IOSR Journal of Economics and Finance (IOSR-JEF)
e-ISSN: 2321-5933, p-ISSN: 2321-5925.Volume 7, Issue 5 Ver. III (Sep. - Oct. 2016), PP 63-73
www.iosrjournals.org

Volume Effects in Standard & Poor's 500 Prices

Yiwei Lian[1]
[1](Department of Applied Mathematics, Columbia University, USA)

Abstract: *This study investigates stock trading volume volatility and its relationship to stock prices. The paper utilizes AR(1)-GARCH(1,1) models to forecast the volatility of the Standard & Poor's 500 ^GSPC stock prices, based on the trends found in evaluating the trade volume volatility over the past two decades. Using such information, it is possible to forecast the stock prices by developing a Markov-switching model. The examination of the relationship between these two features produces results that can further our knowledge in estimating the stock market and the overall American economy.*
Keywords: *GARCH, Markov-switching model, Normal distribution, T-student distribution, Volatility forecasting*

I. Introduction

The ability to predict asset prices is one of the most important results of financial economics research. Early research focused on the attempt to predict future prices based on historical prices alone. Now a crucial part of modern financial economics is the modeling of future prices of equities, bonds, or other derivative securities. One factor that can be considered in the prediction of prices is trade volume. Volume measures the quantity of shares that were traded during a given time. For instance, on the New York Stock Exchange (NYSE), the average daily volume for Standard and Poor's 500 (^GSPC) was 602,024,456 shares.

^GSPC is ticker symbol for United States stock market index Standard & Poor's 500, more commonly known as S&P 500. It is based off the market capitalizations of 500 of the United States' largest companies registered in NYSE and NASDAQ. The index serves as telling data for the overall state of the country's stock market, thus an important indicator for companies and stock investors. Hence, the goal of this paper is to characterize the volume effects of ^GSPC relative returns. Based on investigations of the volume effect, we can also provide valuable and applicable analysis for investors who are interested in ^GSPC stock.

Daily volume on a stock can fluctuate any day depending on many factors such as the amount of new information about the stock that is available, whether options contracts are to expire soon, whether if trading day is a full or half day, etc. The element that correlates the most to stock value is the new information provided. This information can be given by a press release or a regular earnings announcement provided by the company. It can also be by thirdparty communication such as a court ruling or release by a regulatory agency pertaining to the company. For example, GSPC has an average trading volume of 2 billion shares per day. During the month of September in 2008, as result of the stock market crash, the Lehman Brothers' collapsed and the news led to trading of 10 billion shares that day, about five times theaverage, and a drop in price of 8\%. The abnormally large volume was due to differences in the investor's view of the stock value after taking into account the new information. Analysis of trade volume and the corresponding price changes associated to these informational releases has been of popular discussions because of the inferences that can be learned from these abnormal trading volumes.

The relationship between trade volume and stock prices are complicated and when understood properly, can lead to many discoveries in the theory of portfolio. The analysis of trade volume and its relationship with stock prices and changes in price is a topic that has been researched for the past forty years. Its rootsare generally credited to the work of Osborne. In his influential work, he modeled price changes with respect to a diffusion process that had variance dependent on the number of trades on that particular issue. With this, he began his research that looked into the possible relationship between returns and the trade volume. Before this topic is discussed and analyzed, we must try to increase research in this area and hopefully answer the question: How can we use trade volume patterns to improve the prediction of stock price?

In this paper, we propose a new Markov switching model, using known data of volume change for the underlying asset. A number of researchers have recently become interested in modeling economic and financial time series with respect to occasional, discrete shifts in parameters. When the Markov-switching regression framework of Goldfeld and Quandt (1973) and Cosslett and Lee (1985) is applied to a time-series autoregression, it results in a model that allows for nonlinear dynamics and sudden changes in the variability of a series, yet it is still very tractable for rational expectations econometrics. This approach has provided some new insight in stock price forecasting.

First, we preview the AR-GARCH model used in the log return process. With thisknowledge, we proceed to discuss how to use trade volume as a predictive measure of future price changes, and also how trade volume can allow us to create a Markov switching model. We can then show the theoretical models that support empirical results, on a GSPC index. Finally, we can analyze sample price and volume data around the most recent quarter of earnings announcements.

Any empirical analysis of trading activity in the market must start with a measure of volume. There are many books on trading activity in financial markets and measures of volume have been proposed and analyzed. Some studies of aggregate trading activity use the total number of shares traded as a measure of volume (see Epps and Epps (1976), Gallant, Rossi, and Tauchen (1992), Hiemstra and Jones (1994), and Ying (1966)).

II. Time Series and Data

In this study, we examine the daily ^GSPC stock price activity over a 26-year period, from 1990 to 2016. The collection of ^GSPC daily adjusted closing price was from Yahoo Finance.

2.1 Log Return Process

The adjusted closing price is an accurate indication of the stock's performance. The rate of return can be found by comparing a stock's historical adjusted closing price to its current price.

Furthermore, log return is used to examine log price differences for daily returns of the stock.
Let p_t denote the adjusted closing price of a stock on day t, then the daily percentage change on the day is defined by:

$$r_t = 100 \log \frac{p_t}{p_{t-1}} \tag{1}$$

From the negative and positive returns, we can examine the possibility of large losses and large gains.

