

**MOELING THE VOLATILITY OF US EXCESS STOCK RETURNS:  
THE RELATIONSHIP BETWEEN RETURN AND VARIANCE**

by

Arash Jalili  
M.A. in Economics, UBC, 2008

and

Alireza Kanaani  
M.S. in Socioeconomics Engineering, IRPD, 2005

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF FINANCIAL RISK MANAGMENT

In the Financial Risk Managment Program  
of the  
Faculty  
of  
Business Administration

© Arash Jalili & Alireza Kanaani 2009  
SIMON FRASER UNIVERSITY  
Summer 2009

All rights reserved. However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for *Fair Dealing*. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

# Approval

**Name:** Arash Jalili & Alireza Kanaani

**Degree:** Master of Financial Risk Management

**Title of Project:** Modeling the Volatility of US Excess Stock Returns:  
The Relationship between Return and Variance

**Supervisory Committee:**

---

**Dr. Anton Theunissen**  
Adjunct Professor

---

**Dr. Andrey Pavlov**  
Associate Professor  
Academic Chair, Global Asset and Wealth  
Management MBA, Master of Financial Risk  
Management

**Date Approved:** \_\_\_\_\_

## **Abstract**

This paper examines the relationship between monthly US excess stock market return and its volatility using several GARCH models including the GJR GARCH-M and the Nelson's EGARCH-M model. We find Nelson's EGARCH-M model to fit the data the best and to pass most of the diagnostic tests. The relationship between risk and return is found to be negative but insignificant. Also there is strong evidence in favor of asymmetric response of variance to negative and positive residuals using the EGARCH-M model. The monthly conditional volatility is shown to be not as persistent as those found in literature using daily returns.

## Table of Contents

Approval .....	ii
Abstract .....	iii
Table of Contents .....	iv
1. Introduction .....	1
2. Literature Review .....	2
3. Data .....	5
4. Volatility Models .....	6
4.1 Model Specifications .....	6
4.2 Inference and Diagnostic Tests .....	9
5. Empirical Results .....	11
6. Concluding Remarks .....	13
Appendix: Tables and Figures .....	15
References .....	21

## 1. Introduction

Risk in financial markets is often represented by volatility and measured by variance of asset returns. It would be reasonable to assume that volatility is a major factor influencing investors' behaviour. High levels of volatility can lead to unstable market conditions, a decline in investors' confidence and outflow of capital from markets. Therefore, the relationship between risk and returns is of great importance to regulators and investors.

For a long time the finance literature focused on the relationship between risk and return among different securities within in a specific period of time. The general conclusion was that investors require higher expected return (a larger premium) for riskier assets. By late 1970's, new research began to emerge where the focus was to model the relationship between return and volatility across time. This line of research examined whether the investors seek greater premium for period of times when the security carries more risk. It seems natural to assume that there is a positive relationship between the return and volatility. However, as Glosten et al argue, a larger risk premium may not be required because periods with high volatility coincide with periods when investors are better fit to carry the risk. Also investors may want to save relatively more during times when future is perceived to be riskier. If all the assets available are risky, then the price of these assets may be bid up, reducing their risk premium. It should be noted that the latter argument is not as strong since it assumes a world without risk free investment opportunities (Glosten et al 1993, 1).

The results from previous research regarding the relationship between return and volatility have been mixed. French, Schwert, and Stambaugh (1987) find a positive relationship between expected excess return and conditional variance whereas Pagan and Hong (1991), Breen, Glosten, and Jagannathan (1989), and Nelson (1991) find a negative relation. Chan, Karolyi, and Stulz (1992) find no significant relationship for US but find significance for a portfolio of world markets.

Most of the volatility models used in these papers and other such literature are based on Autoregressive Conditional Heteroscedasticity (ARCH) model introduced by Engle (1982), Generalized ARCH (GARCH) model introduced by Bollerslev (1987), or variations and extensions of these. Empirical evidence shows that the distribution of daily stock market returns is far away from a normal distribution. More specifically, the distribution of stock returns exhibit asymmetry, leptokurtosis, and volatility clustering. The GARCH model accounts for time variation of volatility by allowing the current conditional variance to be a function of past

variances and past squared errors (Xu 1999, 142). Theoretically, such specification will allow the GARCH model to better fit the empirical data.

Despite its specifications, the GARCH model cannot explain the asymmetry in distribution of returns. The asymmetry would allow the unexpected positive return to result in less volatility than an unexpected negative return. The Exponential GARCH (EGARCH) model of Nelson and GJR GARCH model both allow for such asymmetry. As will be discussed below, several papers find these models to be superior in fitting US stock returns time series.

In this paper, several models are fit to data and diagnostic tests are performed to find the best suit for modeling volatility of monthly stock returns excess of risk free rate of return and to examine the relationship of variance with returns. We start with the simple GARCH and GARCH-M models. GJR GARCH-M is then used to compare our results for recent data to that of Glosten et al (1993). The GJR model does not include an ARMA process and thus we will test a GJR GARCH model with ARMA process rather than an in-mean process. Given its empirical superiority, the EGARCH-M model of Nelson is also tested.

We find that the EGARCH-M model introduced by Nelson (1991) results in the most significant coefficients. In addition, this model does best when subjected to sign tests proposed by Engle and Ng (1993). We do not find a significant relationship between return and variance using any of the models. Also, the volatility does not seem have any one-period persistence for monthly series.

The rest of the paper is as follows. The next section reviews the literature regarding GARCH models and the corresponding empirical investigations. Section three presents data, examines the distribution of returns and provides descriptive statistics of monthly US excess stock returns. Section four describes the models used in this study and section five presents the results and a discussion of them. Finally section six concludes.

