.169
		Normality Test	222.4875
		Pr > ChiSq	<.0001

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DH	1	0.000905	0.000391	2.31	0.0207	
ARCH1	1	0.0971	0.0136	7.13	<.0001	
GARCH1	1	0.9029	0.0136	66.26	<.0001	
TDFI	1	0.1510	0.0204	7.39	<.0001	Inverse of t DF
HET DT	1	7.4127E-6	2.2157E-6	3.35	0.0008	

Table 4.21. Analysis Results for Utilities Sector – IGARCH Model

Utilities Sector			
Integrated GARCH Estimates			
SSE	0.3525126	**Observations**	2516
MSE	0.0001401	**Uncond Var**	.
Log Likelihood	8139.26909	**Total R-Square**	0.0015
SBC	-16247.216	**AIC**	-16270.538
MAE	0.00803907	**AICC**	-16270.522
MAPE	102.923986	**HQC**	-16262.074
		Normality Test	109.6404
		Pr > ChiSq	<.0001

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Variable Label**
DT	1	0.000821	0.000302	2.71	0.0067	
AR5	1	0.0386	0.0177	2.18	0.0295	
ARCH1	1	0.0791	0.004771	16.58	<.0001	
GARCH1	1	0.9209	0.004771	193.04	<.0001	
TDFI	1	1.0537E-8	$<10^{-20}$	-	<.0001	Inverse of t DF

The summary statistics given in Table 4.1.1 shows the statistically significant Day-of-the-Week effects in returns and volatility for each sector. The effect on returns is computed as Percent Change in Return = [Estimated coefficient of day dummy in the regression portion of Model 2.1.1/Sample mean return]$\times 100.$ The effect on volatility is defined as Percent Change in Volatility = [Estimated Coefficient of day dummy in GARCH formulation of Model 2.1.1/Estimated unconditional volatility] $\times 100.$ In the regression formulation no intercept term was fitted and hence all dummy variables for the five days were included in the model. Note percent change in volatility for the IGARCH Model cannot be computed because the unconditional variance for this model is infinity.

The statistically significant dummy variables d_k all had positive coefficients, suggesting that the corresponding days had higher returns than the other days of the week, which acted as the base-line return in the estimated regression model. This is similar to the results Berument and Kiymaz (2001) obtained, where all the significant dummy variables has positive coefficients. Thus, Monday, for example, was not associated with returns higher than the baseline-level. So is Tuesday and Friday for Customer Discretionary Sector. This sector showed higher than base-line return for Wednesdays and Thursdays. Tuesday had a positive effect on returns on five out of the ten sectors, with the highest effect at 14% for the Healthcare Sector. Wednesdays affected seven out of the ten sectors producing higher than base-line returns, the highest being an almost 17% increase for the Information Technology Sector. Thursdays affected six of the ten sectors, increasing their returns while Friday affected only four of the sectors. The reasons why certain days had more impact on some sectors and not on others is a question that needs insight into the trading strategies and how various markets react to events and is best left to researchers with more familiarity with such issues. One major observation that can be made based on this research results is that Monday had no positive effect on the returns of any sector and Wednesday seems to affect the returns positive for most sectors. This is somewhat similar to the results obtained by Berument and Kiymaz (2001) who studied the S&P 500 returns (aggregated over all sectors) from January 1973 through October 1997 and found lowest returns on Monday and highest on Wednesday. They, however, found a different pattern when data from October 1987 to October 1997 were studied.

As for volatility, six of the ten sectors had higher volatility on Tuesdays, with Telecommunications sector showing a 29% increase in volatility on Tuesdays. Mondays Wednesdays, Thursdays and Fridays did not increase the volatility level over the base-line. This is contrary to the results of Berument and Kiymaz (2001) who found higher volatility on Fridays. However, when the above authors studied the data for the period January 1973 through October 1987, they found highest volatility on Tuesdays. The difference in the results may be due to the time period under study. The period over which the present research was conducted included the recession of 2008/2009 which may have changed the way the market reacts to economic shocks.