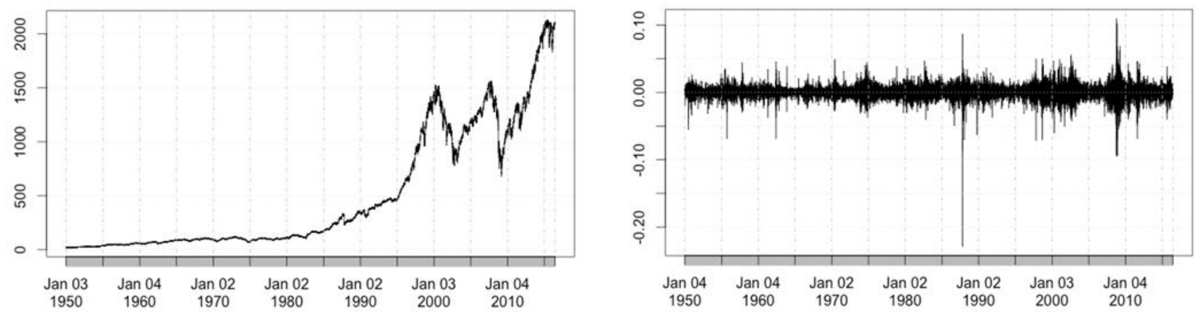

Fig. 1: Time plots of ^GSCP stock from 1950-01-03 to 2016-06-17. The upper panel is for adjusted closing price, and the lower panel is for positive daily log returns.

From the plot, we observe that daily returns of the stock clearly depict volatility clustering. That is, periods of large returns are clustered and distinct from periods of small returns, which are also clustered.If we measure such volatility in terms of variance, then it is natural to think that variance changes with time, thus reflecting the clusters of large and small returns.

Fig.1 shows the time plots of adjusted closing price and daily log returns of GSPC stock from January 4, 1950 to July, 2016.The left plot shows that ^GSPC stock price has skyrocketed over 40 times since 1990.The right plot shows ^GSPC daily log return time series.

We also observe that there are more pronounced peaks than one would expect from Gaussian data. Table 1 summarizes the basic statistical characteristics of the whole ^GSPC stock negative daily log return series. Note that the expected ^GSPC log returns during the test period is 0.0002684437. The skewness and kurtosis measures are highly significant, and those indicate substantial departures from normality.

Mean	**Range**	**Std dev**	**Skewness**	**Kurtosis**	**Observations**
-0.09	(-28.69, 73.12)	3.05	2.55	70.82	5035

Table 1: Summary statistics of the GSPC daily log returns July 3, 1995 - July 2, 2015

Since the possibility of time-varying variance and non-normal behavior are noticed, we provide a formal test to check the normality of the return process.

2.1.1 Test For Normality

In studying the financial time series, one common assumption is that the process follows normal distribution. However, it is barely true in the real stock return series. Our study shows that the ^GSPC stock returns are not normally distributed.

We begin by forming a QQ-plot of the ^GSPC daily log returns sample set against the normal distribution, in order to confirm that an assumption of normality is unrealistic, and that the innovation process has fat tails or is leptokurtic -- see Fig 2.

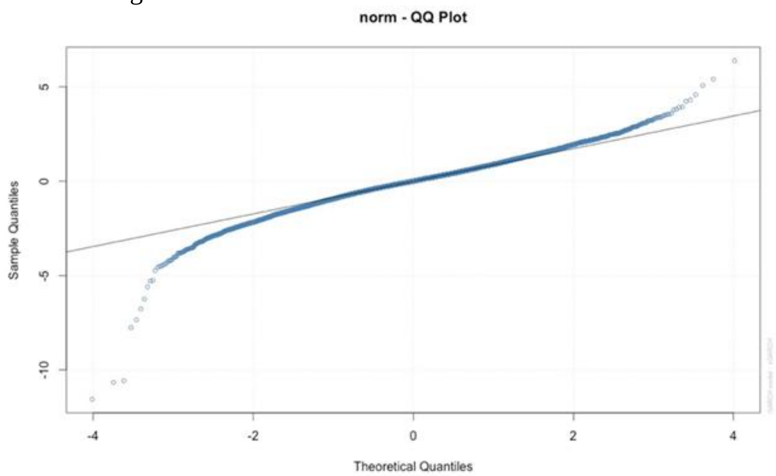

Fig. 2: Quantile-quantile plot of ^GSPC negative log returns from 1950-01-03 to 2016-06-17 against the normal distribution.

To confirm our observation, we also use the Jarque-Bera test for normality. The Jarque-Bera Test tells us the higher value represents the non-normality of the rate of returns. From the last line in the chart, the critical value is very large, so the distributions of the rates of returns are not normal distribution.

The Jarque-Bera test results:
Test Results:
PARAMETER: Sample Size: 6709
STATISTIC: LM: 20766.935, ALM: 20817.635
P VALUE: Asymptotic: < 2.2e-16

We also test the empirical distribution of the daily returns (y-axis) against the t distribution using Q-Q plot. From Fig. 2, one can see clearly that the t distribution is a much better fit and the empirical distribution of the daily returns has lighter tails than the t distribution.