## **2. Literature Review**

Many of the conventional econometric models in 1970's assumed a constant volatility that did not depend on past information. Some of the researches such as Breusch and Pagan (1977) introduced tests to find heteroskedastic disturbances in regression models. However, most of these researchers were primarily concerned with the assumption of homoscedasticity in Ordinary Least Squares (OLS) estimation and the inefficiency that would result if these assumptions were violated.

Engle (1982) introduced the ARCH process that is serially uncorrelated and has a zero mean. ARCH includes non-constant variances conditional on past and constant unconditional variances. Engle then introduced a regression model with disturbances following the ARCH process and used the Maximum Likelihood (ML hereafter) function to estimate the variances of inflation in UK (1987).<sup>1</sup> Prior to Engle, researchers had used variance measures such as the absolute value of the first difference of inflation or a moving variance around a moving mean. These measures made simplistic assumptions about the mean of the distribution and did not allow for a stochastic variance. Engle found fourth-order linear ARCH process to be highly significant using the ML estimation (1002).

Bollerslev introduced the GARCH model in 1987 to allow for more flexible lag structure. While in the ARCH process the conditional variance is a function of past sample variances, the GARCH process allows for the lagged conditional variances as well (309). Bollerslev examined the US inflation rate for periods 1948-1983 using the GARCH model. He found that a GARCH (1, 1) model provides a better fit than an ARCH (8) model and exhibits a more reasonable lag structure (322).

Nelson (1991) introduced an extension to ARCH models that can deal with a number of limitations that arise in application of regular GARCH models. GARCH models by assumption fail to capture the asymmetric impact of negative versus positive news on stock returns which has been a common finding in empirical researches since 1970s. In fact GARCH models structurally ignore the sign of the excess return and only capture the impact of its magnitude. In addition, non-negativity restrictions that are imposed on the parameters of GARCH models, while result in difficulties in estimating the model coefficients, also distort the expected volatility behavior by precluding the oscillatory randomness in its process. Lack of capability in interpreting the persistency of shocks in future volatility is another drawback of regular GARCH models. This has been a major concern in the research about stock volatility and its correlation with stock return. The nature of the shocks (transitory or persistent) and the convergence process of shocks can have thoroughly different influence on the risk premium structure of a stock. E-GARCH (exponential GARCH) method introduced by Nelson to develop a model of risk premium on the CRSP Value-Weighted Market Index in 1962-87 (daily) tries to overcome the above limitations. Nelson findings provide high evidence for persistency of daily excess return shocks on

---

<sup>1</sup> Engle also estimates the model using OLS but finds Maximum Likelihood to be more efficient and to have smaller standard errors.

conditional volatility. His results also indicated upward impact of both good and bad news on the conditional variance although the impacts caused by negative shocks were greater.

Glosten et al (1993) approached the above drawbacks by developing a modified version of GARCH-M in which different impact of positive and negative returns on conditional volatility are incorporated in the model by using a dummy variable. Testing for seasonal patterns and the effect of risk-free rate of return were other two major developments in GJR-GARCH model. Their results brought more support for negative relation between conditional expected return and conditional volatility and they showed that this relationship becomes stronger when monthly data are used and the risk-free interest rate is incorporated in the model. On the other hand their findings on the persistency of the shocks do not agree with the results of similar research done on daily data such as Nelson (1991).

Xu (1999) conducted a research on Shanghai stock market daily data from 1992 to 1995 to compare the explanatory power of GARCH, EGARCH, and GJR-GARCH. The results show little or no sign of leverage effect which is attributed to the presence of regulatory intervention by Chinese government during that period.

In more recent researches more interest is attracted towards the volatility prediction models and many of the papers also focused on evaluating the predictability power of different generations of volatility forecasters. Brooks and Persauds (2003) did a comprehensive study on the historic development of return volatility forecasting models, which provides a valuable resource on the related literature. They also examined the power of linear GARCH type forecasting models with models that are based on multivariate approaches on a number of UK financial statistics. According to their results multivariate techniques do not have significant superiority over the standard linear GARCH models

Awartani and Corradi (2005) compared the out of sample predictive ability of different generations of GARCH models in various horizons using daily observations of S&P-500. Their results strongly advocated the predictive power of asymmetric GARCH models when compared with standard GARCH (1, 1) although the asymmetric models act far better in one-step ahead forecasting but not in longer horizon predictions.

Bluhm (2000) studied the German stock market volatility by using two different approaches to forecast market return conditional variance: 1) univariate time series techniques and 2) volatility implied option pricing models based on ARCH techniques. The results were mixed based on the error measurement methodology and the forecasting horizons. However ARCH-type models showed superiority in tests that were based on Value at Risk (VaR) objectives.



In another study by Marcucci (2005) GARCH, EGARCH and GJR-GARCH models compared with a set of Regime-Switching GARCH models in which parameters are allowed to switch between low and high volatility regimes. The evaluation is based on out-of-sample performance in forecasting volatility on S&P100 index on a one-day, one-week, two-week and one-month horizons. A feature of Marcucci work that makes it unique is attributed to the degree of flexibility he injects to the models. Apart from controlling for low and high volatility to capture non-normality characteristics, models are estimated based on both Gaussian and fat-tailed distributions such as student's t and GED. Additionally to deal with time-varying conditional kurtosis, degree of freedom in student's t distribution is allowed to change in different setting. The results were ambiguous: at short horizons consisting of one-day and one-week forecasting, Regime-Switching models demonstrated better performance but as the time horizon gets longer the standard GARCH models move to a superior position.

One of the recent studies, Lee et al. (2009), employed Skewed Generalized Error Distribution based GARCH (SGED-GARCH), a version of GARCH that first proposed by Theodossiou (2000) to examine volatility process in China stock market. This model adds more adjustments to the standard model to allow for capturing skewness and fat-tailed features of volatility distribution. The model outperformed standard GARCH (1, 1) in out of sample forecasting in different time-horizons.