5. CONCLUSIONS

This thesis examined the ten sectors of S&P 500 indices for the presence of the day-of-the-week effect on returns and volatility. Period of the study spans from the February 2005 to the February 2015. None of the sectors has been observed a significance change in return or volatility on Monday but a clear day-of-the-week effect on Tuesday, both on returns and volatility. The effect of each day of the week differs across the type of sector studied.

Overall, the results obtained in this study points to Tuesday as having the most influence on returns and volatility. One would have expected Monday to have a significant positive effect on volatility because investors would have had no chance to react to financial information that occurred from Friday closing to opening of trading on Monday. Results of the current study shows that there is a one-day delay in this hypothesized effect of information accumulation over the weekend. It may be that the sector indices of the S&P 500 do not react the way individual stocks would react to build-up of information over the weekend. Companies included in the S&P 500 index are financially stable and may not be influenced by market shocks immediately as would individual stocks of smaller or newer companies and are affected only after the rest of the stock market reacts to an incident. Further studies on this are needed to come to a definitive conclusion as to why Tuesday seems to be associated with high volatility.

Future analysis may look into using more general GARCH models rather than GARCH (1, 1) and also study volatility and returns over different time periods. Another suggestion would be to use a GARCH model that incorporate fractional integration (fractional unit root) rather than a unit root as is the case with the IGARCH Model.

APPENDIX A.

TABLE OF COMPUTATIONS FOR RETURN AND VOLATILITY CHANGE

Table A.1. Return and Volatility Change Computations in GARCH Stationary Model

Sector		**Day of the Week**				
		Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Customer Discretionary	$Return(\frac{coefficient}{mean\ abs\ return})$			$\frac{0.001146}{0.00951183}$	$\frac{0.000976}{0.00951183}$	
	$Variance(\frac{coefficient}{unconditional\ variance})$		$\frac{0.0000075705}{0.00003731}$			
Customer Staples	$Return(\frac{coefficient}{mean\ abs\ return})$		$\frac{0.000643}{0.00603414}$	$\frac{0.000675}{0.00603414}$	$\frac{0.000786}{0.00603414}$	
	$Variance(\frac{coefficient}{unconditional\ variance})$					
Energy	$Return(\frac{coefficient}{mean\ abs\ return})$		$\frac{0.001214}{0.01239142}$			$\frac{0.001396}{0.01239142}$
	$Variance(\frac{coefficient}{unconditional\ variance})$					
Financial	$Return(\frac{coefficient}{mean\ abs\ return})$		$\frac{0.000866}{0.01324785}$	$\frac{0.001232}{0.01324785}$		
	$Variance(\frac{coefficient}{unconditional\ variance})$		$\frac{0.0000059636}{0.00010999}$			
Health Care	$Return(\frac{coefficient}{mean\ abs\ return})$		$\frac{0.001014}{0.00717815}$	$\frac{0.000846}{0.00717815}$	$\frac{0.000957}{0.00717815}$	
	$Variance(\frac{coefficient}{unconditional\ variance})$		$\frac{0.0000070436}{0.00004369}$			
Industrials	$Return(\frac{coefficient}{mean\ abs\ return})$			$\frac{0.000948}{0.00946853}$	$\frac{0.000933}{0.00946853}$	$\frac{0.000910}{0.00946853}$
	$Variance(\frac{coefficient}{unconditional\ variance})$		$\frac{0.0000064196}{0.0000444}$			
Information Technology	$Return(\frac{coefficient}{mean\ abs\ return})$		$\frac{0.001050}{0.00940127}$	$\frac{0.001596}{0.00940127}$	$\frac{0.000873}{0.00940127}$	
	$Variance(\frac{coefficient}{unconditional\ variance})$					
Materials	$Return(\frac{coefficient}{mean\ abs\ return})$			$\frac{0.001443}{0.01140074}$		$\frac{0.001522}{0.01140074}$
	$Variance(\frac{coefficient}{unconditional\ variance})$		$\frac{0.0000067738}{0.01140074}$			
Telecommunication Services	$Return(\frac{coefficient}{mean\ abs\ return})$				$\frac{0.000873}{0.00866861}$	
	$Variance(\frac{coefficient}{unconditional\ variance})$		$\frac{0.0000097584}{0.00003329}$			
Utilities	$Return(\frac{coefficient}{mean\ abs\ return})$		$\frac{0.001042}{0.00805862}$			$\frac{0.000972}{0.00805862}$
	$Variance(\frac{coefficient}{unconditional\ variance})$					

APPENDIX B.