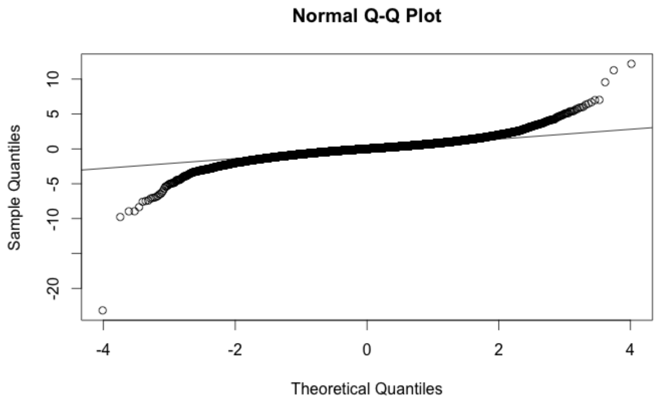

Fig. 3: Quantile-quantile plot of returns from against the t distribution.

2.1.2 Test For Normality

In finance literature, testing for zero autocorrelations has been used as a tool to verify the efficiency of the market hypothesis. Since applying extreme value theory on a data set suggests that the time series are highly uncorrelated with a common cumulative distribution function, we need to check the correlations of the ^GSPC returns.We begin by considering the autocorrelation function of a time series $\{r_t\}$. The correlation between $r_t$and its past values r_{t-1}. is called the lag-l. autocorrelation of $r_t$and is commonly denoted by ρ_l. Under the weakly stationary assumption, we assume ρ_l is a function of l only, i.e.:

$$\rho_l = \frac{Cov(r_t, r_{t-1})}{\sqrt{Var(r_t) Var(r_{t-1})}} = \frac{Cov(r_t, r_{t-1})}{Var(r_1)} = \frac{\gamma_l}{\gamma_0} \quad (2)$$

where the property $Var(r_t) = Var(r_1) = \gamma_0$for a weakly stationary series is used.
For a given sample of returns $\{r_t\}_{t=1}^T$, let $\bar{r} = \frac{\sum_{t=1}^T r_t}{T}$is the sample mean. The lag-l sample autocorrelation of $\{r_t\}$can be represented as:

$$\widehat{\rho_t} = \frac{(\sum_{t=l+1}^T (r_t - \bar{r})(r_{t-1} - \bar{r}))/(T-l-1)}{\frac{\sum_{t=1}^T (r_t - \bar{r})^2}{T-1}}, 0 \leq l \leq T-1 \quad (3)$$

If a time series is not autocorrelated, then estimates of $\hat{\rho}_l$will not be significantly different from 0.

Fig.4 shows the sample autocorrelation coefficient $\hat{\rho}_l$plotted against different lags l (measured in days), along with the 95% confidence band around zero for ^GSPC negative daily log returns, for the period July 3, 1995 to July 2, 2015. The dashed lines represent the upper and lower 95% confidence bands $\pm\frac{1.96}{\sqrt{T}}$, where the time length for our GSPC returns is $T = 5036$ days.Fig. 4 shows a small autocorrelation in ^GSPC daily log price changes. Even in the cases where the autocorrelations are outside the confidence bands, the autocorrelation coefficients are quite small, less than 5%.

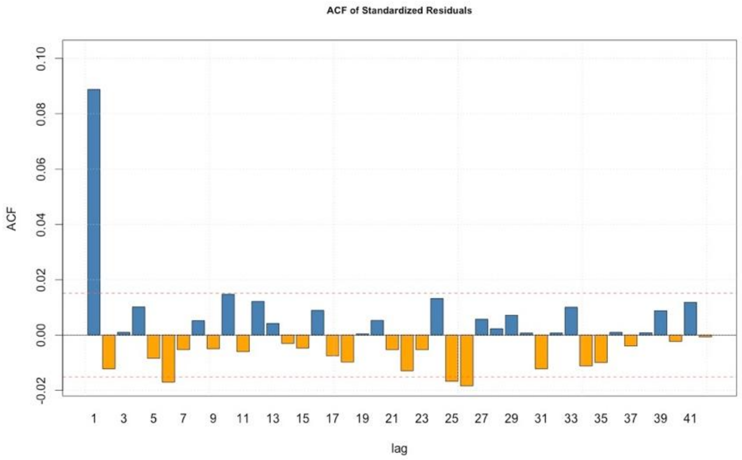

Fig. 4: Sample autocorrelation coefficients up to 42 lags for ^GSPC returns from 1950-01-03 to 2016-06-17.

The Autocorrelation Coefficients (ACF) and Partial Autocorrelation Coefficients (PACF) are extremely useful as they help us to identify the correct specification for an ARMA model that describes the stochastic process. Particularly, if the process is random noise, all autocorrelation and partial autocorrelation coefficients equal zero. If the process is an AR(p), the PACF will equal zero for all lags $k > p$, while if the process is a MA(q) the ACF will equal zero for all lags $k > q$.

There are extensive literature that conclude that high stock volume is closely related to volatile returns; see for example Gallant, Rossi, and Tauchen [1992], Harris [1987]; Jain and Joh [1988]; Jones, Kaul, and Lipson [1991]; and the survey in Karpoff [1987]. Numerous papers have noted that volume tends to be higher when stock prices are increasing, rather than when prices are falling. However, our plot \ref{fig.volume} clearly contradiction to these results. For example, from 2010 to 2015, the stock price has a significant decreasing trend, however the trading volume tends to be higher and with more peaks in this period. Indeed weak prices on higher relative volume is an indicator on underlying activities such as stock news, analyst downgrade, insider trading, or the fact that hedge funds and stock traders are piling out of the stock ahead of a catalyst.