### **3. Data**

The data in this analysis includes the continuously compounded monthly returns on CRSP value weighted index of equities on the NYSE and risk free rate of return on three months treasuries based on the average of bid/ask prices. The excess return is obtained by subtracting the risk free rate from the stock returns. Both databases are obtained from Wharton data services. The data ranges from January of 1970 to December of 2008, a total of 468 observations. It should be noted that all regressions and tests are performed using the Stata software with the exception of the EGARCH-M model. Stata algorithm does not converge for this model and thus EViews software was used.

Table I in the appendix provides descriptive statistics including the mean, variance, skewness, and kurtosis of the returns. Under assumptions of normality skewness and kurtosis have asymptotic distributions of  $N(0,6/T)$  and  $N(3,24/T)$  where T is the number of observations (Xu 1999, 143). Both empirical distributions of stock returns and risk free rate differ significantly from a normal distribution. The stock return is negatively skewed while the risk free return is

positively skewed indicating that neither distribution is symmetric. Also both distributions are leptokurtic compared to a normal distribution. Judging from the statistics including the chi-squared statistic from test of normality of D'Agostino et al. (1990), the risk free rate distribution tends to be more normal than the stock return. Figure 1 shows the distribution of stock returns, risk free returns, and excess stock returns.

The Ljung-Box LB (12) statistics for the test of 12<sup>th</sup> order autocorrelation of times series is also presented in Table I. The stock return statistics is 10.853 which is lower than 21.026, the 5% critical value from a chi-squared distribution with 12 degrees of freedom. On the other hand, the risk free rate statistics is 3986.838 much higher than the critical value. Therefore the stock returns exhibit little correlation while the risk free returns are highly correlated. One reason for this could be that for monthly stock returns are not as correlated daily stock returns. On the contrary, the risk free rate has much less volatility and depends on the Federal funds rate (which does not change very often). Therefore the degree of autocorrelation would be large even with monthly series.

Although the stock returns are not correlated no conclusion can be drawn about the independence of the data. The Ljung-Box LB<sup>2</sup> (12) statistics for squared return provides a test of inter-temporal dependence in variance. The statistics for stock returns is 11.387 which is lower than 26.217, the 1 % critical value from a chi-squared distribution with 12 degrees of freedom. The statistics for risk free rate is 3512.250. Once again the stock returns exhibit no dependence while the risk free rate shows great dependence across time. The distribution of next squared risk free rate of returns depends on past as well as current squared returns (Xu 1999, 144). To summarize, the data for excess returns is independent but non-normal and leptokurtic.

## 4. Volatility Models

### 4.1 Model Specifications

Let  $p_t$  indicate the stock price at time  $t$ , then the return on stock can be defined as  $r_t = \log(p_t) - \log(p_{t-1})$ . Now let's assume that when investors make decision at time  $t-1$  they are aware of the information available at time  $t-1$  which is gathered in the information set  $I_{t-1}$ . Then the realized stock return at time  $t$  may be specified as:

$$r_t = \mu_t + \varepsilon_t \tag{1}$$

Where  $\mu_t$  is expected return at time  $t$  conditional on the information available at time  $t-1$  (i.e.  $\mu_t = E(r_t | I_{t-1})$ ) and  $\varepsilon_t$  is the error term which follows a normal standard distribution.

The development of the variance models has been in line with the process of finding explanation for stylized facts that were discovered in regard with the behavior of stock return volatility. ARCH models are among the first generation of models to explain stock return volatility. The ARCH process was first proposed by Engel (1982) in response to the empirical finding that large errors today are followed by large errors in future dates. This phenomenon is known as volatility clustering and is a major stylized fact that can be seen in many financial variables' time series including stock returns. ARCH (q) models relate current volatility to the amount of volatility in the proceeding successive  $q$  periods:

$$\sigma_t^2 = b_0 + \sum_{i=1}^q g_i \varepsilon_{t-i}^2 \quad (2)$$

$\sigma_t^2$  is the conditional variance of the stock return at time  $t$ . One must assume  $b_0 > 0$ ,  $g_i \geq 0$  ( $i = 1..q$ ) to ensure the variance would always be positive. A common interpretation of equation (2) is that current volatility is the long run average volatility adjusted with the news received in the past  $q$  periods. It is reasonable to expect that more recent news has greater influence than older one and thus we expect that for  $i > j$ ,  $g_i > g_j$ .

A major problem with equation (2) from an econometrician's point of view is that for large values of  $q$  (which frequently happens in empirical researches) a large number of coefficients should be estimated subject to inequality restrictions. However equation (2) is in fact a distributed lag model for  $\sigma_k^2$  and therefore the sum of the lagged values of  $\varepsilon_i^2$  can be replaced by an equivalent lagged value of  $\sigma_k^2$ . This will lead us to the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models that were first proposed by Bollerslev (1992). The GARCH ( $p, q$ ) model can be defined as:

$$\sigma_t^2 = b_0 + \sum_{i=1}^q g_i \varepsilon_{t-i}^2 + \sum_{j=1}^p b_j \sigma_{t-j}^2 \quad (3)$$

with  $b_0 > 0$ ,  $g_i, i = 1, \dots, q$ , and  $b_j, j = 1, \dots, p$ .

GARCH-M and ARMA GARCH model are alternative models that create more reasonable results in some circumstances. GARCH in-mean or GARCH-M model is derived from solving a system of equations consisting of equation (3) and equation (4) which is created by adding an estimate of the conditional variance to equation (1):

$$r_t = a_0 + a_1 \sigma_t^2 + \varepsilon_t \quad (4)$$

ARMA GARCH models are another form of extension to GARCH process. The only difference is replacing the mean equation (equation (4)) with an ARMA process:

$$r_t = a_0 + \sum_{i=1}^p a_{2,i} r_{t-i} + \sum_{j=1}^q a_{3,i} \varepsilon_{t-j} \quad (5)$$

Standard ARCH and GARCH models have a number of limitations that result in their failure to capture all the features of real patterns of volatility process of stock returns. One of the significant stylized facts that has not been explained by early generations of ARCH and GARCH models is the asymmetric impact of good and bad news on the behavior of volatility. In fact in ARCH and GARCH models it is assumed that only the magnitude and not the sign of the unexpected return may determine the future volatility pattern. Such specification is not in line with empirical findings. This limitation is addressed in Exponential GARCH (EGARCH) models and threshold GARCH models such as GJR-GARCH that employ more flexible approach in volatility specification.