SAS CODE FOR MODELLING

```
/* Following code is for the GARCH stationary model */

options ls=78 nodate;

data sp;
filename sp '\file_path';
infile sp dlm=',';

length date $10;
input weekday date $ totalRetuan priceReturn;
Day=_n_;

data sp; set sp;
dm=0; dt=0; dw=0; dh=0; df=0;

if weekday=1 then dm=1;
if weekday=2 then dt=1;
if weekday=3 then dw=1;
if weekday=4 then dh=1;
if weekday=5 then df=1;

data sp; set sp;
retain pclose;
If Day = 1 then pclose=priceReturn;
else do;
return = log(priceReturn/pclose);
output;
pclose = priceReturn;
end;

data sp; set sp;
absReturn = abs(Return);

proc univariate data=sp;
var Return absReturn;

proc autoreg data=sp;
model Return = dm dt dw dh df /nlag=(1 2 3 4 5 6 7 8 9 10 11 12) dist=t
garch=(q=1, p=1, type=stationary) maxiter=1000 noint;
hetero dm dt dw dh;

proc gplot data=sp;
plot Return*Day /vref=0 haxis=0 to 2550 by 100 vaxis=-0.2 to 0.2 by
0.01;
title "Sector";
symbol1 i=joint;
run;
```

```
/* Following code is for the IGARCH model */

options ls=78 nodate;

data sp;
filename sp '\file_path';
infile sp dlm=',';

length date $10;
input weekday date $ totalRetuan priceReturn;
Day=_n_;

data sp; set sp;
dm=0; dt=0; dw=0; dh=0; df=0;

if weekday=1 then dm=1;
if weekday=2 then dt=1;
if weekday=3 then dw=1;
if weekday=4 then dh=1;
if weekday=5 then df=1;

data sp; set sp;
retain pclose;
If Day = 1 then pclose=priceReturn;
else do;
return = log(priceReturn/pclose);
output;
pclose = priceReturn;
end;

data sp; set sp;
absReturn = abs(Return);

proc univariate data=sp;
var Return absReturn;

proc autoreg data=sp;
model Return = dm dt dw dh df /nlag=(1 2 3 4 5 6 7 8 9 10 11 12) dist=t
garch=(q=1, p=1, type= type=integrated, noint) maxiter=1000 noint;
hetero dm dt dw dh;

proc gplot data=sp;
plot Return*Day /vref=0 haxis=0 to 2550 by 100 vaxis=-0.2 to 0.2 by
0.01;
title "Sector";
symbol1 i=joint;
run;
```

BIBLIOGRAPHY

[1] Amado, C. and Terasvirta, T. (2013). "Modeling volatility by variance decomposition," *Journal of Econometrics*, 175, 142-153.

[2] Baillie, R.T. and DeGennaro, R.P. (1990). "Stock Returns and Volatility," *Journal of Quantitative and Financial Analysis*, 25, 203-214.

[3] Berument, H. and Kiymaz, H. (2001). "The Day of the Week Effect on Stock Market Volatility," *Journal of Economics and Finance*, 25(2), 181-193.

[4] Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31, 307-327.

[5] Campbell, J. Y., and Hentschel, L. (1992). "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics*, 31, 281-318.

[6] Chan, K.C., Karolyi, A., and Stulz, R.M. (1992). "Global financial markets and the risk premium on U.S. equity," *Journal of Financial Economics*, 32, 137-167.

[7] Choudhry, T. (2000). "Day of the week effect in emerging Asian stock markets: evidence from the GARCH model," *Applied Financial Economics*, 10, 235-242.