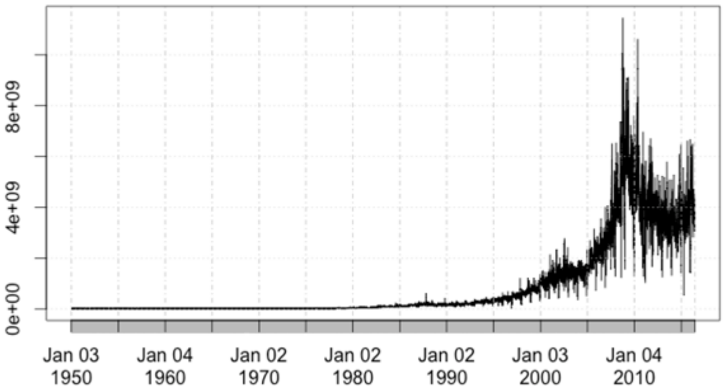

Fig. 5: Daily volume of ^GSPC.

Next table is Box-Ljung test, which helps us check whether the rate of returns has the ARCH effect or not, the null hypothesis is the rate of returns doesn't have ARCH effect, while the alternative hypothesis is opposite. See Forsberg and Bollerslev (2002). Box-Ljung test results:
X-squared = 15.558,df = 1, p-value = 8.001e-05

2.2 The Volume Process

As Beaver noted, volume is a useful tool in determining how much variation exists with the announcement of new information. Anything that causes investors to react can be described as information, whether or not it truly has any fundamental impact on the underlying value of the company. Sometimes, information on one company can even effect the price and volume of another unrelated company due to the mere similarity of their ticker symbol.

Here we denote $\{v_t\}$as the process of the volume for ^GSPC. Fig. 6 shows the volume process. Since we will use the information about increasing/decreasing of volume in our Markov Switching model we must also plot the change-of-volume process $c_t := v_t - v_{\{t-1\}}$, see graph.

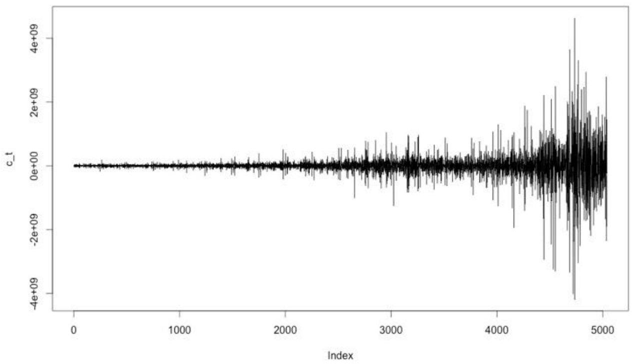

Fig. 6: Change of volume process

Many individuals in finance believe that volume is heavy when the market is going up, and light when it is going down. Karpoff discussed this idea in some detail, citing past works by Epps [10,11] which showed that for both the stock and bond markets, the ratio of volume to absolute price change was larger for trades on upticks than on downticks. This trend is also seen when evaluating daily intervals.

To investigate the relationship between the stock price and the change of volume, we compute the correlation of these two time series. More precisely, we use a Monte-Carlo significance test, which is described as following:

1) use ccf() to compute the cross-correlation between $\{p_t\}$ and $\{c_t\}$.
2) repeat the following steps, say, 1000 times.
   2a) randomly reorder the values of one of the time series, say $\{r_t\}$. Call the randomly reordered series$\{r_t^{'}\}$
   2b) use ccf() to compute the cross-correlation between $\{r_t\}$ and $\{v_t\}$. Store that cross-correlation.
3) the 1000 cross-correlation estimates computed in step 2 are all estimating cross-correlation 0, conditional on the data. A two-tailed test then is: if the cross-correlation computed in step 1 is outside the (0.025, 0.975)

quantiles of the empirical distribution of the cross-correlations computed in step 2, then, reject the null hypothesis that x and y are uncorrelated, with size 0.05.

Our test results show that the absolute price and volume correlation existed for both equity and futures markets across all time intervals, although the correlation was weak occasionally. The weakness in correlation, however, can be attributed to the fact that short selling is often more difficult than buying a stock. This asymmetry causes lower volume in accordance with price reductions.

III. Methodology

3.1 AR-GARCH Model

Let $\{r_t\}$ the the log return process, and $\{v_t\}$ the associated volume process.

Let $\Im_t = \sigma\big((r_s,, v_s) s \leq t\big)$ be the σ-algebra generated by all historical information (based on the log returns as well as the volume) up to time t. Consequently, we obtain a filtration $\{\Im_t, t \leq T\}$ generated by the log return process $\{r_t\}$, and the volume $\{v_t\}$.

We denote the conditional mean of the log return as:

$$\mu_t = \mathrm{E}(r_t | \Im_{t-1}) \tag{4}$$

And the conditional variance is denoted as:

$$\sigma_t^2 = Var(r_t | \Im_{t-1}) = \mathrm{E}((r_t - u_t)^2 | (\Im_{t-1})^2 | \Im_{t-1}) \tag{5}$$

The random variable σ_t is also called the volatility of r_t. Moreover, one can see that σ_t^2 is a predictable process. The fact that large absolute returns tend to be followed by large absolute returns (whatever the sign of the price variations) is hardly compatible with the assumption of constant conditional variance. This phenomenon is called conditional heteroscedasticity, i.e. $\sigma_t^2$is not a constant. Note that σ_t^2 is measurable with respect to $\Im_{t-1}$, so it can be represented as a function of $\{(r_s, v_s)\ s < t\}$.