EGARCH first proposed by Nelson (1991) defines the conditional volatility process as follows:

$$\log(\sigma_t^2) = b_0 + \sum_{i=1}^q \alpha_i F(\eta_{t-i}) + \sum_{j=1}^p b_j \log(\sigma_{t-j}^2) \quad (6)$$

Where  $\eta_t = \varepsilon_t / \sigma_t$  and  $F(\eta_t) = \theta \eta_t + \delta (|\eta_t| - E|\eta_t|)$ ,  $\theta$  and  $\delta$  are coefficients, and  $\eta_t$  is a standard normal variable. In order to have a stationary process, it is required that the sum of the parameters equal to one.

Replacement of  $\varepsilon_t^2$  with the function  $F$  allows EGARCH to incorporate the feature of asymmetry. Also note that using log form for conditional variance guarantees its positivity for all choices of parameters and the restriction of non-negativity which was the case in standard GARCH is relaxed. Non-negativity constraint on the parameters not only causes problems in estimating the model parameters but also rules out the random behavior of volatility through the imposed implication that any increase in  $\varepsilon_t^2$  always leads to increase in volatility in next periods (Nelson, 1991). It should be noted that in this analysis, the EGARCH equation is simplified to the following:

$$\log(\sigma_t^2) = b_0 + \sum_{i=1}^q [\gamma_i \eta_t + \alpha_i (|\eta_t| - E|\eta_t|)] + \sum_{j=1}^p b_j \log(\sigma_{t-j}^2)$$

in order to clarify presentation of the results.

Glosten et al. (1993) extended the standard GARCH model to capture the asymmetry feature by using dummy variables. Their general model which is named GJR-GARCH specifies conditional volatility process as:

$$\sigma_t^2 = b_0 + \sum_{i=1}^q g_i \varepsilon_{t-i}^2 + \sum_{j=1}^p b_j \sigma_{t-j}^2 + \sum_{i=1}^q G_i D_{t-i} \varepsilon_{t-i}^2 \quad (7)$$

The dummy variable  $D_{t-1}$  takes a value of one if the error at  $t-1$  is positive and zero otherwise.

## 4.2 Inference and Diagnostic Tests

All inference in this analysis is based on robust or quasi-maximum likelihood estimates of standard errors. Such estimations are robust to symmetric non-normality in the residuals. The procedure used by our software package Stata is based on the full Huber/White/sandwich formulation. As opposed to the Bollerslev and Wooldridge (1992) methodology used by Glosten et al. (1993) and EViews, the Stata procedure also calculated the second derivative of the log-likelihood function (StataCorp LP 2009).

Both Stata and EViews use the Berndt, Hall, Hall, and Hauman (1974) optimization algorithm (hereafter referred to as BHHH). This algorithm was introduced to deal with non-linearity feature in optimization problems. It uses the outer product of score in place of Hessian in an iteration process defined as below to estimate the coefficient of the regression model:

$$\beta_{k+1} = \beta_k - \theta_k \left[ \frac{1}{N} \sum_{i=1}^N \frac{\partial \ln F_i}{\partial \beta}(\beta_k) \frac{\partial \ln F_i}{\partial \beta}(\beta_k)' \right]^{-1} \frac{\partial \ln F_i}{\partial \beta}(\beta_k)$$

where  $F = \sum_{i=1}^N F_i$  is the objective function (i.e. the negative of likelihood function).

This algorithm is among the most commonly used methods in MLE estimations. BHHH algorithm only requires the computation of the score and also relaxes the requirement of calculating the second derivatives for finding out the optimal value. Another major feature of BHHH process that has contributed to its popularity in econometric computations relates to its flexibility to deal with non-positive definite Hessians which is a usually major problem in ML estimations. In fact, since the sum of outer product of the scores is always positive definite or positive semi-definite there is no concern about non-positive definiteness of Hessians (Wooldridge, 2001).

Various diagnostic tests are used to assess the specifications of different aspects of our models. First, we test the residuals of models for excess skewness and kurtosis. Properly specified GARCH-M model should reduce the excess skewness and kurtosis in data. We test for excess kurtosis and skewness using the test of normality of D'Agostino et al. (1990).

Next, we test whether the squared standardized residuals  $(\varepsilon_t/\sqrt{v_{t-1}})^2$  are independent and identically distributed. We use four test proposed by Engle and Ng (1993): the Sign-Bias test, the Negative-Sign-Bias test, the Positive-Sign-Bias test, and a joint test of all three.

In the Sign-Bias test, the residuals are regressed on a constant and a dummy variable denoted  $D_t^-$  that takes a value of one if the previous error  $(\varepsilon_{t-1})$  is negative and zero otherwise. The statistics for this test is the t-statistic for the coefficient on  $I_t^-$ . This test focuses on the different

impacts that negative and positive innovations have on future volatility which are not predicted by the model (Engle and Ng 1993, 10).

For the Negative-Sign-Bias, the squared standardized residuals are regressed on a constant and  $D_t^- \varepsilon_{t-1}$ . The test statistics is the t-statistics for the coefficient on  $D_t^- \varepsilon_{t-1}$ . This test focuses on the different impacts of large and small negative innovations on future volatility which are not predicted by the model. In other words, it is a test of whether larger negative innovations are correlated with larger biases in predicted volatility.