[8] Corhay, A., and Rad, A.T. (1994). "Statistical Properties of Daily Returns: Evidence from European Stock Markets," *Journal of Business Finance and Accounting*, 21, 271-282.

[9] Cross, F. (1973). "The Behavior of Stock Prices on Fridays and Mondays," *Financial Analysts Journal*, 29: 67-69.

[10] Ederington, L. H., and Lee, J.H. (1993). "How Markets Process Information: News Releases and Volatility," Journal *of Finance*, 48: 1161-1191.

[11] Edirisinghe, A. (2011). *Modeling Hourly Electricity Prices: A Structural Time Series Approach Incorporating Modified GARCH Innovations [dissertation]*, Missouri University of Science and Technology, Rolla, MO.

[12] Engle, R. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

[13] French, K. R. (1980). "Stock Returns and The Weekend Effect," *Journal of Financial Economics*, 8, 55-69.

[14] French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3-29.

[15] Gibbons, M. and Hess, P. (1981). "Day of the Week Effects and Asset Returns," *Journal of Business*, 54, 579-596.

[16] Harvey, C., and R. Huang. (1991). "Volatility in Foreign Exchange Futures Markets," *Review of Financial Studies*, 4, 543-570.

[17] Jaffe, J. and Westerfield, R. (1985). "The Weekend Effect in Common Stock Returns: The International Evidence," *Journal of Finance*, 40, 433-454.

[18] Glosten, L.R., Jagannathan, R., and Runkle, D.E. (1993). "On the Relation between the Expected value and the Volatility of the nominal Excess Return on Stocks," *The Journal of Finance*, 48(5), 1779-1801.

[19] Kendall M. G. (1953). "The Analysis of Economic Time-Series-Part I: Prices," *Journal of the Royal Statistical Society*, Series A (General), 116(1), 11-34.

[20] Keim, D. B., and Stambaugh, F. (1984). "A Further Investigation of Weekend Effects in Stock Returns," *Journal of Finance*, 39, 819-840.

[21] Lakonishok, J., and Levi, M. (1982). "Weekend Effects on Stock Returns: A Note," *Journal of Finance*, 37, 883-889.

[22] Lindner, A. M. (2009). "Stationary, Mixing, Distributional Properties and Moments of GARCH (*p*, *q*) Processes," Handbook *of Financial Time Series*, Springer-Verlag, Berlin Heidelberg.

[23] Osborn, D.R. and Smith, J.P. (1989). "The Performance of Periodic Autoregressive Models in Forecasting Seasonal U.K. Consumption," *Journal of Business and Economic Statistics*, 7, 117-127.

[24] Rodriguez, W.K. (2012). "Day of the week effect in Latin American stock markets," *Revista de Análisis Económico*, 27(1), 71-89.

[25] Rogalski, R. J. (1984). "New Findings Regarding Day-of-the-Week Returns over Trading and Non-Trading Periods: A Note," *Journal of Finance*, 35, 1603-1614.

[26] Solnik, B. and Bousquet, L. (1990). "Day-of-the-Week Effect on the Paris Bourse," *Journal of Banking and Finance*, 14, 461-468.

[27] Theodossiou, P., and Lee, U. (1995). "Relationship between Volatility and Expected Returns Across International Stock Markets," *Journal of Business Finance and Accounting*, 22, 289- 300.

[28] Thilakaratne, M. and Samaranayake, V.A. (2013). "Bootstrap Prediction Intervals for Conditional Heteroscedastic Models with Seasonally Varying Unconditional Variance," *ASA Proceedings of the Joint Statistical Meetings, American Statistical Association (Montreal, Canada),* 2967-2981.

VITA

Juan Liu was born in Changsha, Hunan, China which has been her home before moving to the United States for education. She earned her B.S. and M.S. in Computer Science in Hunan Normal University, Changsha, Hunan, China. Then she went to become a graduate student in Mathematics Department and in August, 2015 received her M.S. in Applied Mathematics from Missouri University of Science and Technology, Rolla, Missouri, USA.