To account for the very specific natureof financial series (price variations or log-returns, interest rates, etc.), one usually denote:

$$r_t - u_t = \sigma_t \varepsilon_t \tag{6}$$

where ε_t is a white noise process with zero mean, unit variance and they are uncorrelated.

Different classes of models can be distinguished depending on the specification adopted for σ_t, such asthe Conditionally heteroscedastic (GARCH)processes and the Exponential GARCH model.

In this paper, AR-1 model is used to simulate the conditional mean. Moreprecisely, the AR-1 model is defined by:

$$\mu_t = \phi_0 + \{\phi_1\} r_{\{t-1\}} \tag{7}$$

where ϕ_0 and ϕ_1 are two constants.

To estimate the conditional variance, we use the GARCH(1,1) model. Bollerslev proposed the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model where α, β, γ are constants. Here ε_t is the weak random noise, with zero mean, unit variance and they are uncorrelated.

The GARCH (generalized autoregressive conditionally heteroscedastic) model is by far the most commonly used model in financial time series analysis. It models volatility clustering by using values of the past observations squared and variances to model variance at time t.

For the GARCH class model, it is rather difficult to give the order of model. Indeed some studies have found that the predictive effect of higher order model is not necessarily better than the low order model, see Hansen and Hansen, P. R., Lunde, A.(2005) and Bollerslev, T., Chou, R.Y., Kroner, K.F (1992). Because of the above considerations and the computational complications, we use the GARCH(1,1) and the EGARCH(1,1) model in this paper. Additionally, the distribution of error $\varepsilon_t$must be assumed with volatility models. In the second part of the paper, we analyze the empirical distribution of the daily returns by taking $\varepsilon_t$to be the normal distribution and the t distribution, respectively.

IV. Markov-Swtiching

In 1959, Osborne hypothesized that securities prices could be modeled as a log-normal distribution with the variance term dependent on the trading volume. Seven years later, in 1966, Ying produced a paper [6] which applied a series of statistical tests to a six-year daily series of price and volume. Ying normalized the trading volume by the number of shares outstanding to avoid any biases from issues with larger number of outstanding shares. His main conclusions were:

(i) A small volume is usually accompanied by a decrease in price; (ii) A large volume is usually accompanied by an increase in price; (iii) A large increase in volume is usually accompanied by a large price change; (iv) A large volume is usually followed by a rise in price; (v) If the volume has decreased (increased) five straight trading days, the price will tend to fall (rise) over the next four trading days.

Using trade volume log return data from January 1st, 1990 to July 7th, 2016, split the data into two groups based on whether the majority of points in every group of 10 points are greater than or equal to 0 or less than 0.

From the log-volume process, a window of length 5 were chosen, which divide the total data into 1000 windows. We now define a new time series $\{s_t\}$, which takes value of 1 or -1 on each time window. More precisely, for each window $L \in \{1, \cdots, 1000\}$, for $k = 1, \cdots, 5$, we take $s_{\{5L+k\}} = -1$, if the volume $c_t \leq 0\, fort \in \{5L+1, \cdots, 5L+5\}$; and $s_{\{5L+k\}} = 1$ if there exists $k \in \{1, \cdots, 5\}$, such that $c_{\{5L+k\}} > 0$. We thus define the binary process $\{r_t, s_t\}$. Note that the sign of s_t also divides the associated time index into two sets $\mathcal{A}^+$ and $\mathcal{A}^-$. We define $\{r_t^+\}$ and $\{r_t^-\}$, respectively, with $n \in \mathcal{A}^+$and $m \in \mathcal{A}^-$.

Now we use the two data sets $\{r_n^+, \text{n} \in \mathcal{A}^+\}$and $\{r_m^-, \text{m} \in \mathcal{A}^-\}$ to fit the AR-GARCH(1,1) model. And we get the two sets of coefficients:

Data	φ_0	φ_1	α	β	γ
$\{r_n^+\}$	φ_0^+	φ_1^+	α^+	β^+	γ^+
$\{r_n^-\}$	φ_0^-	φ_1^-	α^-	β^-	γ^-

Table 2: AR(1)-GARCH(1,1)

Now we use a dummy variable δ, which takes value 0, if $s_t = 1$; and takes value 1 if $s_t = -1$. Now the AR(1)-GARCH(1,1) model can be written as:

$$\mu_{\{t\}} = \varphi_0 + \varphi_1 r_{\{t-1\}} \tag{8}$$

where $\varphi_0 = (1-\delta)\varphi_0^+ + \delta\varphi_0^-$, $\varphi_1 = (1-\delta)\backslash varphi_1^+ + \delta\varphi_1^-$, with $\delta = 0$ or $\delta = 1$.
Moreover, we have:

$$\begin{cases} r_{\{t\}} = \mu_t + \eta_t \\ \eta_{\{t\}} = \sigma_{\{t\}}\epsilon_t \\ \{\sigma_{\{t\}}\}^2 = \alpha + \beta\{\eta_{\{t-1\}}\}^2\} + \gamma\{\sigma_{\{t-1\}}^2\} \end{cases} \tag{9}$$

where $\alpha = (1-\delta)\alpha^+ + \delta\alpha^-$, $\beta = (1-\delta)\beta^+ + \delta\beta^-$, and $\gamma = (1-\delta)\gamma^+ + \delta\gamma^-$.