For the Positive-Sign-Test, the squared standardized residuals are regressed on a constant and  $D_t^+ \varepsilon_{t-1}$ .  $D_t^+$  is a dummy variable that takes a value of one if previous error ( $\varepsilon_{t-1}$ ) is positive and zero otherwise. The test statistic is the t-statistic for the coefficient of  $D_t^+ \varepsilon_{t-1}$ . This test examines whether larger positive innovations are correlated with larger biases in predicted volatility (Glosten et al. 1993, 15).

The model can also be subjected to all of these test at once by regressing the squared standardized residuals on  $D_t^-$ ,  $D_t^- \varepsilon_{t-1}$ , and  $D_t^+ \varepsilon_{t-1}$  and testing that all coefficients are equal to zero. This would be  $TR^2$  (T being the number of observations) or the LM statistics for the regression (Engle and Ng 1993, 12). The LM statistic follows a chi-squared distribution with three degrees of freedom (Xu 1999, 147).

We also use Ljung-Box LB (12) and  $LB^2$  (12) to test for 12<sup>th</sup> order serial correlation of standardized residuals (autocorrelation test) and squared standardized residuals (independence test) respectively.<sup>2</sup> Although there was no autocorrelation or independence in the data, we perform these tests to assess whether the GARCH models could reduce correlation and dependence of excess returns.

Finally, we run two additional regressions. One is adopted from Pagan and Sabau (1987). The squared residuals are regressed against the conditional variance:

$$\varepsilon_t^2 = a + b\sigma_t^2 + v_t$$

The  $R^2$  from this regression is a measure of power of conditional variance to explain the squared residuals and can be used to evaluate the adequacy of the models (Xu 1999, 147). In other words, the  $R^2$  parameter computes the correlation between residual and variance. If the correlation is low, then the model's performance would be questionable. The second regression consists of

---

<sup>2</sup> These tests are used in many studies. For example, Hsieh (1989) used the tests when applying GARCH models to daily exchange rates, Engle and Ng (1993) used them when measuring the impact of news on volatility, and Kearns and Pagan (1993) applied these to their analysis of monthly Australian stock returns.

regressing the conditional variance on its past value to test the persistence in variance across models. The slope and the t-statistic of this regression are reported.

## 5. Empirical Results

The objective of this paper is to examine the role of different GARCH models in explaining volatility of excess return and determining the relationship between risk and return. We include the GARCH (2, 1) and ARMA GJR GARCH (2, 1) in our tests in order to compare the fit of GARCH-M (2, 1) and GJR GARCH-M (2, 1) models against different specifications. The standard model selection criteria introduced by Schwarz (1978) is used to determine the orders the GARCH model. This methodology provides consistent order-estimation for linear ARMA models and therefore would not necessarily be useful for GARCH-M or EGARCH models (Xu 1999, 148). Models with  $p$  and  $q$  less than three are considered. The SBC statistics for the GARCH model is presented in Table II. The SBC selects  $p = 2$  and  $q = 1$ . For lack of better indicator and for sake of comparison, other models are also tested with the same specifications.<sup>3</sup>

Table III presents the estimates for all the models. A comparison of ARMA GARCH (2, 1) and GARCH-M (2, 1) models indicates that the two models provide the same fit for the data. Previous conditional volatilities are highly significant while the arch term is significant at 95% level for the former and at 99% level for the latter. For both models, the mean variables have no significance. Based on results for the GARCH-M model, the relationship between risk and return is negative but not significant. It should be noted both positive and negative innovations result in the increase in conditional variance ( $g_1$  is positive). Also time periods with relatively large variances are associated with relatively smaller returns ( $a_1$  is negative). This result is in contrast to Glosten et al. (1993) finding.

The GJR-GARCH-M model allows the positive and negative innovations to have different effect on the conditional variance. This is done by estimating the parameter  $G_1$ . With the exception of intercept of the mean equation, none of the variables estimated by this model is significant. In this analysis, similar to results of Glosten et al., a negative innovation increases the conditional variance for the next period, while a positive innovation decreases conditional variance. However the result obtained here is of no statistical significance.

A comparison of GJR-GARCH-M and ARMA-GJR-GARCH models will reveal that once again there is not much difference between the two specifications. None of estimates in the

---

<sup>3</sup> We also tested the models using different order specifications. The conclusions of this paper remain the same under different specifications.

variance equation is significant. Also the signs and magnitude of variables across models are very similar. However, the ARMA model fits the data better with regards to the mean equation and results in highly significant coefficients. Once again, the relationship between return and variance is found to be negative but not significant. Glosten et al (1993) found the same result when testing their simple GJR model.

Finally, the EGARCH-M model seems to fit data the best. The relationship between variance and return is found to be negative but not significant. However, most of coefficients in the variance equation are at least significant at 95% level. The second lag of conditional variance is highly significant along with the arch and asymmetric term. The alpha term is positive and the gamma (asymmetry) term is negative. Given that the absolute value of gamma is greater than the alpha, a negative shock would increase variance while a positive shock of the same magnitude would decrease variance. This result is in line with that of Glosten et al. (1993).

Table IV presents the diagnostics tests performed on the models. None of standardized residuals seem to follow a normal distribution. However, all models reduce the skewness of residuals. Furthermore, GARCH, GARCH-M, and EGARCH-M models show some reduction in Kurtosis of the residuals although not significant. Based on the normality test, none of the models is able to create residuals that are normally distributed. Nonetheless the best results were obtained from GARCH and GARCH-M models.

The sign test statistics help us to determine whether we can predict the squared standardized residuals by some variables observed in previous periods which are not included in the model being used. If such variables can predict the residuals, then the model is misspecified. The conventional GARCH and GARCH-M models exhibit strong bias for all sign tests. However, there is no significant sign bias, negative bias or positive bias in the any of the other models. In addition, the joint statistics of the EGARCH-M model estimation is lower than 6.251, the 10% critical value from a chi-squared distribution with three degrees of freedom. This is the only model that passes this test.

The Ljung-Box tests statistics of standardized and squared standardized residuals are lower than 21.026, the critical value for 95% level from a chi-squared distribution with 12 degrees of freedom. The reader should be reminded that the original data did not show any autocorrelation or dependency. Nevertheless, the LB (12) and  $LB^2$  (12) statistics are reduced considerably from the values for the excess returns (refer to Table I) indicating that GARCH models do reduce the 1<sup>st</sup> and 2<sup>nd</sup> order dependence of the data.



The ARMA GJR-GARCH model has the highest  $R^2$  value (3.8%) from the Pagan and Sabau's regression. Therefore this model is superior to the rest in explaining squared returns. The EGARCH model has the lowest  $R^2$  value (1.4%). Large variances must be correlated with large squared residuals. Since the EGARCH model performs the worse, different specifications of the model would be plausible.

Finally, the estimated persistence of variance for one period is highest for GJR models. Also these coefficients are highly significance at 99% level. However, no estimation shows major persistence in volatility. The highest coefficient of past conditional variance is for the ARMA GJR-GARCH model and it is only 0.312. Regarding this matter, we have come to the same conclusion as Glosten et al. (1993).

Our results were also tested using the EViews software. The EViews estimation and diagnostic test results for the GJR GARCH-M model are provided in Table V for comparison. As it can be seen, with the exception of  $g_1$  (coefficient for the past residual), all coefficients are the nearly the same. However, the t-statistics are quite different and the EViews estimation results in more significant coefficients than that of Stata. This is due to the fact that Stata uses Huber/White/sandwich robust errors while the EViews software uses the Bollerslev and Wooldridge (1992) methodology. The results for the diagnostic tests are almost identical with negligible variations in sign bias tests.

## 6. Concluding Remarks

This paper has followed the previous literature in modeling the volatility and examining the relationship between risk and return. We have used Engle's GARCH-M model, Glosten et al GJR-GARCH-M model, and the Nelson's EGARCH in our analysis. The GJR and the EGARCH models have been two of the most successful in empirical studies. The analysis follows GJR and extends the discussion to tests of monthly rather than daily series. The results in this paper can be summarized as follows:

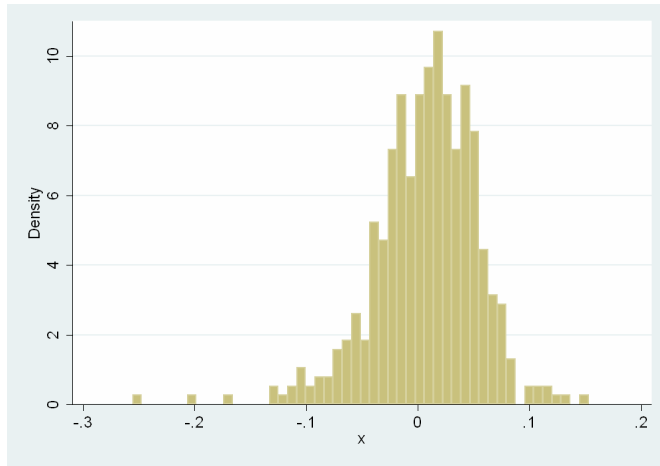
- a) The relationship between mean and conditional variance is negative but not statistically significant.
- b) The EGARCH model fits the data the best compared to GARCH and GJR models however it fails the Pagan and Sabau  $R^2$  test.
- c) Negative shocks are associated with an increase in variance and positive shocks are associated with a decrease in variance.

d) The conditional volatility of monthly excess return is not highly persistent.

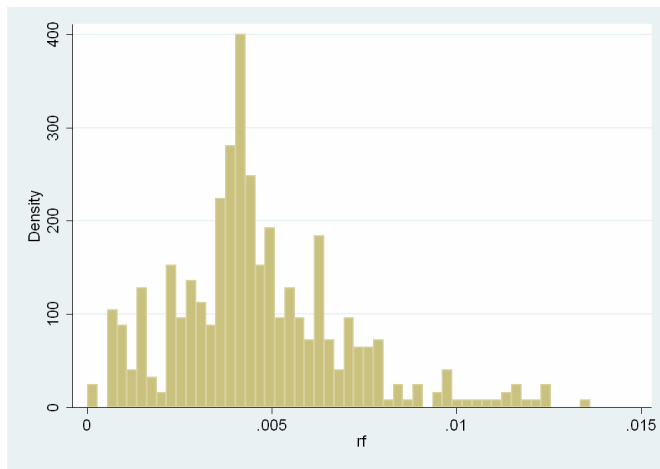
Our results differ from those of Glosten et al. (1993) in that they find a significant coefficient for their asymmetric dummy variable and thus conclude that negative and positive innovations have different impact on variance. More specifically, negative residuals increase conditional variance while positive residuals decrease conditional variance. We have come to the same conclusion by using the EGARCH-M model.

Future research could focus on using monthly data for exchanges other than that of US and extending this analysis to incorporate other GARCH-M models. In fact following GJR to incorporate risk free rate directly into the variance equation and use dummy variables to deseasonalize the residuals could result in a significant relationship between mean and variance. Also most of similar GARCH analysis done to this date has focused mostly on equity and foreign exchange markets. It would be plausible to apply more of these methodologies to other markets such as the commodity markets.