Now that we have obtain a Markov switching model, by defining the following transition probability (or equivalently the Markov matrix):

$$P(\delta_t = j | \delta_{\{t-1\}} = i) = p_{\{i,j\}} \tag{10}$$

for $i, j \in \{0,1\}$.This implies that we also have $p_{\{i,0\}} + p_{\{i,1\}} = 1$.
Note that the expected length the system is going to stay in state j can be calculated from the transition probabilities. Let D_j denote the number of periods the system is in state j.Application of the chain rule and the Markovproperty yield for the probability to stay $k$periods in state j is:

$$P(D_j = k) = p_{\{jj\}}^{\{k-1\}}(1 - p_{\{jj\}}) \tag{11}$$

which implies for the expected duration ofthat state is:

$$\mathbb{E}(D_j) = \sum_{k=0}^{\infty} kP(D_j = k) = \frac{1}{1-p_{\{jj\}}} \tag{12}$$

Thus in our paper, we take:

$$p_{\{00\}} = 1 - \frac{1}{\mathcal{A}^+}, p_{\{11\}} = 1 - \frac{1}{\mathcal{A}^-} \tag{13}$$

where $|\mathcal{A}^\pm|$ denotes the number of states in $\mathcal{A}^\pm$.

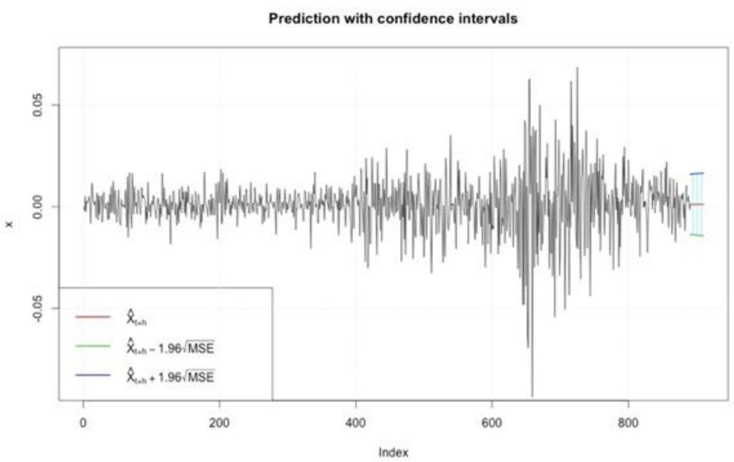

Fig. 7: Group One Data Projected Confidence Interval

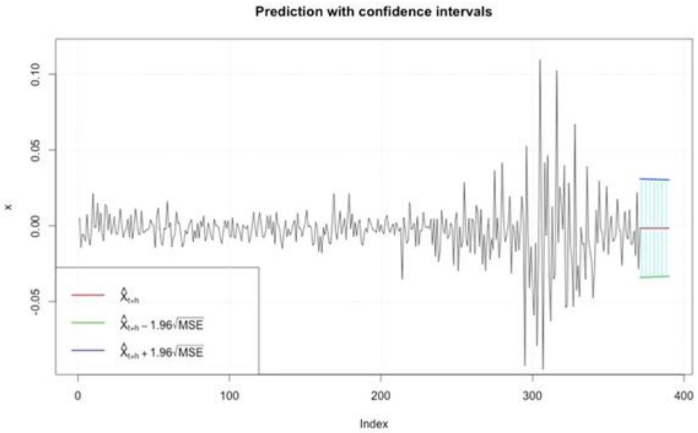

Fig. 8: Group Two Data Projected Confidence Interval

Using the probability distributions of these two groups, randomly sample between the two using the probabilities to forecast price data. After finding epsilon for the two groups, we can analyze the data.
Find the AR(1) model and GARCH(1,1) model with the data. With that, epsilon can be found.