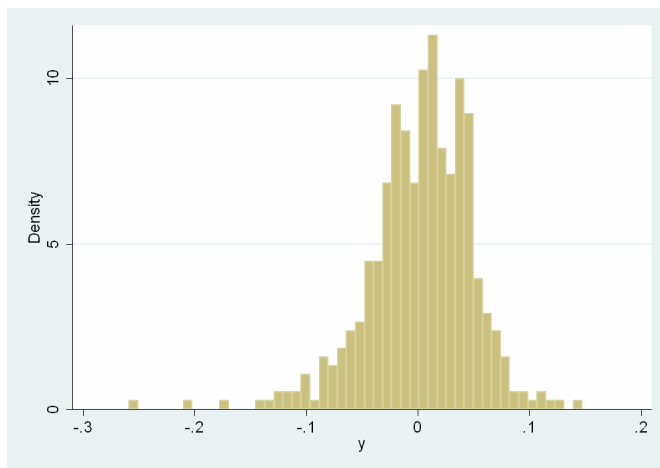
## Appendix: Figures and Tables



(a) Monthly Stock Returns



(b) Monthly Risk Free Rate of Return



(c) Monthly Excess Monthly Returns

Figure 1: Return Distributions

**Table I**  
Descriptive Statistics

Variables	Mean	Std. Dev.	Skewness	Kurtosis	LB(12)	LB <sup>2</sup> (12)	Adj. chi2 Test of normality
Stock returns	0.007	0.047	-0.888	6.205	10.853 (0.542)	11.387 (0.496)	68.34
Risk free rate of return	0.005	0.002	0.868	4.318	3986.838 (0)	3512.25 (0)	48.6
Excess stock return	0.003	0.047	-0.886	6.087	10.799 (0.546)	12.297 (0.422)	67.2

**Table II**  
Comparison of SBC statistics for the GARCH (p, q) model with different p and q

GARCH		q		
		1	2	3
p	1	-1534.3535	-1528.8372	-1518.2181
	2	<b>-1539.7982</b>	-1538.4756	-1532.251
	3	-1539.3215	-1539.1465	-1530.2427

**Table III**  
Model Estimates

With  $r$  being the excess continuously compounded monthly return on the CRSP value-weighted index of equities on the NYSE, the models are defined by:

$$r_t = a_0 + a_1\sigma_t^2 + a_2r_{t-1} + a_3\epsilon_{t-1} + \epsilon_t;$$

$$\sigma_t^2 = \text{Var}_{t-1}(\epsilon_t); \text{ conditional variance of } y_t \text{ given information set at } t-1$$

$$D_{t-1} = 1 \text{ if } \epsilon_{t-1} > 0, 0 \text{ otherwise}$$

$$\sigma_t^2 = b_0 + b_1\sigma_{t-1}^2 + b_2\sigma_{t-2}^2 + g_1\epsilon_{t-1}^2 + G_1\epsilon_{t-1}^2D_{t-1}$$

For Nelson's EGARCH-M Model:

$$H_t = b_0 + b_1H_{t-1} + b_2H_{t-2} + \alpha \left| \frac{\epsilon_{t-1}}{\sigma_{t-1}^2} \right| + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}^2}, \text{ where } H_t = \log(\sigma_t^2)$$

The t-statistics are shown in the parenthesis.

	GARCH	GARCH-M	GJR-GARCH-M	GJR-GARCH	EGARCH-M
$r$					
$a_0$	0.005 (2.010)	0.006 (1.109)	0.009 (2.805)	0.005 (2.207)	0.007 (2.559)
ARCHM					
$a_1$		-0.977 (-0.432)	-2.001 (-1.859)		-1.864 (-1.264)
ARMA					
$a_2$	-0.749 (-1.374)			-0.970 (-50.178)	
$a_3$	0.817 (1.719)			0.999 (399.019)	
ARCH					
$b_0$	0.000 (2.538)	0.000 (2.035)	0.001 (1.027)	0.001 (1.372)	-2.038 (-2.343)
$b_1$	-0.068 (-5.066)	-0.069 (-5.113)	0.065 (0.729)	0.087 (0.903)	0.168 (1.638)
$b_2$	0.859 (13.433)	0.877 (16.336)	0.397 (1.227)	0.397 (1.735)	0.537 (4.528)
$g_1$	0.119 (2.215)	0.112 (3.019)	0.509 (1.372)	0.454 (1.563)	
$G_1$			-0.575 (-1.517)	-0.515 (-1.719)	
$\gamma$					-0.371 (-5.037)
$\alpha$					0.248 (2.493)

**Table IV**  
Diagnostic Tests

Skewness and Kurtosis are the estimated kurtosis and skewness of the standardized residuals from the mean equation.

For the Sign-Bias, Negative-Sign-Bias, and Positive-Sign-Bias tests, the slope coefficient and t-statistics from the regression of the squared standardized residuals on respectively 1) a dummy variable that takes the value one if the previous residual is negative and zero otherwise, 2) the product of this dummy variable and the previous residual, and 3) The product of a dummy variable that takes the value one if the previous residual is positive and zero otherwise and the previous residual. The Joint test is the LM-statistics from the regression of squared standardized residuals on all the above three regressors.

LB(12) and LB2 (12) are the Ljung-Box test statistics for the 12th order serial correlation of standardized and squared standardized residuals respectively.

The R2 is the coefficient of determination from the regression  $\epsilon_t^2 = a + b\sigma_t^2$ .

AR(1) Coefficient is the slope coefficient in the regression of conditional variance on its value at time t-1.

The t-statistics are shown in the parenthesis.

	GARCH	GARCH-M	GJR-GARCH-M	GJR-GARCH	EGARCH-M
Skewness	-0.623	-0.677	-0.839	-0.797	-0.771
Kurtosis	4.760	4.718	6.128	5.376	5.542
Normality test	38.725	41.337	64.660	55.158	55.18
Sign-Bias	0.717 (3.983)	0.681 (3.806)	0.447 (2.107)	0.392 (1.999)	0.303 (1.520)
Negative-Sign-Bias	-6.937 (-2.389)	-6.857 (-2.404)	0.181 (0.053)	-0.054 (-0.017)	.777 (0.240)
Positive-Sign-Bias	-13.256 (-3.460)	-12.279 (-3.253)	-6.423 (-1.464)	-5.996 (-1.429)	-5.235 (-1.270)
Joint test <sup>1</sup>	16.944	15.291	7.567	6.449	4.955
LB(12)	4.521	8.415	7.141	6.337	7.159
LB <sup>2</sup> (12)	6.827	9.013	5.617	4.798	6.623
R <sup>2</sup>	0.016	0.017	0.017	0.038	0.0139
AR(1) Coefficient	-0.048 (-1.043)	-0.09 (-1.957)	0.160 (3.483)	0.312 (6.993)	.0915 ( 1.98)

[1] The chi-squared statistics for probability levels 0.1, 0.05, and 0.01 are 6.251, 7.815, and 11.345 respectively.