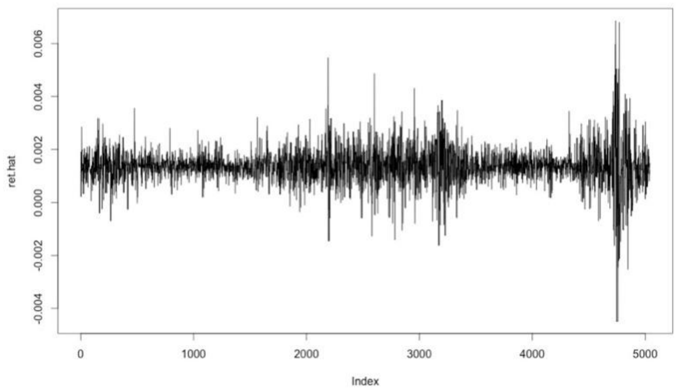

Fig. 9: AR(1) of log return

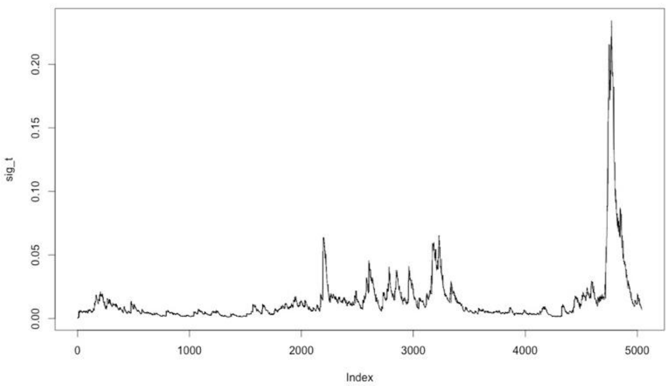

Fig. 10: GARCH(1,1) of log return

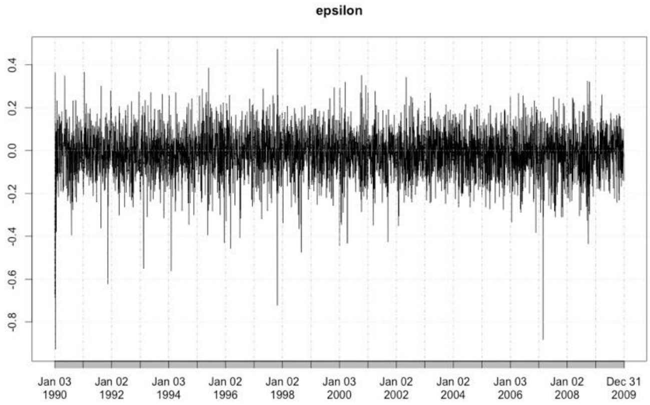

Fig. 11: Epsilon

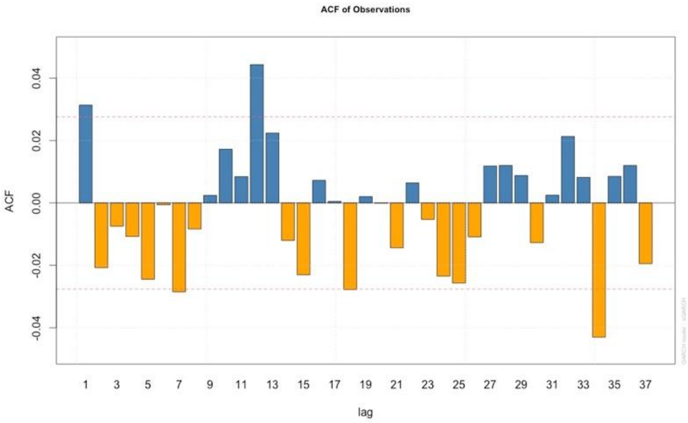

Fig. 12: ACF of Epsilon For Group One Data

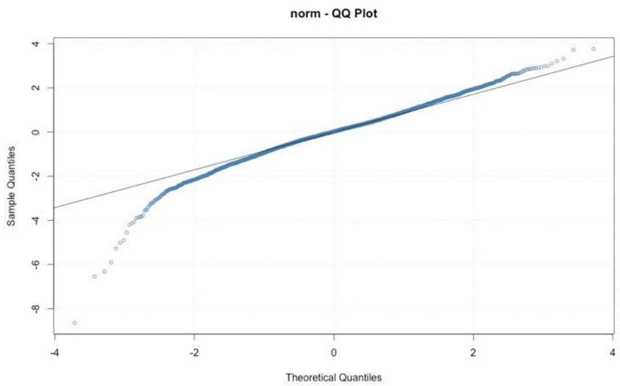

Fig. 13: QQ-Norm of Epsilon

V. Model Fitting and VaR Estimation for GSPC

In this paper, Markov-Switching GARCH models is used to estimate and forecast the different rates of returns under the different error distributions. We then compare the results and choose the appropriate model to forecast the conditional variance.

5.1 Selection of ARMA (p, q) model

First step is selection of suitable ARMA (p, q) model for GSPC daily return. By observing the autocorrelation and partial autocorrelation, the rough p and q can be acquired. After comparing the value of AIC and BIC, the more accurate p and q will be picked up. Finally, we take p=1, q=0. From the following function:

$$r_t = \varphi_0 u_t + \varphi_1 r_{\{t-1\}} \tag{14}$$

Our results on the log volume data c_t implies that we will split it into two sets with index $\mathcal{A}^{\pm}$, such that $|\mathcal{A}^{+}| = 2815$, $|\mathcal{A}^{-}| = 2225$.

The estimated parameters are $\varphi_0^+ = 0.002820801, \varphi_0^- = -0.07980724$; and $,\varphi_0^- = -0.003067779$, $\varphi_0^- = -0.1008321$.

5.2 Result of the Two GARCH Models for GSPC

Compare MA (1)-GARCH (1, 1) mode under different error terms distributions: normal and student t. All of the estimated parameters have been shown in the Table:

Data	φ_0	φ_1	α	β	γ
$\{r_n^+\}$	0.002820801	-0.07980724	5.386722e-07	5.332175	0.9417135
$\{r_n^-\}$	-0.003067779	-0.1008321	2.205628e-06	0.09660167	0.8929065

Table 3: AR(1)-GARCH(1,1)

Based on the assumption of 5% significance level, for GARCH (1, 1) model, when the error term is normal distribution, all of the estimated parameters are significant, while the error term is student-t distribution, the estimated is not significant.