**Table V\***  
Comparison of Stata and EViews Outputs for GJR GARCH-M Model

Panel A: Coefficient Estimation		
	EViews	Stata
r		
a <sub>0</sub>	0.009 (2.905)	0.009 (2.805)
ARCHM		
a <sub>1</sub>	-2.000 (-1.408)	-2.001 (-1.859)
ARCH		
b <sub>0</sub>	0.001 (2.485)	0.001 (1.027)
b <sub>1</sub>	0.065 (1.337)	0.065 (0.729)
b <sub>2</sub>	0.397 (2.928)	0.397 (1.227)
g <sub>1</sub>	-0.065 (-1.984)	0.509 (1.372)
G <sub>1</sub>	-0.574 (3.005)	-0.575 (-1.517)

Panel B: Diagnostic Tests		
	EViews	Stata
Sign-Bias	.442 (2.08)	0.447 (2.107)
Negative-Sign-Bias	.178 (0.050)	0.181 (0.053)
Positive-Sign-Bias	-6.447 (-1.470)	-6.423 (-1.464)
Joint Test	7.800	7.567
LB(12)	7.142	7.141
LB <sup>2</sup> (12)	5.616	5.617
R <sup>2</sup>	0.017	0.017
AR(1) Coefficient	.160 (3.49)	0.160 (3.483)

\* Refer to Table III & IV for an explanation of coefficients and diagnostic tests



## References

- Awartani, Basel M. A., and Valentina Corradi. "Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries" *International Journal of Forecasting*, 21 (2005): 167– 183.
- Bluhm, H. H. W., & Yu, J. "Forecasting volatility: Evidence from the German stock market." Working Paper, University of Auckland (2000).
- Bollerslev, Tim. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, 31 (1986): 307-327
- Bollerslev, T., R.Y. Chou and K.F. Kroner. "ARCH modelling in finance: A review of the theory and empirical evidence." *Journal of Econometrics*, 52 (1992): 5–59.
- Breen, William, Lawrence R. Glosten, and Ravi Jagannathan. "Economic significance of predictable variations in stock index returns." *Journal of Finance*, 44 (1989): 1177-1189.
- Berndt, E., B. Hall, R. Hall, and J. Hausman, (1974), "Estimation and Inference in Nonlinear Structural Models", *Annals of Social Measurement*, Vol. 3, 653-665.
- Brooks, C., and G. Persaud. "Volatility Forecasting for Risk Management." *Journal of Forecasting*, 22 (2003): 1–22
- Chan, K.C., G.A. Karolyi, Rene M. Stulz. "Global financial markets and the risk premium on U.S. equity." *Journal of Financial Economics*, 32 (1992): 137-167.
- D'Agostino, R. B., A. J. Belanger, and R. B. D'Agostino Jr. "A suggestion for using powerful and informative tests of normality." *American Statistician*, 44 (1990) 316-321.
- Engle, Robert F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica*, vol 50, No 4 (July 1982): 987-1007
- Engle, Robert F. and Victor K. Ng. "Measuring and testing the impact of news on volatility." *Journal of Finance*, vol. 48 (1993): 1749-1778
- French, Kenneth R., G. William Schwert, and Robert F. Stambaugh. "Expected stock returns and volatility," *Journal of Financial Economics*, 19 (1987) 3-29.
- Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. "On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance*, 48 (1993): 1779–1801.
- Hsieh, David A. "Heteroscedasticity in Daily Foreign-Exchange Rates." *Journal of Business and Economic Statistics*, vol 7, no 3 (July 1989): 307-317.
- Kearns, P. and A.R. Pagan. "Australian Stock Market Volatility: 1875-1987." *The Economic Record*, vol 69, no 205 (June 1993) 163-178.
- Lee, Ming-Chih, Yen-Hsien Lee, and Hung-Chun Liu. "Forecasting China Stock Markets Volatility via GARCH Models Using Skewed GED Distribution." *Journal of Money Investment and Banking*, 7 (2009).

- Marcucci, Juri. "Regime Switching GARCH Models." *Studies in Nonlinear Dynamics & Econometrics*, Volume 9, issue 4 (2005).
- Nelson, D.B. "Conditional heteroskedasticity in asset returns: A new approach." *Econometrica*, 59 (1991): 347–370.
- Pagan, Adrian R. and Y. S. Hong. "Nonparametric estimation and the risk premium, in William Barnett, James Powell, and George Tauchen, ed.: *Nonparametric and Semiparametric Methods in Econometrics and Statistics*." *Cambridge University Press*, (1991): 51-57.
- Pagan, Adrián R. and Hernán Sabau. "On the inconsistency of the MLE in certain heteroscedastic regression models." *Estudios Económicos*, vol 6, issue 2 (1991): 159-172
- StataCorp LP. "Stata Manual – Stata Help for arch." <http://www.stata.com/help.cgi?arch>. (2009)
- Wooldridge, Jeffrey M., (2001), "Econometric Analysis of Cross Section and Panel Data", *MIT Press*, 374-375.
- Xu, Jiangang. "Modeling Shanghai Stock Market Volatility." *Annals of Operations Research*, 87 (1999): 141-152