5.3 Markov-Switching Model

More precisely, since $|\mathcal{A}^+| = 2815$, so $p_{\{00\}} = 0.9996446$; $|\mathcal{A}^-| = 2225$, so $p_{\{11\}} = 0.9995504$. This implies that the Markov matrix is:

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} 0.9996446 & 0.0003554 \\ 0.0004496 & 0.9995504 \end{pmatrix}$$

Coefficients for $\{r_n^+\}$ and $\{r_n^-\}$ are used to generate simulated stock price log return data with the Markov-switching model. The first data is standard deviation of the observed stock price log return, then the later data points are added on based on calculations by two GARCH(1,1) equations with the corresponding coefficients. The equations are chosen each time with the four conditional probabilities $p_{00}, p_{01}, p_{10}, p_{11}$.

5.4 ARCH-LM Test

Engle's (1982) Autoregressive Conditional Heteroscedasticity-Lagrange Multiplier (ARCH-LM) test is the undisputed standard test to detect ARCH.

ARCH-LM test is used to check the model results and select the lag equals to 2, 5 and 10 in the following table. see Engle (2001).

Data	GARH – normal	GARCH – student t	MS - normal	MS – student t
Lag 2	0.6104	0.4479	6.545e-13	0.008884
Lag 5	0.8668	0.7177	< 2.2e-16	< 2.2e-16
Lag 10	0.9736	0.8517	< 2.2e-16	< 2.2e-16

Table 4: The p-value for ARCH-LM test of GARCH and Markov-Switching (MS) models

VI. Out-of-sample Forecasts

In order to acquire the appropriate model to forecast the conditional variance, this paper use out-of-sample to calculate the root mean square error (RMSE), and the detail can be got from Forsberg and Bollerslev's paper (2002).

Root Mean Square Error (RMSE) measures the difference between the true values and estimated values, and accumulates all these difference together as a standard for the predictive ability of a model. The criterion is the smaller value of the RMSE, the better the predicting ability of the model. This article uses this method to determine which model has the best forecasting performance.(http://en.wikipedia.org/wiki/Root-mean-square_deviation):

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} r_t^2 - \sigma_t^2} \tag{15}$$

where r_t is observed values and σ_t is the predicted value of conditional variance at time t, T is the number of forecasts.

Here, use GSPC as an example to explain the process in details. The GSPC stock return includes 5039 observations, 20 years, and reserve the last 5 years as out-of-sample, including 1672 observations.

Put in-sample data into a window, so the length of window is fixed which equals to 5039. First, pick up the observations from 1 to 5039 into this fixed window and use MS-GARCH models to estimate and forecast. In this way, I get the first prediction conditional variance. This process is called as one-step-ahead forecast. Second, repeat the first step except pick up the observations from 2 to 5040 in the fixed window, and get the second prediction conditional variance. Next, repeat the first step except pick up the observations from 3 to 5041 in the fixed window, and get the third prediction conditional variance. Repeat this step 1672 times. We call such a process as multi-step-ahead forecast. Finally, use the 1672 prediction conditional variance to calculate the RMSE by the formula (6).

GARCH (1,1) – Normal	0.01380118
GARCH (1,1) – Student t	0.01373756
MS – GARCH(1,1) – Normal	0.009939033
MS – GARCH (1,1) – Student t	0.00993895

Table 5: The result of RMSE

Other tests that can be added are mean error (ME), mean absolute error (MAE), and the smoothed mean absolute percentage error (SMAPE).

Criteria	GARH – normal	GARCH – student t	MS - normal	MS – student t
ME	0.00891-324	0.00882757	-0.0003878219	-0.0003879767
MAE	0.01031166	0.0100516	0.00694761	0.006947021
SMAPE	10.69132	10.58459	1.000999	1.001291

Table 6: Forecast evaluation criteria

VII. Conclusion

The S&P 500 Index is a basket of 500 stocks that are considered to be widely held.

It's weighted by the market value, and its performance is thought to be representative of the stock market as a whole.

This paper use different volatility models to analyze and forecast the conditional variance. It simultaneously chooses the normal distribution and the student-t distribution as the error terms distribution. Table 5 illustrates which model has the smallest RMSE for different applications. Our main finding is that MS-GARCH(1,1)-student t is the best appropriate one to forecast the conditional variance for GSPC.

Overall, our findings have implications for investors, financial institutions, and futures exchanges.

References

[1]. Cross, F. (1973). The behavior of stock prices on Fridays and Mondays. Financial analysts journal, 29(6), 67-69.

[2]. Campbell, J. Y. (1993).Understanding risk and return (No. w4554).National Bureau of Economic Research.

[3]. Epps, T. & Epps, M. (1976). The stochastic dependence of stochastic price changes and transaction volume:Implications for the mixture of distribution hypothesis, Econometrica, 44, 305-321.

[4]. Ronald Gallant; Peter E. Rossi; George Tauchen. The Review of Financial Studies, Vol. 5, No. 2. (1992), pp. 199-242.

[5]. Ronald Gallant; Peter E. Rossi; George Tauchen, Stock Prices and Volume, The Review of Financial Studies, Vol. 5, No. 2. (1992), pp. 199-242.

[6]. Ying, C.C. (1966) Stock Market Prices and Volumes of Sales, Econometrica, 34, 676-686.

[7]. Have Aluminium And Copper Futures Markets Become More Volatile Over Time? Modelling and Simulation Society of Australia and New Zealand Inc.